



Faculté des Sciences Exactes et Informatique
Département de Mathématiques

N° d'ordre :

N° de séries :

Mémoire de fin d'études

Présenté pour l'obtention du diplôme de

Master

Spécialité : Mathématiques.

Option : Probabilités et Statistique.

Thème

*Modèles Stochastiques et leurs Prévisions
par des Processus ARMA*

Présenté par :

- *Rim Harrati.*
- *Horiya Tibigui.*

Devant le jury :

Président	: Yakoubi Fatima	Maitre Assistante A	Université de Jijel
Encadreur	: Sellami Nawel	Maitre Assistante A	Université de Jijel
Examineur	: Ghouil Djoweyda	Maitre Assistante A	Université de Jijel

Remerciements

Tout d'abord, nous remercions Dieu pour sa générosité, qui nous a permis d'accomplir ce travail grâce à lui, sa louange étant avant tout.

*Nous aimerions également remercier notre superviseuse **Sellami Nawal** pour le temps que vous avez passé avec nous et pour les informations précieuses qu'elle nous a fournies avec intérêt et compréhension.*

Nous adressons également nos sincères remerciements aux membres des jurys pour avoir bien voulu examiner et juger ce travail.

Finalement, nous remercions tous ceux qui ont contribué de près ou de loin à la réalisation de ce mémoire.

Dédicaces

Je dédie mon mémoire :

*A ma chère mère **Fatima Harrati***

*A mon cher père **Bachir Harrati***

Qui m'ont soutenu et encouragé pendant ces années d'études.

*A mes frères **Nassim, Mohammed, Zaki, Ramzi et Anis**, qui ont été toujours
présent pour moi.*

*A ma chère binôme, **Houria***

Pour comprendre et sympathiser avec moi.

*A mes chères amies, **Fatima zohra Bouchoura, Boudjerda Yasmina, Ayda
Bouzaraa, Aicha Boulmalh, Mofida Bouhama, Naouel Hanous, Mohammed
Saoudi, Mohammed Bouzaout.***

Pour les aider et me soutenir dans les moments difficiles.

A toute ma famille,

A tous mes autres ami(e)s,

A tous ceux que j'aime et ceux qui m'aiment.



Dédicaces

Je dédie ce travail :

*A ma chère mère **Ghania Haid***

*A mon cher père **Ahmed Tibigui***

Qui m'ont soutenu et encouragé pendant ces années d'études.

*A mon frère **Abd Ellah** et mes belles sœurs **Souad, Karima, Yamina, Hlima, Samira, Sara**, qui m'ont toujours encouragé, et à qui je vous souhaite plus de succès, de joie et de bonheur*

*A ma chère binôme, **Rim***

Merci pour ta patience, ta tolérance, et pour les bons moments qu'on a partagé

*A mes chères amies, **Lamia Boulouar, Hayat Houche, Aicha Boulmalh, Mofida Bouhama, Mohammed Saoudi, Mahammed Bouzaout.***

Merci Pour les très bons moments qu'on avait ensemble.

A toute ma famille,

A tous mes autres ami(e)s,

Que dieu nous garde si tendres et aimants les un envers les autres



Table des matières

Notations	v
Résumé	vi
Introduction	vii
1 Concepts de base des séries temporelles	1
1.1 Vocabulaire et exemples	1
1.1.1 Objectifs principaux	2
1.1.2 méthodologie de box-jenkins	3
1.1.3 les composantes principales d'une séries chronologique	4
1.1.4 les modèles de décomposition déterministe	6
1.1.5 Elimination de la partie déterministe(T_t et S_t)	6
1.2 Quelques définitions	7
1.2.1 Opérateurs définis sur une série chronologique	9
1.2.2 Les fonctions auto-covariances et auto-corrélations	9
1.3 Processus ARMA	11
1.3.1 Processus MA(q)	11
1.3.2 Le processus AR(p)	13
1.3.3 Auto-covariance, Auto-corrélation et équation de Yule-Walker.	15

1.3.4	Le processus ARMA(p,q)	17
1.4	Processus ARIMA et SARIMA	19
1.4.1	processus ARIMA	19
1.4.2	processus SARIMA	20
2	Estimation des paramètres des modèles ARMA	21
2.1	Estimation des paramètres AR(p) par les équations de Yule-Walker	21
2.2	Estimation des paramètres de modèle ARMA par la méthode du maximum de vraisemblance	23
2.2.1	La méthode du maximum de vraisemblance :	24
2.2.2	L'estimation de la fonction de vraisemblance d'un processus AR(p) :	25
2.2.3	L'estimation de la fonction de vraisemblance d'un processus MA(q)	27
2.2.4	L'estimation de la fonction de vraisemblance d'un processus ARMA(p,q)	29
2.3	Test de compatibilité du modèle	31
2.3.1	test de la stationnarité ADF	32
2.3.2	Test de normalité	33
2.3.3	Test sur les résidus	34
2.4	Choix d'un modèle	34
2.5	Critères d'erreur	35
2.6	Prévisions	35
2.7	Prévision dans les ARMA(p,q)	36
2.7.1	Prévision à l'aide d'un modèle AR(p)	36
2.7.2	Prévision à l'aide d'un modèle MA(q)	36
2.7.3	Prévision à l'aide d'un processus ARMA(p,q)	37

3	Application	38
3.1	Analyse préliminaire de la série (dialyse)	39
3.1.1	Représentation graphique de la série dialyse	40
3.1.2	corrélogramme et corrélogramme partiel de la série dialyse	40
3.2	Différentiation de la série dialyse	41
3.2.1	Représentation graphique de la série dialysediff	41
3.2.2	corrélogramme et corrélogramme partiel de la série dialysediff	42
3.3	L'estimation et validation du modèle	43
3.3.1	Estimation des modèles identifiés	43
3.3.2	Validation des résidus	44
3.4	choix de modèle	49
3.4.1	Test de normalité pour les résidus du modèle AR(2)	49
3.5	Prévision	50
3.5.1	Le graphe de prévision	50
	Conclusion	51

Liste des tableaux

1.1	Résumé des propriétés des fonctions d'auto-corrélations simples et partielles	19
3.1	Nombre des séances d'hémodialyse	38
3.2	estimation des paramètres des modèles MA(1) AR(2) ARMA(2,1)	43
3.3	Résultat de test de Box-pierce	44
3.4	la table des critères du choix de meilleur modèle	49
3.5	la table des prévisions	50

Table des figures

1.1	La température mensuelle moyenne à Nottingham,1920-1939	2
1.2	Le nombre de passagers (en milliers) dans les transports	3
1.3	La concentration atmosphérique de CO2 à Maunaloa	3
1.4	Méthodologie de Box et Jenkins	4
1.5	Exemple d'un tendance linéaire	5
1.6	Exemple d'un saisonnalité	5
1.7	Exemple d'une fluctuations irrégulières	6
3.1	Procédure de construction de la méthode Box-Jenkins.	39
3.2	le graphe de la série dialyse	40
3.3	Le corrélogramme (acf) de la série dialyse	40
3.4	Le corrélogramme partiel (pacf) de la série dialyse	41
3.5	le graphe de la série dialysediff	41
3.6	Le corrélogramme (acf) de la série dialysediff	42
3.7	Le corrélogramme partiel (pacf) de la série dialysediff	42
3.8	le graphe de la série résiduelle pour le modèle MA(1)	44
3.9	le corrélogramme de la série résiduelle pour le modèle MA(1)	45
3.10	le corrélogramme partiel de la série résiduelle pour le modèle MA(1)	45
3.11	le graphe de la série résiduelle pour le modèle AR(2)	46

3.12 le corrélogramme de la série résiduelle pour le modèle AR(2)	46
3.13 le corrélogramme partiel de la série résiduelle pour le modèle AR(2)	47
3.14 le graphe de la série résiduelle pour le modèle ARMA(2,1)	47
3.15 le corrélogramme de la série résiduelle pour le modèle ARMA(2,1)	48
3.16 le corrélogramme partiel de la série résiduelle pour le modèle ARMA(2,1) .	48
3.17 l'histogramme des résidus pour le modèle AR(2)	49
3.18 Graphe de la prévision	50

NOTATIONS

T	Le nombre d'observations
T_t	la tendance
S_t	La saisonnalité
ε_t	Le résidu
<i>i.i.d</i>	Indépendant identiquement distribués
WN	Bruit blanc "White Noise"
E	Espérance
Var	Variance
Cov	Covariance
$Corr$	Corrélation
$ARMA$	Auto regressive moving average "Auto-régressif moyenne mobile"
AR	Auto regressive "Auto-régressif"
MA	Moving average "Moyenne mobile"
$ARIMA$	Auto-régressif intégré moyenne mobile
$SARIMA$	Auto-régressif moyenne mobile intégré saisonnier
ACF	Fonction d'auto-corrélation
$PACF$	Fonction d'auto-corrélation partielle
X^*	Prévisions

Résumé

L'objectif principal de ce travail consiste en la prévision des futures valeurs des processus stochastiques par des modèles : Auto-régressifs, moyenne mobile, auto-régressifs moyenne mobile et auto-régressifs moyenne mobile intégrés. Pour ce faire, on cherche tout d'abord un modèle adéquat pour la série chronologique étudiée et on procède à une analyse de celle-ci dans le but de déterminer ses caractéristiques et de voir si celles-ci correspondent bien à celles d'un des modèles mentionnés. En outre, on procède à l'estimation des paramètres du modèle considéré et on teste sa compatibilité avec l'un des modèles mentionnés. le modèle obtenu est alors utilisé pour le calcul des prévisions des processus stochastiques considérés.

Mots clés :

Séries chronologiques, processus stochastiques, la fonction d'auto-corrélation, la fonction d'auto-corrélation partielles, le modèle ARIMA, Fonction de vraisemblance, Fonction de prévision.

Introduction

Depuis toujours, l'homme a voulu prédire l'avenir, que ce soit pour prendre de meilleures décisions ou simplement pour satisfaire sa curiosité. Les premières tentatives étaient basées sur l'astrologie ou autres superstitions. Toute fois, ces méthodes de prédiction ne peuvent pas être considérées comme rigoureuses aux yeux de la science. Heureusement, la science a énormément évolué au cours des dernières décennies : les méthodes de prévisions qui en découlent ne font pas exception.

Une méthode de prévision très populaire est basée sur l'étude rigoureuse de séries chronologiques (temporelles). Cette méthode permet de prédire, par exemple, de nombreux phénomènes naturels et financiers. Une série chronologique est constituée de valeurs observées à des intervalles de temps réguliers. Par exemple, les débits annuels sur un cours d'eau ou encore les valeurs mensuelles de titres boursiers sont des séries chronologiques. À la base, l'étude formelle des séries chronologiques consiste à trouver un modèle mathématique qui explique le mieux possible les données observées. A partir de ce modèle, il est possible de faire de la prévision, ce pendant, la justesse des prévisions dépend fortement de la qualité du modèle choisi. Il est donc primordial de trouver des modèles qui reflètent le mieux possible la réalité afin de minimiser les erreurs de prévisions.

Dans ce mémoire, nous proposons la méthodologie de Box-Jenkins pour la modélisation et la prédiction en se basant sur des modèles auto-régressif (AR), les modèles moyenne mobile (MA), les modèles ARMA,...etc, ce mémoire est constitué de la présente introduction, de trois chapitres. Il est organisé comme suit.

Le premier chapitre, consacré à la présentation des outils indispensables pour l'étude de la méthodologie de Box-Jenkins et d'introduire la présentation de deux types de processus stochastique linéaire, à savoir, les modèles linéaires stationnaires dans lesquels on trouve le modèle auto-régressif (AR), le modèle moyenne mobile (MA) et le modèle auto-régressif moyenne mobile (ARMA) et celui du modèle linéaire non stationnaire ou l'on présente le modèle auto-régressif moyenne mobile intégré (ARIMA) et le modèle auto-régressif moyenne mobile intégré saisonnier (SARIMA) cette présentation suppose qu'on définisse

essentiellement des notions sur l'analyse de ces modèles, et en particulier la notion de stationnarité, la causalité et l'inversibilité, ainsi que la fonction d'auto-corrélation (ACF) et la fonction d'auto-corrélation partielle (PACF).

Dans le deuxième chapitre, on a abordé l'estimation des paramètres du modèle ARMA en utilisant deux méthodes les plus usuelles, la première étant celle de Yule-Walker (Y-W) et la seconde est celle du maximum de vraisemblance (MMV), ainsi que la validation. En outre, on a destiné à la prévision en utilisant les processus AR, MA et ARMA.

Dans le troisième chapitre, on a proposé une application numérique sur des données réelles en utilisant le logiciel R.

Chapitre 1

Concepts de base des séries temporelles

La théorie des séries chronologiques est appliquée dans nos jours dans de vastes domaines tels que l'économétrie, la médecine et la démographie. On s'intéresse à l'évolution au cours du temps d'un phénomène, dans le but d'expliquer puis prévoir ce phénomène dans le futur.

1.1 Vocabulaire et exemples

Définition 1.1. Une série chronologique (temporelle) provient de la réalisation d'une famille de variables aléatoires $(X_t)_{t \in \Theta}$, nous la noterons

$$(X_t)_{t \in \Theta} \quad \text{ou} \quad (X_t, t \in \Theta)$$

où l'ensemble Θ est appelé espace de temps qui peut être :

– **Discret** (nombre de voyageurs SNCF quotidien, température maximale,...). Dans ce cas, $\Theta \subset \mathbb{Z}$.

Les dates des observations sont le plus souvent équidistantes : par exemple relevés mensuels, trimestriels... Ces dates équidistantes sont alors indexées par des entiers :

$t = 1, 2, \dots, T$ et T est le nombre d'observations. On dispose donc des observations des variables X_1, X_2, \dots, X_T issues de la famille $(X_t)_{t \in \Theta}$ où $\Theta \subset \mathbb{Z}$ (le plus souvent $\Theta = \mathbb{Z}$).

Ainsi si h est l'intervalle de temps séparant deux observations et t_0 l'instant de la première observation, on a le schéma suivant :

$$\begin{array}{cccc} t_0 & t_0 + h & \dots & t_0 + (T-1)h \\ X_{t_0} & X_{t_0+h} & \dots & X_{t_0+(T-1)h} \end{array}$$

$$X_1 \quad X_2 \quad \dots \quad X_T$$

– **Continu** (signal radio,...). L'indice de temps est à valeurs dans un intervalle de \mathbb{R} et on dispose (au moins potentiellement) d'une infinité d'observations issues d'un processus $(X_t)_{t \in \Theta}$ où Θ est un intervalle de \mathbb{R} .

Un tel processus est dit à temps continu.

Exemple 1.2. On peut songer par exemple à l'évolution du nombre de voyageurs utilisant le train, à l'accroissement relatif mensuel de l'indice des prix ou encore à l'occurrence d'un phénomène naturel. Prenant les exemples suivants :

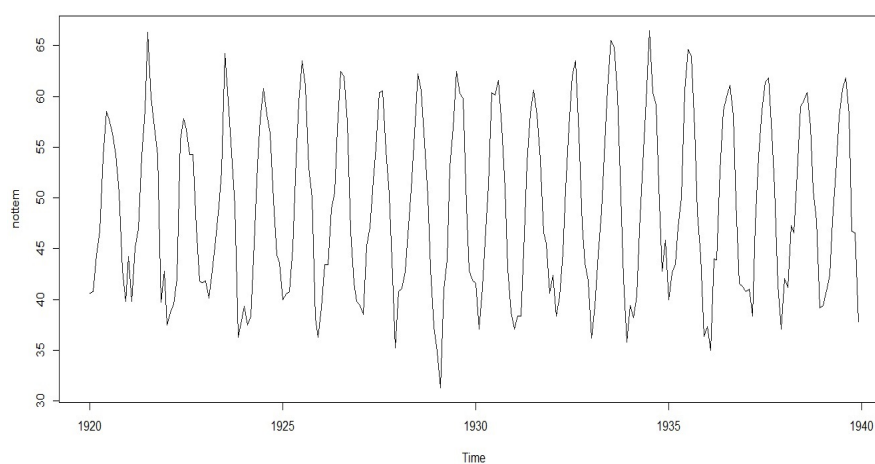


FIGURE 1.1 – La température mensuelle moyenne à Nottingham, 1920-1939

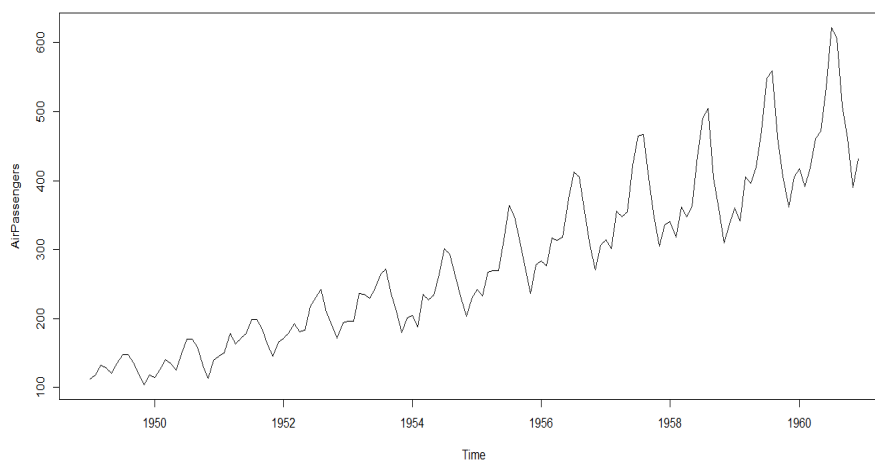


FIGURE 1.2 – Le nombre de passagers (en milliers) dans les transports

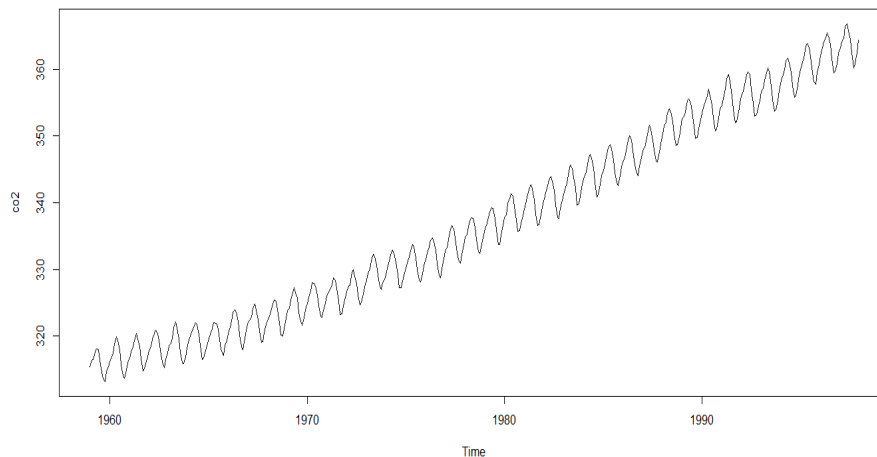


FIGURE 1.3 – La concentration atmosphérique de CO2 à Maunaloa

1.1.1 Objectifs principaux

L'étude d'une série chronologique permet d'analyser, de décrire et d'expliquer un phénomène au cours du temps et d'en tirer des conséquences pour des prises de décision (marketing...).

Cette étude permet aussi de faire un contrôle, par exemple pour la gestion des stocks, le contrôle d'un processus chimique. Mais un objectif principal de dans l'étude d'une série chronologique est la prévision qui consiste à prévoir les valeurs futures

X_{T+h} ($h = 1, 2, 3, \dots$) de la série chronologique à partir de ses valeurs observées jusqu'au temps T : X_1, X_2, \dots, X_T .

1.1.2 méthodologie de box-jenkins

Box-Jenkins (1976) ont promu une méthodologie consistant à modéliser les séries temporelles univariées au moyen des processus *ARMA*. Ces processus sont parcimonieux et constituent une bonne approximation de processus plus généraux pourvu que l'on se restreigne au cadre linéaire. Les modèles ARMA donnent souvent de bons résultats en prévision et sont bénéficiés de la vague de scepticisme quant à l'intérêt des gros modèles économétriques. **Les étapes de la méthodologie**

La méthodologie du Box et Jenkins comporte essentiellement les étapes suivantes :

– La représentation graphique de la série et différenciation des données pour les stationnariser.

- Visualisation des ACF et des $PACF$ empiriques pour identifier l'ordre p et q appropriés.
- Estimation des paramètres du(des) modèle(s) sélectionné(s).
- Diagnostique et tests d'adéquation du modèle.
- Prévision : La dernière étape consiste à la prévision des valeurs futures à travers le modèle retenu.

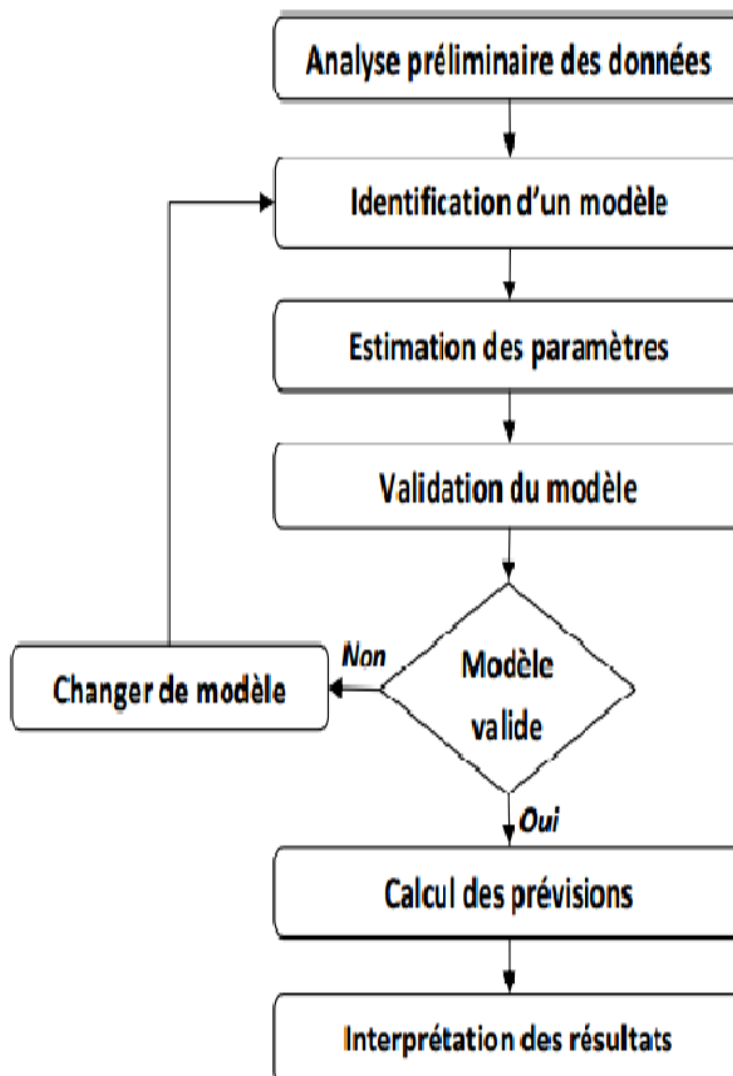


FIGURE 1.4 – Méthodologie de Box et Jenkins

1.1.3 les composantes principales d'une série chronologique

On considère qu'une série chronologique $(X_t)_{t \in \mathbb{Z}}$ est la résultante de différentes composantes fondamentales :

1. **la tendance** (T_t) : (ou trend) Représente l'évolution à long terme de la série étudiée, elle traduit le comportement "moyen" de la série. Cette tendance peut-être linéaire, polynômial ou exponentielle. Par exemple la figure 1.5 représente une tendance qui augmente d'une façon linéaire

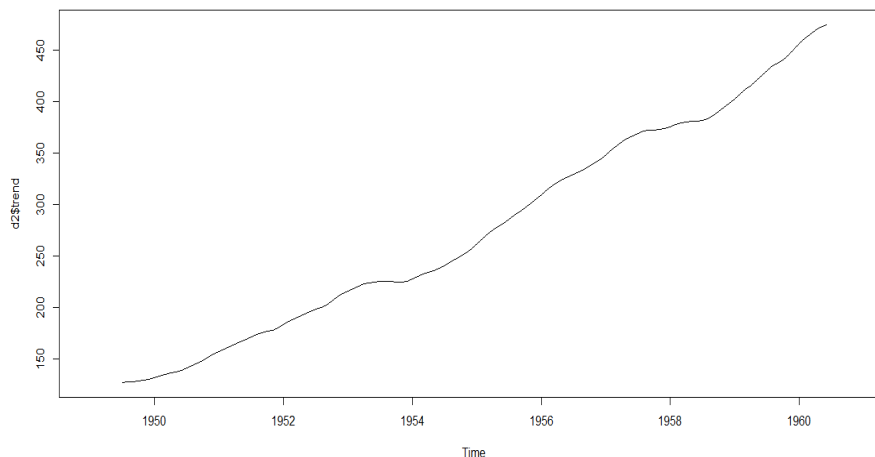


FIGURE 1.5 – Exemple d'un tendance linéaire

2. **la saisonnalité** (S_t) : (la composante saisonnière) Correspond à un phénomène qui se répète dans un intervalle de temps réguliers(périodique). Par exemple la figure 1.6 représente des cycles réguliers au cours du temps et de même amplitude

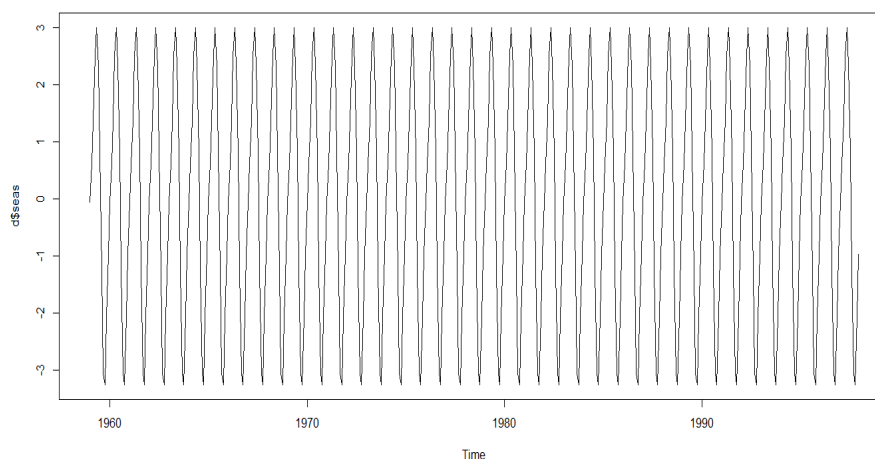


FIGURE 1.6 – Exemple d'un saisonnalité

3. **le résidu** (ε_t) :(la composante résiduelle) Correspond à des fluctuations irrégulières, en général de faible intensité mais de nature aléatoire on parle aussi d'aléas. Par exemple la figure 1.7 représente des fluctuations irrégulières

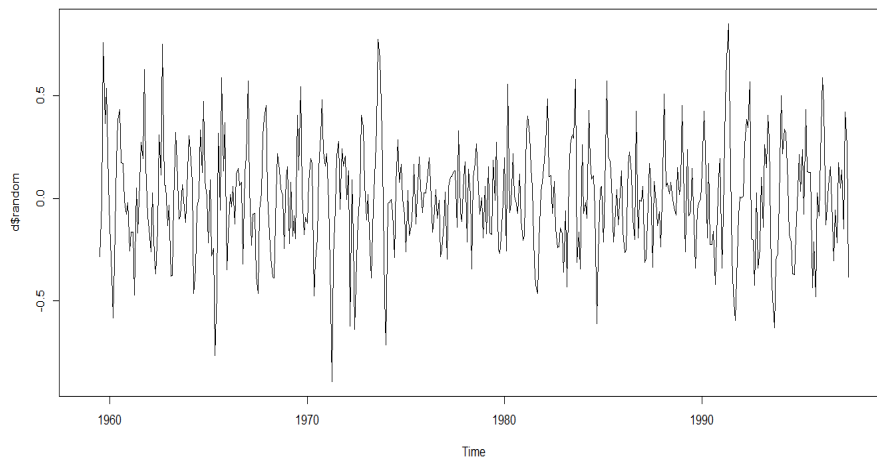


FIGURE 1.7 – Exemple d'une fluctuations irrégulières

1.1.4 les modèles de décomposition déterministe

1) le modèle additif :

Dans un modèle additif ,on suppose que les trois composantes (tendance T_t ,variations saisonnières S_t et variations accidentelles ε_t) sont indépendantes les une des autres.

$$X_t = T_t + S_t + \varepsilon_t$$

2) le modèle multiplicatif :

1° forme de modèle multiplicatif : On suppose que les variations saisonnières dépendent de la tendance .

$$X_t = T_t \times S_t + \varepsilon_t$$

2° forme de modèle multiplicatif : On suppose que les variations saisonnières et les variations accidentelles dépendent de la tendance.

$$X_t = T_t \times S_t \times \varepsilon_t$$

3) le modèle mixte :

$$X_t = (T_t + S_t) \times \varepsilon_t$$

1.1.5 Elimination de la partie déterministe(T_t et S_t)

Il existe plusieurs méthodes pour éliminer (T_t et S_t). En présentant deux méthodes :
— la méthode moyenne mobile pour éliminer la tendance (T_t).

— la méthode des différences pour éliminer la tendance (T_t) et la saisonnalité (S_t).

Cette élimination peut s'opérer au moyen de différentiation de différents types :

$-\nabla^d = (1 - B)^d$ pour éliminer une tendance de degré d .

$-\nabla_D = (1 - B^s)^D$ pour éliminer une saisonnalité de période s .

Dans ce travail, on utilise la méthode des différences pour éliminer la tendance (T_t) et la saisonnalité (S_t).

1.2 Quelques définitions

Définition 1.3. Un *processus stochastique* est une famille $(X_t)_{t \in \mathbb{Z}}$ de variables aléatoires définies sur un même espace de probabilité. L'indice t est souvent interprété comme le temps.

Définition 1.4. On dit que $(\varepsilon_t)_t$ est un **Bruit Blanc faible** si il possède les propriétés suivantes :

$$\begin{cases} E(\varepsilon_t) = 0 & \forall t \in \mathbb{Z} \\ E(\varepsilon_t^2) = \sigma_\varepsilon^2 \\ Cov(\varepsilon_s, \varepsilon_t) = 0 & \text{si } t \neq s \end{cases}$$

et on note : $\varepsilon_t \sim WN(0, \sigma_\varepsilon^2)$

Définition 1.5. On dit que $(\varepsilon_t)_t$ est un **Bruit Blanc fort** s'il est un bruit blanc faible et que les variables aléatoires $(\varepsilon_t)_t$ sont indépendantes et identiquement distribuées (*i.i.d*)

.

et on note : $\varepsilon_t \sim i.i.d(0, \sigma_\varepsilon^2)$

Définition 1.6. On dit que $(X_t)_{t \in \mathbb{Z}}$ est un **processus linéaire** de moyenne m s'il peut être écrit sous forme :

$$X_t = m + \sum_{k=-\infty}^{+\infty} b_k \varepsilon_{t-k}$$

où $(\varepsilon_t)_t \sim i.i.d(0, \sigma^2)$ et $\sum_{k=-\infty}^{+\infty} b_k^2 < +\infty$

Remarque. Si $(\varepsilon_t)_t \sim WN(0, \sigma_\varepsilon^2)$ alors $(X_t)_t$ est un **processus linéaire général** de moyenne m .

Une des grandes questions dans l'étude de séries temporelles est de savoir si celles-ci suivent un processus stationnaire.

On entend par là le fait que la structure du processus supposé évolue ou non avec le temps. Si la structure reste la même, le processus est dit alors stationnaire.

Définition 1.7. Le processus $(X_t)_t$ est **stationnaire au sens stricte (fort)** si pour tout (t_1, t_2, \dots, t_n) avec $t_i \in T$; $i = 1, \dots, n$, et si pour tout $h \in T$ avec $t_{i+h} \in T$ alors : la loi conjointe de $(X_{t_1}, X_{t_2}, \dots, X_{t_n})$ est la même loi conjointe de $(X_{t_1+h}, \dots, X_{t_n+h})$.

Définition 1.8. $(X_t)_{t \in \mathbb{Z}}$ est **un processus stationnaire du second ordre** (ou un processus faiblement stationnaire) s'il vérifie :

$$\begin{cases} E(X_t) = m < +\infty & \forall t \in \mathbb{Z} \\ E(X_t^2) < +\infty & \forall t \in \mathbb{Z} \\ \text{Cov}(X_t, X_{t+h}) = \gamma_h & \forall t \in \mathbb{Z} \quad (\text{ne dépend pas de } t) \end{cases}$$

où γ_h est la fonction d'auto-covariance du processus.

Exemple 1.9. Soit $(X_t)_{t \in \mathbb{Z}}$ un processus défini par :

$$\forall t \in \mathbb{Z}; X_t = \varepsilon_t + \theta \varepsilon_{t-1}$$

où $\theta \in \mathbb{R}^*$ et $\varepsilon \sim WN(0, \sigma_\varepsilon^2)$

alors $(X_t)_{t \in \mathbb{Z}}$ est un processus stationnaire car :

1.

$$\begin{aligned} E(X_t) &= E(\varepsilon_t + \theta \varepsilon_{t-1}) \\ &= 0 < +\infty \end{aligned}$$

2.

$$\begin{aligned} E(X_t^2) &= E((\varepsilon_t + \theta \varepsilon_{t-1})^2) \\ &= \sigma_\varepsilon^2(1 + \theta^2) < +\infty \end{aligned}$$

3.

$$\begin{aligned} \text{Cov}(X_t, X_{t+h}) &= E(X_t X_{t+h}) \\ &= \begin{cases} \sigma_\varepsilon^2(1 + \theta^2) & \text{si } h = 0 \\ \sigma_\varepsilon^2 \theta & \text{si } |h| = 1 \\ 0 & \text{si } |h| > 1 \end{cases} \end{aligned}$$

Définition 1.10. Un processus est dit **causal** s'il existe une suite $\{a_k\}_k$ réelle telle que $\sum_{k=0}^{+\infty} |a_k| < +\infty$ et que :

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

Définition 1.11. Un processus est dit **inversible** s'il existe une suite $\{b_k\}_k$ réelle telle que $\sum_{k=0}^{+\infty} |b_k| < +\infty$ et que :

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

1.2.1 Opérateurs définis sur une série chronologique

- Opérateurs de retard et d'avance :

L'opérateur de retard **B** se définit de la manière suivante :

$$B(X_t) = X_{t-1} \quad \text{et} \quad B^n(X_t) = X_{t-n} \quad n = 1, 2, \dots$$

B est linéaire et inversible tel que $B^{-1} = F$ défini par :

$$\forall t \in \mathbb{Z}; F X_t = X_{t+1}$$

F est appelé l'opérateur d'avance.

- Opérateur de différence d'ordre s :

On définit l'opérateur ∇_s de différence d'ordre s comme l'opérateur linéaire tel que :

$$\begin{aligned} \nabla_s X_t &= X_t - X_{t-s} \\ &= (1 - B^s) X_t \end{aligned}$$

1.2.2 Les fonctions auto-covariances et auto-corrélations

Définition 1.12. *Moyenne empirique :*

On définit la moyenne empirique par :

$$\bar{X}_T = \frac{1}{T} \sum_{i=1}^T X_i$$

Définition 1.13. La fonction d'auto-covariance

La fonction d'auto-covariance d'un processus stationnaire $(X_t)_{t \in \mathbb{Z}}$ est défini par :

$$\begin{aligned}\gamma_h &= \text{Cov}(X_t, X_{t-h}) \\ &= E[(X_t - E(X_t))(X_{t-h} - E(X_{t-h}))] \quad \forall h \in \mathbb{Z}\end{aligned}$$

L'expression de γ_h est estimée par l'auto-covariance empirique :

$$\forall h \in \{0, \dots, T-1\}, \hat{\gamma}_h = \frac{1}{T} \sum_{i=1}^{T-h} (X_{i+h} - \bar{X}_T)(X_i - \bar{X}_T)$$

où \bar{X}_T désigne la moyenne de la série calculée.

Proposition 1.14.

$$\sqrt{T}(\bar{X}_T - m) \xrightarrow[T \rightarrow +\infty]{L} \mathcal{N}(0, s^2)$$

$$\text{où } s^2 = \sum_{h \in \mathbb{Z}} |\sigma_X(h)|$$

Définition 1.15. La fonction d'auto-corrélation (ACF)

La fonction d'auto-corrélation d'un processus stationnaire $(X_t)_{t \in \mathbb{Z}}$ est défini par :

$$\begin{aligned}\forall h \in \mathbb{Z}; \quad \rho_h &= \frac{\text{Cov}(X_t, X_{t-h})}{\sqrt{\text{Var}(X_t)}\sqrt{\text{Var}(X_{t-h})}} \\ &= \frac{\gamma_h}{\gamma_0}\end{aligned}$$

L'expression de ρ_h est estimée par l'auto-corrélation empirique :

$$\forall h \in \{0, \dots, T-1\}, \hat{\rho}_h = \frac{\hat{\gamma}_h}{\hat{\gamma}_0}$$

Proposition 1.16. on a :

$$\hat{\rho}_h = \begin{pmatrix} \hat{\rho}_1 \\ \vdots \\ \hat{\rho}_h \end{pmatrix}$$

et

$$\rho_h = \begin{pmatrix} \rho_1 \\ \vdots \\ \rho_h \end{pmatrix}$$

alors :

$$\sqrt{T}(\hat{\rho}_h - \rho_h) \xrightarrow[T \rightarrow +\infty]{L} \mathcal{N}(0, \Sigma)$$

où :

$$\forall i, j \in \{1, \dots, h\}, \Sigma_{i,j} = \sum_{l=1}^{+\infty} [\rho_{l+i} + \rho_{l-i} - 2\rho_l\rho_i] - [\rho_{l+j} + \rho_{l-j} - 2\rho_l\rho_j]$$

Théorème 1.17. *La fonction d'auto-covariance (resp. d'auto-corrélation) d'un processus X_t stationnaire vérifie les propriétés suivantes :*

- $\gamma_0 = Cov(X_t, X_t) = E[(X_t - E(X_t))^2] = Var(X_t) = \sigma_X^2 > 0$
- $|\gamma_h| \leq \gamma_0$
- $\gamma_h = \gamma_{-h}$ (fonction paire)
(resp. $\rho_0 = 1, |\rho_h| \leq \rho_0, \rho_h = \rho_{-h}$)

Définition 1.18. *La fonction d'auto-corrélation partielle (PACF)*

La fonction d'auto-corrélation partielle mesure la corrélation entre X_t et X_{t-h} l'influence des autres variables $X_{t-1}, \dots, X_{t-h+1}$ ayant été retirée

Corollaire 1.19. *Soit $(X_t)_t$ un processus stationnaire, l'auto-corrélation partielle d'ordre h notée ϕ_{hh} , et définie comme suit :*

$$\phi_{hh} = \frac{|R_h^*|}{|R_h|}$$

avec :

$$R(h) = \begin{bmatrix} 1 & \rho_1 & \dots & \rho_{h-1} \\ \rho_1 & 1 & \dots & \rho_2 \\ \vdots & \vdots & \dots & \vdots \\ \rho_{h-1} & \rho_{h-2} & \dots & 1 \end{bmatrix}$$

$R^*(h)$ obtenue en remplaçant la dernière colonne de $R(h)$ par le vecteur $[\rho_1, \rho_2, \dots, \rho_h]'$

$$R^*(h) = \begin{bmatrix} 1 & \rho_1 & \dots & \rho_1 \\ \rho_1 & 1 & \dots & \rho_2 \\ \vdots & \vdots & \dots & \vdots \\ \rho_{h-1} & \rho_{h-2} & \dots & \rho_h \end{bmatrix}$$

1.3 Processus ARMA

1.3.1 Processus MA(q)

Définition 1.20. *On dit que la suite $(X_t)_{t \geq 0}$ est un processus à moyenne mobile d'ordre q ($q > 0$), notée $MA(q)$ si :*

$$\forall t \in \mathbb{Z} : X_t = \varepsilon_t + \sum_{i=1}^q \theta_i \varepsilon_{t-i}$$

où : $q \geq 1$, $(\theta_i)_{1 \leq i \leq q}$ réels et $(\varepsilon_t)_t$ est un Bruit Blanc .

on utilise généralement la notation suivante :

$$X_t = \Theta(B)\varepsilon_t, \text{ où } \Theta(B) = 1 + \sum_{i=1}^q \theta_i B^i.$$

Les propriétés d'un processus MA(q)

1) **La stationnarité** : Soit $(X_t)_{t \in \mathbb{Z}}$ un processus $MA(q)$, alors X_t est stationnaire .

Démonstration. On a :

(1)

$$\begin{aligned} E(X_t) &= E\left(\varepsilon_t + \sum_{i=1}^q \theta_i \varepsilon_{t-i}\right) \\ &= E(\varepsilon_t) + \sum_{i=1}^q \theta_i E(\varepsilon_{t-i}) \\ &= 0 < +\infty \end{aligned}$$

car : $(\varepsilon_t)_t \sim WN(0, \sigma_\varepsilon^2)$

(2)

$$\begin{aligned} \gamma_h &= Cov(X_t, X_{t+h}) \\ &= E(X_t X_{t+h}) - E(X_t)E(X_{t+h}) \\ &= E(X_t X_{t+h}) \\ &= E\left(\sum_{i=0}^q \theta_i \varepsilon_{t-i} \sum_{k=0}^q \theta_k \varepsilon_{t+h-k}\right) \\ &= E\left(\sum_{i=0}^q \sum_{k=0}^q \theta_i \theta_k \varepsilon_{t-i} \varepsilon_{t+h-k}\right) \\ &= \sum_{i,k=0}^q \theta_i \theta_k E(\varepsilon_{t-i} \varepsilon_{t+h-k}) \\ &= \begin{cases} \sum_{i=0}^{q-|h|} \theta_i \theta_{i+h} \sigma_\varepsilon^2 & \text{si } i = k - h \\ 0 & \text{sinon} \end{cases} \\ &= \begin{cases} \sigma_\varepsilon^2 \sum_{i=0}^{q-|h|} \theta_i \theta_{i+h} & \text{si } 0 \leq h \leq q \\ 0 & \text{si } h > q \end{cases} \end{aligned}$$

donc γ_h ne dépend pas de t .

$$(3) E(X_t^2) = \text{Var}(X_t) = \gamma_0 = \sigma_\varepsilon^2 \sum_{i=0}^q \theta_i^2 < +\infty$$

Donc tout MA(q) est automatiquement un processus stationnaire. ■

2) **La causalité** : Tout processus MA(q) est causal par définition.

3) **inversibilité** : Un processus MA(q) est inversible si et seulement si son polynôme $\Theta(z)$ est tel que :

$$\Theta(z) \neq 0 \text{ avec } z \in \mathbb{C} \text{ tel que } |z| \leq 1 .$$

ou bien : $\Theta(z) = 0$, pour tout $z \in \mathbb{C}$ tel que $|z| > 1$.

La fonction d'auto-covariance d'un processus MA(q) :

La fonction d'auto-covariance d'un processus MA(q) est donnée par :

$$\gamma_h = \text{Cov}(X_t, X_{t+h}) = \begin{cases} \sigma_\varepsilon^2 \sum_{i=0}^{q-|h|} \theta_i \theta_{i+h} & \text{si } 0 \leq h \leq q \\ 0 & \text{si } h > q \end{cases}$$

La fonction d'auto-corrélation d'un processus MA(q)

La fonction d'auto-corrélation d'un processus MA(q) est donnée par :

$$\rho_h = \begin{cases} 1 & \text{si } h = 0 \\ \frac{\theta_h + \theta_1 \theta_{h+1} + \dots + \theta_{q-h} \theta_q}{1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2} & \text{si } 0 \leq h \leq q \\ 0 & \text{si } |h| > q \end{cases}$$

La fonction d'auto-corrélation partielle d'un MA(q)

Proposition 1.21. *Les auto-corrélations partielles sont solutions d'une équation de récurrence linéaire simple d'ordre q. Elles décroissent vers 0 de manière exponentielle .*

1.3.2 Le processus AR(p)

Définition 1.22. *On dit que la suite $\{X_t; t \geq 0\}$ est un **processus auto-régressif** d'ordre p ($p \geq 0$) ,notée AR(p) si :*

$$\forall t \in \mathbb{Z} : X_t = \varepsilon_t + \sum_{i=1}^p \varphi_i X_{t-i}$$

où $(\varphi_1, \dots, \varphi_p)' \in \mathbb{R}^p$ sont des constants et $\varphi_p \neq 0$.

on utilise généralement la notation suivante :

$$\Phi(p)X_t = \varepsilon_t$$

où $\Phi(p) = 1 - \sum_{i=1}^p \varphi_i B^i$.

Les propriétés d'un processus AR(p)

- 1) **la stationnarité et la causalité** : Un processus auto-régressif AR(p) est causal et stationnaire si et seulement si son polynôme $\Phi(z)$ est tel que :
 $\Phi(z) \neq 0$ avec $z \in \mathbb{C}$ tel que $|z| \leq 1$.
 où bien : $\Phi(z) = 0$, pour tout $z \in \mathbb{C}$ tel que $|z| > 1$.
- 2) **l'inversibilité** : Tout processus AR(p) est inversible par définition.

Exemple 1.23. *Le processus AR(2) suivant est stationnaire et causal :*

$$X_t = -\frac{1}{6}X_{t-1} + \frac{1}{6}X_{t-2} + \varepsilon_t ; \text{ où } (\varepsilon_t)_t \sim WN(0, \sigma_\varepsilon^2)$$

En effet, on peut le récrire comme :

$$\left(-\frac{1}{6}B^2 + \frac{1}{6}B + 1\right)X_t = \varepsilon_t$$

Ainsi, on a :

$$\begin{aligned} \Phi(z) &= -\frac{1}{6}z^2 + \frac{1}{6}z + 1 \\ &= \frac{1}{6}(z^2 - z - 6) \\ &= \frac{1}{6}(z + 2)(z - 3) \end{aligned}$$

les racines sont à l'extérieur du cercle unité, donc le processus est stationnaire et causal.

Fonction d'auto-corrélation partielle d'un processus AR(p)

Proposition 1.24. *Les auto-corrélations partielles d'un processus AR(p) sont nulles à partir du rang $p+1$.*

L'écriture $AR(\infty)$ d'un processus $MA(q)$

Soit $(X_t)_{t \in \mathbb{Z}}$ un processus $MA(q)$ si $(X_t)_t$ est sous forme canonique (ie :inversible et causal), alors il admet une écriture $AR(\infty)$.

$$\varepsilon_t = \Phi^{-1}(B)X_t = X_t + \sum_{i=1}^{+\infty} \pi_i X_{t-i}$$

C'est-à-dire :

$$X_t = \varepsilon_t - \sum_{i=1}^{+\infty} \pi_i X_{t-i}$$

où $(\pi_i)_{i \in \mathbb{N}}$ est une suite réelle et $\sum_{i=1}^{+\infty} |\pi_i| < \infty$ et,

$$\pi_i = - \sum_{j=1}^q \theta_j \pi_{i-j} \quad \text{pour tous } i > 1 \text{ avec } \pi_0 = 1 .$$

L'écriture $MA(\infty)$ d'un processus $AR(q)$

Soit $(\varepsilon_t)_{t \in \mathbb{Z}}$ un bruit blanc faible de variance σ_ε^2 , si $(X_t)_{t \in \mathbb{Z}}$ un processus $AR(p)$ sous forme canonique (ie :inversible et causal), alors il admet une écriture $MA(\infty)$.

$$X_t = \Phi^{-1}(B)\varepsilon_t = \varepsilon_t + \sum_{i=1}^{\infty} \psi_i \varepsilon_{t-i}$$

où $(\psi_i)_{i \in \mathbb{N}}$ est une suite réelle et $\sum_{i=1}^{\infty} |\psi_i| < \infty$. et,

$$\psi_i = \sum_{j=1}^p \varphi_j \psi_{i-j} \quad \text{pour } i > 0 \text{ avec } \psi_0 = 1, \text{ et } \psi_i = 0 \text{ pour } i < 0$$

1.3.3 Auto-covariance, Auto-corrélation et équation de Yule-Walker.

L'auto-covariance :

$$\gamma_h = Cov(X_t, X_{t-h}) = E(X_t X_{t-h}) \quad ; \text{ pour } h \geq 0.$$

on a :

$$X_t = \varphi_1 X_{t-1} + \dots + \varphi_p X_{t-p} + \varepsilon_t.$$

donc :

$$X_t^2 = \varphi_1 X_t X_{t-1} + \dots + \varphi_p X_t X_{t-p} + X_t \varepsilon_t.$$

et

$$\gamma_0 = \varphi_1 \gamma_1 + \dots + \varphi_p \gamma_p + E(X_t \varepsilon_t).$$

or

$$E(X_t \varepsilon_t) = E[\underbrace{(\varphi_1 X_{t-1} + \dots + \varphi_p X_{t-p})}_{=0 \text{ car } \varepsilon_t \perp \varepsilon_{t-1}; \forall t > 0} \varepsilon_t] + E(\varepsilon_t^2).$$

d'ou :

$$\gamma_0 = \varphi_1 \gamma_1 + \dots + \varphi_p \gamma_p + \sigma_\varepsilon^2$$

si $h > 0$ on procède de la même façon :

$$X_t X_{t-h} = \varphi_1 X_{t-1} X_{t-h} + \dots + \varphi_p X_{t-p} X_{t-h} + \varepsilon_t X_{t-h}.$$

$$\gamma_h = \varphi_1 \gamma_{h-1} + \dots + \varphi_p \gamma_{h-p} + E(X_{t-h} \varepsilon_t).$$

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

donc :

$$\gamma_h = \begin{cases} \varphi_1 \gamma_{h-1} + \dots + \varphi_p \gamma_{h-p} + \sigma_\varepsilon^2 & \text{si } h = 0 \\ \varphi_1 \gamma_{h-1} + \dots + \varphi_p \gamma_{h-p} & \text{sinon} \end{cases}$$

Les auto-corrélations :

À partir de la relation de récurrence de γ_h on déduit celle sur ρ_h telle que :

$$\rho_h = \frac{\gamma_h}{\gamma_0}$$

$$\rho_h = \varphi_1 \rho_{h-1} + \dots + \varphi_p \rho_{h-p} \quad ; \forall h \geq 0.$$

donc :

$$\rho_h = \begin{cases} \varphi_1 \rho_{h-1} + \dots + \varphi_p \rho_{h-p} + \sigma_\varepsilon^2 & \text{si } h = 0 \\ \varphi_1 \rho_{h-1} + \dots + \varphi_p \rho_{h-p} & \text{sinon} \end{cases}$$

ces dernières équations sont appelées **équation de Yule -Walker**.

pour $h > 0$, les γ_h et les ρ_h vérifient une relation de récurrence d'ordre p et :

$$1 = \varphi_1 \rho_1 + \dots + \varphi_p \rho_p + \frac{\sigma_\varepsilon^2}{\gamma_0}.$$

$$\implies \gamma_0 = \frac{\sigma_\varepsilon^2}{1 - (\varphi_1 \rho_1 + \dots + \varphi_p \rho_p)}.$$

les équations de Yule-Walker pour $h = 1, \dots, p$; peuvent s'écrire :

$$\begin{pmatrix} 1 & \rho_1 & \dots & \rho_{p-1} \\ \rho_1 & 1 & \dots & \rho_{p-2} \\ \vdots & & \ddots & \\ \rho_{p-1} & \rho_{p-2} & \dots & 1 \end{pmatrix} \begin{pmatrix} \varphi_1 \\ \varphi_2 \\ \vdots \\ \varphi_p \end{pmatrix} = \begin{pmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_p \end{pmatrix}$$

$$R_p^X \boldsymbol{\varphi} = \boldsymbol{\rho}.$$

tel que R_p^X : la matrice d'auto-corrélation, et $\boldsymbol{\varphi} = (\varphi_1, \varphi_2, \dots, \varphi_p)'$, $\boldsymbol{\rho} = (\rho_1, \rho_2, \dots, \rho_p)'$.

1.3.4 Le processus ARMA(p,q)

Définition 1.25. On dit qu'un processus stationnaire $(X_t)_{t \in \mathbb{Z}}$ est un processus ARMA (Auto-régressive Moving Average) d'ordre (p, q) , noté ARMA(p,q) si :

-il existe $\boldsymbol{\varphi} = (\varphi_1, \dots, \varphi_p)' \in \mathbb{R}^p$ avec $\varphi_p \neq 0$ et $\boldsymbol{\theta} = (\theta_1, \dots, \theta_q)' \in \mathbb{R}^q$ avec $\theta_q \neq 0$ tel que :

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

on utilise généralement la notation suivante :

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

où $\Phi(B)X_t = 1 - \sum_{i=1}^p \varphi_i B^i$ et $\Theta(B) = 1 + \sum_{i=1}^q \theta_i B^i$

Remarque. -Un processus AR(p) est un processus ARMA(p,0). -Un processus MA(q) est un processus ARMA(0,q).

propriétés d'un processus ARMA(p,q)

- 1) **la stationnarité et causalité** : Un processus ARMA(p,q) est stationnaire et causal si le polynôme $\Phi(z)$ a toutes ses racines à l'extérieur du cercle unité, autrement dit :
 $\Phi(z) = 0$ pour tout $z \in \mathbb{C}$ tel que $|z| > 1$.
- 2) **l'inversibilité** : On dit que le processus ARMA(p,q) est inversible si :
 $\Theta(z) = 0$ pour tout $z \in \mathbb{C}$ tel que $|z| > 1$.

L'écriture $MA(\infty)$ et $AR(\infty)$ d'un processus ARMA(p,q)

Proposition 1.26. Soit $(X_t)_{t \in \mathbb{Z}}$ un processus ARMA(p,q) canonique (ie :inversible et causal) :

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

alors :

1. il admet une écriture $MA(\infty)$:

$$\begin{aligned} X_t &= \Phi^{-1}(B)\Theta(B)\varepsilon_t \\ &= \varepsilon_t + \sum_{i=1}^{+\infty} \psi_i \varepsilon_{t-i} \end{aligned}$$

où $(\psi_i)_{i \in \mathbb{N}}$ sont des réels .

En posant $\psi_i = 0$ pour $i < 0$, $\theta_0 = 1$, on a :

$$\forall i \in \mathbb{N} : \psi_i - \sum_{j=1}^p \varphi_j \psi_{i-j} = \theta_i$$

2. il admet une écriture $AR(\infty)$:

$$\begin{aligned} \varepsilon_t &= \Theta^{-1}(B)\Phi(B)X_t \\ &= X_t + \sum_{i=1}^{+\infty} \pi_i X_{t-i} \iff X_t = \varepsilon_t - \sum_{i=1}^{+\infty} \pi_i X_{t-i} \end{aligned}$$

où $(\pi_i)_{i \in \mathbb{N}}$ sont des réels .

En posant $\pi_i = 0$ pour $i < 0$, on a :

$$\forall i \in \mathbb{N} : \pi_i + \sum_{j=1}^q \theta_j \pi_{i-j} = -\varphi_i$$

Les propriétés des auto-covariances

La fonction d'auto-covariance d'un processus ARMA(p,q) est donnée par :

$$\gamma_h = \sum_{i=1}^p \varphi_i \gamma_{h-i} \quad , \text{pour } h \geq q + 1$$

et

$$\gamma_h - \sum_{i=1}^p \varphi_i \gamma_{h-i} = \sigma_\varepsilon^2 \sum_{j=h}^q \theta_j \psi_{j-h} \quad 0 \leq h \leq q$$

où les ψ_k correspondent aux coefficients de la forme $MA(\infty)$ de $(X_t)_t$;

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

La fonction d'auto-corrélation d'un processus ARMA(p,q)

La fonction d'auto-corrélation d'un processus ARMA(p,q) est donnée par :

$$\frac{\gamma_h - \sum_{i=1}^p \varphi_i \gamma_{h-i}}{\gamma_0} = 0 \quad ; \text{pour } h \geq q + 1$$

où

$$\rho_h - \sum_{i=1}^p \varphi_i \rho_{h-i} \quad ; \text{pour } h \geq q + 1$$

et

$$\rho_h - \sum_{i=1}^p \varphi_i \rho_{h-i} = \frac{\sigma_\varepsilon^2}{\gamma_0} \sum_{j=h}^q \theta_j \psi_{j-h} \quad 0 \leq h \leq q$$

Dans le tableau 1.1, Box-Jenkins [2], proposent un récapitulatif sur les formes des fonctions d'auto-corrélation et d'auto-corrélation partielles théoriques des processus AR(p), MA(q) et ARMA (p,q).

Processus	Fonction auto-corrélation simple	Fonction auto-corrélation partielle
Bruit blanc	Nulle pour $k < 0$	Nulle pour $k > 0$
AR(p)	Décroissance exponentielle	Pics significatifs pour les p premiers retards. autres coefficients sont nuls pour des retards $> p$.
MA(q)	Pics significatifs pour les q premiers retards. autres coefficients sont nuls pour des retards $> q$.	Décroissance exponentielle
ARMA(p,q)	Décroissance exponentielle	Décroissance exponentielle

TABLE 1.1 – Résumé des propriétés des fonctions d'auto-corrélations simples et partielles

1.4 Processus ARIMA et SARIMA

1.4.1 processus ARIMA

Nous allons mentionné dans ce chapitre comment corriger une série non stationnaire de composantes déterministes telles que la tendance entre autre, nous avons étudié l'opérateur de différenciation pour les détendancialiser

Un processus ARIMA (p,d,q) convient pour modéliser une série temporelle comprenant une tendance de degré d.

Définition 1.27. On dit qu'un processus $(X_t)_{t \in \mathbb{Z}}$ admet une représentation ARIMA (p, d, q) s'il satisfait :

$$\Phi_p(B)(1 - B)^d X_t = \Theta_q(B)\varepsilon_t, \text{ pour tout } t \in \mathbb{Z}$$

avec :

$$\Phi_p(B) = 1 - \varphi_1 B - \varphi_2 B^2 - \dots - \varphi_p B^p$$

$$\Theta_q(B) = 1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q$$

où $(\varepsilon_t)_{t \in \mathbb{Z}} \sim WN(0, \sigma_\varepsilon^2)$, les coefficients $\varphi_1, \varphi_2, \dots, \varphi_p, \theta_1, \theta_2, \dots, \theta_q$ sont réels avec $\varphi_p \neq 0$ et $\theta_q \neq 0$.

Remarque. Si $d = 0$ on obtient un processus ARMA (p, q) .

1.4.2 processus SARIMA

Certaines séries chronologiques présentent des saisonnalités. Nous avons déjà abordé le cas de séries dont les composantes saisonnières et tendanciennes sont déterministes et nous avons procédé par différenciation pour désaisonnaliser et en suite leur appliquer une modélisation ARMA.

Box-Jenkins [2] proposent des modèles SARIMA définis ainsi :

Définition 1.28. Soient p, q, d et $s \geq 0$, un processus $(X_t)_{t \in \mathbb{Z}}$ est un processus SARIMA $(p, d, q)(P, D, Q)_s$ (multiplicatif saisonnier auto-régressif moyenne mobile intégré) s'il vérifie l'équation :

$$\begin{aligned} \nabla^d \Phi_p(B) \nabla_s^D \Phi_P(B^s) X_t &= \Theta_q(B) \Theta_Q(B^s) \varepsilon_t \\ (1 - B)^d \Phi_p(B) (1 - B^s)^D \Phi_P(B^s) X_t &= \Theta_q(B) \Theta_Q(B^s) \varepsilon_t \end{aligned}$$

Où :

$$\Phi_P(B^s) = 1 + \Phi_1 B^s + \dots + \Phi_P B^{sP}$$

$$\Theta_Q(B^s) = 1 + \Theta_1 B^s + \dots + \Theta_Q B^{sQ}$$

les polynômes $\Phi_p(B), \Phi_P(B), \Theta_q(B), \Theta_Q(B)$ sont à coefficient inconnus.

- s correspond à la période du processus SARIMA qu'on peut identifier en regardant l'auto-corrélogramme.
- Les entiers d et D sont choisis de sorte que la série différenciée $= (1 - B)^d (1 - B^s)^D X_t$ soit stationnaire.
- Les ordres p et q s'obtiennent comme pour les modèles ARMA (p, q) (auto-corrélation partielle et simple)
- Les ordres P et Q en regardant les ordres de s de l'auto-corrélogramme.

Chapitre 2

Estimation des paramètres des modèles ARMA

L'estimation des paramètres d'un modèle de séries chronologique est une des étapes importantes de la modélisation. Les méthodes d'estimations des modèles ARMA ont été largement développée et expérimentés. Des travaux récents de Box-Jenkins [2], Brockwell et Davis [3], Hamilton [7] ont conduit à certaines méthodes d'estimations que nous présentons dans ce chapitre.

Il existe plusieurs méthodes, dont la méthode des moments, la méthode des moindres carrés et la méthode du maximum de vraisemblance, la méthode de Yule-Walker. Dans notre travail on utilise deux, la première est celle de Yule-Walker et la deuxième est celle du maximum de vraisemblance.

2.1 Estimation des paramètres AR(p) par les équations de Yule-Walker

considérons un processus auto-régressif d'ordre p causal et de moyenne nulle défini par :

$$X_t = \sum_{i=1}^p \varphi_i X_{t-i} + \varepsilon_t \quad (2.1)$$

où

$$(\varepsilon_t)_t \sim WN(0, \sigma_\varepsilon^2)$$

L'idée de l'estimateur de Yule-Walker est de trouver l'équation reliant les paramètres φ_i et l'auto-covariance utilisant l'hypothèses de processus causal.

Commençons d'abord par multiplier chaque membre de l'équation (2.1) par X_{t-h} et prenons l'espérance des deux cotés de l'équation ,on obtient :

$$\begin{aligned}\gamma_h &= E\left[\sum_{i=1}^p \varphi_i X_{t-i} X_{t-h} + \varepsilon_t X_{t-h}\right] \\ &= \sum_{i=1}^p \varphi_i E[X_{t-i} X_{t-h}] + E[\varepsilon_t X_{t-h}] \\ &= \sum_{i=1}^p \varphi_i \gamma_{h-i} + E[\varepsilon_t X_{t-h}]\end{aligned}$$

Comme nous savons que $(X_t)_t$ est causal :

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

Alors on obtient :

$$\begin{aligned}E[\varepsilon_t X_{t-h}] &= E\left[\varepsilon_t \sum_{k=0}^{\infty} a_k \varepsilon_{t-h-k}\right] \\ &= E\left[\sum_{k=0}^{\infty} a_k \varepsilon_t \varepsilon_{t-h-k}\right] \\ &= \sum_{k=0}^{\infty} a_k E[\varepsilon_t \varepsilon_{t-h-k}]\end{aligned}$$

Pour $h = 0$, on a :

$$E[\varepsilon_t \varepsilon_{t-k}] = \begin{cases} \sigma_\varepsilon^2 & \text{si } k = 0 \\ 0 & \text{sinon} \end{cases}$$

Pour $h > 0$, on aura toujours que $E[\varepsilon_t \varepsilon_{t-h-k}] = 0$ car $(\varepsilon_t)_t \sim WN(0, \sigma_\varepsilon^2)$. On obtient donc le système d'équation suivant :

$$\gamma_h = \begin{cases} \sum_{k=0}^p \varphi_i \gamma_{h-k} & \text{si } h = 1, \dots, p \\ \sum_{k=0}^p \varphi_i \gamma_{h-k} + \sigma_\varepsilon^2 & \text{si } h = 0 \end{cases} \quad (2.2)$$

Ce système peut s'écrire sous forme matricielle :

en notant par $\boldsymbol{\gamma}_p$ le vecteur $(\gamma_1, \gamma_2, \dots, \gamma_p)'$ et par $\boldsymbol{\varphi}_p$ le vecteur $(\varphi_1, \varphi_2, \dots, \varphi_p)'$, et en formant la matrice $p \times p$ dont l'élément (i,j) vaut $\gamma_{|i-j|}$ c'est -à-dire :

$$R_p^X = \begin{pmatrix} \gamma_0 & \gamma_1 & \dots & \gamma_{p-1} \\ \gamma_1 & \gamma_0 & \dots & \vdots \\ \vdots & \vdots & \dots & \vdots \\ \gamma_{p-1} & \dots & \dots & \gamma_0 \end{pmatrix} \quad (2.3)$$

Alors le système (2.2) peut se réécrire :

$$\begin{cases} \boldsymbol{\gamma}_p = R_p^X \boldsymbol{\varphi}_p \\ \gamma_0 = \boldsymbol{\varphi}_p' \boldsymbol{\gamma}_p + \sigma_\varepsilon^2 \end{cases}$$

Ces équations sont appelées les équations de Yule-Walker, elles créent un lien entre les paramètres $\boldsymbol{\varphi}_p$ du processus auto-régressif et la fonction d'auto-corrélation définie par :

$$\hat{\rho}_h = \frac{\hat{\gamma}_h}{\hat{\gamma}_0}$$

Si on veut estimer les paramètres $\varphi_1, \dots, \varphi_p$ du modèle, on peut se servir des équations de Yule-Walker en utilisant la fonction d'auto-covariance empirique $\hat{\boldsymbol{\gamma}}_p$ définie au chapitre précédent comme estimateur de $\boldsymbol{\gamma}_p$. On aura donc un système d'équations qui permettra d'estimer le vecteur $\boldsymbol{\varphi}_p$. Le système se réécrit de la façon suivante :

$$\begin{cases} \hat{\boldsymbol{\gamma}}_p = \hat{R}_p^X \hat{\boldsymbol{\varphi}}_p \\ \hat{\gamma}_0 = \hat{\boldsymbol{\varphi}}_p' \hat{\boldsymbol{\gamma}}_p + \hat{\sigma}_\varepsilon^2 \end{cases} \quad (2.4)$$

Pour obtenir un estimateur de $\boldsymbol{\varphi}_p$ il faut donc résoudre (2.4) en inversant la matrice \hat{R}_p^X :

$$\hat{\boldsymbol{\varphi}}_p = (\hat{R}_p^X)^{-1} \hat{\boldsymbol{\gamma}}_p \quad (2.5)$$

puis en calculant :

$$\hat{\sigma}_\varepsilon^2 = \hat{\gamma}_0 - \hat{\boldsymbol{\varphi}}_p' \hat{\boldsymbol{\gamma}}_p.$$

Proposition 2.1. *Si $(X_t)_t$ est un processus stationnaire et causal de la forme (2.1) avec $(\varepsilon_t)_t \sim i.i.d(0, \sigma_\varepsilon^2)$ et si $\hat{\boldsymbol{\varphi}}_p$ est l'estimateur de Yule-Walker de $\boldsymbol{\varphi}_p$, obtenu à partir de T observations X_1, \dots, X_T du processus auto-régressif, alors lorsque $T \rightarrow \infty$:*

$$\sqrt{T}(\hat{\boldsymbol{\varphi}}_p - \boldsymbol{\varphi}_p) \rightarrow \mathcal{N}_p(0, \sigma_\varepsilon^2 (R_p^X)^{-1})$$

où R_p^X est la matrice $p \times p$ d'auto-covariance donnée par (2.3)

de plus pour $T \rightarrow \infty$:

$$\hat{\sigma}_\varepsilon^2 \xrightarrow{P} \sigma_\varepsilon^2$$

2.2 Estimation des paramètres de modèle ARMA par la méthode du maximum de vraisemblance

Rappelons tout d'abord, le principe de la méthode du maximum de vraisemblance. Considérons une variable aléatoire X , d'une loi de probabilité P_θ , et densité f_θ qui

dépendent du paramètre θ (θ appartient à une famille paramétrique $\theta = (\theta_1, \theta_2, \dots, \theta_k) \in \mathbb{R}^k$), à estimer en fonction des observations disponibles de la variable X (i.e. l'échantillon (x_1, x_2, \dots, x_T)).

le principe de la méthode du maximum de vraisemblance consiste à déterminer la valeur du paramètre θ en fonction des observation (x_1, x_2, \dots, x_T) et qui assure la plus grande probabilité l'apparition de ces observations .

Définition 2.2. Soit X une variable aléatoire de loi P_θ , on note $f(X, \theta)$ la densité de P_θ et $f(x_1, x_2, \dots, x_T; \theta)$ la densité empirique correspondante. On appelle fonction de vraisemblance associée a un échantillon x_1, x_2, \dots, x_T la fonction définie par :

$$L(\theta) = L(x_1, x_2, \dots, x_T; \theta) = \prod_{i=1}^T f(x_i, \theta) = f(X, \theta)$$

L'estimateur du maximum de vraisemblance (EMV) est la valeur $\hat{\theta}_{EMV}$ qui maximise $\ln L(\theta)$ telle que :

$$\begin{aligned} \hat{\theta} &= \arg \max_{\theta \in \Theta} f(X, \theta) \\ &= \arg \max_{\theta \in \Theta} \ln L(x_1, x_2, \dots, x_T; \theta) \end{aligned}$$

Il est possible de déduire cet estimateur en dérivant $\ln L(\theta)$ par rapport à chacun des paramètres $(\theta_1, \dots, \theta_k)$ et de résoudre le système à k équations

$$\begin{cases} \frac{\partial \ln L(\theta)}{\partial \theta_j} = 0 \\ \frac{\partial^2 \ln L(\theta)}{\partial \theta_j^2} < 0 \end{cases}$$

2.2.1 La méthode du maximum de vraisemblance :

La méthode de Yule-Walker ne permet pas d'estimer de la façon efficace les paramètres d'un modèle contenant une partie MA parce qu'elle mène à un système d'équations non linéaires.

Dans cette section, nous allons appliquer, comme méthode alternative, la méthode du maximum de vraisemblance pour estimer les paramètres d'un modèle ARMA.

Considérons un modèle canonique (ie :causal et inversible) ARMA(p, q) de moyenne nulle

$$X_t = \varphi_1 X_{t-1} + \dots + \varphi_p X_{t-p} + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q} \quad (2.6)$$

$$(\varepsilon_t)_t \sim WN(0, \sigma_\varepsilon^2)$$

où $\theta_0 = 1$ et dans lequel on supposera que :

$$1 - \varphi_1 z - \dots - \varphi_p z^p \neq 0 \quad \text{pour } |z| \leq 1$$

et

$$1 + \theta_1 z + \dots + \theta_q z^q \neq 0 \quad \text{pour } |z| \leq 1$$

Notre objectif est à présent d'estimer les vecteurs colonnes $\boldsymbol{\varphi}_p = (\varphi_1, \dots, \varphi_p)'$ et $\boldsymbol{\theta} = (\theta_1, \dots, \theta_q)'$, σ_ε^2 .

Dans un premier temps, supposons que l'on connaisse la distribution de X_t pour tout t . Nous supposons que le processus (X_t) est gaussien, cette hypothèse nous permettra d'écrire simplement la fonction de vraisemblance du processus gaussien telle que :

$$L(R_T^X) = \frac{1}{(2\pi\sigma_\varepsilon^2)^{T/2} |\det R_T^X|^{1/2}} \exp\left(-\frac{1}{2} X_T' (R_T^X)^{-1} X_T\right)$$

où R_T^X est la matrice d'auto-covariance dont l'élément (i, j) est donné par $E(X_i X_j) = \gamma_{|i-j|}^X$.

Or, nous avons vu au chapitre précédent que la fonction d'auto-covariance pouvait être exprimée, par la relation :

$$\gamma_h^X = \sigma_\varepsilon^2 \sum_{k=0}^{\infty} \psi_k \psi_{k+h}$$

Proposition 2.3. *Si $(X_t)_t$ est un processus causal et inversible de la forme (2.6) avec $\varepsilon_t \sim i.i.d(0, \sigma_\varepsilon^2)$, et si $\beta = (\boldsymbol{\varphi}_p; \boldsymbol{\theta}_q)$, l'estimateur du maximum de vraisemblance $\hat{\beta} = (\hat{\boldsymbol{\varphi}}_p; \hat{\boldsymbol{\theta}}_q)$ est asymptotiquement normal.*

$$\sqrt{T}(\hat{\beta} - \beta) \rightarrow \mathcal{N}(0, V(\beta))$$

où la matrice variance-covariance $V(\beta)$ est une fonction du paramètre β .

2.2.2 L'estimation de la fonction de vraisemblance d'un processus AR(p) :

Supposons que la série chronologique X_1, X_2, \dots, X_T soit identifiée comme étant un modèle AR(p) défini tel que :

$$X_t - \varphi_1 X_{t-1} - \dots - \varphi_p X_{t-p} = \varepsilon_t \tag{2.7}$$

où $(\varepsilon_t)_{t \in \mathbb{Z}} \sim WN(0, \sigma_\varepsilon^2)$ est un bruit blanc de moyenne 0 et de variance σ_ε^2

Soit $\boldsymbol{\varphi} = (\varphi_1, \dots, \varphi_p)'$ le vecteur des paramètres du modèle et R_T^X la matrice d'auto-covariance du vecteur $X = (X_1, X_2, \dots, X_T)'$.

La densité jointe du vecteur X est définie par :

$$f(X/\boldsymbol{\varphi}, \sigma_\varepsilon^2) = (2\pi\sigma_\varepsilon^2)^{-\frac{T}{2}} (\det(R_T^X)^{-1})^{\frac{1}{2}} \exp\left(\frac{-1}{2\sigma_\varepsilon^2} X'(R_T^X)^{-1} X\right) \quad (2.8)$$

d'après l'axiome de probabilité conditionnelle, on peut écrire la relation (2.8) sous la forme :

$$f(X/\boldsymbol{\varphi}, \sigma_\varepsilon^2) = f(X_{p+1}, X_{p+2}, \dots, X_T/X^{(p)}, \boldsymbol{\varphi}, \sigma_\varepsilon^2) f(X^{(p)}/\boldsymbol{\varphi}, \sigma_\varepsilon^2)$$

Où $X^{(p)} = (X_1, X_2, \dots, X_p)'$ et $f(X^{(p)}/\boldsymbol{\varphi}, \sigma_\varepsilon^2)$ est la densité jointe des p premières observations de X telle que :

$$f(X^{(p)}/\boldsymbol{\varphi}, \sigma_\varepsilon^2) = (2\pi\sigma_\varepsilon^2)^{-\frac{p}{2}} (\det(R_p^X)^{-1})^{\frac{1}{2}} \exp\left(\frac{-1}{2\sigma_\varepsilon^2} (X^{(p)})'(R_p^X)^{-1} X^{(p)}\right)$$

où R_p^X est la matrice de auto-covariance de X_1, X_2, \dots, X_p .

La densité jointe $f(X_{p+1}, X_{p+2}, \dots, X_T/X^{(p)}, \boldsymbol{\varphi}, \sigma_\varepsilon^2)$, peut être exprimée telle que :

$$f(X_{p+1}, X_{p+2}, \dots, X_T/X^{(p)}, \boldsymbol{\varphi}, \sigma_\varepsilon^2) = (2\pi\sigma_\varepsilon^2)^{-\frac{T-p}{2}} \exp\left(\frac{-1}{2\sigma_\varepsilon^2} \sum_{t=p+1}^T (X_t - \varphi_1 X_{t-1} - \dots - \varphi_p X_{t-p})^2\right)$$

Finalement la fonction de vraisemblance $L_T(X/\boldsymbol{\varphi}, \sigma_\varepsilon^2)$ s'écrit telle que :

$$L_T(X/\boldsymbol{\varphi}, \sigma_\varepsilon^2) = (2\pi\sigma_\varepsilon^2)^{-\frac{T}{2}} (\det(R_p^X)^{-1})^{\frac{1}{2}} \exp\left(\frac{-1}{2\sigma_\varepsilon^2} S(\boldsymbol{\varphi})\right)$$

où :

$$S(\boldsymbol{\varphi}) = \sum_{i=1}^p \sum_{j=1}^p m_{ij}^{(p)} X_i X_j + \sum_{t=p+1}^T (X_t - \varphi_1 X_{t-1} - \dots - \varphi_p X_{t-p})^2$$

d'après ([7]) les éléments $m_{ij}^{(p)}$ de la matrice inverse $(R_p^X)^{-1}$ peuvent être calculer par :

$$m_{ij}^{(p)} = \sum_{k=0}^{i-1} \varphi_k \varphi_{k+j+i} - \sum_{k=p+1-j}^{p+i-j} \varphi_k \varphi_{k+j-i} \quad 1 \leq i \leq j \leq p$$

où $\varphi_0 = -1$

On calcule les valeurs $m_{ij}^{(p)}$ pour $i > j$ puisque $(R_p^X)^{-1}$ est symétrique i.e :

$$m_{ij}^{(p)} = m_{ji}^{(p)}.$$

Le logarithme de la fonction de vraisemblance est :

$$\ln L_T(X/\boldsymbol{\varphi}, \sigma_\varepsilon^2) = \frac{-T}{2} \ln(2\pi\sigma_\varepsilon^2) + \frac{1}{2} \ln(\det(R_p^X)^{-1}) - \frac{S(\boldsymbol{\varphi})}{2\sigma_\varepsilon^2}$$

Le problème d'estimation des paramètres par la méthode du maximum de vraisemblance est résolu à l'aide des équations de Yule-Walker ([1],[2]).

2.2.3 L'estimation de la fonction de vraisemblance d'un processus MA(q)

Supposons que la série chronologique X_1, X_2, \dots, X_T , soit identifiée comme étant un modèle MA(q) défini tel que :

$$X_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q} \quad (2.9)$$

où $(\varepsilon_t)_t$ est un bruit blanc de loi $\mathcal{N}(0, \sigma_\varepsilon^2)$.

Soit $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_q)'$ le vecteur des paramètres du modèle. Alors la densité jointe du vecteur $X = (X_1, X_2, \dots, X_T)'$ peut être exprimée telle que :

$$f(X/\boldsymbol{\theta}, \sigma_\varepsilon^2) = (2\pi\sigma_\varepsilon^2)^{-\frac{T}{2}} (\det(R_T^X))^{-\frac{1}{2}} \exp\left(\frac{-1}{2\sigma_\varepsilon^2} X'(R_T^X)^{-1} X\right)$$

où R_t^X est la matrice des auto-covariance de $X = (X_1, X_2, \dots, X_T)'$ pour un processus MA(q).

et

$$S(\boldsymbol{\theta}) = X'(R_T^X)^{-1} X = \frac{1}{\sigma_\varepsilon^2} \sum_{k=-\infty}^T \hat{\varepsilon}_k^2 \quad ; \quad \hat{\varepsilon}_k = E(\varepsilon_k / X_1, X_2, \dots, X_T)$$

La forme quadratique dans la fonction de vraisemblance est égale à la somme des carrés exact $S(\boldsymbol{\theta}) = \sum_{k=-\infty}^T \hat{\varepsilon}_k^2$. Il a été établi dans ([1],[2]) que pour T assez grand, la fonction de vraisemblance est dominée par le terme $\frac{1}{2\sigma_\varepsilon^2} S(\boldsymbol{\theta})$, et ainsi les estimateurs des paramètres $\theta_1, \theta_2, \dots, \theta_q, \sigma_\varepsilon^2$ sont obtenus en maximisant la log-vraisemblance ou ce qui revient au même, en minimisant la somme des carrés $S(\boldsymbol{\theta})$.

De plus, il a été établi dans ([1],[2],[6]) que la somme des carrés exacte $S(\boldsymbol{\theta})$ est équivalente à la somme des carrés $\sum_{t=1-q}^T \hat{\varepsilon}_t^2$, appelée somme des carrés inconditionnelle.

A l'aide de l'équation (2.9) et pour $t = 1, 2, \dots, T$, on obtient un système d'équations. En posant $\varepsilon = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_T)'$ et $\varepsilon^{(0)} = (\varepsilon_{1-q}, \varepsilon_{2-q}, \dots, \varepsilon_0)'$ le vecteur des q valeurs initiales, ce système d'équations peut être écrit sous forme matricielle suivante :

$$X = L_\theta \varepsilon + F \varepsilon^{(0)} \quad (2.10)$$

où L_{θ} est une matrice de dimension $T \times T$ triangulaire inférieure et F est une matrice de dimension $T \times q$, définies telles que :

$$L_{\theta} = \begin{pmatrix} 1 & 0 & \dots & \dots & \dots & \dots & \dots & 0 \\ -\theta_1 & 1 & 0 & \dots & \dots & \dots & \dots & 0 \\ -\theta_2 & \theta_1 & 1 & 0 & \dots & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ -\theta_q & \theta_{q-1} & \dots & -\theta_1 & 1 & 0 & \dots & 0 \\ 0 & 0 & \dots & \dots & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & 0 \\ 0 & \dots & \dots & \dots & \dots & \dots & 0 & 1 \end{pmatrix}$$

et

$$F = (\Theta', 0')'$$

où

$$\Theta = - \begin{pmatrix} \theta_q & \theta_{q-1} & \dots & \theta_1 \\ 0 & \theta_q & \dots & \theta_2 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \theta_q \end{pmatrix}$$

La distribution jointe du vecteur $(\varepsilon', \varepsilon^{(0)'})$ de taille $T \times q$ peut s'exprimer telle que :

$$f(\varepsilon', \varepsilon^{(0)'}/\sigma_{\varepsilon}^2) = (2\pi\sigma_{\varepsilon}^2)^{-\frac{(T+q)}{2}} \exp\left(\frac{-1}{2\sigma_{\varepsilon}^2}(\varepsilon'\varepsilon + \varepsilon^{(0)'}\varepsilon^{(0)})\right)$$

Considérons la transformation linéaire $(\varepsilon, \varepsilon^{(0)}) \rightarrow (X, \varepsilon^{(0)})$ de jacobien unité avec $\varepsilon = L_{\theta}^{-1}(X - F\varepsilon^{(0)})$. Alors la distribution jointe de $(X, \varepsilon^{(0)})$ est :

$$f(X, \varepsilon^{(0)}/\sigma_{\varepsilon}^2) = (2\pi\sigma_{\varepsilon}^2)^{-\frac{(T+q)}{2}} \exp\left(\frac{-1}{2\sigma_{\varepsilon}^2}S(\theta, \varepsilon^{(0)})\right)$$

où

$$S(\theta, \varepsilon^{(0)}) = (X - F\varepsilon^{(0)})'(L'_{\theta})^{-1}L_{\theta}^{-1}(X - F\varepsilon^{(0)}) + \varepsilon^{(0)'}\varepsilon^{(0)} \quad (2.11)$$

On remarque que $S(\theta, \varepsilon^{(0)})$ dépend du vecteur des valeur initiales $\varepsilon^{(0)}$ à déterminer.

Pour cela, notons par $\hat{\varepsilon}^{(0)}$ le vecteur qui minimise $S(\theta, \varepsilon^{(0)})$. En en utilisant la méthode des moindre carrés généralisée on trouve :

$$\hat{\varepsilon}^{(0)} = D^{-1}F'(L'_{\theta})^{-1}L_{\theta}^{-1}X \quad (2.12)$$

où

$$D = I_q + F'(L'_{\theta})^{-1}L_{\theta}^{-1}F \quad (2.13)$$

La fonction de vraisemblance inconditionnelle est donnée par :

$$L_T(\boldsymbol{\theta}, \sigma_\varepsilon^2 / X) = (2\pi\sigma_\varepsilon^2)^{-\frac{T}{2}} (\det D)^{-\frac{1}{2}} \exp\left[\frac{-1}{2\sigma_\varepsilon^2} \sum_{t=1-q}^T [\varepsilon_t]^2\right]$$

Pour déterminer q valeurs initiales $(\varepsilon_{1-q}, \dots, \varepsilon_{-1}, \dots, \varepsilon_0)$ en utilisant le principe de l'algorithme aller-retour proposé par Box-Jenkins ([1],[2],[6]).

2.2.4 L'estimation de la fonction de vraisemblance d'un processus ARMA(p,q)

Supposons que la série chronologique X_1, X_2, \dots, X_T , soit identifiée comme étant un modèle ARMA(p,q) défini tel que :

$$X_t - \sum_{i=1}^p \varphi_i X_{t-i} = \varepsilon_t + \sum_{i=1}^q \theta_i \varepsilon_{t-i} \quad (2.14)$$

La relation (2.14) peut être écrit sous la forme matricielle comme suit :

$$L_\varphi X = L_\theta \varepsilon + F e_*$$

L_θ est la matrice triangulaire inférieur de dimension $T \times T$ avec 1 sur la diagonale principale, $-\theta_1$ sur la première sub diagonale, $-\theta_2$ est la deuxième sub diagonale , $\theta_i = 0$ pour $i > q$ et L_φ est une matrice de $T \times T$ qui a la même forme de L_θ mais avec φ_i à la place de θ_i .

$e'_* = (X^{(0)}, \varepsilon^{(0)})' = (X_{1-p}, \dots, X_0, \varepsilon_{1-q}, \dots, \varepsilon_0)'$ est le vecteur de dimension $p + q$ des vecteurs initiales et F est la matrice définie par :

$$F = \begin{pmatrix} \tilde{\Phi} & \Theta \\ 0 & 0 \end{pmatrix}$$

Avec

$$\tilde{\Phi} = \begin{pmatrix} \varphi_p & \varphi_{p-1} & \dots & \dots & \varphi_1 \\ 0 & \varphi_p & \dots & \dots & \varphi_1 \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \dots & \varphi_p \end{pmatrix}$$

et

$$\Theta = - \begin{pmatrix} \theta_q & \theta_{q-1} & \dots & \dots & \theta_1 \\ 0 & \theta_q & \dots & \dots & \theta_2 \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \dots & \theta_q \end{pmatrix}$$

La matrice Ω est la matrice d'auto-covariance de e_* qui s'exprime sous la forme suivante :

$$\Omega = \begin{pmatrix} \sigma_\varepsilon^{-2} J_p & C' \\ C & I_q \end{pmatrix}$$

Où $J_p = E[X^{(0)} X'^{(0)}]$ est la matrice de dimension $p \times p$ de terme général $\gamma_{|i-j|}^{(0)}$, et où $\sigma_\varepsilon^2 C = E[\varepsilon^{(0)} X'^{(0)}]$ une matrice ayant pour terme général

$$E[\varepsilon_{i-q}, X_{j-q}] = \begin{cases} \sigma_\varepsilon^2 \psi_{j-i-p-q} & \text{si } j - i - p - q \geq 0 \\ 0 & \text{sinon} \end{cases}$$

Les ψ_k sont les coefficients de la forme moyennes mobiles infinies d'un ARMA(p,q), i.e

$$X_t = \psi(B)\varepsilon_t = \varphi^{-1}(B)\theta(B)\varepsilon_t = \sum_{k=0}^{\infty} \psi_k \varepsilon_{t-k}$$

Puisque $\varepsilon = L_\theta^{-1}(L_\varphi X - F e_*)$ et le vecteur e_* sont indépendants, alors la distribution jointe de X et e_* est

$$f(X, e_*/\varphi, \theta, \sigma_\varepsilon^2) = (2\pi\sigma_\varepsilon^2)^{-\frac{(T+p+q)}{2}} (\det \Omega)^{-\frac{1}{2}} \exp\left[\frac{-1}{2\sigma_\varepsilon^2} S(\varphi, \theta, e_*)\right]$$

où

$$S(\varphi, \theta, e_*) = (L_\varphi X - F e_*)'(L'_\theta)^{-1} L_\theta^{-1} (L_\varphi X - F e_*) + e_*' \Omega^{-1} e_*$$

D'après la méthode des moindres carrés généralisée ([2]), nous pouvons écrire

$$S(\varphi, \theta, e_*) = S(\varphi, \theta) + (e_* - \hat{e}_*)' D (e_* - \hat{e}_*)$$

où

$$\begin{aligned} S(\varphi, \theta) &= S(\varphi, \theta, \hat{e}_*) = \hat{\varepsilon}' \hat{\varepsilon} + \hat{e}_*' \Omega^{-1} \hat{e}_* \\ &= \sum_{t=1}^T [\varepsilon_t]^2 + \hat{e}_*' \Omega^{-1} \hat{e}_* \end{aligned} \quad (2.15)$$

Par ailleurs

$$\hat{e}_* = E(e_*/X^{(0)}, \varepsilon^{(0)}) = [e_*] = D^{-1} F' (L'_\theta)^{-1} L_\theta^{-1} L_\varphi X \quad (2.16)$$

où

$$D = \Omega^{-1} + F'(L'_\theta)^{-1} L_\theta^{-1} F$$

De plus

$$\hat{\varepsilon} = E(\varepsilon/X^{(0)}, \varepsilon^{(0)}) = [\varepsilon] = L_\theta^{-1}(L_\varphi X - F\hat{\varepsilon}_*)$$

La log-vraisemblance d'un ARMA(p,q) peut être exprimé telle que :

$$\ln L_T(X/\varphi, \theta, \sigma_\varepsilon^2) = \frac{-T}{2} \ln(2\pi\sigma_\varepsilon^2) - \frac{1}{2} \ln(\det \Omega)(\det D) - \frac{1}{2\sigma_\varepsilon^2} S(\varphi, \theta)$$

Il est évident que pour calculer la somme des carrés inconditionnelle $S(\varphi, \theta)$ on a besoin de déterminer $p + q$ valeurs initiales .

$(X_0, X_{-1}, \dots, X_{1-p}, \varepsilon_0, \varepsilon_1, \dots, \varepsilon_{1-q})$ on utilisant le principe de l'algorithme aller-retour proposé par Box-Jenkins ([1],[2],[3],[6]).

2.3 Test de compatibilité du modèle

Une fois les paramètres du modèle sont estimés on procède à un test de vérification pour le modèle estimé (reproduit le modèle qui a engendré les données). Pour cela les résidus obtenus à partir du modèle estimé sont utilisés pour vérifier s'ils se comportent comme des erreurs bruit blanc à l'aide d'un test "portemanteau" (test global qui permet de tester l'hypothèse d'indépendance des résidus ie : $Cov = (\varepsilon_t, \varepsilon_{t-h}) = 0$ pour tout h). Le test le plus usuel est le test proposé par Box-Pierce. Ce test s'écrit :

$$\begin{cases} H_0 : \rho_1 = \rho_2 = \dots = \rho_h = 0 \\ H_1 : \exists i \neq j \text{ tel que } \rho_i \neq \rho_j \neq 0 \end{cases}$$

Pour effectuer ce test, on utilise la statistique Q définie par :

$$Q = T \sum_{h=1}^s \hat{\rho}_h^2$$

$\hat{\rho}_h^2$ étant la corrélation empirique entre les résidus, distants de h.

Sous l'hypothèse d'indépendance, cette statistique suit asymptotiquement une loi de χ^2 avec s-p-q degrés de liberté. Si l'hypothèse n'est pas rejetée on dira que le modèle estimé est bonne approximation du vrai modèle. Dans le cas contraire on rejettera le modèle estimé et on procédera à identifier un nouveau modèle.

En échange on utilise aussi le test de Ljung-Box, la statistique de Ljung-Box est définie

par :

$$Q = T(T + 2) \sum_{h=1}^s \frac{\hat{\rho}_h^2}{T - h}$$

T : le nombre d'observations

s : nombre de résidus auxquels on test l'auto corrélations.

$\hat{\rho}_h^2$: coefficients d'auto corrélations (pour le décalage h)

Q : test de portemanteau statistique.

Si la valeur de Q pour la série chronologique (l'échantillon) dépasse la valeur critique d'une distribution Chi-deux avec s degré de liberté alors au moins une valeur de $\hat{\rho}^2$ est différente de zéro au degré significative précisé, ce qui mène à rejeter l'hypothèse d'indépendance des résidus et par conséquent le rejet de modèle, sinon le modèle constitue une bonne représentation pour le processus stochastique.

2.3.1 test de la stationnarité ADF

•**Test de Dickey Fuller simple** : Dickey et Fuller sont les premiers qui ont fournit un ensemble d'outils statistiques formels pour détecter la présence d'une racine unitaire dans un processus auto-régressif du premier ordre, ce test permet de tester l'hypothèse

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

Ce test est regroupée en 4 cas :

$$X_t = \rho X_{t-1} + \varepsilon_t; \quad H_0 : \rho = 1$$

$$X_t = \alpha + \rho X_{t-1} + \varepsilon_t; \quad H_0 : \alpha = 0 \quad \text{et} \quad \rho = 1$$

$$X_t = \alpha + \rho X_{t-1} + \varepsilon_t; \quad H_0 : \alpha \neq 0 \quad \text{et} \quad \rho = 1$$

$$X_t = \alpha + \beta t + \rho X_{t-1} + \varepsilon_t; \quad H_0 : \alpha = 0, \beta = 0 \quad \text{et} \quad \rho = 1$$

•**Test de Dickey Fuller augmenté** : Dickey et Fuller en 1981 étendent ensuite cette procédure de test à des processus autorégressifs d'ordre p, il s'agit alors des tests ADF «Augmented Dickey-Fuller ».

Ce test permet de tester

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

Ce test peuvent être regroupés en 4 cas :

$$X_t = \rho X_{t-1} + \sum_{i=1}^p \alpha_i X_{t-i} + \varepsilon_t \quad H_0 : \rho = 1$$

$$X_t = \alpha + \rho X_{t-1} + \sum_{i=1}^p \alpha_i X_{t-i} + \varepsilon_t \quad H_0 : \alpha = 0 \quad \text{et} \quad \rho = 1$$

$$X_t = \alpha + \rho X_{t-1} + \sum_{i=1}^p \alpha_i X_{t-i} + \varepsilon_t \quad H_0 : \alpha \neq 0 \quad \text{et} \quad \rho = 1$$

$$X_t = \alpha + \beta t + \rho X_{t-1} + \sum_{i=1}^p \alpha_i X_{t-i} + \varepsilon_t \quad H_0 : \alpha = 0, \beta = 0 \quad \text{et} \quad \rho = 1$$

Pour simplifier, on écrira les trois cas général :

$$\Delta X_t = \phi X_{t-1} + \sum_{i=1}^p \alpha_i \Delta X_{t-i} + \varepsilon_t \quad \text{avec} \quad \phi = \rho - 1$$

$$\Delta X_t = \alpha + \phi X_{t-1} + \sum_{i=1}^p \alpha_i \Delta X_{t-i} + \varepsilon_t$$

$$\Delta X_t = \alpha + \beta t + \phi X_{t-1} + \sum_{i=1}^p \alpha_i \Delta X_{t-i} + \varepsilon_t$$

2.3.2 Test de normalité

Il s'agit de tester que les résidus estimés $\hat{\varepsilon}_t$ suivent une loi normale, c'est-à-dire ne présentent pas d'asymétrie (Skewness) ni d'aplatissement (kurtosis).

Le coefficient de Skewness est donné par :

$$\beta_1^{1/2} = \frac{\mu_3}{\mu_2^{3/2}}$$

et le coefficient de kurtosis est donné par :

$$\beta_2 = \frac{\mu_4}{\mu_2^2}$$

où

$$\mu_k = \frac{1}{T} \sum_{t=1}^T (\hat{\varepsilon}_t - \bar{\hat{\varepsilon}})^k$$

est le moment centré d'ordre k de la variable $\hat{\varepsilon}_t$:

Si la distribution est normale et le nombre d'observations est grand, alors :

$$\beta_1^{1/2} \sim \mathcal{N}(0, \sqrt{6/T}) \quad \text{et} \quad \beta_2 \sim \mathcal{N}(3, \sqrt{24/T})$$

On construit alors les statistiques :

$$\nu_1 = \frac{\beta_1^{1/2}}{\sqrt{6/T}} \quad \text{et} \quad \nu_2 = \frac{\beta_2 - 3}{\sqrt{24/T}}$$

qui suivent chacune une $\mathcal{N}(0, 1)$. Le test de Jarque Bera permet de tester simultanément l'absence d'asymétrie et l'absence d'aplatissement (ie : la normalité des résidus). La statistique de test est donnée par :

$$JB = \frac{T}{6}\beta_1 + \frac{T}{24}(\beta_2 - 3)^2$$

Cette statistique suit, sous l'hypothèse nulle de normalité, une loi du χ^2

2.3.3 Test sur les résidus

Les tests, concernant le bruit blanc, ont pour but de vérifier si l'estimateur des résidus est bien cohérente avec les hypothèses relatives au bruit blanc.

L'analyse de la fonction d'auto-corrélation permet de vérifier si on est en présence d'un bruit blanc.

Lorsque T est assez grand, les auto-corrélations d'un bruit blanc sont approximativement indépendantes et de loi $N(0, \frac{1}{T})$. Ainsi 95% des auto-corrélations devraient se trouver dans l'intervalle

$$[-1.96/\sqrt{T}; 1.96/\sqrt{T}]$$

2.4 Choix d'un modèle

Revenons sur le choix de l'ordre d'un ARMA(p,q). Nous avons vu précédemment que l'étude des auto-corrélations et auto-corrélations partielles peuvent permettre de sélectionner un certain nombre des modèles, une fois les paramètres de ces modèles sont estimés on peut sélectionner celui qui minimise les critères suivants :

• **AIC** : (Akaike Information Contrite), qui sera généralement préféré si l'objectif de l'étude est de faire de la prévision, et qui est défini par :

$$AIC(p, q) = -2 \log L + 2(p + q)$$

• **BIC** : (Bayésien Information Contrite), sera quant à lui généralement préféré si l'objectif de l'étude est d'ajuster à la série observée, et est défini par :

$$BIC(p, q) = -2 \log L + (p + q) \log T$$

Où $\log L$ est la log-vraisemblance du modèle ARMA(p,q) estimé

Les modèles ayant la plus petite valeur du critère devront être choisis.

2.5 Critères d'erreur

Dans l'échantillon initial, (X_1, \dots, X_t) on considère seulement une partie d'observations. Le reste sera à prévoir par le modèle. On peut alors considérer les erreurs :

-Racine quadratique moyenne :

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (X_i^* - X_i)^2}$$

-Quadratique moyenne :

$$MSE = \frac{1}{n} \sum_{i=1}^n (X_i^* - X_i)^2$$

-Absolue :

$$MAE = \frac{1}{n} \sum_{i=1}^n |X_i^* - X_i|$$

2.6 Prévisions

On ce qui concerne la prévision on est censé de connaître une future valeur désirée la plus proche possible de la vraie valeur X_{T+h} ; $h > 0$ au moment où on est à l'instant T. on cherche donc à prévoir la valeur que va prendre le processus à l'instant $T + h$, $h \geq 1$.

Il s'agit de calculer les prévisions du modèle ARMA estimé, à savoir $X_T^*(h)$ la prévision de X_{T+h} sachant l'ensemble d'information disponible en T noté :

$$X_T^*(h) = E(X_{T+h}/X_T, X_{T-1}, \dots, X_1)$$

Dans le cadre de la prévision de l'ARMA on choisit de prendre comme prédicteur une combinaison linéaire de ces X_i ; $1 \leq i \leq T$. Autrement dit :

$$X_T^*(h) = a_1 X_1 + a_2 X_2 + \dots + a_t X_T$$

où a_1, a_2, \dots sont réelles Ce qui est équivalent à une combinaison linéaire des ε_{t-i} ; $i \geq 0$ du fait que pour toutes t, X_t s'écrit comme combinaison linéaire des ε_{t-i} ; $i \geq 0$.

on a :

$$X_t = \sum_{i=0}^{+\infty} \psi_i \varepsilon_{t-i}$$

ce qui permet d'écrire :

$$\begin{aligned} X_T^*(h) &= \sum_{k=0}^{+\infty} \psi_k^* \varepsilon_{T-k} \\ &= \sum_{i=h}^{+\infty} \psi_i^* \varepsilon_{T+h-i} \end{aligned}$$

Dans le cadre choisit il est bien naturel de prendre celle de ces combinaisons linéaires qui est la plus proche de la variable aléatoire X_{T+h} . Dans notre travail on prend la combinaison linéaire qui minimise l'erreur moyenne quadratique définie par :

$$E[X_{T+h} - X_T^*(h)]^2 = E\left(X_{T+h} - \sum_{i=h}^{+\infty} \psi_j^* X_{T+h-i}\right)^2$$

et pour définir le prédicteur il suffit de calculer les valeurs des $\psi_i^*, i \geq h$.

2.7 Prévision dans les ARMA(p,q)

2.7.1 Prévision à l'aide d'un modèle AR(p)

Le modèle s'écrit,

$$X_t = \varphi_1 X_{t-1} + \dots + \varphi_p X_{t-p} + \varepsilon_t$$

où

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

La prévision optimale à la date $T+1$, faite à la date T est $X_{T+1}^* = E(X_{T+1}/X_T, X_{T-1}, \dots, X_1)$.

Aussi,

$${}_T X_{T+1}^* = \varphi_1 X_T + \dots + \varphi_p X_{T+1-p}$$

De façon analogue, $X_{T+h} = \varphi_1 X_{T+h-1} + \dots + \varphi_p X_{T+h-p} + \varepsilon_{T+h}$, et donc

${}_T X_{T+1}^* = E(X_{T+1}/X_T, X_{T-1}, \dots, X_1)$ est donné de façon récursive par :

$${}_T X_{T+h}^* = \begin{cases} \varphi_1 \cdot {}_T X_{T+h-1}^* + \dots + \varphi_{h-1} \cdot {}_T X_{T+1}^* + \varphi_h X_T + \dots + \varphi_p X_{T+h-p} & \text{pour } h \leq p \\ \varphi_1 \cdot {}_T X_{T+h-1}^* + \dots + \varphi_p \cdot {}_T X_{T+h-p}^* & \text{pour } h > p \end{cases}$$

2.7.2 Prévision à l'aide d'un modèle MA(q)

Soit le modèle MA(q) :

$$X_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}$$

La prévision optimale à la date $T+1$, faite à la date T est $X_{T+1}^* = E(X_{T+1}/X_T, X_{T-1}, \dots, X_1) = X_{T+1}^* = E(X_{T+1}/\varepsilon_T, \varepsilon_{T-1}, \dots, \varepsilon_1)$. Aussi,

$${}_T X_{T+1}^* = 0 + \theta_1 \varepsilon_T + \dots + \theta_q \varepsilon_{T+1-q}$$

De façon analogue, X_{T+h} est estimé par ${}_T X_{T+h}^* = E(X_{T+h}/X_T, X_{T-1}, \dots, X_1) = X_{T+h}^* = E(X_{T+h}/\varepsilon_T, \varepsilon_{T-1}, \dots, \varepsilon_1)$, est donc

$${}_T X_{T+h}^* = \begin{cases} \theta_h X_T + \dots + \theta_q X_{T+h-q} & \text{pour } h \leq q \\ 0 & \text{pour } h > q \end{cases} \quad (2.17)$$

2.7.3 Prévision à l'aide d'un processus ARMA(p,q)

$(X_t)_t$ est un processus ARMA(p,q), alors

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

Donc,

$$X_{T+h} = \sum_{i=1}^p \varphi_i X_{T+h-i} + \varepsilon_{T+h} + \sum_{j=1}^q \theta_j \varepsilon_{T+h-j}$$

On peut noter que pour $h > q$

$${}_{T+1} X_{T+h}^* = \begin{cases} \varphi_{1 \cdot T} X_{T+h-1}^* + \dots + \varphi_{h-1 \cdot T} X_{T+1}^* + \varphi_p X_{T+h-p} & \text{pour } h \leq p \\ \varphi_{1 \cdot T} X_{T+h-1}^* + \dots + \varphi_p X_{T+h-p} & \text{pour } h > p \end{cases}$$

Chapitre 3

Application

Nous intéressons à appliquer la méthode de Box-Jenkins (voir la figure 3.1) sur des données réelles notée "dialyse" qui représente le nombre des séances d'hémodialyse tirée de l'hôpital Ibn Sina d'Adrar, les données considérées sont mensuelles et la période retenue pour l'étude entre janvier 2011 et août 2014.

Mois /années	2011	2012	2013	2014
janvier	582	606	776	875
février	510	555	709	786
mars	573	588	769	912
avril	417	623	729	992
mai	545	673	651	871
juin	625	638	749	901
juillet	656	706	771	925
août	554	619	744	854
septembre	545	661	758	
octobre	593	707	777	
novembre	529	707	730	
décembre	592	738	778	

TABLE 3.1 – Nombre des séances d'hémodialyse

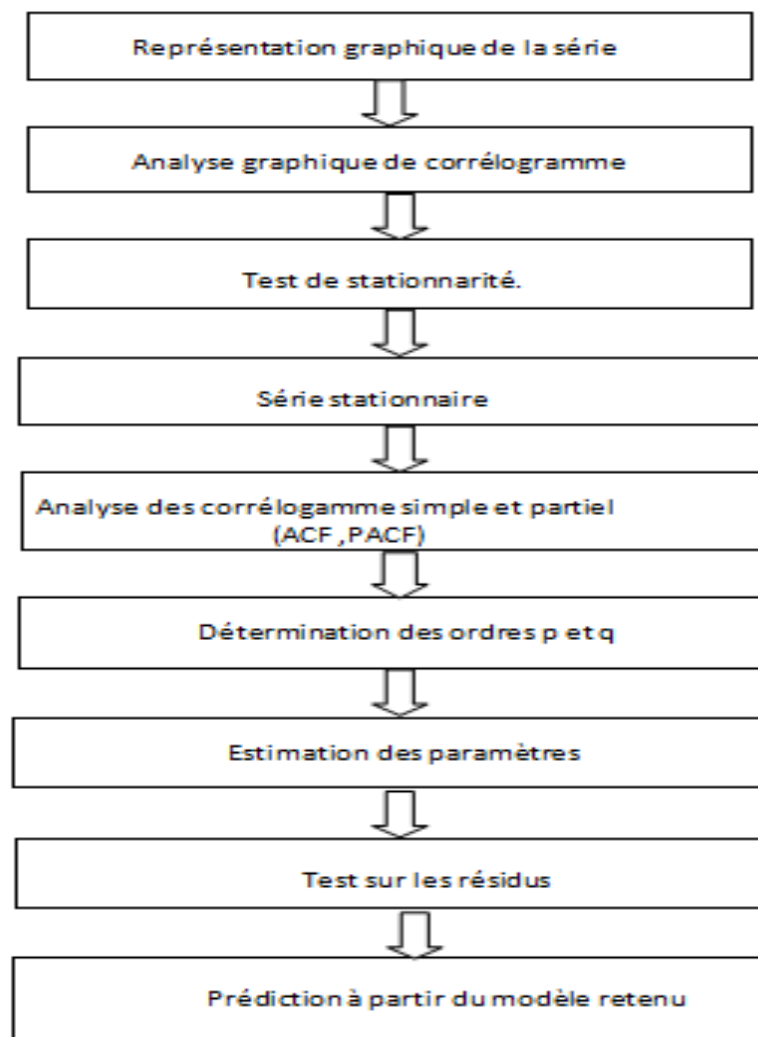


FIGURE 3.1 – Procédure de construction de la méthode Box-Jenkins.

3.1 Analyse préliminaire de la série (dialyse)

La série (dialyse) contient 44 observations (janvier 2011 ; août 2014, données mensuelles).

3.1.1 Représentation graphique de la série dialyse

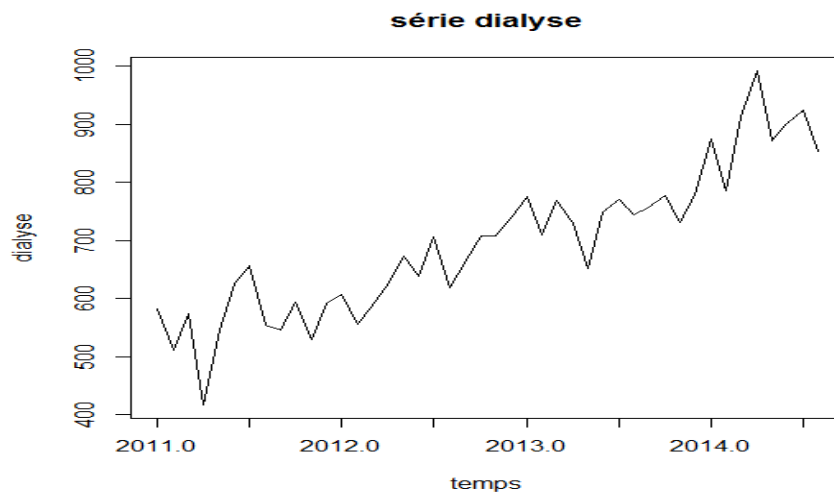


FIGURE 3.2 – le graphe de la série dialyse

Le graphe de la série dialyse montre que la série est non stationnaire caractérisée par une tendance. On utilise aussi le test de racine unitaire (test ADF) pour confirmer la non stationnarité .

La série a une racine unitaire car la p-value = 0.159 > 0.05, (la probabilité critique affectées à la constante est supérieur à 0.05).

3.1.2 corrélogramme et corrélogramme partiel de la série dialyse

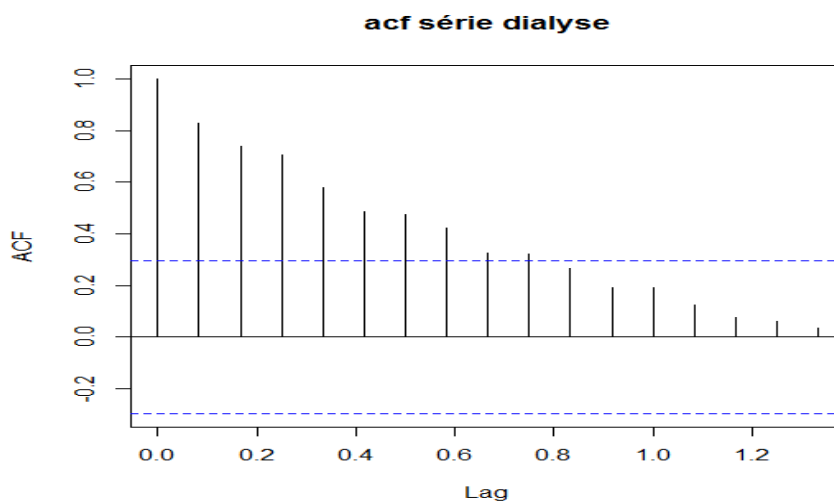


FIGURE 3.3 – Le corrélogramme (acf) de la série dialyse

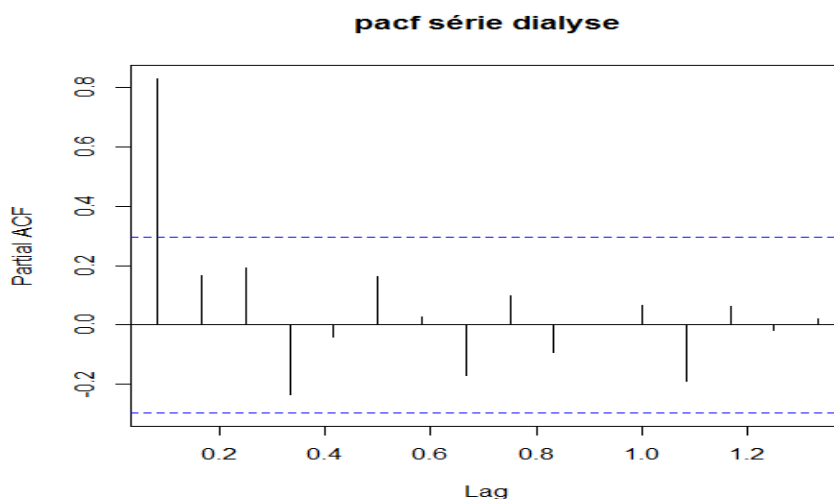


FIGURE 3.4 – Le corrélogramme partiel (pacf) de la série dialyse

D'après les graphes de corrélogramme et corrélogramme partiel, nous remarquons que la série dialyse est non stationnaire, puisque la fonction d'auto-corrélation ne décroît pas de manière rapide.

3.2 Différentiation de la série dialyse

Notation : On note la série différenciée de la série dialyse par (dialysediff).

3.2.1 Représentation graphique de la série dialysediff

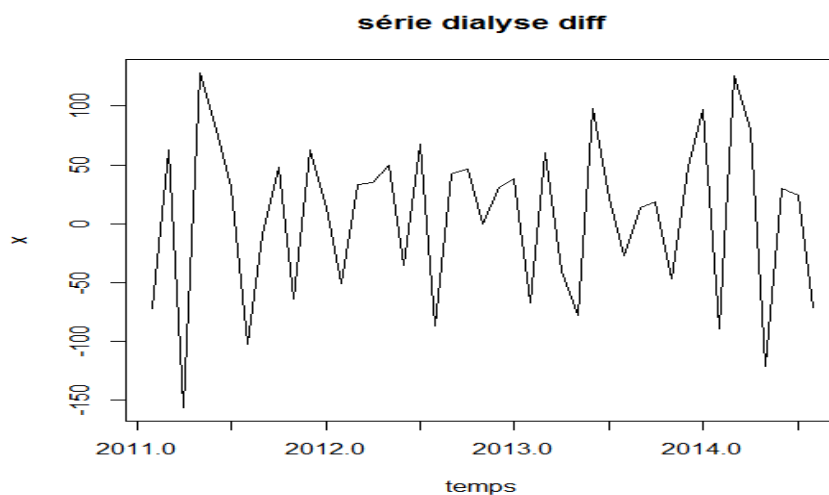


FIGURE 3.5 – le graphe de la série dialysediff

Le graphe de la série dialysediff montre que la série est stationnaire. On utilise aussi le test de racine unitaire (test ADF) pour confirmer la stationnarité .

cette test donne un p-value = $0.01 < 0.05$ (la probabilité critique affectées à la constante est inférieur à 0.05).

3.2.2 corrélogramme et corrélogramme partiel de la série dialysediff

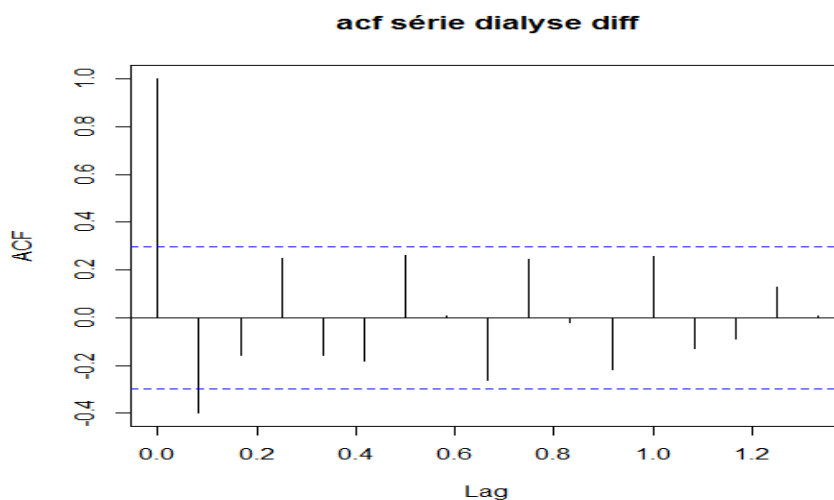


FIGURE 3.6 – Le corrélogramme (acf) de la série dialysediff

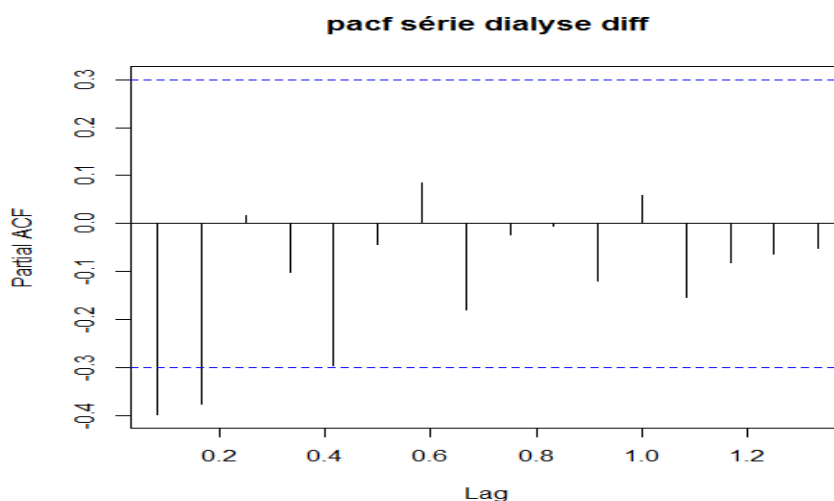


FIGURE 3.7 – Le corrélogramme partiel (pacf) de la série dialysediff

Détermination des modèles possibles :

S'appuyant sur le corrélogramme et corrélogramme partiel des fonctions d'auto-corrélations et d'auto-corrélations partielles

-le corrélogramme simple de la fonction d'auto-corrélation nous fournit la valeur de q .

-le corrélogramme partiel de la fonction d'auto-corrélation partielle nous donne la valeur de p .

ainsi on déduit les modèles suivants :

- Premier modèle : MA(1)
- Deuxième modèle : AR(2)
- Troisième modèle : ARMA(2,1)

3.3 L'estimation et validation du modèle

Après avoir identifié l'ordre des processus MA et AR, il convient d'estimer les paramètres du modèle, puis de vérifier à la base du certain nombre de tests statistiques que l'estimation du modèle est valable pour faire les prévisions.

3.3.1 Estimation des modèles identifiés

		Avec constante				Sans constante	
		Coeff		Cte		Coeff	
		Estimation	P(>)	Estimation	P(>)	Estimation	P(>)
ARMA(2,1)	AR(2)	-0.7277	0.02173	16.3085	0.12972	-0.8369	0.00544
		-0.4476	0.00308			-0.4387	0.00101
	MA(1)	0.2208	0.57925			0.4221	0.26257
AR(2)		-0.5648	7.4e-05	15.6233	0.08076	-0.5121	0.000381
		-0.3921	0.00514			-0.3460	0.015111
MA(1)		-1.3192844	<2e-16	10.4102940	<2e-16	-0.5256	5.82e-06

TABLE 3.2 – estimation des paramètres des modèles MA(1) AR(2) ARMA(2,1)

Remarque. *En comparant toujours les p-values avec la valeur 0.05 (5%). les modèles où les p-values sont supérieures à 0.05 seront écartés, si un des paramètres (ou des constantes) ne soit pas significativement différent de 0 le modèle sera alors rejeté.*

Au regard des résultats d'estimation, on constate d'une part que les modèles ARMA(2,1) et AR(2) avec constante seront rejetés car les constantes ne sont pas significativement différentes de 0 (p-value supérieures à 0.05), mais le modèle MA(1) sera accepté car la constante est significativement différente de 0, d'autre part le modèle ARMA(2,1) sans constante sera rejeté car p-value supérieur à 0.05 et les modèles AR(2) et MA(1) sans constantes restent candidats dans la mesure où les coefficients sont significativement différents de zéro (les p-values sont inférieures à 0.05)

3.3.2 Validation des résidus

Pour tester la stationnarité des résidus, on fait le test de portmanteau (Box-Pierce) et on trace les corrélogrammes et corrélogrammes partielle.

	χ^2	p-value
MA(1)	13.495	0.3341
AR(2)	7.4032	0.8299
ARMA(2,1)	6.644	0.8802

TABLE 3.3 – Résultat de test de Box-pierce

- Pour le modèle MA(1)

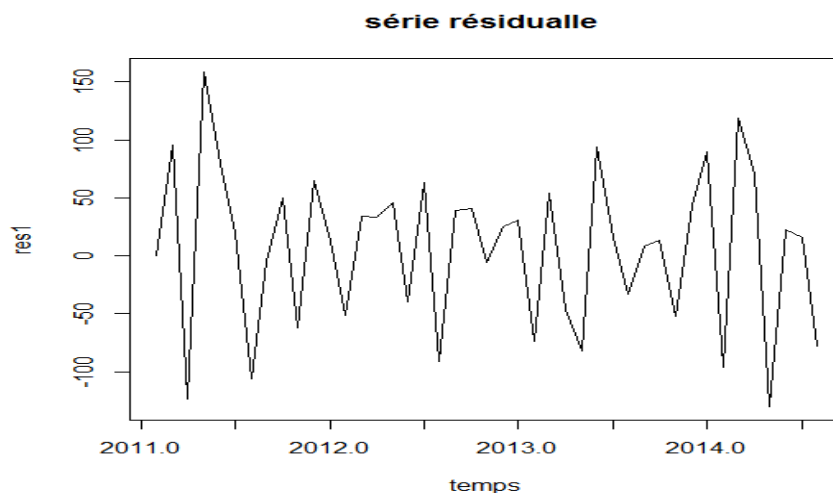


FIGURE 3.8 – le graphe de la série résiduelle pour le modèle MA(1)

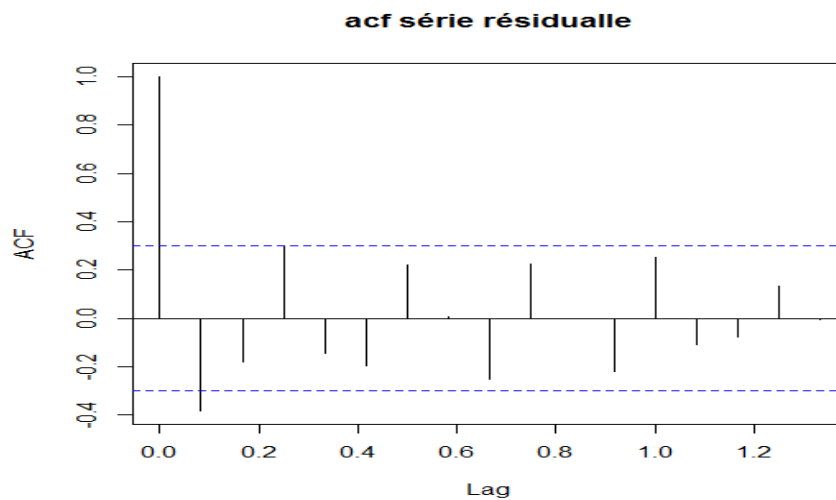


FIGURE 3.9 – le corrélogramme de la série résiduelle pour le modèle MA(1)

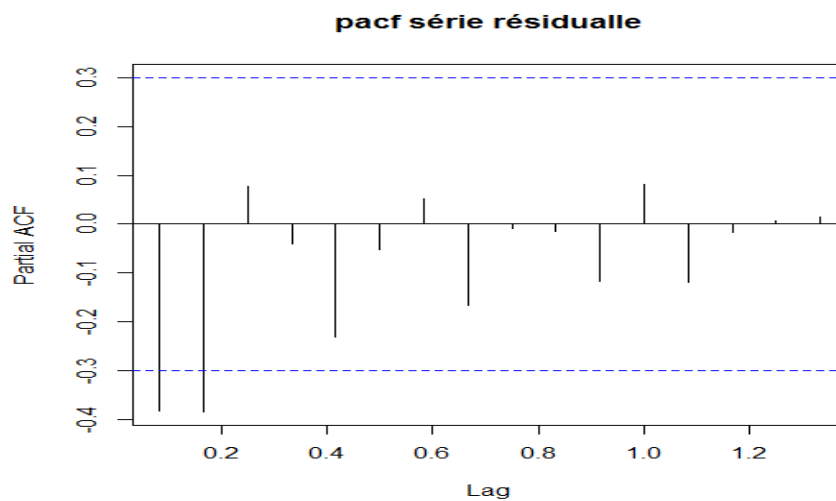


FIGURE 3.10 – le corrélogramme partiel de la série résiduelle pour le modèle MA(1)

$Q_h = 13.495$, $Khi\ deux(h) = Khi\ deux(12) = 21.03$ (au seuil 5%). Et $p\text{-value} = 0.3341 > 0.05$ (accepter H_0), alors les pics ne sont pas significatifs. Donc on accepte l'hypothèse nulle, absence d'auto-corrélation du résidu, ce qui implique qu'il peut être assimilé à un bruit blanc. Alors le modèle est valide.

- Pour le modèle AR(2)

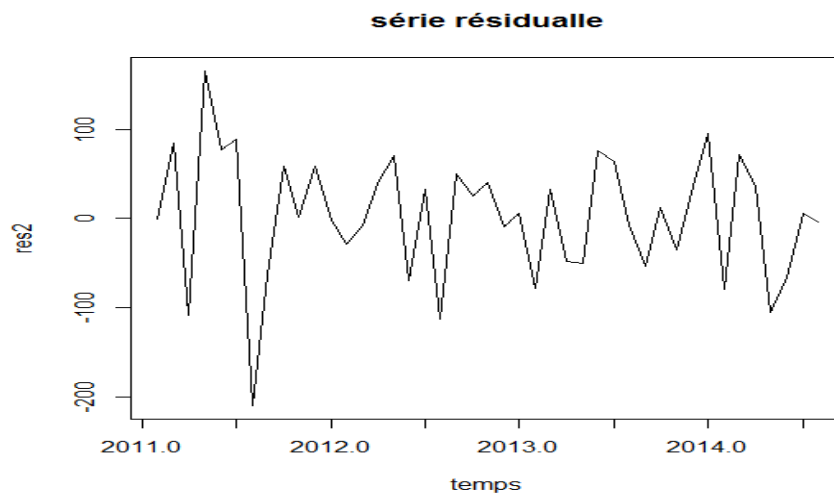


FIGURE 3.11 – le graphe de la série résiduelle pour le modèle AR(2)

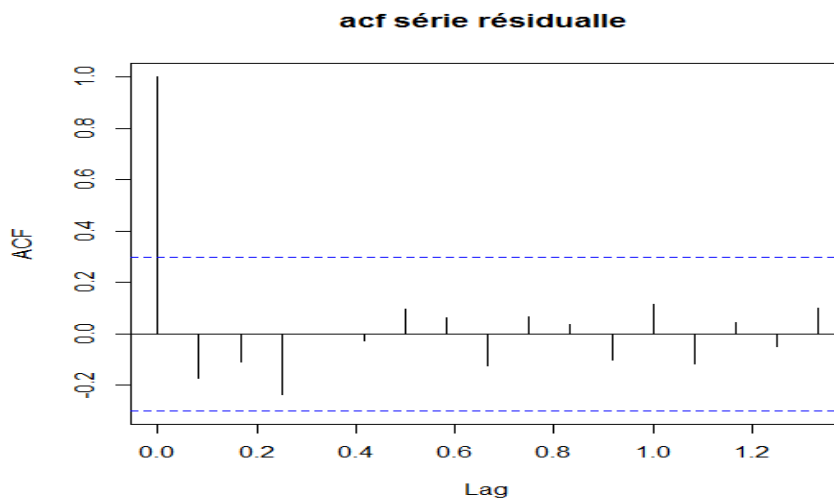


FIGURE 3.12 – le corrélogramme de la série résiduelle pour le modèle AR(2)

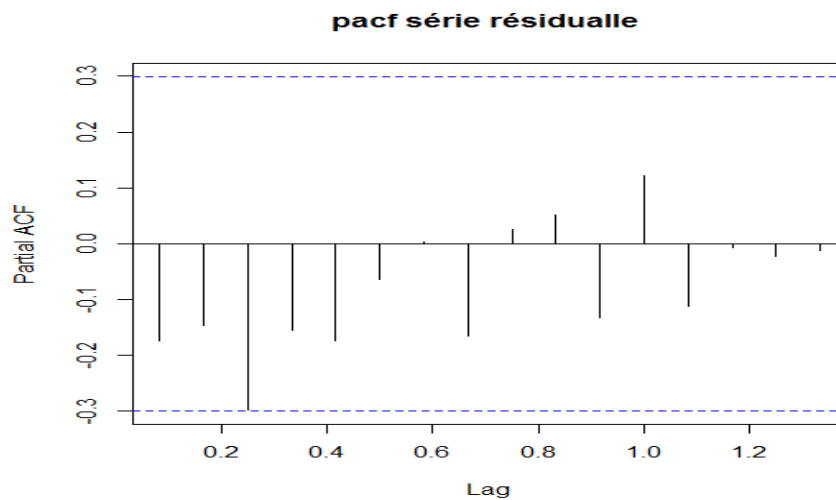


FIGURE 3.13 – le corrélogramme partiel de la série résiduelle pour le modèle AR(2)

$Q_h=7.4032 < \text{Khi deux}(h)=\text{Khi deux}(12)=21.03$ (au seuil 5%). Et $p\text{-value}=0.8299 > 0.05$ (accepter H_0) alors les pics ne sont pas significatifs. Donc on accepte l'hypothèse nulle, absence d'auto-corrélation du résidu, ce qui implique qu'il peut être assimilé à un bruit blanc. Alors le modèle est valide.

- Pour le modèle ARMA(2,1)

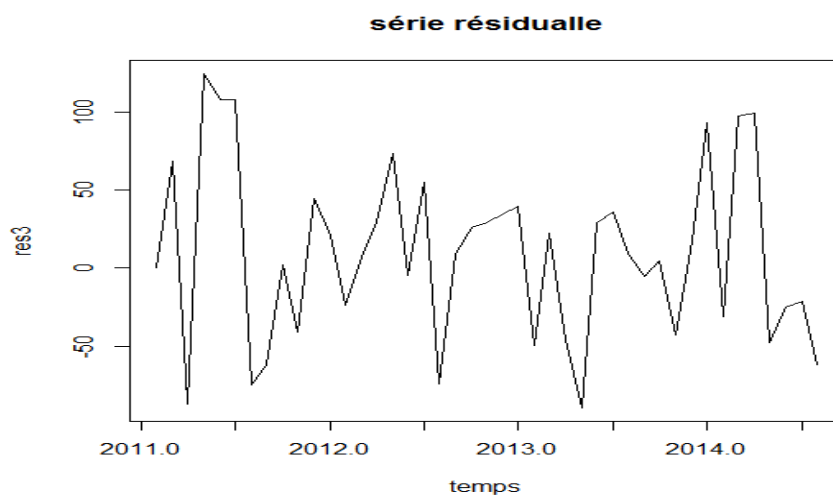


FIGURE 3.14 – le graphe de la série résiduelle pour le modèle ARMA(2,1)

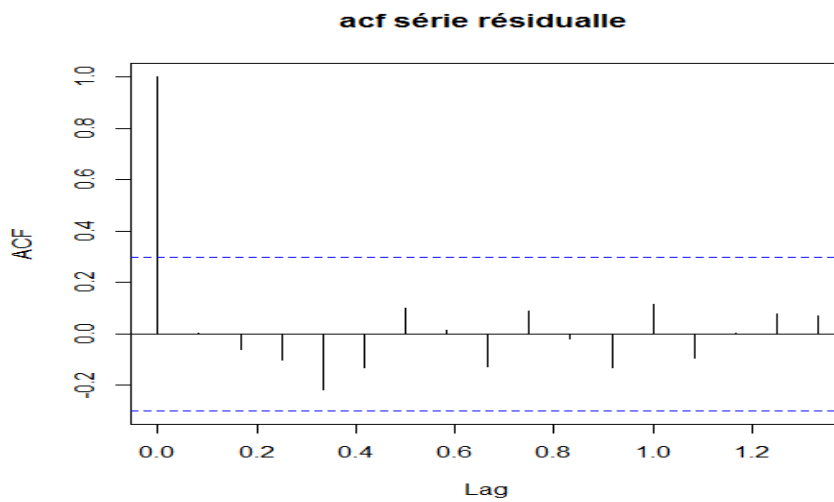


FIGURE 3.15 – le corrélogramme de la série résiduelle pour le modèle ARMA(2,1)

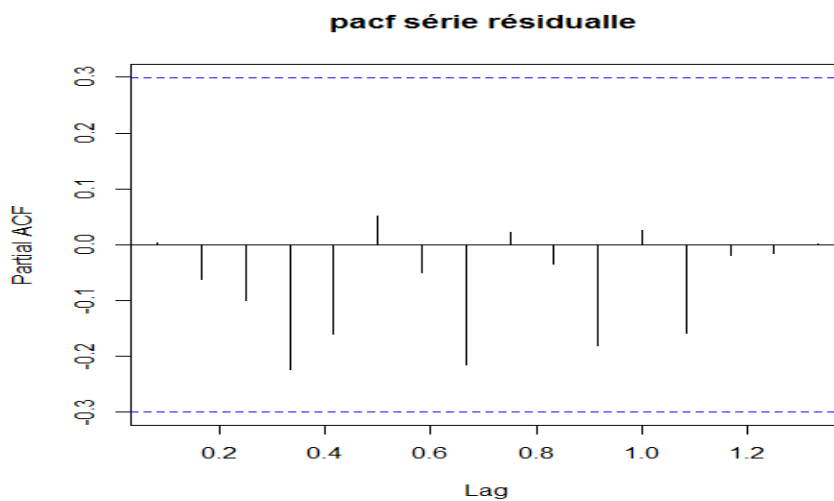


FIGURE 3.16 – le corrélogramme partiel de la série résiduelle pour le modèle ARMA(2,1)

$Q_h=6.644 < \text{Khi deux}(h)=\text{Khi deux}(12)=21.03$ (au seuil 5%). Et $p\text{-value}=0.8802 > 0.05$ (accepter H_0). Donc on accepte l'hypothèse nulle, absence d'auto-corrélation du résidu, ce qui implique qu'il peut être assimilé à un bruit blanc. Alors le modèle est valide.

3.4 choix de modèle

Pour comparer les modèles entre eux, on cherche à minimiser les critères AIC et BIC des trois modèles.

	AIC	BIC	$\hat{\sigma}^2$
MA(1)	482.473	492.864	3567
AR(2)	477.716	483.760	3430
ARMA(2,1)	479.2471	487.264	3408

TABLE 3.4 – la table des critères du choix de meilleur modèle

D'après ces critères, il ressort que le modèle AR(2) dispose d'une qualité supérieure car les critères d'information de ce modèle sont minimales

3.4.1 Test de normalité pour les résidus du modèle AR(2)

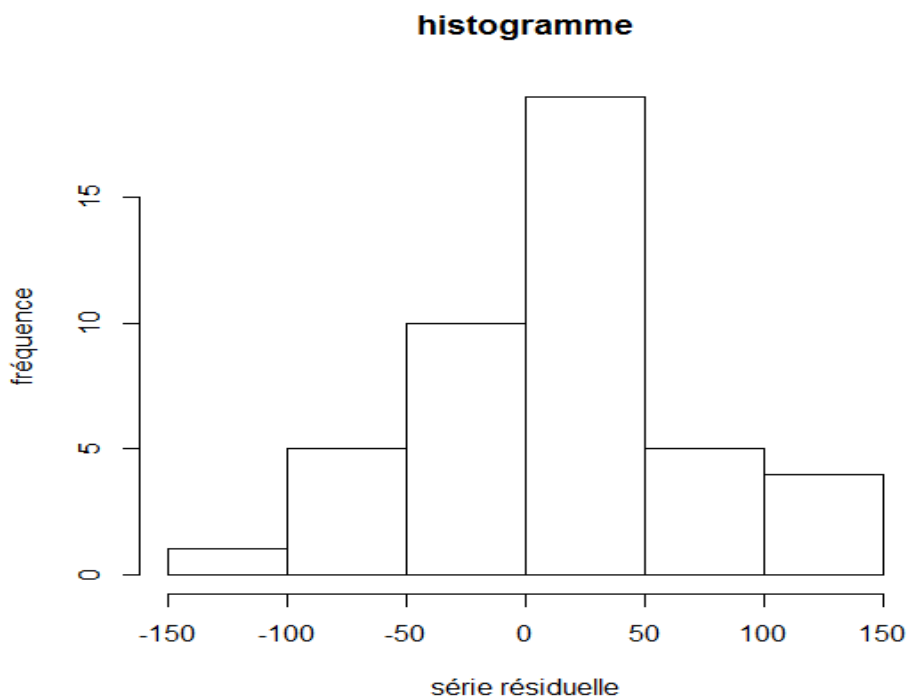


FIGURE 3.17 – l'histogramme des résidus pour le modèle AR(2)

Pour tester la normalité des résidus, on utilise le test de jarque-bera :
 La statistique $\mathbf{JB}=0.49805 < \chi^2_{1-\alpha}(2)$ ($0.49805 < 5.991$) avec $\alpha = 0.05$; alors on accepte l'hypothèse H_0 de normalité des résidus au seuil α du modèle AR(2).

3.5 Prévision

Nous allons maintenant tester les performances du modèle. Les valeurs prévues pour les 12 mois (septembre 2014 - août 2015).

Mois	Valeurs réelles	80%	95%
septembre 2014	911	[805,959]	[765,1000]
octobre 2014	911	[806,977]	[761,1022]
novembre 2014	867	[786,968]	[738,1016]
décembr 2014	959	[779,984]	[724,1038]
janvier 2015	890	[773,995]	[715,1053]
février 2015	828	[764,998]	[702,1061]
mars 2015	926	[757,1007]	[691,1073]
avril 2015	936	[751,1014]	[861,1084]
mai 2015	890	[744,1020]	[671,1092]
juin 2015	916	[738,1026]	[662,1102]
juillet 2015	900	[732,1032]	[653,1111]
août 2015	892	[727,1037]	[645,1119]

TABLE 3.5 – la table des prévisions

3.5.1 Le graphe de prévision

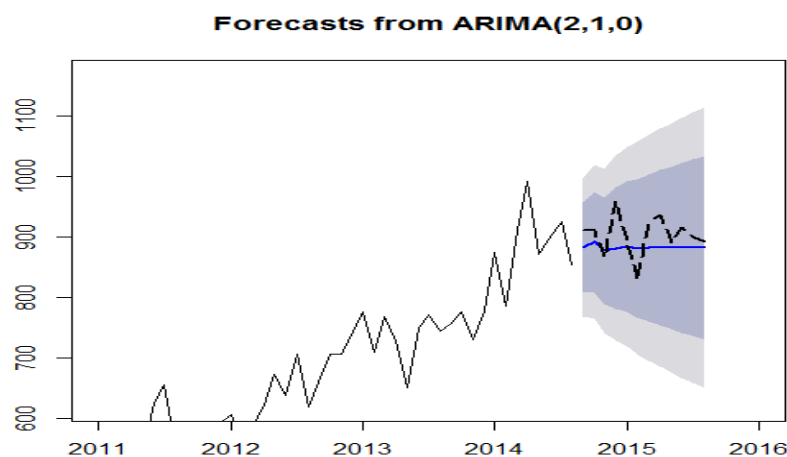


FIGURE 3.18 – Graphe de la prévision

Conclusion Générale et perspective

En guise de conclusion générale, nous allons tenter d'établir une synthèse globale sur le travail qui à été réalisé dans ce mémoire. L'objectif principal de ce mémoire est de développer des modèles statistique qui peuvent être utilisé pour prédire le nombre de séances d'hémodialyse d'une série temporelle réelles, qui sera valable pour des prédictions futures.

Les résultats de cette étude confirment la capacité des modèles auto-régressives dans le domaine de la prédiction des séries temporelles. Ce qui est vérifié par les critères de performances, l'utilisation de la méthode de Box-Jenkins pour construire les différents modèles : AR, MA, ARMA qui peuvent être utilisés pour résoudre le problème de la prédiction les résultats de la prédiction ont été acceptable.

Les principaux objectifs à atteindre dans une perspective à court terme de notre étude seront :

1. Prédire le nombre de séances d'hémodialyse en utilisant les méthodes de lissages exponentielle.
2. Étendre cette études à la prédiction des séries appartenant à d'autres domaines (finance, commerce,...etc).

Bibliographie

- [1] **Azencot R. et Dacunha-castelle D.** Séries d'observations irrégulières , Modélisation et prévision, Masson,Paris(1984)
- [2] **Box G. E. P and Jenkins G. M.** Time series analysis, Forecasting and control. Holden Day, San Francisco (1976).
- [3] **Brockwell. P. J. and Davis R. A.** Time series : "Theory and methods". Springer-verlag, New York, 2nd Ed (1996).
- [4] **Bouzgou H et Kherchi I.** Modélisation des séries temporelles par les processus auto-regressifs (juin 2016).
- [5] **Dakhmouche M .** Approximation of the exact sum of squares function of a gaussian ARMA process,Survey in Approximation Theory,(2009)(to appear).
- [6] **Dakhmouche M .** Statistique des processus linéaires : test et estimation.Thèse de doctorat (2006).
- [7] **Hamilton J. D.** Time series analysis. Princeton University Press, Princeton (1994).
- [8] **Hamrita M.S et Kairouna I.** Introduction aux modèles de séries temporelles.
- [9] **Julien Jacques.**introduction aux séries temporelles.
- [10] **Bouazza M.T.** La méthodologie de Box Jenkins et applications(2017).
- [11] **Hannan E.J. and Deisteler M.** the Statistical Theory of linear systems, John Wiley, New York (1988).
- [12] **Peter B.F .** Analysis of time series. John Wiley et Sons, Inc (2000).
- [13] **von Sachs R et Van Bellegem S.** STAT 2414 Séries Chronologiques(2005).
- [14] **Fortier S.** Les modèles MA,AR et ARMA multidimensionnels : estimation et causalité.
- [15] **Zetili S.**Inférence statistique dans les processus ARMA fractionnaires.
- [16] **Zoghbi F Amira S.** Modélisation et prévision de la pluviométrie.Apport et Transferts des Eaux.cas :Barrage Béni Haroun(2016).

-
- [17] **Yahia Dj.** Séries temporelles et d' adéquation pour un modele GARCH(1,1)(2015).

Annex

Le programme d'application sous R

```
>ns=c(582,510,573,417,545,625,656,554,545,593,529,592,606,555,588
,623,673,638,706,619,661,707,707,738,776,709,769,729,651,749,771
,744,758,777,730,778,875,786,912,992,871,901,925,854,911,911,867
,959,890,828,926,936,890,916,900,892)
>ns
# entré le vecteur des donner
>dialyse=ts(ns,2011,c(2014,8),f=12)
>dialyse
# transformer les donner vers une série "dialyse"
>plot(dialyse,xlab="temps",main="série dialyse")
# tracer le graphe de de la série "dialyse"
>acf(dialyse,main="acf série dialyse")
# tracer le corrélogramme de la série "dialyse"
>pacf(dialyse,main="pacf série dialyse")
# tracer le corrélogramme partiel de la série "dialyse"
>library(tseries)
# pour lancer le package "tseries"
>adf.test(dialyse)
# test de racine unitaire pour la série "dialyse"
> x=diff(dialyse)
>x
# la différentiation de la série "dialyse"
>plot(x,xlab="temps",main="série dialyse diff")
# tracer le graphe de la série "x"
> acf(x,main="acf série dialyse diff")
# tracer le corrélogramme de la série "x"
```

```
> pacf(x,main="pacf série dialyse diff")
# tracer le corrélogramme partiel de la série "x"
>library(tseries)
>adf.test(x)
# test de racine unitaire pour la série "x"
>library(FitARMA)
# pour lancer le package "FitARMA"
>s1=arma(x,order=c(0,1),include.intercept=FALSE)
>s1
# estimer les paramètres du modèle ARMA(0,1) sans constante
>summary(s1)
# description des résultats
>t1=arma(x,order=c(0,1),include.intercept=TRUE)
>t1
# estimer les paramètres du modèle ARMA(0,1) avec constante
>summary(t1)
# description des résultats
>s2=arma(x,order=c(2,0),include.intercept=FALSE)
>s2
# estimer les paramètres du modèle ARMA(2,0) sans constante
>summary(s2)
# description des résultats
>t2=arma(x,order=c(2,0),include.intercept=TRUE)
>t2
# estimer les paramètres du modèle ARMA(2,0) avec constante
>summary(t2)
# description des résultats
>s3=arma(x,order=c(2,1),include.intercept=FALSE)
>s3
# estimer les paramètres du modèle ARMA(2,1) sans constante
>summary(s3)
# description des résultats
>t3=arma(x,order=c(2,1),include.intercept=TRUE)
>t3
# estimer les paramètres du modèle ARMA(2,0) avec constante
```

```
>summary(t3)
# description des résultats
>res=t2$residuals
# donner la série des résids d'un modèle ARMA(2,0)
>plot(res,xlab="temps",main="série résiduelle")
# tracer le graphe de la série résiduelle
>acf(res,main="acf série résiduelle ")
# tracer le corrélogramme de la série résiduelle
>pacf(res,main="pacf série résiduelle ")
# tracer le corrélogramme partiel de la série résiduelle
>Box.test(res)
# le test de bruit blanc
>jarque.bera.test(res)
le test de normalité des résidus
>hist(res2,xlab="série résiduelle"
ylab="fréquence",main="histogramme")
#l'histogramme des résidus
>library(forecast)
# lancer le package "forecast"
>f1=auto.arima(dialyse,trace=T,ic=c("bic","aic"),allowdrift=T,allowmean=F)
>f1
# pour choix le meilleur modèle
>forecast(f1,h=12)
# la prévision
train.data=window(dialyse,start=c(2011,1),end=c(2014,8))
train.data
test.data=window(dialyse,start=c(2014,9))
test.data
fit=arima(train.data,order=c(2,1,0))
fit
library(TSPred)
# lancer le package "TSPred"
plotarimapred(test.data,fit,xlim=c(2011,2016),range.percent=0.05)
# tracer le graphe de prévision
```