

République Algérienne Démocratique et Populaire  
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

**UNIVERSITÉ MOHAMED KHIDER, BISKRA**

FACULTÉ des SCIENCES EXACTES et des SCIENCES de la NATURE et de la VIE

**DÉPARTEMENT DE MATHÉMATIQUES**



Mémoire présenté en vue de l'obtention du Diplôme :

**MASTER en Mathématiques**

Option : **Statistique**

Par

**ROUBA Saliha**

Titre :

# Modélisation Statistique d'une Série Chronologique

Membres du Comité d'Examen :

Pr.	BRAHIMI BRAHIM	UMKB	Président
Dr.	YAHIA DJABRANE	UMKB	Encadreur
Dr.	SOLTANE LOUIZA	UMKB	Examinatrice

Juin 2019

## DÉDICACE

*Je dédie ce humble travail à*

*ma très chère mère :*

*Fatma Hantat*

*et à mon très chère père :*

*Rouba Moberk*

*Vous m'avez comblé avec tendresse et affection tout au long de mon parcours, vous n'avez cessé de me soutenir et de m'encourager durant toutes les années de mes études, toujours été présentes à mes côtés pour me consoler quand il fallait.*

*En ce jour mémorable pour moi ainsi que pour vous, reçoit ce travail en signe de ma vive reconnaissance et mon profond estime, puisse le tout puissant te donner santé , bonheur et longue vie afin que je puisse vous combler à mon tour.*

## REMERCIEMENTS

Je tiens à remercier toutes les personnes qui ont contribué au succès de mon parcours et qui m'ont aidée lors de la rédaction de ce mémoire.

Je voudrais dans un premier temps remercier mon directeur de mémoire Dr. Yahia Djabrane, pour sa patience, sa disponibilité et surtout ses judicieux conseils qui ont contribué à alimenter ma réflexion.

Je remercie également Pr. Brahim Brahimi Président du Jury et Mme Soltane L. l'examinatrice de ce travail.

Je remercie également toute l'équipe pédagogique de notre département de mathématiques de l'université de Mouhamed Khider et les intervenants professionnels responsables de ma formation pour avoir assuré la partie théorique de celle-ci.

# Table des matières

Remerciements	ii
Table des matières	iii
Table des figures	vi
Liste des Tableaux	vii
Introduction	1
<b>1 Généralités</b>	<b>3</b>
1.1 Composantes d'une série chronologique . . . . .	4
1.1.1 Tandance . . . . .	4
1.1.2 Saisonnalité . . . . .	5
1.1.3 Résidus . . . . .	5
1.2 Opérateurs de différenciation . . . . .	5
1.3 Processus stochastique . . . . .	6
1.3.1 Processus bruit blanc . . . . .	7
1.4 Stationnarité . . . . .	7
1.5 Non stationnarité . . . . .	8
1.5.1 Non stationnarité stochastique . . . . .	8
1.5.2 Non stationnarité déterministe . . . . .	8

1.6	Autocorrélation et Autocorrélation partielle . . . . .	9
1.6.1	Fonction d'autocorrélation . . . . .	9
1.6.2	Estimation . . . . .	10
1.6.3	Fonction d'autocorrélation partielle . . . . .	11
1.7	Théorème de Wold . . . . .	11
1.8	Critères d'information . . . . .	12
1.9	Tests sur les séries temporelles . . . . .	12
1.9.1	Test de Ljung-Box . . . . .	13
1.9.2	Test de Dickey Fuller . . . . .	13
1.9.3	Test de normalité des erreurs . . . . .	14
1.9.4	Tests de Skewness et Kurtosis . . . . .	15
1.9.5	Test Global de Jarque et Berra . . . . .	16
<b>2</b>	<b>Les modèles d'une série chronologique</b>	<b>18</b>
2.1	Modèle autorégressif . . . . .	18
2.1.1	Représentation stationnaire . . . . .	19
2.1.2	Estimation des paramètres . . . . .	19
2.1.3	Autocorrélation d'un AR(p) . . . . .	20
2.2	Modèle moyenne mobile . . . . .	20
2.2.1	Autocorrélation d'un MA(q) . . . . .	21
2.3	Les modèles ARMA . . . . .	22
2.3.1	Autocorrélation d'un ARMA(p,q) . . . . .	23
2.4	Les modèles ARIMA . . . . .	23
2.5	La méthode de Box et Jenkins . . . . .	24
2.5.1	L'identification . . . . .	24
2.5.2	L'estimation . . . . .	26
2.5.3	La validation . . . . .	27
2.5.4	La prvision . . . . .	28

<b>3 Application</b>	<b>31</b>
3.1 Identification du modèle . . . . .	33
3.1.1 Description de la série . . . . .	33
3.1.2 Etude de la stationnarité de la série . . . . .	33
3.1.3 Stationnarisation de la série . . . . .	34
3.2 Estimation . . . . .	35
3.2.1 Critère d'information . . . . .	36
3.3 Validation . . . . .	36
3.3.1 Test de Student des paramètres . . . . .	36
3.3.2 Test de l'ACF pour les résidus . . . . .	38
3.3.3 Test de normalité pour les résidus . . . . .	39
3.4 Privision . . . . .	39
<b>Conclusion</b>	<b>41</b>
<b>Bibliographie</b>	<b>42</b>
<b>Annexe : Notations et abriviations</b>	<b>43</b>

# Table des figures

3.1	Représentation graphique de la série temporelle. . . . .	33
3.2	ACF et PACF de la série temporelle . . . . .	34
3.3	Représentation de la série après la différentiation. . . . .	35
3.4	ACF et PACF de du série après la différentiation . . . . .	35
3.5	Représentation de l'acf et pacf des résidus du modèle $ARMA(1,1)$ . . . . .	38
3.6	Normalité des résidus du modèle $ARMA(1,1)$ . . . . .	39

# Liste des tableaux

3.1	Série des prix de pétrole : Juillet 1988 - Juillet 2008 (Source : <a href="http://yahoo.finance">http://yahoo.finance</a> ) . . . . .	32
3.2	Critères d'information por les 4 modèles . . . . .	36
3.3	Les valeurs d'autocorrélation du modèle ARMA(1,1) de la série . . . . .	38
3.4	Privision des prix de pétrole (Out 2008 - Jul 2009) . . . . .	40

# Introduction

*L'étude des séries temporelles, ou séries chronologiques correspond à l'analyse statistique d'observations régulièrement espacées dans le temps dans le but d'expliquer et prévoir le phénomène étudié dans le futur.*

*La série chronologique dépend principalement de temps en tant que facteur indépendant déterminant et expliquer du phénomène et l'utilisation de valeurs variables dans les périodes précédentes. Une série chronologique (ou chronique, temporelle) est un ensemble d'observations numériques d'une grandeur effectuées à intervalles réguliers au cours du temps (année, trimestre, mois, jour, heure,.....).*

*On représente généralement les séries temporelles sur des graphiques de valeurs (ordonnées) en fonction du temps (abscisses). Lorsqu'une série est stable autour de sa moyenne, on parle de série stationnaire. Inversement, on trouve aussi des séries non stationnaires.*

*L'analyse des séries temporelles est un domaine de la statistique, de l'économétrie, d'astronomie....etc. Cette dernière a connue un grand développement depuis la parution du livre de Box-Jenkins (1970) où les principales propriétés des processus stationnaires autorégressif moyenne mobile (ARMA) ont été décrites avec les méthodes d'identification, d'estimation et de validation.*

*Dans ce mémoire, nous essayons de donner une synthèse générale sur les séries temporelles, leurs propriétés, et quelques exemples d'applications, dont nous proposons une*

*modélisation de la série des prix de pétrole entre Juillet 1988 et Juillet 2013, en utilisant la méthode dite de Box-Jenkins.*

*Ce travail est composé en trois chapitres comme suit :*

- Le premier chapitre est un expose sur les préliminaires et quelques notations, définitions et tests sur les séries chronologiques.*
- Le deuxième chapitre est partagé en deux parties, dans la première partie nous représentons différentes modèles : AR, MA, ARMA, ARIMA, ... en étudiant leurs propriétés (stationnarité, fonction d'autocorrélation simple et partielle,.....). Dans la deuxième partie on présente la méthode d'analyse et de prévision des séries chronologiques de Box-Jenkins.*
- Finalement, dans le troisième chapitre, nous appliquons la méthode de Box-Jenkins sur des données réelles des prix de pétrole.*

# Chapitre 1

## Généralités

Dans ce chapitre on donne certains notations et définitions dont on aura besoin dans la suite du chapitre et de mémoire.

*La série chronologique dépend principalement de temps entant que facteur indépendant déterminant et expliquer du phénomène et l'utilisation de valeurs variables dans les périodes précédentes. Elle sera notée prochainement par :*

$$X_t := (X_t)_t = (X_t(\omega))_t, \quad X \text{ est une v.a. et } t \in \mathbb{R}.$$

Il y a nombreux domaines d'application des séries chronologiques :

- Démographie : analyse de l'évolution d'une population.
- Energie : prévision de la consommation d'électricité.
- Finance et économétrie : évolution des indices boursiers, des productions agricoles ou industrielles.
- Médecine : analyse d'electroencephalogrammes d'electrocardiogrammes.
- Science de la terre et de l'espace : évolution des taches solaires, phénomènes d'avalanches.
- Métiorologie : analyse de données climatique.
- Traitement de signal : signaux de communications de radars.

- Assurance : analyse des sinistres.
- Agriculture : la quantité d'un produit.
- Traitement des données : mesures successives de position ou de direction d'un objet mobile (trajectographie).

## 1.1 Composantes d'une série chronologique

Une série  $X_t$  est composée de trois parties comme suit :

### 1.1.1 Tandance

Il s'agit d'un terme de la série qui traduit l'évolution à moyen terme du phénomène. elle sera estimée sous forme paramétrique (linéaire, polynomiale, logarithmique, exponentielle, etc) :

- a. Tandance lineaire :** la tendance la plus simple est linéaire on peut estimer les paramètres au moyen de la méthode des moindres carrés c'est une régression simple :

$$T_t = a + bt$$

- b. Tandance polynomiale d'ordre q :** on peut ajuster la série par un polynôme d'ordre  $q$ . les paramètres peuvent être estimés au moyen de la méthode de moindres carrés c'est une régression avec  $q$  variables explicatives :

$$T_t = b_0 + b_1t + b_2t^2 + \dots + b_qt^q$$

- c. Tandance logistique :** la fonction logistique permet de modéliser des processus ne pouvant dépasser une certaine valeur  $c$  :

$$T = \frac{c}{1 + be^{-at}}$$

### 1.1.2 Saisonnalité

Elle représente des effets périodiques de période connue  $p$  qui se reproduisent de façon plus ou moins identique d'une période à l'autre elle est notée par  $S_t$   $t = 1, \dots, T$ . Elle est généralement supposée périodique :  $S_{t+p} = S_t$  d'une période  $p$ .

### 1.1.3 Résidus

Les résidus notée  $\varepsilon_t$  est la partie non structurée du phénomène. Elle est modélisée par une suite de variables aléatoires  $\varepsilon_t$ , centrées, non corrélées et de même variance, on parle donc de Bruit Blanc.

## 1.2 Opérateurs de différenciation

**Définition 1.2.1** On appelle l'opérateur retard l'opérateur noté  $L$  qui à tout processus  $(X_t)_{t \in \mathbb{Z}}$  associe le processus  $(Y_t)_{t \in \mathbb{Z}}$  défini par

$$\forall t \in \mathbb{Z} \quad Y_t = LX_t = X_{t-1}.$$

L'opérateur  $L$  est inversible et linéaire son inverse est  $L^{-1} = D$  défini par

$$\forall t \in \mathbb{Z} \quad DX_t = X_{t+1}$$

avec, l'opérateur  $D$  est appelé **opérateur avance**.

**Définition 1.2.2** L'opérateur  $\Delta_s$  est défini par :

$$\Delta_s X_t = X_t - X_{t-s}$$

$s$  : la période de saisonnalité.

**Définition 1.2.3** De même, l'opérateur de différenciation est l'opérateur  $\nabla$  qui à tout processus  $X_t$ , on associe le processus  $Y_t$  défini par :

$$\forall t \in \mathbb{Z} \quad \nabla X_t = Y_t = (1 - L)X_t = X_t - X_{t-1}$$

**Proposition 1.2.1**

1.  $L^0 X_t = X_t$
2.  $L \circ L \dots \circ L = L^n$  et  $L^n X_t = X_{t-n}$
3. l'opérateur retard est distributif :  $L(aX_t + bY_t) = aX_t + bY_t$
4.  $\nabla^n X_t = (1 - L)^n X_t$

On peut définir un polynôme retard tout polynôme qui s'écrit en fonction de l'opérateur retard. On donne un exemple, soit le processus  $X_t$  défini par :

$$X_t = \theta_1 X_{t-1} + \theta_2 X_{t-2} + \varepsilon_t$$

ce processus peut s'écrire comme suit :

$$X_t = \theta_1 L X_t + \theta_2 L^2 X_t + \varepsilon_t = (\theta_1 L + \theta_2 L^2) X_t + \varepsilon_t$$

d'où

$$(1 - \theta_1 L - \theta_2 L^2) X_t = \varepsilon_t \implies \phi(L) X_t = \varepsilon_t,$$

avec  $\phi(L) = (1 - \theta_1 L - \theta_2 L^2)$  est le polynôme retard associé au processus  $X_t$ .

## 1.3 Processus stochastique

Considérons un espace de probabilité  $(\Omega, A, P)$  et un espace probabilisé  $(\Omega', A', P)$  et un ensemble d'indices  $\Gamma$ .

**Définition 1.3.1** *Un processus stochastique est une famille de variables aléatoires  $(X_t)$  définies sur  $(\Omega, A, P)$  indexées par  $t \in \Gamma \{X_t, t \in \Gamma\}$  à valeur dans  $A$ . Un processus est dite causal s'il existe une suite  $\{a_k\}$  réelle telle que  $\sum_{k=0}^{\infty} |a_k| < \infty$  et que*

$$X_t = \sum_{k=0}^{\infty} a_k \varepsilon_{t-k}.$$

**Définition 1.3.2** *Un processus est dite inversible s'il existe une suite  $\{b_k\}$  réelle telle que  $\sum_{k=0}^{\infty} |b_k| < \infty$  et que*

$$\varepsilon_t = \sum_{k=0}^{\infty} b_k X_{t-k}.$$

*Une autre façon de dire qu'un processus est inversible est de dire qu'il possède une représentation  $AR(\infty)$ .*

### 1.3.1 Processus bruit blanc

**Définition 1.3.3** *Le processus  $\varepsilon_t$  est un bruit blanc si :*

- i)  $E(\varepsilon_t) = 0$  pour tout  $t$ .*
- ii)  $var(\varepsilon_t) = E(\varepsilon_t^2) = \delta^2$  pour tout  $t$ .*
- iii)  $cov(\varepsilon_t, \varepsilon_{t-h}) = E(\varepsilon_t \varepsilon_{t-h}) = 0$  pour tout  $t$  et  $h \neq 0$ .*

## 1.4 Stationnarité

Une des grandes questions dans l'étude de séries temporelles est de savoir si celles-ci suivent un processus stationnaire. Une série chronologique de réalisations d'une grandeur aléatoire, à un pas de temps donné, est dite stationnaire si ses réalisations sont issues d'un même processus stochastique dont les paramètres (moyenne, variance, autocorrélation...) restent constants au cours du temps.

**Définition 1.4.1** *On dit que le processus  $(X_t)$  est strictement stationnaire si la probabilité*

ne dépend pas de l'instant  $t$  :

$$\forall h, \forall n, \quad P(X_1, \dots, X_n : t_1, \dots, t_n) = P(X_1, \dots, X_n : t_1 + h, \dots, t_n + h).$$

**Définition 1.4.2** On dit que le processus  $(X_t)$  est faiblement stationnaire si ses propriétés statistiques ne dépend pas de  $t$ .

i)  $E(X_t) = u$  independ de  $t$ .

ii)  $cov(X_t, X_{t+h})$  independante de  $t$ .

## 1.5 Non stationnarité

Un processus dont l'espérance et la variance ne se stabilisent pas simultanément au cours du temps est qualifié de **non stationnaire**. On peut distinguer deux phénomènes à l'origine de la non stationnarité d'un processus, selon sa nature déterministe ou stochastique.

### 1.5.1 Non stationnarité stochastique

On dit que le processus  $Y_t$  est caractérisé par une non stationnarité stochastique, ou encore que le processus  $Y_t$  est *DS* si le processus différencié une fois  $(1 - L)Y_t$  est stationnaire.

On parle aussi de processus intégré d'ordre 1 :

$$(1 - L)Y_t = Z_t \text{ est stationnaire ce qui implique } Y_t = Y_{t-1} + Z_t.$$

De manière générale, on dit que le processus  $Y_t$  est un processus intégré d'ordre  $d$  si le processus différencié  $d$  fois  $(1 - L)^d Y_t$  est stationnaire.

### 1.5.2 Non stationnarité déterministe

On dit que le processus  $Y_t$  est caractérisé par une non stationnarité déterministe, ou encore que le processus  $Y_t$  est *TS* s'il peut s'écrire

$$Y_t = f(t) + Z_t$$

$f(t)$  : une fonction qui dépend du temps.

$Z_t$  : un processus stationnaire.

Ainsi, ce processus est rendu stationnaire en lui enlevant sa tendance déterministe :  $Y_t -$

$f(t) = Z_t$  est stationnaire.

## 1.6 Autocorrélation et Autocorrélation partielle

### 1.6.1 Fonction d'autocorrélation

Pour une série stationnaire  $(X_t)$ , on définit la fonction d'autocovariance pour tout  $t$  par :

$$\gamma_X(h) = cov(X_t, X_{t-h}) = E(X_t X_{t-h}) - E(X_t)E(X_{t-h})$$

**Remarque 1.6.1** *La fonction d'autocovariance est paire :*

$$\forall h \in \mathbb{Z} \quad \gamma_X(h) = \gamma_X(-h) \quad \text{avec} \quad \gamma_X(0) = var(X_t).$$

**Définition 1.6.1** *La fonction d'autocorrélation d'un processus  $(X_t, t \in \mathbb{Z})$  notée  $\rho(h)$  ou  $\rho_h$  de moyenne  $E(X_t) = u$ , est définie par*

$$\forall h \in \mathbb{Z} : \rho(h) = \rho_h = \frac{\gamma(h)}{\gamma(0)}$$

avec  $\rho(0) = 1$  et  $|\rho(h)| < 1$  (donc  $|\rho(h)| \leq \rho(0)$ ).

On appelle coefficient d'autocorrélation d'ordre 1 (resp. d'ordre  $k$ ) le coefficient de corrélation linéaire  $\rho(1)$  (resp.  $\rho(h)$ ) calculé entre la série  $X_t$  et cette série décalée d'une période 1 (resp.  $k$  période). On définit alors, la matrice de corrélation (de dimension  $n$ ) de la manière

suivante :

$$R(n) = \begin{pmatrix} 1 & \rho(1) & \rho(2) & \dots & \rho(n-1) \\ \rho(1) & 1 & \rho(1) & \dots & \rho(n-2) \\ & & \cdot & & \\ & & \cdot & & \\ & & \cdot & & \\ \rho(n-1) & \rho(n-2) & \dots & \rho(1) & 1 \end{pmatrix}$$

puisque la fonction  $\rho(h)$  est de type positif, on a la propriété suivante :

$$\det(R(n)) \geq 0, \quad \forall n \in \mathbb{N}^*$$

Ainsi, on a les contraintes suivante :

$$\det(R(1)) \geq 0, \quad \det(R(2)) \geq 0 \iff \rho(1)^2 \leq 1.$$

**Remarque 1.6.2** Les fonctions  $\varphi(h)$  et  $\gamma(h)$  sont symétriques :  $\forall h \in \mathbb{Z} : \gamma(h) = \gamma(-h)$  et  $\varphi(h) = \varphi(-h)$ .

## 1.6.2 Estimation

L'estimateur de la fonction d'autocorrélation, noté  $\hat{\rho}(h)$  ou  $\hat{\rho}_h$  obtenu pour un échantillon de  $T$  réalisations du processus  $(X_t, t \in \mathbb{Z})$  est donné par  $\forall h \in \mathbb{Z} :$

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

où  $\hat{\gamma}(h)$  désigne l'estimateur de la fonction d'autocovariance :

$$\hat{\gamma}(h) = \frac{1}{T-h} \sum_{t=1}^{T-h} (X_t - \bar{X}_t)(X_{t-h} - \bar{X}_{t-h})$$

avec  $\bar{X}_{t-k} = \frac{1}{T-h} \sum_{t=1}^{T-h} X_t$ .

### 1.6.3 Fonction d'autocorrélation partielle

**Définition 1.6.2** *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. Les coefficient d'autocorrélation partielle d'ordre  $h$  d'un processus stationnaire se calcul de manière suivante :*

$$r(h) = \frac{|R(h)^*|}{|R(h)|}$$

et  $R(h)^*$  est la matrice  $R(h)$  dans laquelle on a remplacé la colonne  $h$  par

$$\begin{pmatrix} \rho_1 \\ \rho_2 \\ \cdot \\ \cdot \\ \rho_h \end{pmatrix}$$

## 1.7 Théorème de Wold

Le théorème de Wold (1948) est fondamental pour l'analyse de série temporelles stationnaire. Il est donné par

**Théorème 1.7.1** *Tout processus  $(X_t)_{t \in \mathbb{Z}}$  centré et stationnaire peut s'écrire sous la forme :*

$$X_t = \sum_{j=0}^{\infty} \theta_j \varepsilon_{t-j} + l_t$$

$\theta_j$  : les paramètres sont des réels tel que  $\theta_0 = 1$  et  $\sum_{j=0}^{\infty} \theta_j^2 < \infty$

$\varepsilon_t$  : Bruit Blanc

$l_t$  : un processus déterministe  $cov(\varepsilon_t, l_t) = 0 \quad \forall s, t \in \mathbb{Z}$ . O aura dans le cas réel ( $l_t = 0$ )  $X_t = \psi(l)\varepsilon_t$ .

## 1.8 Critères d'information

Deux critères importantes dans les applications des séries temporelles, sont

**Définition 1.8.1** *AIC (Akaike Information Critèrion) :*

$$AIC = -2\ln V + 2k$$

$k$  : le nombre de paramètres.

$2k$  : représente la pénalité.

$V$  : est la vraisemblance.

Le modèle à retenir est celui qui montre l' $AIC$  le plus faible, l' $AIC$  utilise le principe du maximum de vraisemblance. Il pénalise les modèles comportant trop de variables, et évite le sur-apprentissage. Hurvich et Tsai, préconisent d'utiliser l' $AIC$  corrigé lorsque le nombre de paramètre  $k$  est grand par rapport au nombre d'observations

**Définition 1.8.2** *BIC (Bayesian Information Criterion) :*

$$BIC = -2\ln V + k\ln(n)$$

$k$  : le nombre de paramètres libres du modèle.

$n$  : le nombre de données.

$\ln(n)$  : le terme de pénalité.

Le  $BIC$  utilise le principe du maximum vraisemblance. Il pénalise les modèles comportant trop de variables, et évite le sur-apprentissage.

## 1.9 Tests sur les séries temporelles

Dans cette dernière partie, nous donnons quelques tests utiles dans l'analyse des séries temporelles :

### 1.9.1 Test de Ljung-Box

Si on cherche à tester les hypothèses :

$$\begin{cases} H_0 : \rho_1 = \rho_2 = \dots = \rho_r = 0 \text{ (pas d'autocorrélation des erreurs d'ordre 1)} \\ H_1 : \text{l'un au moins des } \rho_i \neq 0 \text{ (il y a autocorrélation des erreurs d'ordre entre 1 et } r \text{)}. \end{cases}$$

Pour l'effectuer on récupère les résidus  $\varepsilon_t$  du modèle de base et on construit

$$\varepsilon_t = \rho_1 \varepsilon_{t-1} + \rho_2 \varepsilon_{t-2} + \dots + \rho_r \varepsilon_{t-r} + u_t$$

Ljung-Box ont montré que sous l'hypothèse  $H_0$  la variable

$$Q = n(n+2) \sum_{i=1}^{i=r} \frac{\rho_i^2}{n-i}$$

suit sous l'hypothèse  $H_0$  une loi du  $\chi^2$  à  $r$  degrés de liberté. Par la suite on mettra  $Q$  cette statistique car elle est beaucoup plus utilisée que la statistique de Box-Pierce. Ce test est le plus utilisé des tests d'autocorrélation d'ordre supérieur à 1.

### 1.9.2 Test de Dickey Fuller

#### Test de Dickey Fuller simple

Dickey et Fuller sont les premiers à fournir un ensemble d'outils statistiques formels pour détecter la présence d'une racine unitaire dans un processus autorégressif du premier ordre, ce test permet de tester l'hypothèse :

$$\begin{cases} H_0 : \text{Le modèle est de racine unitaire} \\ H_1 : \text{Le modèle n'est pas de racine unitaire} \end{cases}$$

Ce test est regroupée en 4 cas :

$$\begin{array}{ll}
 y_t = py_{t-1} + \varepsilon_t & H_0 : p = 1 \\
 y_t = \alpha + py_{t-1} + \varepsilon_t & H_0 : \alpha = 0 \text{ et } p = 1 \\
 y_t = \alpha + py_{t-1} + \varepsilon_t & H_0 : \alpha \neq 0 \text{ et } p = 1 \\
 y_t = \alpha + \beta t + py_{t-1} + \varepsilon_t & H_0 : \alpha = 0 \text{ et } \beta = 0 \text{ et } p = 1.
 \end{array}$$

### Test de Dickey Fuller augmenté

Cette procédure de test à des processus autorégressifs d'ordre  $p$ , il s'agit alors des tests *ADF*. Ce test permet de tester

$$\left\{ \begin{array}{l}
 H_0 : \text{Le modèle est de racine unitaire} \\
 H_1 : \text{Le modèle n'est pas de racine unitaire}
 \end{array} \right.$$

Ce test peuvent être regroupés en 4 cas :

$$\begin{array}{ll}
 y_t = py_{t-1} + \sum_{i=1}^p \alpha_i y_{t-i} + \varepsilon_t & H_0 : p = 1 \\
 y_t = \alpha + py_{t-1} + \sum_{i=1}^p \alpha_i y_{t-i} + \varepsilon_t & H_0 : \alpha = 0 \text{ et } p = 1 \\
 y_t = \alpha + py_{t-1} + \sum_{i=1}^p \alpha_i y_{t-i} + \varepsilon_t & H_0 : \alpha \neq 0 \text{ et } p = 1 \\
 y_t = \alpha + \beta t + py_{t-1} + \sum_{i=1}^p \alpha_i y_{t-i} + \varepsilon_t & H_0 : \alpha = 0 \text{ et } \beta = 0 \text{ et } p = 1
 \end{array}$$

Pour simplifier, on écrira :

$$\begin{array}{ll}
 \Delta y_t = \varphi y_{t-1} + \sum_{i=1}^p \alpha_i \Delta y_{t-i} + \varepsilon_t & \text{avec } \varphi = p - 1 \text{ "Modèle [1]"} \\
 \Delta y_t = \alpha + \varphi y_{t-1} + \sum_{i=1}^p \alpha_i \Delta y_{t-i} + \varepsilon_t & \text{"Modèle [2]"} \\
 \Delta y_t = \alpha + \beta t + \varphi y_{t-1} + \sum_{i=1}^p \alpha_i \Delta y_{t-i} + \varepsilon_t & \text{"Modèle [3]"}
 \end{array}$$

### 1.9.3 Test de normalité des erreurs

L'hypothèses de test sont :

$$\left\{ \begin{array}{l}
 H_0 : \text{les erreurs suivent une loi Normale} \\
 H_1 : \text{les erreurs ne suivent pas une loi Normale}
 \end{array} \right.$$

Sous l'hypothèse  $H_1$  la loi des erreurs est donc inconnue. On caractérise la loi normale

$N(m, \delta^2)$  par le fait :

- \_ qu'elle est symétrique  $\implies$  le moment centré d'ordre 3 est nul :  $\mu_3 = 0$
- \_ que le moment centré d'ordre 4 est  $\mu_4 = 3\mu_2^2 = 3(\delta^2)^2 \implies$  sa Kurtosis  $K = \mu_4/\delta^4 = 3$

Les hypothèses du test peuvent alors s'écrire

$$\left\{ \begin{array}{l} H_0 : \mu_3 = 0 \text{ et } \mu_4 = 3\delta^4 \\ H_1 : \text{l'une au moins de ces deux propriétés n'est pas vérifiée} \end{array} \right.$$

Le moment  $\delta^2$  est estimé à l'aide des résidus par  $s^2 = \sum_t \varepsilon_t^2 / (n - k)$

Le moment  $\mu_3 = E[(x - E(x))^3] = E[x^3]$  est estimé à l'aide des résidus par  $\mu_3 = \sum \varepsilon_t^3 / n$

Le moment  $\mu_4 = E[(x - E(x))^4] = E[x^4]$  est estimé à l'aide des résidus par  $\mu_4 = \sum \varepsilon_t^4 / n$

### 1.9.4 Tests de Skewness et Kurtosis

#### Test de Skewness

On définit le coefficient de symétrie par

$$\alpha_3 = \frac{\mu_3}{\delta^3}$$

estimé par

$$\alpha_3 = \frac{\sum \varepsilon_t^3 / n}{s^3}$$

qui suit asymptotiquement une loi  $N(0, 3!/n)$  sous l'hypothèse  $H_0$

$$\left\{ \begin{array}{l} H_{01} : \mu_3 = 0 \implies \alpha_3 = 0 \\ H_{11} : \mu_3 \neq 0 \implies \alpha_3 \neq 0 \end{array} \right.$$

sous l'hypothèse  $H_0$  l'estimateur  $\hat{\alpha}_3$  de  $\alpha_3$  suit asymptotiquement une loi  $N(0, 3!/n)$ .  $t = \sqrt{n/6}(\hat{\alpha}_3)$  suit asymptotiquement une  $N(0, 1)$ .

Si la décision est  $H_{11}$  le teste est terminé car la loi n'étant pas symétrique elle ne peut être normale. Si la décision est  $H_{01}$  on passe au test suivant.

### Test de Kurtosis

On définit le coefficient d'aplatissement (Kurtosis) :

$$\alpha_4 = \frac{\mu_4}{\delta^3}$$

estimé par

$$\hat{\alpha}_4 = \frac{\sum e_t^4/n}{s^4}$$

qui suit asymptotiquement une loi  $N(3, 4!/n)$  sous l'hypothèse  $H_0$

$$\left\{ \begin{array}{l} H_{02} : \mu_4 = 3\delta^4 \implies \alpha_4 = 3 \\ H_{12} : \mu_4 \neq 3\delta^4 \implies \alpha_4 \neq 3 \end{array} \right.$$

Sous l'hypothèse  $H_0$  l'estimateur  $\hat{\alpha}_4$  de  $\alpha_4$  suit asymptotiquement une loi  $N(3, 4!/n)$ , sa variable centrée et réduite  $t = \sqrt{n/24}(\hat{\alpha}_4 - 3)$  suit asymptotiquement une  $N(0, 1)$ .

Si la décision est  $H_{02}$  on a donc les deux propriétés vérifiées, on décide alors normalité des erreurs. Si on décide  $H_{12}$  le test est déterminé car la loi n'étant pas un coefficient d'aplatissement égal à 3 nous ne sommes pas dans le cadre de la loi normale.

### 1.9.5 Test Global de Jarque et Berra

Il teste globalement les hypothèses :

$$\left\{ \begin{array}{l} H_0 : \mu_3 = 0 \text{ et } \mu_4 = 3\delta^4 \\ H_1 : \text{l'une au moins de ces deux propriétés n'est pas vérifiée} \end{array} \right.$$

Sous l'hypothèse  $H_0$  vraie la variable aléatoire  $S$  somme des carés des deux précédents résultats centrés réduits suit donc le loi du  $\chi^2$  à deux degrés de liberté :

$$S = \frac{n}{6}\hat{\alpha}_3^2 + \frac{n}{24}(\hat{\alpha}_4 - 3)^2.$$

Si  $S$  est inférieur à la borne du  $\chi_2^2$  on décide  $H_0$  sinon on décide  $H_1$ .

# Chapitre 2

## Les modèles d'une série chronologique

### 2.1 Modèle autorégressif

**Définition 2.1.1** *On dira que le processus  $X_t$  est un modèle autorégressif d'ordre  $p$ , s'il existe un Bruit Blanc  $\{\varepsilon_t\}$ , et s'il peut s'écrire sous la forme suivant :*

$$X_t = \sum_{j=1}^p \theta_j X_{t-j} + \varepsilon_t$$

$\theta_j : (j = 1, \dots, p)$  les termes du modèle :

$$X_t - \sum_{j=1}^p \theta_j X_{t-j} = \varepsilon_t \iff (1 - \theta_1 L - \dots - \theta_p L^p) X_t = \varepsilon_t \iff \phi(L) X_t = \varepsilon_t$$

### 2.1.1 Représentation stationnaire

Si le polynôme  $\phi$  a toutes ses racines de module différent de 1, on peut inverser l'opérateur  $\phi(L)$ . On en déduit que l'équation admet une solution unique, avec une écriture  $MA(\infty)$  :

$$X_t = \phi(L)^{-1}\varepsilon_t = \sum_{j=-\infty}^{\infty} f_j \varepsilon_{t-j}$$

On peut alors montrer que l'on a  $\sum_{j=-\infty}^{\infty} |f_j| < \infty$  et donc la représentation est stationnaire.

### 2.1.2 Estimation des paramètres

On dispose d'une observation  $\{x_0, \dots, x_T\}$  d'un processus stationnaire  $X_t$  suivant un modèle  $AR(p)$ , c'est à dire

$$X_t = \sum_{j=1}^p \theta_j X_{t-j} + \varepsilon_t$$

$t \in \mathbb{Z}$ ,  $X_0 = x_0$  et  $\theta_1, \dots, \theta_p$  des paramètres inconnus.

Pour estimer ces paramètres il y a beaucoup de méthodes :

- Maximum vraisemblance : l'estimation d'un modèle  $AR(p)$  par la méthode du maximum de vraisemblance est délicate car la fonction de vraisemblance est très complexe et n'est pas dérivable analytiquement. Cette difficulté provient de l'interdépendance des valeurs, ainsi que du fait que les observations antérieures ne sont pas toutes disponibles pour les premières valeurs.

- Méthode des moments : utilisation des équations de Yule-Walker.

### 2.1.3 Autocorrélation d'un AR(p)

Sous l'hypothèse de stationnarité, on a  $E(X_t) = 0$ , il est clair que la fonction d'autocovariance d'un modèle  $AR(p)$

$$\begin{aligned}\gamma_x(h) &= \text{cov}(X_t, X_{t-h}) = E(X_t X_{t-h}) \\ E(X_t X_{t-h}) &= \theta_1 E(X_{t-1} X_{t-h}) + \dots + \theta_p E(X_{t-p} X_{t-h}) + E(\varepsilon_t X_{t-h})\end{aligned}$$

$$E(\varepsilon_t X_{t-h}) = \begin{cases} \delta_\varepsilon^2 & \text{si } h = 0 \\ 0 & \text{sinon} \end{cases}$$

$$\gamma_x(0) = \sum_{j=1}^p \theta_j \gamma_x(j) + \delta_\varepsilon^2 \quad \text{et} \quad \gamma_x(h) = \sum_{j=1}^p \theta_j \gamma_x(h-j) \quad , h \geq 1 \quad (2.1)$$

On divise (2.1) par  $\gamma_x(0)$ , nous obtenons la fonction d'autocorrélation :

$$\varphi_x(h) = \frac{\gamma_x(h)}{\gamma_x(0)} = \frac{\sum_{j=1}^p \theta_j \gamma_x(h-j)}{\gamma_x(0)} = \sum_{j=1}^p \theta_j \frac{\gamma_x(h-j)}{\gamma_x(0)} = \sum_{j=1}^p \theta_j \varphi_x(h-j) \quad , h \geq 1 \quad (2.2)$$

(2.1) et (2.2) s'appelle le système de **Youle-Walker** :

$$\begin{cases} \gamma_x(h) = \sum_{j=1}^p \theta_j \gamma_x(h-j) \\ \varphi_x(h) = \sum_{j=1}^p \theta_j \varphi_x(h-j) \end{cases}$$

## 2.2 Modèle moyenne mobile

**Définition 2.2.1** On appelle modèle moyenne mobile d'ordre  $q$ , un modèle stationnaire  $X_t$  sous forme :

$$X_t = c + \varepsilon_t + \sum_{j=1}^q \theta_j \varepsilon_{t-j} \quad , t \in \mathbb{Z}$$

$\theta_j$  : réels non nul et  $\varepsilon_t$  est  $BB(0, \delta^2)$ .

**Remarque 2.2.1** Un  $MA(q)$  est toujours stationnaire, de moyenne  $c$ .

### 2.2.1 Autocorrélation d'un MA(q)

La fonction d'autocovariance d'un  $MA(q)$  est :

$$\begin{aligned} \gamma_x(h) &= cov(X_t, X_{t-h}) \\ &= E((\varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2} + \dots + \theta_q\varepsilon_{t-q})(\varepsilon_{t-h} + \theta_1\varepsilon_{t-h-1} + \dots + \theta_q\varepsilon_{t-h-q})) \\ &= \begin{cases} (1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2)\delta_\varepsilon^2 & h = 0 \\ (-\theta_k + \theta_1\theta_{k+1} + \dots + \theta_{q-k}\theta_q)\delta_\varepsilon^2 & 0 < h \leq q \\ 0 & h < q \end{cases} \end{aligned}$$

d'où la fonction d'autocorrélation est

$$\rho_h = \frac{\gamma_h}{\gamma_0} = \begin{cases} 1 & h = 0 \\ \frac{-\theta_h + \theta_1\theta_{h+1} + \dots + \theta_{q-h}\theta_q}{1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2} & 0 < h \leq q \\ 0 & h < q \end{cases}$$

**Remarque 2.2.2** La fonction d'autocorrélation d'un modèle  $MA(q)$  s'annule au delà du rang ( $h < q$ ).

**Remarque 2.2.3** Un modèle  $X_t$  est un  $MA(\infty)$  s'il peut s'écrire sous la forme  $X_t = \sum_{j=-\infty}^{+\infty} \theta_j \varepsilon_{t-j}$ . On dit alors que  $X_t$  est un modèle linéaire, si de plus  $\theta_j = 0, \forall j < 0$  on dit que  $X_t$  est causal.

**Propriété 2.2.1** Pour un  $MA(\infty)$ , on a les propriétés suivantes :

1.  $E(X_t) = 0, \quad \forall t$
2.  $var(X_t) = \delta_\varepsilon^2 \sum_{j=-\infty}^{+\infty} \theta_j^2$
3.  $cov(X_t, X_{t-h}) = \gamma_x(h) = \delta_\varepsilon^2 \sum_{j=-\infty}^{+\infty} \theta_j \theta_{j+|h|}$

**Exemple 2.2.1**  $MA(1)$  :  $X_t = \varepsilon_t - \theta\varepsilon_{t-1} = (1 - \theta L)\varepsilon_t$  avec  $\varepsilon_t$  est  $BB(0, \delta^2)$

La racine de  $1 - \theta z = 0$  est  $z = 1/\theta$ .

— Si  $|\theta| = 1$  alors la forme  $AR(\infty)$  n'existe pas, la représentation est stationnaire, causale mais non inversible, donc non canonique.

— Si  $|Z| < 1$  et donc  $|\theta| < 1$  alors :

$$\varepsilon_t = (1 - \theta L)^{-1} X_t = \sum_{i=0}^{\infty} \theta^i L^i X_t = X_t + \sum_{i=1}^{\infty} \theta^i X_{t-i}$$

et la forme  $AR(\infty)$  s'écrit :

$$Y_t = - \sum_{i=1}^{\infty} \theta^i X_{t-i} + \varepsilon_t$$

La représentation est alors inversible, en plus d'être stationnaire et causale, elle est donc canonique.

— Si  $|z| < 1$  et donc  $|\theta| < 1$  alors :

$$\varepsilon_t = (1 - \theta L)^{-1} X_t = - \sum_{i=0}^{\infty} \theta^{-i} L^{-i} X_t = - \left( \frac{X_{t+1}}{\theta} + \frac{X_{t+2}}{\theta^2} + \dots \right)$$

Dans ce cas, on peut toujours (tant que  $|\theta| \neq 1$ ) se ramener à une représentation canonique quitte à changer la représentation et surtout à changer de  $BB$  et à inverser les racines.

## 2.3 Les modèles ARMA

Les processus *ARMA* généralisent simultanément les modèles *AR* purs et *MA* purs.

**Définition 2.3.1** *Un processus stationnaire  $X_t$  admet une représentation  $ARMA(p, q)$  minimale :*

$$X_t - \sum_{j=1}^p \theta_j X_{t-j} = \varepsilon_t - \sum_{i=1}^q \alpha_i \varepsilon_{t-i} \iff \phi(L)X_t = \Theta(L)\varepsilon_t$$

s'il satisfait les conditions suivantes :

1.  $\theta_j \neq 0, \alpha_i \neq 0$ .
2. les polynômes  $\phi$  et  $\Theta$  toutes leur racines de module strictement supérieur à 1.
3.  $\phi$  et  $\Theta$  n'ont de racine commune.
4.  $\varepsilon_t$  et un bruit blanc, de variance  $\delta_\varepsilon^2 \neq 0$ .

### 2.3.1 Autocorrélation d'un ARMA(p,q)

**Proposition 2.3.1** Soit  $(X_t)$  un processus ARMA(p,q), alors les autocorrélation  $\gamma(h)$  statisfont

$$\gamma(h) - \sum_{i=1}^p \phi_i \gamma(h-i) = \delta^2 [\theta_h + h_1 \theta_{h+1} + \dots + h_{q-h} \theta_q] \quad \text{pour } 0 \leq h \leq q$$

**Propriété 2.3.1** Soit  $(X_t)$  un processus ARMA(p,q), alors les autocorrélation  $\gamma(h)$  satisfont

$$\gamma(h) - \sum_{i=1}^p \theta_i \gamma(h-i) = 0 \quad \text{pour } h \geq q+1$$

où les  $h_i$  correspondent aux coefficients de la forme MA( $\infty$ ) de  $(X_t)$  :

$$X_t = \sum_{j=0}^{\infty} h_j \varepsilon_{t-j}.$$

**Preuve.** La forme ARMA(p,q) de  $(X_t)$  est

$$X_t - \sum_{j=1}^p \theta_j X_{t-j} = \varepsilon_t - \sum_{i=1}^q \alpha_i \varepsilon_{t-i}$$

En multipliant par  $X_{t-h}$ , où  $h \geq q+1$ , et en prenant l'espérance, on obtient le resultat. ■

## 2.4 Les modèles ARIMA

**Définition 2.4.1** Pour  $d \geq 1$ , le processus  $X_t$  est un ARIMA(p,d,q) si le processus  $Y_t = (1-L)^d X_t$  est un processus ARMA(p,q) de moyenne nulle.  $X_t$  satisfait donc une

équation de la forme :

$$\Phi(L)(1 - L)^d X_t = \Theta(L)\varepsilon_t.$$

$\Phi$  et  $\Theta$  : des polynômes de degrés respectifs  $p$  et  $q$

$L$  : l'opérateur de retard

$$\Phi(z) = 1 - \phi_1 z - \dots - \phi_p z^p$$

$$\Theta(z) = 1 + \theta_1 z + \dots + \theta_q z^q$$

**Remarque 2.4.1** *On dit aussi que le processus  $ARIMA(p, d, q)$  est un processus  $ARMA(p, q)$  de moyenne nulle "intégré"  $d$  fois.*

## 2.5 La méthode de Box et Jenkins

Dans cette dernière section nous présentons une méthode très utilisée dans la modélisation des séries chronologiques. Il s'agit de la méthode ou l'algorithme de Box et Jenkins (1970), dont les étapes sont les suivantes :

### 2.5.1 L'identification

L'identification consiste à spécifier les trois paramètres  $p, d, q$  du modèle  $ARIMA(p, d, q)$ . La stationnarité du modèle est d'abord testée. Étude graphique, de corrélogrammes et test de Dickey-Fuller augmenté sont tour à tour effectués. Si la série n'est pas stationnaire, il convient de la transformer (en général par différenciation) pour obtenir une série stationnaire. L'ordre d'intégration  $d$  est obtenu par le nombre de fois que la série initiale a été différenciée pour obtenir la stationnarité. Test de Dickey-Fuller augmenté, analyses la fonction d'autocorrélation et la fonction d'autocorrélation partielle pour déterminer les paramètres  $p, q$  du modèle.

La fonction d'autocorrélation est constituée par l'ensemble des d'autocorrélations  $\rho_k = cov(X_t, X_{t-k}), k$  dans  $\{1, \dots, K\}$ ,  $K$  étant le décalage maximum admissible pour que le

coefficient d'autocorrélation ait un sens. Le coefficient d'autocorrélation d'ordre  $k$ ,  $\rho_k$ , peut être estimé par :

$$r_k = \frac{\sum_{t=k+1}^n (X_t - \bar{X}_1) (X_{t-k} - \bar{X}_2)}{\sqrt{\sum_{t=k+1}^n (X_t - \bar{X}_1)^2 \sum_{t=k+1}^n (X_{t-k} - \bar{X}_2)^2}}$$

avec

$$\bar{X}_1 = \frac{1}{n-k} \sum_{t=k+1}^n X_t, \quad \bar{X}_2 = \frac{1}{n-k} \sum_{t=k+1}^n X_{t-k}$$

Sous l'hypothèse  $H_0(\rho_k = 0)$ , la statistique  $t_c = \frac{|r_k|}{\sqrt{1-r_k^2}}$  suit une loi de student à  $(n-2)$  degrés de liberté. Si la valeur calculée de  $t_c$  est supérieure au quantile d'ordre  $\alpha/2$  d'une loi de student à  $n-2$  degrés de liberté  $t_c < t_{1-\alpha/2}(n-2)$ , alors l'hypothèse  $H_0$  est rejetée au seuil  $\alpha$  (test bilatéral).

La fonction d'autocorrélation partielle désigne l'ensemble des autocorrélations entre les variables entre  $X_t$  et  $X_{t-k}$ , l'influence de la variable  $X_{t-k-i}$  étant contrôlée pour  $i < k$ .

Plusieurs logiciels de traitement des séries temporelle (Eviews, SPSS, R, STATA....) construisent des corrélogrammes qui affichent des intervalles de confiance à 95% permettant de déterminer les coefficients statistiquement significatifs à prendre en compte. Certaines règles sont en général retenues pour interpréter les corrélogrammes :

Les valeurs d'une fonction d'autocorrélation d'un processus autorégressif  $AR(p)$ , décroissent exponentiellement avec des alternances possibles de valeurs positives et négatives. La fonction d'autocorrélation partielle, quant à elle, présente  $p$  pics aux  $p$  premières valeurs du corrélogramme d'autocorrélation partielle. Le corrélogramme d'une fonction d'autocorrélation d'une moyenne mobile,  $MA(q)$ , présente  $q$  pics aux  $q$  première valeurs. La fonction d'autocorrélation partielle d'une  $MA(q)$  présente des valeurs exponentiellement décroissantes. Si la fonction d'autocorrélation décroît lentement, il faut différencier la série avant l'identification du modèle.

### 2.5.2 L'estimation

La procédure *ARIMA* du modèle R, SPSS, EVIEUS.... permet selon un algorithme rapide d'estimation du maximum de vraisemblance d'estimer les coefficients du modèle que vous avez identifié au préalable en fournissant les paramètres  $p, q$  et  $d$ . L'exécution de la procédure ajoute de nouvelles séries chronologiques représentant les valeurs ajustées ou prédites par le modèle, les résidus données courant. Ces séries pourront être utilisées dans une nouvelle itération de type identification-estimation-diagnostic.

L'estimation de paramètre se réalise par différentes méthodes :

#### Estimation de Yule-Walker

Dans le cas d'un  $AR(p)$ , on utilise les équations de Yule-Walker :

$$\begin{bmatrix} \rho(1) \\ \cdot \\ \cdot \\ \cdot \\ \rho(p) \end{bmatrix} = \begin{bmatrix} \rho(0) & \rho(0) & \dots & \dots & \rho(p-1) \\ \rho(1) & \rho(0) & \dots & \dots & \rho(p-2) \\ \cdot & \cdot & & & \cdot \\ \cdot & \cdot & & & \cdot \\ \rho(p-1) & \rho(p-1) & & & \rho(p-2) \end{bmatrix} \begin{bmatrix} \phi_1 \\ \cdot \\ \cdot \\ \cdot \\ \phi_p \end{bmatrix}$$

pour déterminer  $\hat{\phi}_i, i = 1, \dots, p$  en fonction de  $\hat{\rho}(i)$  estimés.

#### Maximum de Vraisemblance

Une méthode populaire pour estimer les paramètres d'un modèle est le maximum vraisemblance. La fonction de vraisemblance associée à un échantion  $X_1, \dots, X_N$  i.i.d d'un loi dont la densité est  $f(x, \theta)$  avec  $\theta = (\theta_1, \dots, \theta_k) \in \mathbb{R}^+$ , est défini par :

$$L(\theta) = \prod_{i=1}^N f(x, \theta)$$

L'estimateur du maximum de vraisemblance est la valeur  $\hat{\theta}$  qui maximise  $L(\theta)$  par rapport à chacun des paramètres  $\theta = (\theta_1, \dots, \theta_k)$  et de résoudre le système à  $k$  équations :

$$\frac{\partial L(\theta)}{\partial \theta_i} = 0 \quad , i = 1 \dots k$$

$$L(x, \phi, \theta, \delta^2) = \frac{1}{(2\pi\delta^2)^{N/2}} \frac{1}{\sqrt{|\Omega(\theta, \phi)|}} \exp \left\{ \frac{1}{2\delta^2} X^N \Omega^{-1}(\theta, \phi) X \right\}$$

$L(x, \phi, \theta, \delta^2)$  : la fonction de vraisemblance.

$\delta^2$  : la variance.

$\Omega(\theta, \phi)$  : la matrice de variance covariance de  $X$ .

Le log -vraisemblance est :

$$LnL(x, \phi, \theta, \delta^2) = -\frac{T}{2} \ln(2\pi) - \frac{T}{2} \ln \delta^2 - \frac{T}{2} \ln(\det(\Omega(\theta, \phi))) - \frac{1}{2\delta^2} X \Omega(\theta, \phi)^{-1} X$$

### 2.5.3 La validation

Quand on finit d'estimer les paramètres du modèle, on examine les résultats de cette estimation. En premier lieu, les coefficients du modèle doivent être significativement différents de 0. Si un coefficient n'est pas significativement différent de 0, il convient d'envisager une nouvelle spécification éliminant l'ordre du modèle *AR* ou *MA* non valide.

En deuxième lieu, coefficient de détermination doit être traité de la même manière que les autres modèles ordinaires. La statistique de student représente le ration du coefficient à estimer sur son écart type.

On troisième lieu, l'analyse des résidus s'effectue à partir de deux critères à respecter :

- Dans le cas où la moyenne n'est pas nulle il convient d'ajouter une constante au modèle.
- Les résidus constituons un bruit blanc, les statistiques  $Q$  et  $Q^*$  de Box-Pierce et de Ljung-Box permet de tester cette hypothèse. Si le résidu n'est pas un Bruit Blanc, le modèle est incomplète et il manque au moins un ordre à un processus.

### 2.5.4 La prevision

La prevision est la dernière étape de méthode de Box-Jenkins. Etant donnée une série stationnaire  $(X_t)$ , observée entre 1 et  $T$ , on cherche à faire de la prevision à horizon  $h$  et donc à prévoir  $X_{T+1}, \dots, X_{T+h}$ . Tous les processus  $AR$ ,  $MA$  et  $ARMA$  seront supposés sous forme canonique, et n'avoir aucun racine unité. Aussi, toutes les racines des polynôme autorégressifs  $\Phi$  et des polynômes moyenne-mobiles  $\Theta$  auront leurs racines à l'extérieur du disque unité. Ainsi, pour tous les processus  $X_t$  tels que  $\Phi(L)X_t = \Theta(L)\varepsilon_t$ ,  $\varepsilon_t$  sera l'innovation du processus  $X_t$ .

#### Prevision d'un modèle $AR(p)$

Le modèle s'écrit :

$$X_t = \theta_1 X_{t-1} + \dots + \theta_p X_{t-p} + \varepsilon_t \quad \text{ou} \quad \Phi(L)X_t = \varepsilon_t$$

La prevision optimale à la date  $T + 1$ , faite à la date  $T$  est

$$X_{T+1}^* = E(X_{T+1}/X_T, X_{T-1}, \dots) = X_{T+1}^* = \theta_1 X_T + \dots + \theta_p X_{T-p}$$

car  $(\varepsilon_t)$  est l'innovation. De façon analogue

$$X_{t+h} = \theta_1 X_{t+h-1} + \dots + \theta_p X_{t+h-p} + \varepsilon_{t+h}$$

Donc,  $X_{T+h}^* = E(X_{T+h}/X_T, X_{T-1}, \dots)$ , de façon réursive par

$$X_{T+h}^* = \begin{cases} \theta_1 X_{t+h-1} + \dots + \theta_{h-1} X_{t+1} + \theta_h X_T \dots + \theta_p X_{t+h-p} & \text{pour } h \leq p \\ \theta_1 X_{t+h-1} + \dots + \theta_p X_{t+h-p} & \text{pour } h < p \end{cases}$$

### Prévisions d'un modèle MA(q)

Soit le modèle  $MA(q)$  :

$$X_t = \varepsilon_t + \phi_1\varepsilon_{t-1} + \dots + \phi_q\varepsilon_{t-q} = \Theta(L)\varepsilon_t$$

La prévision optimale à la date  $T + 1$  est

$$\hat{X}_{t+1} = \phi_1\varepsilon_t + \dots + \phi_q\varepsilon_{t+1-q}$$

Donc

$$\hat{X}_{T+h} = \begin{cases} \phi_h\varepsilon_T + \dots + \phi_q\varepsilon_{T+h-q} & \text{pour } h \leq q \\ 0 & \text{pour } h < q \end{cases}$$

Toutefois, cette méthode présente le désavantage d'estimer  $X_{T+h}$  à partir des résidus passés, priori non observables, et non pas du passé de la variable.

### Prévision d'un modèle ARMA(p,q)

On supposera là aussi que l'on s'est ramené à un processus centré ( $X_t$ ), satisfaisant

$$\Phi(L)X_t = \Theta(L)\varepsilon_t$$

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

et donc

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

Alors,

$$\hat{X}_{t+h} = \begin{cases} \theta_1 \hat{X}_{T+h-1} + \dots + \theta_{h-1} \hat{X}_{T+1} + \theta_{1h} X_T + \dots + \theta_p \hat{X}_{T+h-p} & \text{pour } h \leq p \\ \theta_1 \hat{X}_{T+h-1} + \dots + \theta_p \hat{X}_{T+h-p} & \text{pour } h < p \end{cases}$$

### Prévision d'un modèle ARIMA

Le processus  $(X_t)$  satisfaisant une équation de la forme  $\Phi(L)(1-L)^d X_t = \Theta(L)\varepsilon_t$  avec les conditions initiales  $Z = (X_{-1}, \dots, X_{-p-d}, \varepsilon_{-1}, \dots, \varepsilon_{-q})$ . Posons alors  $\Psi(L) = \Phi(L)(1-L)^d$ .

La forme *ARIMA*( $p, d, q$ ) peut s'écrire

$$X_t = \sum_{i=1}^{p+d} \psi_i X_{t-i} + \varepsilon_t + \sum_{j=1}^q \phi_j \varepsilon_{t-j}$$

et donc

$$X_{t+h} = \sum_{i=1}^{p+d} \psi_i X_{t+h-i} + \varepsilon_{t+h} + \sum_{j=1}^q \phi_j \varepsilon_{t+h-j}$$

Notons  $\hat{X}_{T+h}$  la prévision faite à la date  $T$

$$\hat{X}_{T+h} = \sum_{i=1}^{p+d} \psi_i \hat{X}_{t+h-i} + 0 + \sum_{j=1}^q \phi_j \hat{\varepsilon}_{t+h-j}$$

où

$$\begin{cases} \hat{X}_{t+h-i} = X_{t+h-i} & \text{pour } i \geq h \\ \hat{\varepsilon}_{t+h-j} = \begin{cases} 0 & \text{pour } j < h \\ \varepsilon_{t+h-j} & \text{pour } j \geq h \end{cases} \end{cases}$$

En particulier, pour  $h \geq q$ , on obtient une relation de récurrence de la forme

$$\hat{X}_{T+h} = \sum_{i=1}^{p+d} \psi_i \hat{X}_{t+h-i}.$$

# Chapitre 3

## Application

Il existe de nombreux façons de construire et de former des modèles, ainsi que des tests permettant de vérifier sa validité ils diffèrent en termes d'utilisation et de caractéristique. En général ils sont basées sur la stationnairité de la série chronologique, le plus important de ces méthodes est methode de Box-Jenkins. La série chronologique dépend principalement du temps en tant que facteur indépendant déterminent et expliquer du phénomène et l'utilisation de valeurs variables dans les période précédentes. Dans ce chapitre nous appliquons pratiquement la methode de Box-Jenkins sur une série des données réelles des prix de pétrole (en \$) en Europe de Juillet 1988 à Juillet 2013 (Source : Cryer, Chan, 2008), dont le bute est de choisir le modèle approprié et faire la privision.

La série est donnée dans le tableau suivant :

	<i>Jan</i>	<i>Fév</i>	<i>Mar</i>	<i>Avr</i>	<i>Mai</i>	<i>Jui</i>	<i>Jul</i>	<i>Aou</i>	<i>Sép</i>	<i>Oct</i>	<i>Nov</i>	<i>Déc</i>
1988							14.91	14.89	13.18	12.41	13.02	15.31
1989	17.17	16.89	18.7	20.32	18.63	17.67	17.62	16.77	17.77	18.91	18.73	19.84
1990	21.25	19.81	18.39	16.61	16.35	15.1	17.17	27.17	34.9	36.02	33.07	28.27
1991	23.57	19.54	19.08	19.18	19.19	18.17	19.4	19.77	20.5	22.21	21.11	18.41
1992	18.16	18.05	17.63	18.92	19.89	21.16	20.24	19.74	20.27	20.26	19.21	18.14
1993	17.39	18.47	18.79	18.67	18.51	17.65	16.78	16.7	16.01	16.61	15.2	13.73
1994	14.29	13.8	13.82	15.23	16.19	16.76	17.6	16.89	15.9	16.49	17.19	15.93
1995	16.55	17.11	17.01	18.65	18.35	17.31	15.85	16.1	16.7	16.11	16.86	17.93
1996	17.85	18	19.85	20.9	19.15	18.46	19.57	20.51	22.63	24.16	22.76	23.78
1997	23.54	20.85	19.13	17.56	19.02	17.58	18.46	18.6	18.46	19.87	19.17	17.18
1998	15.19	14.07	13.1	13.53	14.36	21.21	12.08	11.91	13.34	12.7	11.04	9.82
1999	11.11	10.27	21.51	15.29	15.23	15.86	19.08	20.22	22.54	20	24.58	25.47
2000	25.51	27.78	27.49	22.76	27.74	29.8	28.68	30.2	33.14	30.96	32.55	25.66
2001	25.62	27.5	24.5	25.66	28.31	27.85	24.61	25.68	25.62	20.54	18.8	18.71
2002	19.42	20.28	23.7	25.73	25.35	24.08	25.74	26.56	28.4	27.54	24.34	28.33
2003	31.18	32.77	30.61	25	25.86	27.65	28.35	29.89	27.11	29.61	28.75	29.81
2004	31.28	30.86	33.63	33.59	37.57	35.18	38.22	42.74	43.2	49.78	43.11	39.6
2005	44.51	45.48	53.1	51.88	48.65	54.35	57.52	63.98	62.91	58.54	55.24	56.86
2006	62.99	60.21	62.06	70.26	69.78	68.56	73.67	73.23	61.96	57.81	58.76	62.47
2007	53.68	57.56	62.05	67.49	67.21	71.05	76.93	70.76	77.17	82.34	92.41	90.93
2008	92.18	94.99	103.64	109.07	122.8	132.32	132.32					

TAB. 3.1 – Série des prix de pétrole : Juillet 1988 - Juillet 2008 (Source : [http:// yahoo.finance](http://yahoo.finance))

## 3.1 Identification du modèle

### 3.1.1 Description de la série

La série représente les prix de pétrole (241 observations) de Juillet 1988 à Juillet 2008 de min 9.82\$ est et de max 132\$ et les données de cette série sont représentées dans la courbe suivante :

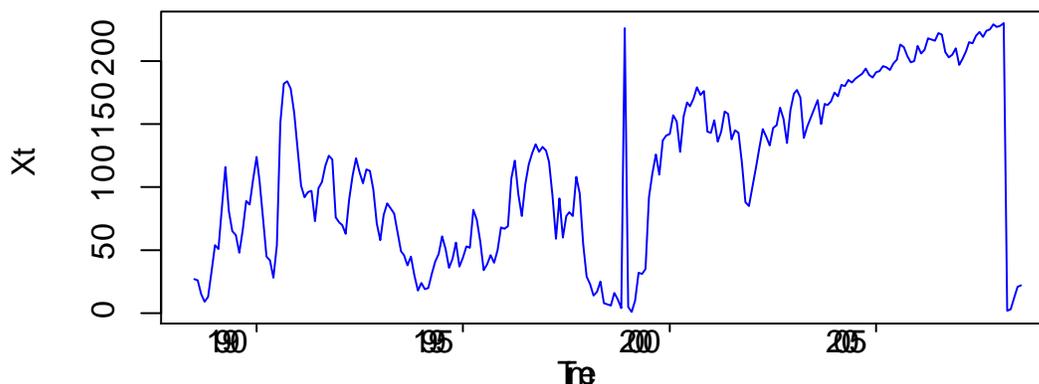


FIG. 3.1 – Représentation graphique de la série temporelle.

A travers cette figure (Fig. 3.1), nous observons que la série n'est pas stationnaire et elle est de modèle mixte. Pour affirmer cette affirmation, nous allons procéder à l'analyse des autocorrélations.

### 3.1.2 Etude de la stationnarité de la série

Pour étudier et analyser une série il faut étudier la forme de la série et connaître les informations statistiques pour choisir le bon modèle approprié. Les données précédentes : graphiques et statistiques ne peut pas donner la réalité de la série et sa caractéristiques, pour ce la on doit faire des tests. On dessine la fonction d'autocorrelation simple et la fonction d'autocorrelation partielle.

Nous remarquons (Fig. 3.2) que les autocorrélations simples sont presque toutes significativement différentes de zero et diminuent lentement. La première, sixième et septième

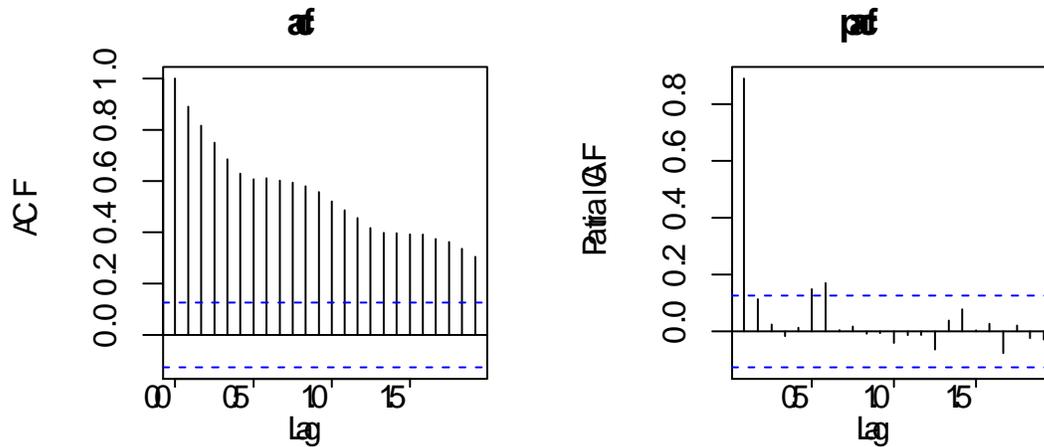


FIG. 3.2 – ACF et PACF de la série temporelle

autocorrelation partielle ( $k = 1, k = 6, k = 7$ ) est significativement différente de zéro au niveau de signification 5%. Ce qui indique que la série n'est pas stationnaire.

### 3.1.3 Stationnarisation de la série

Comme nous l'avons vu, notre série étudiée est un processus non stationnaire. pour stationnariser notre série, nous allons utiliser les différences premières :

$$Y_t = \Delta X_t = X_t - X_{t-1}$$

On obtient cette figure du modèle différencié :

De la figure (Fig. 3.3), on remarque que la courbe de la série est autour de zéro, ce qui indique que la série est stationnaire.

Nous allons maintenant étudiée les fonction d'autocorrélations simples et partielles de notre série.

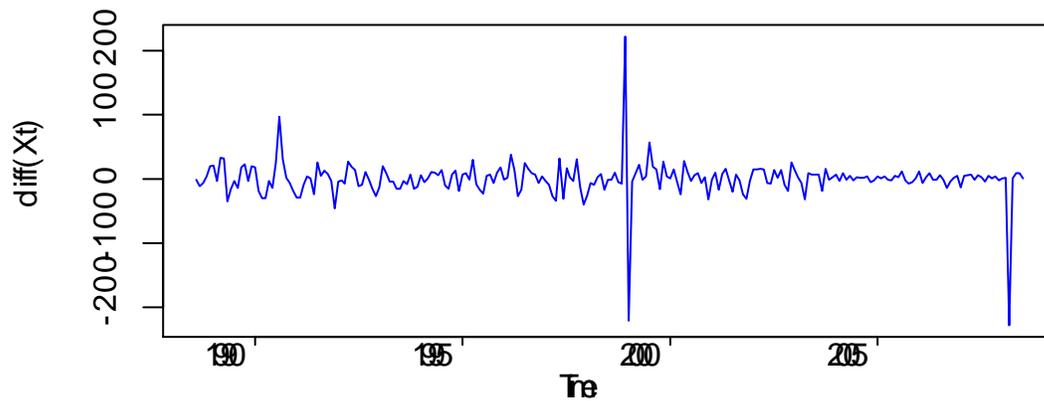


FIG. 3.3 – Représentation de la série après la différentiation.

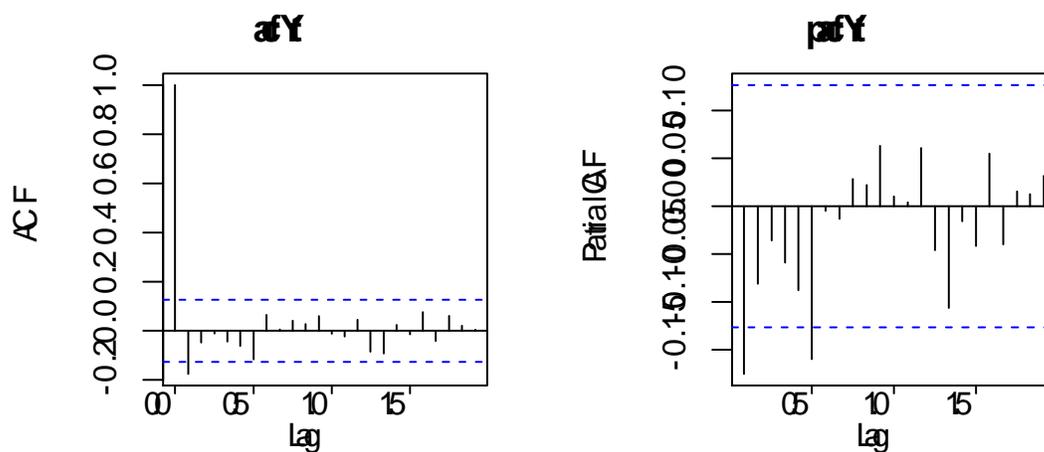


FIG. 3.4 – ACF et PACF de du série après la différentiation

On remarquons (Fig. 3.4) que les autocorrélations simple sont significativement différentes de zéro pour  $K < 2$  et pour les autocorrélations partielles elles sont significativement différentes de zéro pour  $K < 2$ .

## 3.2 Estimation

On estime les paramètre du modèle par la méthode de Yull-Walker sur le progamme  $R$ , on obtient :

$$ARMA(1, 1) : X_t = 0.93X_{t-1} + \varepsilon_t - 0.1567\varepsilon_{t-1} + 105.1268$$

$$ARMA(1, 2) : X_t = 0.9358X_{t-1} + \varepsilon_t - 0.1592\varepsilon_{t-1} + -0.0324\varepsilon_{t-2} + 104.1617$$

$$ARMA(2, 1) : X_t = -1.4882X_{t-1} + -0.5087X_{t-2} + \varepsilon_t - 0.6984\varepsilon_{t-1} + 101.6821$$

$$ARMA(2, 2) : X_t = 0.0138X_{t-1} + 0.8581X_{t-2} + \varepsilon_t + 0.7697\varepsilon_{t-1} - 0.1759\varepsilon_{t-1} + 104.5356$$

### 3.2.1 Critère d'information

Modèle	AIC	Sigma <sup>2</sup>	BIC
<i>ARMA</i> (1, 1)	2313.37	829.50	2327.31
<i>ARMA</i> (1, 2)	2315.19	828.90	2332.61
<i>ARMA</i> (2, 1)	2314.61	826.80	2332.04
<i>ARMA</i> (2, 2)	2315.89	824.10	2336.80

TAB. 3.2 – Critères d'information par les 4 modèles

D'après ces critères, les valides modèles sont *ARMA*(1, 1), *ARMA*(1, 2) et *ARMA*(2, 1) car les critères d'informations de ces modèles sont minimales.

## 3.3 Validation

### 3.3.1 Test de Student des paramètres

Pour vérifier que le paramètre du modèle qui a été estimé est statistiquement différent de 0, on fait ce test

les hypothèses du test sont :

$$\left\{ \begin{array}{l} H_0 : \theta = 0, \quad \text{le coefficient est non significatif} \\ H_1 : \theta \neq 0, \quad \text{le coefficient est significatif} \end{array} \right.$$

Pour  $\alpha = 0.05$ , les paramètres du modèle sont statistiquement différent de zéro

$$\text{On accepte } H_0(|t_i| < |T|) \quad / \quad t_i = \left| \frac{\hat{\theta}}{\hat{\sigma}_{\hat{\theta}}} \right|$$

1) Le modèle *ARMA*(1, 1) est écrit sous la forme :

$$X_t = 0.93X_{t-1} + \varepsilon_t - 0.1567\varepsilon_{t-1} + 105.1268$$

$$T_{N-p-q}^{\alpha} = T_{239}^{0.05} = 1.651254$$

$$t_1 = 35.09$$

$$t_2 = 2.15$$

$$(i = 1, 2) \quad |t_i| < |T|$$

Donc les paramètres sont statiquement significatifs.

2) Le modèle  $ARMA(1, 2)$  est écrit sous la forme :

$$X_t = 0.9358X_{t-1} + \varepsilon_t - 0.1592\varepsilon_{t-1} + -0.0324\varepsilon_{t-2} + 104.1617$$

$$T_{238}^{0.05} = 1.651281$$

$$t_1 = 33.067$$

$$t_2 = 2.23$$

$$t_3 = 0.42$$

On a  $t_3 < T$  alors  $\theta_3$  n'est pas significatives

3) Le modèle  $ARMA(2, 1)$  est écrit sous la forme :

$$X_t = -1.4882X_{t-1} + -0.5087X_{t-2} + \varepsilon_t - 0.6984\varepsilon_{t-1} + 101.6821$$

$$T_{238}^{0.05} = 1.651281$$

$$t_1 = 5.78$$

$$t_2 = 2.13$$

$$t_3 = 3.044$$

Les paramètres sont statiquement significatifs.

En conclusion, d'après les critères et les tests de Studente le modèle valide est  $ARMA(1, 1)$  :

$$X_t = 0.93X_{t-1} + \varepsilon_t - 0.1567\varepsilon_{t-1} + 105.1268$$

### 3.3.2 Test de l'ACF pour les résidus

La figure ci-dessus (Fig. 3.5) montre que la série de résidus est stable, car les coefficients d'autocorrélation se situent principalement dans l'intervalle de confiance, ce qui signifie qu'il existe une totale indépendance entre les erreurs ce que confirment le test de Box Pièce :

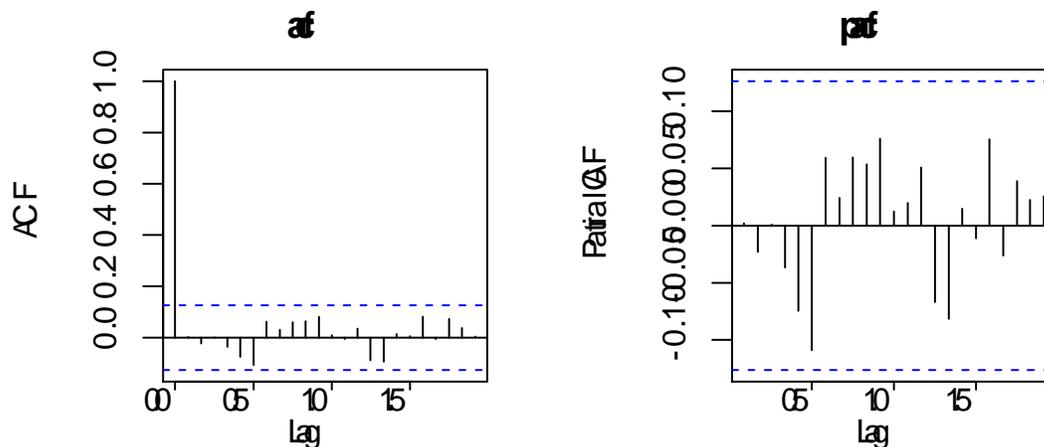


FIG. 3.5 – Représentation de l'acf et pacf des résidus du modèle  $ARMA(1, 1)$

Test de Box Pièce :

<i>Lag</i>	<i>Acp</i>	<i>Lag</i>	<i>Acp</i>
1	0.083	11	0.002
2	0.167	12	-0.023
3	0.250	13	0.001
4	0.333	14	-0.036
5	0.417	15	-0.075
6	0.500	16	-0.107
7	0.583	17	0.063
8	0.667	18	0.030
9	0.750	19	0.061
10	0.833	20	0.064

TAB. 3.3 – Les valeurs d'autocorrélation du modèle  $ARMA(1,1)$  de la série

La statistique de Box Pièce est

$$Q = T \sum_{k=1}^{20} \rho_k^2 = 241 \sum_{k=1}^{20} \rho_k^2 = 1.0141$$

On a  $Q < \chi_{0.05}^2(20) = 10,85$ , donc on accepte l'hypothèse  $H_0$  et que les paramètres d'autocorrélation simple et d'autocorrélation partielle sont corrélés.

### 3.3.3 Test de normalité pour les résidus

Le but de ce test est de s'assurer que la distribution des résidus a une distribution normale, ce qui peut être déterminé en traçant les résidus en utilisant la probabilité de normalité représentée par la forme suivante :

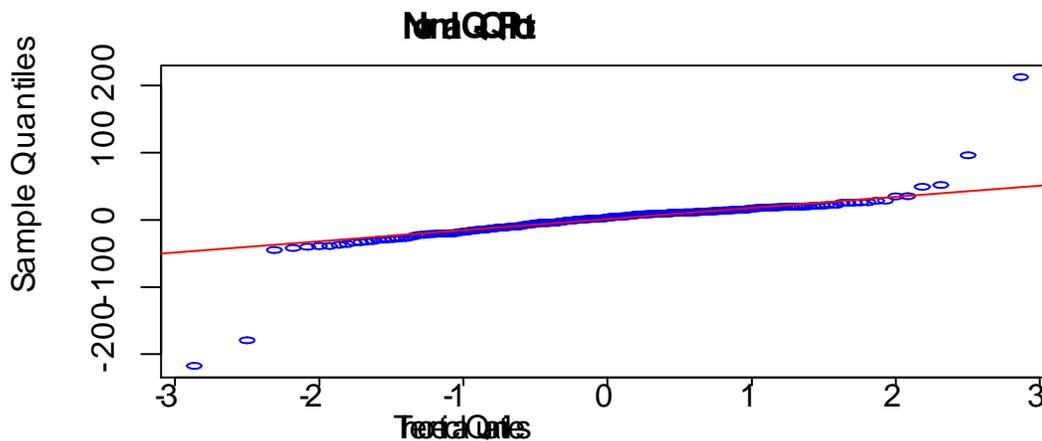


FIG. 3.6 – Normalité des résidus du modèle  $ARMA(1, 1)$

De la figure (Fig. 3.6) des quantiles-quantile (QQ-plot), on remarque que les résidus ont une distribution normale.

## 3.4 Prédiction

Finalement, après avoir testé le modèle estimé et testé sa validité, nous pouvons prédire les périodes à court terme pour l'année suivante : Oct 2008 - Jul 2009. Nous obtenons les

prévisions suivantes :

<i>Mois</i>	<i>Out</i>	<i>Sep</i>	<i>Oct</i>	<i>Nov</i>	<i>Dec</i>	<i>Jan</i>
<i>Privision</i>	33.32	40.47	45.37	49.01	51.94	54.45

<i>Mois</i>	<i>Fev</i>	<i>Mar</i>	<i>Avr</i>	<i>Mai</i>	<i>Jui</i>	<i>Jul</i>
<i>Privision</i>	56.69	58.76	60.69	62.51	64.24	65.88

TAB. 3.4 – Privision des prix de pétrole (Out 2008 - Jul 2009)

# Conclusion

*En conclusion, nous avons discuté dans ce mémoire sur le sujet important de la modélisation des séries chronologiques, notre travail est basé sur l'application de la méthode de Box-Jenkins sur des données réelles qui représente les prix de pétrole en Europe en utilisant les modèles ARIMA à l'aide des test de Ljung-Box, les autocorrélation simple et partielle et les critères d'information AIC et BIC.*

*La prévision a un rôle important dans la prise de décision est que prévoie une ou des valeurs futures de la série sous une autre forme est une vision de ce que seront les variables futures.*

*Le processus de prévision de série chronologique est influencé par une combinaison de facteurs tels que le type du données, la nature et le type de modèle, la taille des données, et les composantes de la série elle-même.*

# Bibliographie

- [1] Aragon, Y. (2011). *Séries temporelles avec R. Méthodes et cas.*, Springer, Paris.
- [2] Bosq D., Lecoutre J.P. (1992). *Analyse et prévision des séries chronologiques*, Masson, Paris.
- [3] Bourbonnais R., Terraza M. (2016). *Analyse des séries temporelles : applications à l'économie et à la gestion*, 4e edt. Dunod.
- [4] Box G., Jenkins G. (1970). *Time series analysis : Forecasting and control*, Holden-Day.
- [5] Bresson G., Pirotte A. (1995). *Économétrie des séries temporelles : théorie et applications*, PUF.
- [6] Cryer J.D., Chan K.S. (2008). *Time Series Analysis With Applications in R*. Springer.
- [7] Dickey D. A., Fuller W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the american statistical association*. 74 (366) : 427–431.
- [8] Gouriéroux C., Montfort A. (1983). *Cours de séries temporelles*, Economica, Paris.
- [9] Shumway R. H., Stoffer D. S. (2017). *Time series analysis and its applications : with R examples*. Springer.
- [10] Woodward W. A., Gray H. L., Elliott A. C. (2017). *Applied time series analysis with R*. CRC press.

## Annexe : Notations et abriviations

<i>ACF</i>	Fonction d'autocorrelation simple
<i>ADF</i>	Test de Dickey Fuller Augmenté
<i>AIC</i>	Critère d'information d'Akaike
<i>AR</i>	Autorégressif
<i>ARIMA</i>	Autorégressif moyenne mobile intégré
<i>ARMA</i>	Autorégressif moyenne mobile
<i>BB</i>	Bruit blanc
<i>BIC</i>	Critère d'information bayésien
<i>Cov</i>	Covariance
<i>DS</i>	Npn stationnarité stochastique
<i>E</i>	Espérance
<i>Var</i>	Variance
<i>MA</i>	Moyenne mobile
<i>PACF</i>	Fonction d'autocorrélation partielle
<i>TS</i>	Non stationnarité déterministe

$T_t$	Tendance
$X_t$	Série temporelle (Processus)
$S_t$	Saisonalité
$\varepsilon_t$	erreurs (bruit blanc)
$R_h(.)$	matrice des corrélations
$\det(R)$	déterminant de la matrice $R$
$L$	Opérateur de retard
$\Delta = (1 - L)$	Opérateur de différentiation
$\bar{X}_t$	Moyenne empirique
$H_0$	Hypothèse simple
$H_1$	Hypothèse alternative
$\alpha_3$	Skewness
$\alpha_4$	Kurtosis
$\chi$	Khi-deux
$\gamma$	Fonction d'autocovariance
$\rho$	Fonction d'autocorrelation