

RÉPUBLIQUE ALGÉRIENNE DÉMOCRATIQUE ET POPULAIRE

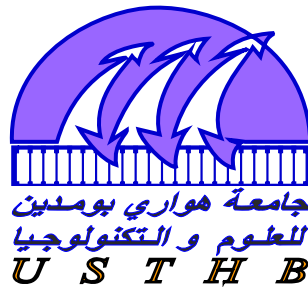
MINISTRE DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA

RECHERCHE SCIENTIFIQUE

UNIVERSITÉ DES SCIENCES ET DE LA TECHNOLOGIE

HOUARI BOUMEDIENNE

DÉPARTEMENT DES MATHÉMATIQUES



Domaine: Statistiques Et Probabilités Appliquées

Mini Projet

Sur Les Séries Temporelles

Résumé Sur:  
La Méthodologie de Box-Jenkins

Proposé Par:

Mme. BENDJEDDOU Sara

Présenté par:

AMZERT Kaouthar

BERREKSI Khadidja

# Table of contents

---

Table of contents	i
Introduction	1
1 Modélisation des séries temporelles	3
1.1 Le Pincipe Des Séries Temporelles . . . . .	3
1.2 Décomposition d'une série temporelle . . . . .	3
1.2.1 La tendance: . . . . .	4
1.2.2 La saisonnalité: . . . . .	4
1.2.3 Le résidu . . . . .	4
1.3 Modélisations de base pour les séries temporelles . . . . .	5
1.3.1 La décomposition additive . . . . .	5
1.3.2 La décomposition multiplicative . . . . .	6
1.3.3 Dynamique autorégressive . . . . .	7
1.4 Modélisation de la partie déterministe . . . . .	7
1.4.1 La Tendance . . . . .	7
1.4.2 La Saisonnalité . . . . .	8

1.4.3	Elimination de la tendance et de la saisonnalité . . . . .	9
1.5	Processus Stationnaire . . . . .	9
1.5.1	Processus Stationnaires Classiques . . . . .	11
1.6	Processus Linéaires . . . . .	11
1.7	Modèles Autorégressifs Moyenne Mobile . . . . .	12
1.8	Décomposition de Wold . . . . .	18
1.9	Processus ARIMA . . . . .	18
1.10	Processus SARIMA . . . . .	19
<b>2</b>	<b>La méthodologie de Box-Jenkins</b>	<b>21</b>
2.1	Identification du modèle . . . . .	21
2.1.1	Estimation de $d$ . . . . .	22
2.1.2	Estimation de $p$ et $q$ . . . . .	22
2.2	Estimation des paramètres du modèle . . . . .	23
2.2.1	La méthode des moments . . . . .	23
2.2.2	La méthode des moindres carrés (MC) . . . . .	23
2.2.3	La méthode du maximum de vraisemblance . . . . .	24
2.3	Validation . . . . .	24
2.3.1	Test de Student sur les coefficients . . . . .	25
2.3.2	Test sur les résidus . . . . .	25
2.3.3	Critères de choix des modèles . . . . .	28
2.3.4	Critère d'information . . . . .	28

2.4	Prévision . . . . .	29
2.4.1	Erreur de prévision . . . . .	29
<b>3</b>	<b>Applications sous R</b>	<b>31</b>
	<b>Bibliography</b>	<b>44</b>

# Introduction

---

Une série temporelle (ou série chronologique) est une séquence spécifique d'observations repérés dans le temps. Leur domaine d'application est très vaste et s'étend de l'astronomie à l'économie et la finance en passant par la biologie, la géophysique, et la médecine,... La théorie et la pratique de l'analyse des séries temporelles se sont développées rapidement depuis l'apparition des travaux fondateurs de Box et Jenkins en 1970. L'utilisation des séries chronologiques peut servir à diverses fins. L'objectif peut être la prédiction de l'avenir basée sur la connaissance du passé, il peut être le contrôle du processus produisant la série, ou tout simplement une brève description des principales caractéristiques de la série.

---

# Chapter 1

## Modélisation des séries temporelles

---

### 1.1 Le Principe Des Séries Temporelles

**Definition 1.1.1** Une Série Temporelle est une suite d'observations indexées par le temps, notée  $X_t$ . Elle représente généralement l'évolution d'un phénomène au cours du temps, telle que :  $X_t := \{X_t, t = 1, \dots, T\}, T \in \mathbb{N}, \mathbb{R}, \mathbb{Z}, \dots$

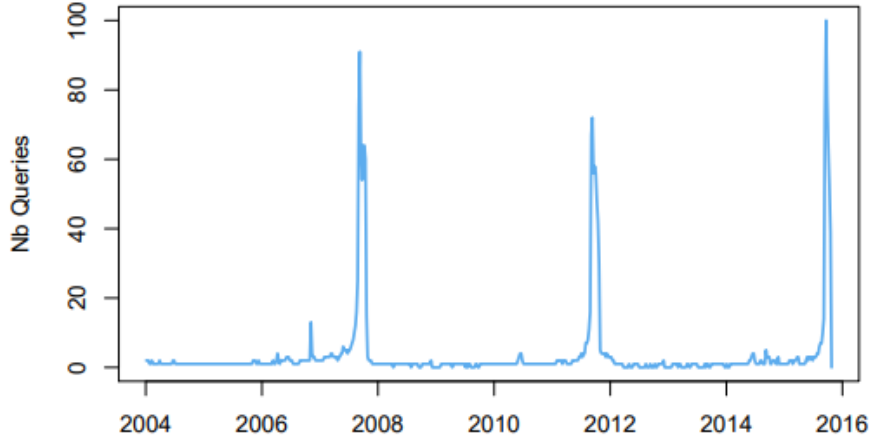
L'indice  $t$  représente une unité de temps (qui peut être le mois, l'année, . . . ).

#### Exemple de série temporelle:

Les séries temporelles sont présentes dans de nombreux domaines d'application. Par exemple, les données du nombre de requêtes google "rugby world cup" par semaine (source <https://www.google.fr/trends/>), représentée dans la Figure 1.1.1.

### 1.2 Décomposition d'une série temporelle

Les séries temporelles sont modélisées selon les trois composantes suivantes :



### 1.2.1 La tendance:

La tendance, notée  $Z_t$ , est une fonction représente le mouvement à long terme de la série étudiée. Il existe trois types de tendance:

- tendance linéaire:  $T_t = a + bt$
- tendance quadratique:  $T_t = a + bt + ct^2$
- tendance logarithmique:  $T_t = \log(t)$

### 1.2.2 La saisonnalité:

La composante saisonnière, notée  $S_t$ , est une fonction périodique de période  $s$ , i.e.  $S_t = S_{t \pm s}$ . Elle est liée au rythme imposé par les saisons.  $S_t$  est entièrement déterminée par ses  $s$  composantes  $\delta_1, \delta_2, \dots, \delta_s$ , et par hypothèse  $\sum_{j=1}^s \gamma_j = 0$ , telle que:

$$S_t := \sum_{j=1}^s \gamma_j S^j, \quad \text{où } S^j = \begin{cases} 1 & \text{si } j = t. \\ 0 & \text{si } j \neq t. \end{cases}$$

### 1.2.3 Le résidu

On appelle résidu (ou erreur) ou partie irrégulière, noté " $\epsilon_t$ ", la variable aléatoire à effet de faible intensité et à courte durée. La figure 1.1 illustre la décomposition d'une série

temporelle réelle. Il s'agit de la production laitière mensuelle aux États-Unis (en millions de livres) de janvier 1994 à décembre 2005.

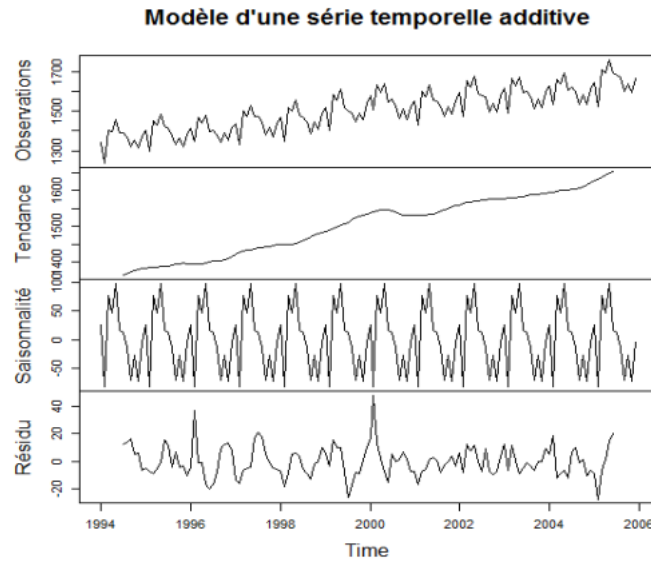


Figure 1.1: Décomposition d'une série temporelle ; le graphe de la série (en haut), la tendance et la saisonnalité (resp.) (au milieu) et le résidu (en bas).

## 1.3 Modélisations de base pour les séries temporelles

### 1.3.1 La décomposition additive

Une des décompositions de base est la suivante:  $X_t = m_t + s_t + \epsilon_t$ ,  $1 \leq t \leq T$

Où:

- $(m_t)_t$  est une composante tendancielle déterministe qui donne le comportement de la variable observée sur le long terme.
- $(s_t)_t$  est une suite périodique qui correspond à une composante saisonnière.
- $(\epsilon_t)_t$  représente une composante irrégulière et aléatoire, le plus souvent de faible amplitude par rapport à la composante saisonnière mais importante en pratique puisque



ce terme d'erreur sera le plus souvent autocorrélé (c'est à dire que la covariance entre  $\epsilon_t$  et  $\epsilon_{t+h}$  sera non nulle).

### 1.3.2 La décomposition multiplicative

On décompose la série temporelle sous la forme  $X_t = m_t s_t \epsilon_t$ ,  $1 \leq t \leq T$ . Les composantes  $(m_t)_t$  et  $(s_t)_t$  sont de la même forme que pour le modèle additif et la composante irrégulière  $(\epsilon_t)_t$  a pour moyenne 1. Par une transformation logarithmique, on se ramène à une décomposition additive. Cette décomposition multiplicative est intéressante lorsqu'on observe une variation linéaire des effets saisonniers comme le montre la Figure 1.2. On peut aussi combiner l'approche additive et l'approche multiplicative en décomposant

$$X_t = m_t s_t + \epsilon_t$$

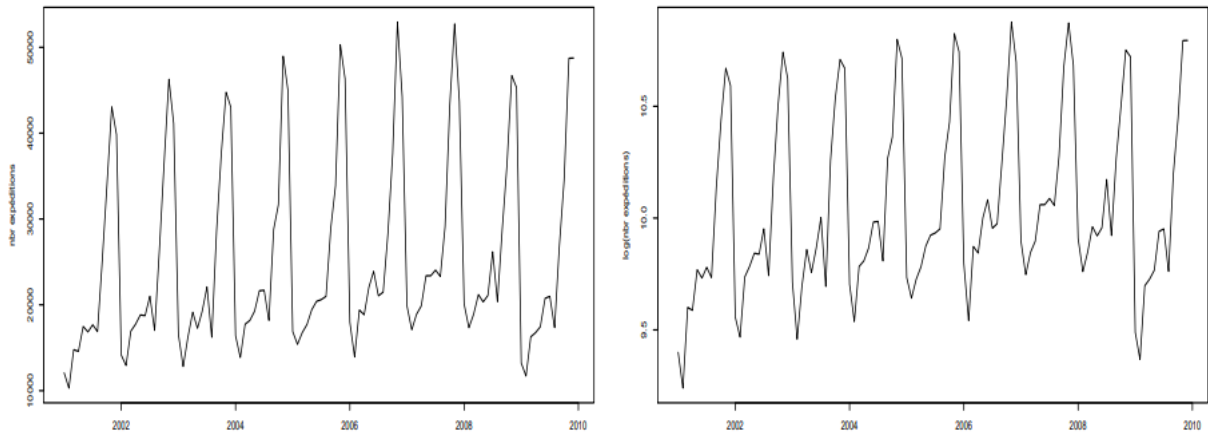


Figure 1.2: Expédition mensuelle de Jus en milliers de bouteilles entre 2001 et 2010 (série initiale à gauche et son logarithme à droite)

### 1.3.3 Dynamique autorégressive

La dynamique (aléatoire) de la série temporelle est basée sur des équations récurrentes du type  $X_t = f(X_{t-1}, \dots, X_{t-p}, \epsilon_t)$ ,  $p + 1 \leq t \leq T$

Où  $f$  est une fonction mesurable qui dépend d'un paramètre inconnu et  $(\epsilon_t)_t$  est une perturbation aléatoire non observée.

## 1.4 Modélisation de la partie déterministe

### 1.4.1 La Tendance

Il existe différents procédés permettant d'analyser puis/ou de corriger la tendance d'une série temporelle.

#### Estimation non paramétrique (Moyenne Mobile)

La moyenne mobile est une méthode simple permettant d'extraire les composantes basses fréquences d'une série temporelle autrement dit sa tendance. Le calcul de la moyenne mobile dépend d'un paramètre  $l$  appelé la largeur de fenêtre. Ce paramètre correspond au nombre d'observations incluses dans le calcul de la moyenne glissante effectuée.

$$\widehat{m}_t = \frac{1}{2l+1} \sum_{i=t-l}^{t+l} m_t$$

#### Estimation paramétrique de la tendance

On procède par régression (linéaire le plus souvent mais potentiellement non-linéaire) pour estimer cette tendance. Par exemple, dans le cas d'un processus  $y$  admettant une tendance polynomiale d'ordre  $k$  :  $y_t = \sum_{j=0}^k a_j t^j + \epsilon_t$ , un estimateur de la tendance pourra être

obtenu ainsi:  $\hat{T} = X\hat{a}$  où  $X$  est la matrice dont les colonnes sont les vecteurs  $(1, \dots, t^j)$  et:

$$\hat{a} = (X'X)^{-1} X'Y$$

avec:  $Y = (y_1, \dots, y_t)$

## Différenciation

Pour nettoyer une série de sa tendance et/ou de sa saisonnalité, nous pouvons procéder par différenciation. Cela fonctionne pour des séries à tendance polynomiale. Notons  $\Delta$  l'opérateur de différenciation:  $\Delta X_t = X_t - X_{t-1}$ . L'opérateur de différenciation d'ordre  $k$  correspondant est:  $\Delta^k y_t = \Delta(\Delta^{k-1} y_t)$ .

### 1.4.2 La Saisonnalité

Les traitements s'appliquant à la tendance s'appliquent également à la saisonnalité avec les variantes suivantes.

#### Estimation non paramétrique (Moyenne Mobile)

La moyenne mobile, en choisissant un paramètre de fenêtre  $l$  égal à la période de la série, permet de désaisonnaliser une série. En effet, par définition, la composante périodique est d'intégrale nulle sur la période.

#### Estimation paramétrique de la saisonnalité

Un modèle paramétrique naturel pour modéliser un processus saisonnier est la décomposition en série de Fourier. Soit un processus  $y_t$  admettant une saisonnalité de période  $\tau$  alors le modèle suivant est généralement proposé:

$$y_t = \sum_{j=1}^q a_j \cos(\omega_j t) + b_j \sin(\omega_j t) + \epsilon_t$$

où  $\omega_j = 2j\pi/\tau$ ,  $q$  est à déterminer par une méthode de sélection de modèle sur les données. Les coefficients  $a_j$  et  $b_j$  sont obtenus par moindres carrés sur les données.

## Différenciation

De même que pour nettoyer un processus de sa tendance, il est possible de le désaisonnaliser par différenciation.

**Proposition 1** soit un processus  $y_t$  admettant une saisonnalité additive de période  $\tau$ , alors le processus  $\Delta_\tau y_t = y_t - y_{t-\tau}$  est un processus désaisonnalisé.

### 1.4.3 Elimination de la tendance et de la saisonnalité

Si la série se décompose sous la forme  $X_t = m_t + s_t + \epsilon_t$  avec  $(m_t)_t$  polynomial et  $(s_t)_t$  de période  $k$ , on peut toujours éliminer la tendance et la saisonnalité en appliquant une moyenne mobile du type  $M = (\mathcal{I} - \mathcal{B}^k)(\mathcal{I} - \mathcal{B})^d$ . C'est le moyen le plus simple de se ramener à une composante résiduelle qui ne présente ni périodicité, ni tendance.

## 1.5 Processus Stationnaire

**Définition 1.5.1** Soit  $\{X_t, t \in \mathbb{Z}\}$  un processus tel que  $\mathbb{E}[X_t^2] < \infty$  pour  $t \in \mathbb{Z}$ . On dira que ce processus est (faiblement) stationnaire (ou stationnaire d'ordre 2) si les 2 conditions suivantes sont vérifiées pour tout  $t$  :

- $\mathbb{E}[X_t] = m$  pour tout  $t \in \{1, \dots, T\}$
- $\mathbb{E}[X_t X_{t-h}] = \gamma(h)$  est indépendant de  $t$ .

On appellera alors fonction d'autocovariance de  $X_t$  la fonction  $\gamma_X$  définie par:

$$\gamma_X(t, k) = \mathbb{E}(X_t X_{t+k}) - \mathbb{E}(X_t)\mathbb{E}(X_{t+k}) \quad \forall t, k \in \mathbb{Z}$$

et fonction d'autocorrélation la fonction  $\rho$  définie par:  $\rho(k) = \frac{\mathbb{E}(X_t X_{t+k}) - \mathbb{E}(X_t)^2}{\sqrt{\mathbb{E}(X_t^2) \mathbb{E}(X_{t+k}^2)}}$

**Remark 1 :**

- On a  $\gamma(t) = \gamma(-t)$  (fonction paire) donc il suffit de connaître la fonction d'autocovariance sur  $[0, T]$ . De même pour  $\rho$ .
- Pour tout  $k$ ,  $|\gamma(k)| \leq \gamma(0)$ .
- La fonction d'autocovariance est définie positive.

Par la suite le terme stationnaire fera référence à la stationnarité faible.

**Definition 1.5.2** Soit  $x_1, \dots, x_n$  une série chronologique.

- La Moyenne empirique est:  $\bar{x}_n = \frac{\sum_{i=1}^n x_i}{n}$
- La fonction d'autocovariance empirique est:  $\hat{\gamma}_h = \frac{1}{n} \sum_{t=1}^{n-|h|} (x_t - \bar{x}_n)(x_{t+|h|} - \bar{x}_n)$ ,  $-n < h < n$ .
- La fonction d'autocorrélation empirique:  $\hat{\rho}_h = \frac{\hat{\gamma}_h}{\hat{\gamma}_0}$ ,  $-n < h < n$

### La Stationnarité au sens strict

Les distributions de probabilité conjointe de  $X_{t_1}, \dots, X_{t_k}$  et de  $X_{t_1+h}, \dots, X_{t_k+h}$  sont les mêmes pour tout  $k \in \mathbb{N}^*$  et pour tous  $t_1, \dots, t_k$  de  $T^k$  et  $h$  de  $\mathbb{Z}$ .

**Remark 2** la stationnarité au sens strict implique la stationnarité au sens large (stationnarité faible) si le processus est du second ordre.

### 1.5.1 Processus Stationnaires Classiques

**Definition 1.5.3** - On dira que le processus  $\epsilon_t$  est un "bruit blanc" s'il forme une suite de variables indépendantes et identiquement distribuées. Il sera dit centré si  $\mathbb{E}[\epsilon_t] = 0$  et réduit si  $\text{Var}[\epsilon_t] = 1$ , On note:  $\epsilon_t \sim \mathcal{BB}(0, \sigma_\epsilon^2)$ .

**Definition 1.5.4** Un processus **gaussien** à temps discret  $\{X_t\}_{t \in \mathbb{Z}}$  est une série temporelle telle que la loi de n'importe quel vecteur extrait est gaussien, i.e.

$\forall n \in \mathbb{N}^*, \forall (t_1, \dots, t_n) \in \mathbb{Z}^\times : (X_{t_1}, \dots, X_{t_n})$ , est un vecteur gaussien.

Un bruit blanc gaussien est une série temporelle gaussienne.

**Definition 1.5.5** Soient  $A, B$  deux v.a.r. non corrélées de moyenne 0 et de variance  $\sigma_\epsilon^2$ , et  $\theta \in [-\pi; \pi]$  une constante. On considère le processus **harmonique** défini par:

$$\forall t \in \mathbb{Z}, X_t = A \cos(\theta t) + B \sin(\theta t)$$

## 1.6 Processus Linéaires

**Definition 1.6.1 (Processus Linéaire)** On dit qu'un processus  $X_t$  est linéaire s'il s'écrit comme suit :

$$X_t = c + \sum_{j \in \mathbb{Z}} \psi_j \epsilon_{t-j}, \quad \forall t \in T,$$

Où :  $c \in \mathbb{R}$ ,  $(\psi_j)_{j \in \mathbb{Z}}$  est une suite de constantes, telle que:  $\sum_{j=-\infty}^{+\infty} |\psi_j| < +\infty$ ,

et :  $\epsilon_t \sim \mathcal{BB}(0, \sigma_\epsilon^2)$ .

**Definition 1.6.2** On dit qu'un **Processus**  $X_t$  est **causal** s'il s'écrit en fonction de  $(\epsilon_k, k \leq t)$  seulement:

$$X_t = \sum_{j=0}^{+\infty} \varphi_j \epsilon_{t-j}, \quad \varphi_j \in \mathbb{R} \text{ et } \epsilon_t \sim \mathcal{BB}(0, \sigma_\epsilon^2).$$

**Definition 1.6.3** On dit qu'un **Proceuss**  $X_t$  est **inversible** si  $\epsilon_t$  s'écrit en fonction de  $(X_k, k \leq t)$  seulement:

$$\epsilon_t = \sum_{j=0}^{+\infty} \phi_j X_{t-j}, \phi_j \in \mathbb{R}$$

## 1.7 Modèles Autorégressifs Moyenne Mobile

Dans cette section, on considère les modèles linéaires proposés par Box et Jenkins et qui sont les plus fréquemment utilisés pour modéliser une série chronologique  $X_t$ , tels que le modèle autorégressif, le modèle à moyenne mobile et le modèle ARMA qui combine ces deux modèles.

**Definition 1.7.1 (Modèle Autorégressif)** On dit que  $X_t$  est un modèle autorégressif ("Auto-Regressive") d'ordre  $p$  ( $p \geq 1$ ), abrégé  $AR(p)$  si :

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + \epsilon_t = \sum_{j=1}^p \phi_j X_{t-j} + \epsilon_t$$

Où :  $\phi_1, \phi_2, \dots, \phi_p$  sont des constantes ( $\phi_p \neq 0$ ), et  $\epsilon_t \sim \mathcal{BB}(0, \sigma_\epsilon^2)$ . On utilise généralement la notation suivante:

$$\Phi(L)X_t = \epsilon_t,$$

avec  $\Phi(L) = 1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p$ .

On dit dans ce cas, qu'un modèle  $AR(p)$  est stationnaire si toutes les racines du polynome caractéristique  $\Phi(z)$  sont de module supérieur strictement à 1. Sous hypothèse de stationnarité, un modèle  $AR(p)$  vérifie les propriétés suivantes :

1. *Esperance mathématique:*  $\mathbb{E}(X_t) = 0$ ,

2. *fonction d'autocovariance :*

$$\gamma_X(h) = \sum_{j=1}^p \phi_j \gamma_X(h-j),$$

3. *La fonction d'autocorrélation:*

$$\rho_X(h) = \sum_{j=1}^p \phi_j \rho_X(h-j).$$

4. La variance :

$$\gamma_X(0) = \frac{\sigma_\epsilon^2}{1 - \sum_{j=1}^p \phi_j \rho_X(j)}$$

Les trois équations 2, 3, 4 s'appellent le système de **Yule-Walker**. **Illustrons ces propriétés** : Les Figures ci-dessous montrent respectivement une réalisation d'un processus  $AR(2)$ , sa fonction de corrélation empirique et sa fonction de corrélation partielle empirique.

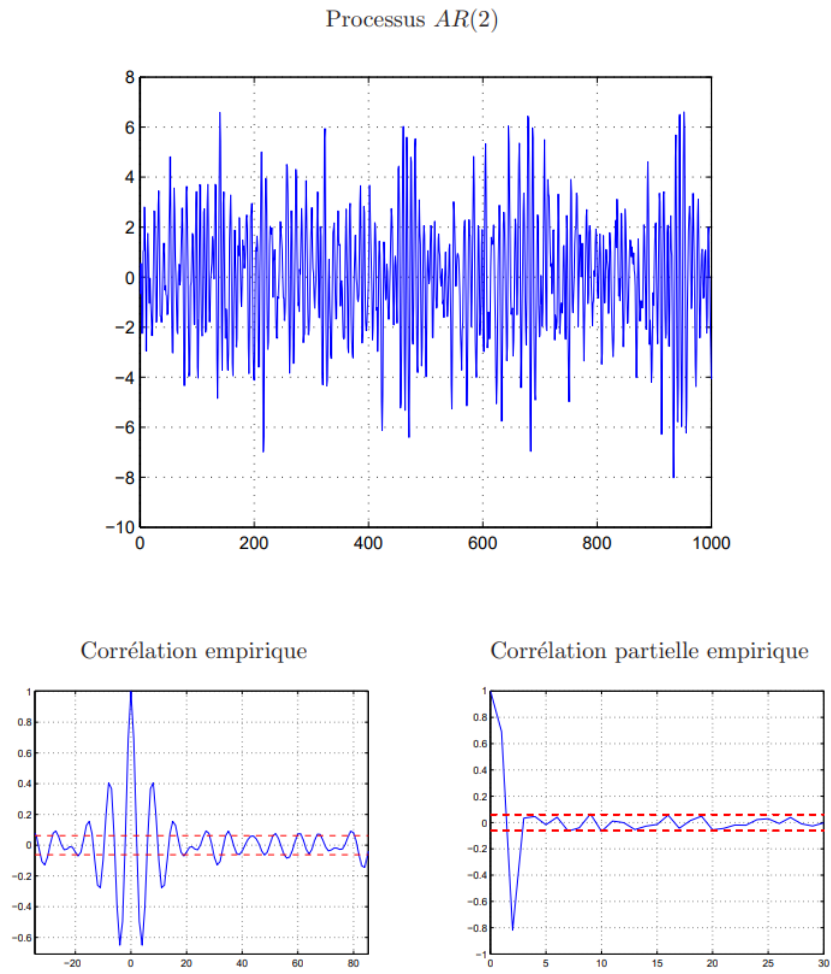


Figure 1.3: Propriétés d'un processus  $AR(2)$

Système de **Yule-Walker** sous forme matricielle.

Les équations de Yule-Walker peuvent être écrites sous forme matricielle, telle que:

$$\rho = \mathcal{R}_p \phi,$$



Où :  $\rho = (\rho_X(1), \rho_X(2), \dots, \rho_X(p))^t$ , et  $\phi = (\phi_1, \phi_2, \dots, \phi_p)^t$ , et :

$$\mathcal{R}_p = \begin{pmatrix} 1 & \rho_X(1) & \cdots & \rho_X(p-1) \\ \rho_X(1) & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \rho_X(1) \\ \rho_X(p-1) & \cdots & \rho_X(1) & 1 \end{pmatrix}$$

Cette dernière s'appelle matrice des autocorrélations, elle est de rang  $p$  et de forme Toeplitz.

**(Modèle Autoregressif d'ordre infini  $AR(\infty)$ )**

On dit qu'un modèle  $X_t$  est  $AR(\infty)$ , s'il peut s'écrire sous la forme :

$$X_t = \sum_{j=-\infty}^{+\infty} \pi_j X_{t-j}, \quad \text{avec } \pi_0 = 1$$

De plus, on peut dire qu'un modèle  $MA(\infty)$  est un modèle linéaire, encore si  $\phi_j = 0, \forall j < 0$ :

$$X_t = \sum_{j=0}^{+\infty} \phi_j \epsilon_{t-j}, \quad \text{avec } \epsilon_t \in \mathbb{N}$$

on dit dans ce cas que le modèle  $X_t$  est causal.

**Definition 1.7.2 fonction d'Autocorrélation Partielle**

La fonction d'autocorrélation partielle de  $X_t$ , notée  $\phi_h$ , est définie par :

$$\phi(h) = \text{corr}[X_t - \mathbb{E}(X_t | X_{t-1}, \dots, X_{t-h+1}) ; X_{t-h} - \mathbb{E}(X_{t-h} | X_{t-1}, \dots, X_{t-h+1})], \quad \forall h \in \mathbb{N}^*.$$

**Definition 1.7.3 Modèle Moyenne Mobile**

On appelle modèle moyenne mobile ("Moving Average") d'ordre  $q$ , abrégé  $MA(q)$ , un modèle stationnaire  $X_t$  qui vérifie :

$$X_t = c + \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \cdots + \theta_q \epsilon_{t-q} = c + \epsilon_t + \sum_{i=1}^q \theta_i \epsilon_{t-i},$$

Où:  $c \in \mathbb{R}, \theta_1, \dots, \theta_q \in \mathbb{R} (\theta_q \neq 0)$  et  $\epsilon_t \sim \mathcal{BB}(0, \sigma_\epsilon^2)$ .

Cette relation est équivalente à l'écriture :

$$X_t = \Theta(L) \epsilon_t,$$

Où:  $\Theta(L) = 1 + \theta_1 L + \theta_2 L^2 + \dots + \theta_q L^q$

Un modèle  $MA(q)$  vérifie les propriétés suivantes :

1. La variance:  $\gamma_X(0) = 1 + (\theta_1^2 + \theta_2^2 + \dots + \theta_q^2)\sigma_\epsilon^2$

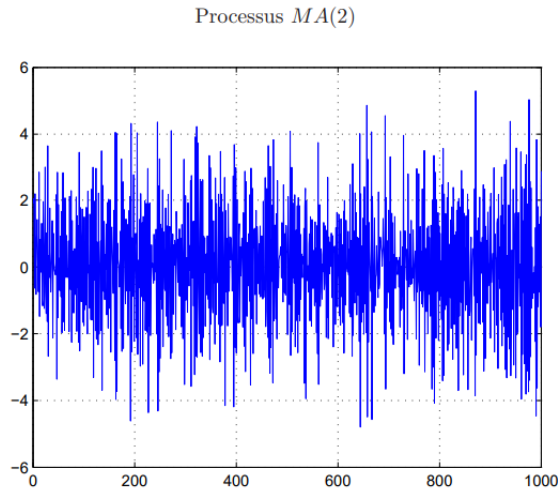
2. La fonction d'autocovariance:

$$\gamma_X(h) = \begin{cases} (\theta_h + \theta_{h+1}\theta_1 + \dots + \theta_q\theta_{q-h})\sigma_\epsilon^2, & \text{si } 1 \leq h \leq q, \\ 0 & \text{si } h > q. \end{cases}$$

3. La fonction d'autocorrélation:

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

La figure ci-dessous montre respectivement une réalisation d'un processus  $MA(2)$ , sa fonction de corrélation empirique et sa fonction de corrélation partielle empirique.



**(Moyenne mobile d'ordre infini  $MA(\infty)$ )**

On dit qu'un modèle  $X_t$  est  $MA(\infty)$ , s'il peut s'écrire sous la forme :

$$X_t = \sum_{j=-\infty}^{+\infty} \varphi_j \epsilon_{t-j}, \quad \epsilon_t \sim \mathcal{BB}(0, \sigma_\epsilon^2)$$

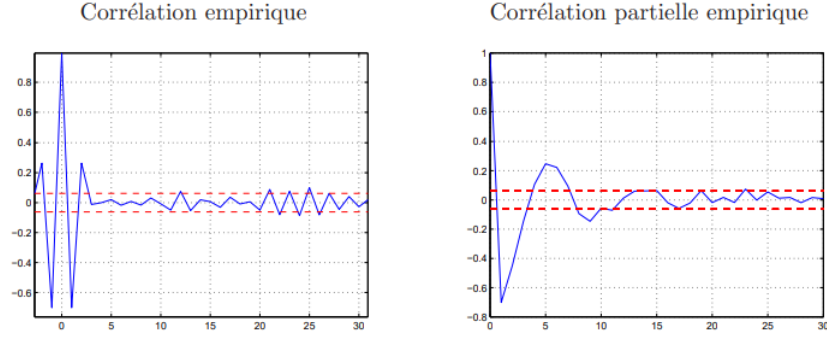


Figure 1.4: Propriétés d'un processus MA(2)

De plus, on peut dire qu'un modèle  $MA(\infty)$  est un modèle linéaire, encore si  $\varphi_j = 0, \forall j < 0$ :

$$X_t = \sum_{j=0}^{+\infty} \varphi_j \epsilon_{t-j}, \quad \text{avec } \epsilon_t \in \mathbb{N}$$

on dit dans ce cas que le modèle  $X_t$  est causal.

**Theorem 1** Tout modèle  $X_t$  centré et stationnaire (ou seconde ordre) peut être représenté sous une forme proche de la forme  $MA(\infty)$ .

**Definition 1.7.4** Processus ARMA On dit qu'un processus stationnaire  $\{X_t \in \mathbb{Z}\}$  admet une représentation  $ARMA(p, q)$  s'il vérifie l'équation :

$$X_t = \varphi_1 X_{t-1} + \varphi_2 X_{t-2} + \cdots + \varphi_p X_{t-p} + \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \cdots + \theta_q \epsilon_{t-q}$$

Où:  $(\epsilon_t)_{t \in \mathbb{Z}}$  est un  $\mathcal{BB}(0, \sigma_\epsilon^2)$  et  $\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q$  sont des coefficients réels, avec  $\phi_p \neq 0$  ainsi que  $\theta_q \neq 0$ . Autrement dit, en utilisant les polynômes en  $B$ .

$$\Phi(B) = I - \varphi_1 B - \varphi_2 B^2 - \cdots - \varphi_p B^p$$

et

$$\Theta(B) = I + \theta_1 B + \theta_2 B^2 + \cdots + \theta_q B^q,$$

un processus  $ARMA(p, q)$  vérifie l'équation:

$$\Phi(B)X_t = \Theta(B)\epsilon_t, \quad \text{pour tout } t \text{ dans } \mathbb{Z}$$

**Nb.** Naturellement un processus  $AR(p)$  est un  $ARMA(p, 0)$  et un processus  $MA(q)$  est un  $ARMA(0, q)$ .

**Definition 1.7.5** La représentation d'un processus ARMA,

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

est dite:

- **minimale** si les polynômes  $\Phi(z)$  et  $\Theta(z)$  n'ont pas de racine commune;
- **causale** si le polynôme  $\Phi(z)$  a toutes ses racines à l'extérieur du disque unité;
- **canonique** si elle est causale et inversible

Disons quelques mots sur le critère de représentation minimale. Si la représentation n'est pas minimale, c'est à dire si les polynômes  $\Phi(Z)$  et  $\Theta(z)$  ont une ou des racine(s) commune(s), alors deux cas peuvent se présenter.

- Soit aucune de ces racines communes n'est sur le cercle unité. Dans ce cas, le processus  $(X_t)_{t \in \mathbb{Z}}$  admet aussi la représentation  $\tilde{\Phi}(B)X_t = \tilde{\Theta}(B)\epsilon_t$ , pour tout  $t \in \mathbb{Z}$ , où les nouveaux polynômes sont obtenus à partir des précédents en enlevant les racines communes.
- Si au moins une des racines communes est sur le cercle unité, alors il peut y avoir plus d'un unique processus stationnaire vérifiant l'équation.

**Écriture MA( $\infty$ ) et AR( $\infty$ ):**

**Proposition 2 Écritures MA( $\infty$ )** Soit  $(X_t)_t \in \mathbb{Z}$  un processus ARMA( $p, q$ ) de représentation minimale et causale. Il admet alors la représentation MA( $\infty$ ):

$$X_t = \Phi^{-1}(B)\Theta(B)\epsilon_t = \epsilon_t + \sum_{i=1}^{+\infty} \psi_i \epsilon_{t-i}$$

où les coefficients  $(\psi_i)_{i \in \mathbb{N}}$  forment une famille absolument sommable et vérifient l'équation de récurrence linéaire :

$$\psi_i - \sum_{j=1}^p \varphi_j \psi_{i-j} = \theta_i, \text{ pour } i \in \mathbb{Z}$$

avec  $\psi_i = 0$ , pour  $i < 0$ ,  $\psi_0 = 1$ ,  $\theta_0 = 1$  et  $\theta_i = 0$  pour  $i > q$ .

**Proposition 3 Écriture  $AR(\infty)$**  Soit  $(X_t)_t \in \mathbb{Z}$  un processus  $ARMA(p, q)$  de représentation minimale et causale. Il admet alors la représentation  $AR(\infty)$ :

$$\epsilon_t = \Theta^{-1}(B)\Phi(B)_t =_t + \sum_{i=1}^{+\infty} \pi_{it-i}$$

où les coefficients  $(\pi_i)_{i \in \mathbb{N}}$  forment une famille absolument sommable et vérifient l'équation de récurrence linéaire :

$$\pi_i + \sum_{j=1}^q \theta_j \pi_{i-j} = -\varphi_i, \text{ pour } i \in \mathbb{Z}$$

avec  $\pi_i = 0$ , pour  $i < 0$ ,  $\varphi_0 = -1$ ,  $\varphi_i = 0$  et  $\theta_i = 0$  pour  $i > p$ .

## 1.8 Décomposition de Wold

Le théorème de Wold (1938) est considéré comme théorème fondamentale dans le domaine des séries temporelles. En vertu de ce théorème, tout processus stochastique  $X_t$  stationnaire au sens faible peut être écrit comme une combinaison linéaire.

**Theorem 2** Tout processus stationnaire d'ordre deux, non déterministe  $\{X_t; t \in \mathbb{Z}\}$  peut être représenté sous la forme:

$$X_t = \sum_{i=0}^{+\infty} \phi_i \epsilon_{t-i} + k_t, \text{ avec } t \in \mathbb{N}$$

Où les paramètres  $\phi_i$  satisfont  $\phi_0 = 1$ ,  $\phi_i \in \mathbb{R}$ ,  $\forall i \in \mathbb{N}^*$ ,  $\sum_{i=0}^{+\infty} \phi_i^2 < \infty$ ,  $\{\epsilon_t\} \sim \mathcal{BB}(0, \sigma_\epsilon^2)$

Le terme  $k_t$  désigne la composante linéaire déterministe telle que  $\text{Cov}(k_t; \epsilon_{t-i}) = 0$ ,  $\forall i \in \mathbb{Z}$

## 1.9 Processus ARIMA

Les processus étant dans la pratique rarement stationnaires, on a introduit une généralisation des processus  $ARMA$  vus précédemment de manière à les étendre à des processus non stationnaires. L'idée générale, essentiellement conçue pour les processus non stationnaires à

tendance polynomiale, est de différencier suffisamment le processus initial afin d'obtenir un processus sans tendance et sur la partie différenciée appliquer un modèle ARMA. Cette classe de modèle est connue sous le nom de processus ARIMA (AutoRegressive Integrated Moving Average).

**Definition 1.9.1** On dit qu'un processus  $\{X_t\}_{t \in \mathbb{Z}}$  admet une représentation ARIMA( $p, d, q$ ) s'il vérifie l'équation :

$$\Phi(B)\nabla^d X_t = \Theta(B)\epsilon_t,$$

pour tout  $t$  dans  $\mathbb{Z}$ , où:

$$\nabla^d = (I - B)^d$$

$$\Phi(B) = I - \varphi_1 B - \varphi_2 B^2 - \cdots - \varphi_p B^p$$

$$\Theta(B) = I + \theta_1 B + \theta_2 B^2 + \cdots + \theta_q B^q$$

Où:  $(\epsilon_t)_{t \in \mathbb{Z}}$  est un  $\mathcal{BB}(0, \sigma_\epsilon^2)$  et  $\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q$  sont des coefficients réels, avec  $\phi_p \neq 0$  ainsi que  $\theta_q \neq 0$ .

## 1.10 Processus SARIMA

Les modèles SARIMA ( $S$  pour Seasonal) permettent de rendre compte des variations saisonnières dans la série considérée, ces variations pouvant elles-mêmes présenter un caractère aléatoire.

**Definition 1.10.1** Le processus  $X_t$  est appelée SARIMA( $p, d, q$ ) $\times$ ( $P, D, Q$ ) $s$  si le processus différencié :

$$Y_t = (1 - B)^d (1 - B^s)^D X_t$$

est un processus ARMA tel que :

$$\Phi(B)\Phi_s(B^s)Y_t = \Theta(B)\Theta_s(B^s)\epsilon_t, \quad \text{soit la période de saisonnalité}$$

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

$$\Phi(z) = 1 - \varphi_1 z - \cdots - \varphi_p z^p$$

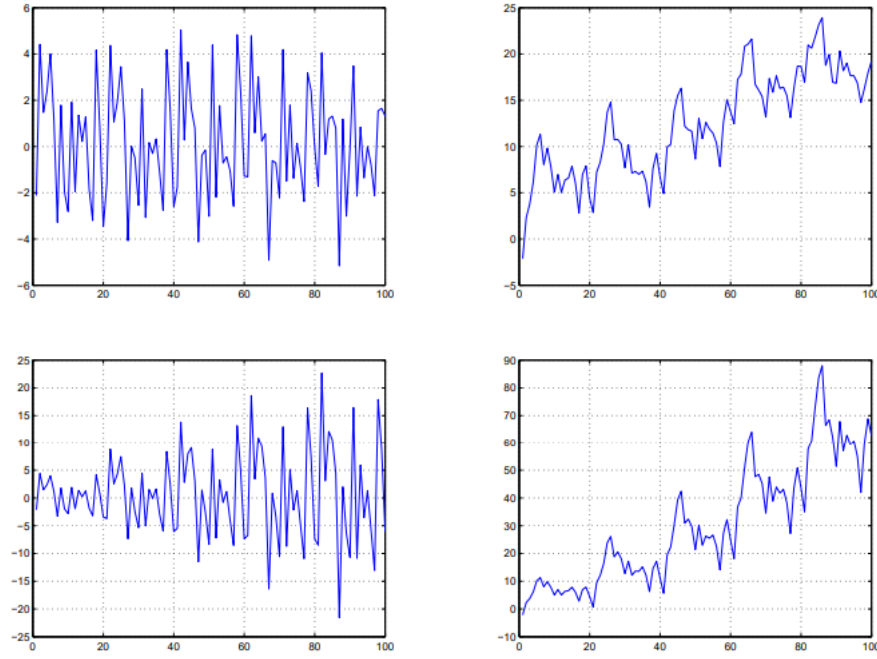


Figure 1.5: De l'ARMA au SARIMA

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

$\Phi_s$  et  $\Theta_s$  sont des polynômes de degrés respectifs  $P$  et  $Q$  :

$$\Phi_s(z) = 1 - \varphi_{s,1}z - \dots - \varphi_{s,P}z^P$$

$$\Theta_s(z) = 1 + \theta_{s,1}z - \dots - \theta_{s,Q}z^Q$$

La Figure 1.5 en dessus représente des parties de réalisations de divers niveaux de processus SARIMA, générés à partir d'une même séquence de bruit blanc. On trouve successivement (de gauche à droite puis de haut en bas) :

- un processus ARMA présentant une composante saisonnière de période 20,
- le même processus intégré une fois, à la période de l'échantillon,
- le processus initial intégré une fois, à la période de la saison,
- le processus a été intégré une fois à la période de l'échantillon, et une fois à la période de la saison

---

# Chapter 2

## La méthodologie de Box-Jenkins

---

*L'approche de Box-Jenkins (1976) consiste en une méthodologie rigoureuse étude systématique des série chronologique à partir de leur caractéristique. L'objectif est de déterminer le modèle le plus adapté à représenter le phénomène étudié. Cette méthodologie suggère une procédure à trois étapes :*

- 1. Identification du modèle.*
- 2. Estimation des paramètres du modèle.*
- 3. Validation du modèle*

### 2.1 Identification du modèle

*L'étape d'identification d'un processus ARMA (choix entre AR, MA et ARMA et choix de  $p$  et  $q$ ) de BJ est basé sur la comparaison des caractéristiques théoriques à leur équivalent empiriques ( c.à.d calculer sur la série observée), les caractéristiques utilisées sont les AC simples et partielles.*



- Si le corrélogramme simple n'a que ses  $q$  premiers terme différents de zéros et que les termes du corrélogramme partiel diminuent lentement, nous pouvons pronostiquer un MA ( $q$ ).
- Si le corrélogramme partiel n'a que ses  $p$  premiers terme différents de zéros et que les termes du corrélogramme simple diminuent lentement, cela caractérise un AR ( $p$ ).
- Si les fonctions d'AC simple et partiel ne paraissent pas tronquées, il s'agit d'un processus ARMA.

### 2.1.1 Estimation de $d$

En pratique si  $\rho_h$  est proche de 1 pour un grand nombre de retard c.à.d que la décroissance est lente donc on a une racine unité et le processus n'est pas stationnaire, il faut différencier pour avoir une série stationnaire.  $d$  représente le nombre de fois nécessaire pour stationnariser la série. Une autre méthode plus objective est d'utiliser le test de Dickey Fuller.

### 2.1.2 Estimation de $p$ et $q$

Si:  $X_t \sim \text{ARIMA}(p, d, q) \implies (1 - L)^d X_t \sim \text{ARMA}(p, q)$  On rappelle l'estimateur de la FAC:

$$\hat{\rho}_h = \frac{\sum_{t=h+1}^T (X_t - \hat{\mu})(X_t - h - \hat{\mu})}{\sum_{t=1}^T (X_t - \hat{\mu})^2}, h = 1, 2, \dots, T - 1.$$

En pratique, on calcule jusqu'à  $H = \frac{T}{4}$ . Selon Bartlett la variance du coefficient d'AC du processus où  $\rho_h = 0$  pour  $h > q$  est:

$$V(\hat{\rho}_h) = \frac{1}{T} \left( 1 + 2 \sum_{j=1}^q \hat{\rho}_j^2 \right)$$

d'où l'intervalle de confiance à 95% des AC:

$$\hat{\rho}_h \pm 1.96 \sqrt{\frac{1}{T} \left( 1 + 2 \sum_{j=1}^q \hat{\rho}_j^2 \right)}$$

alors  $X_t \sim MA(q)$ . (En pratique on prend  $\pm 1.96 \frac{1}{\sqrt{T}}$ ).

La variance des ACP où  $\varphi_{hh} = 0$  pour  $h > p$  est  $\frac{1}{T}$ , d'où l'IC à 95% des ACP est:

$$\varphi_{hh} \pm 1.96 \frac{1}{\sqrt{T}}.$$

## 2.2 Estimation des paramètres du modèle

### 2.2.1 La méthode des moments

Pour le modèle donné, on sait que les autocorrélations  $\rho_h$  et  $\tau_h$ , des processus AR et MA respectivement, dépendent des paramètres  $\xi = (\varphi_1, \dots, \varphi_p, \theta_1, \dots, \theta_q)$  selon des équations théoriques connues:

$$(\rho_1, \dots, \rho_p, \tau_1, \dots, \tau_q) = \mathcal{F}(\varphi_1, \dots, \varphi_p, \theta_1, \dots, \theta_q)$$

On sait facilement estimer les autocorrélations  $\rho_h$  et  $\tau_h$ , il suffit donc d'inverser les équations pour estimer les paramètres:

$$(\hat{\varphi}_1, \dots, \hat{\varphi}_p, \hat{\theta}_1, \dots, \hat{\theta}_q) = F^{-1}(\hat{\rho}_1, \dots, \hat{\rho}_p, \hat{\tau}_1, \dots, \hat{\tau}_q).$$

### 2.2.2 La méthode des moindres carrés (MC)

**Modèle AR(p):**

Soit:

$$X_t = \delta + \varphi_1 X_{t-1} + \dots + \varphi_p X_{t-p} + \epsilon_t \quad \text{où } \epsilon_t \sim \mathcal{BB}(0, \sigma^2).$$

En donnant  $N$  observations  $x_1, \dots, x_N$ , les paramètres  $\delta, \varphi_1, \dots, \varphi_p$  peuvent être estimés par la méthode des moindres carrés en minimisant:

$$S = \sum_{t=p+1}^N (x_t - \delta + \varphi_1 x_{t-1} + \dots + \varphi_p x_{t-p})^2$$

par rapport à  $\delta, \varphi_1, \dots, \varphi_p$ .

Si  $\epsilon_t \sim N \implies$  les estimateurs des MC sont aussi les estimateurs du MV conditionnel sur les  $p$  premières valeurs.

### Modèle MA :

Le problème d'estimation est plus difficile pour le modèle MA, on ne peut pas trouver des estimateurs explicitement on obtient une forme itérative numérique.

### 2.2.3 La méthode du maximum de vraisemblance

Soit le modèle ARMA suivant:

$$X_t = \delta + \sum_{j=1}^p \varphi_j X_{t-j} + \epsilon_t - \sum_{j=1}^q \theta_j \epsilon_{t-j}.$$

avec:  $\epsilon_t \sim BB(0, \sigma^2)$ . Soit:  $\vartheta = (\delta, \varphi_1, \dots, \varphi_p, \theta_1, \dots, \theta_q, \sigma^2)$  le vecteur des paramètres. Soit la série chronologique de taille  $n$ :  $x_1, \dots, x_n$ , l'approche du MV consiste à calculer la densité de probabilité:

$$\mathcal{F}_{X_n, X_{n-1}, \dots, X_1}(x_n, x_{n-1}, \dots, x_1; \vartheta) \quad (*)$$

L'estimateur du MV de  $\vartheta$  est la valeur qui maximise (\*). On doit spécifier la distribution de  $\epsilon_t$ , on suppose que  $\epsilon_t \sim iid \mathcal{N}(0, \sigma^2)$ .

Trouver l'EMV  $\implies$  2 étapes:  $\begin{cases} 1. \text{calculer la fct de } V(*) \\ 2. \text{Trouver les valeurs de } \vartheta \text{ qui maximise cette fonction.} \end{cases}$

## 2.3 Validation

Lors de la détermination des ordres  $p$  et  $q$  du processus ARMA  $(p, q)$ , à l'aide du corrélogramme simple et partiel, on peut être amené à sélectionner plusieurs ordres possibles. Après avoir estimé les différents processus possibles, il reste à les valider. la validation passe par un examen

des coefficients estimés (ils doivent être significativement différent de 0) et par un examen des résidus (les résidus estimés doivent suivre un processus bruit blanc  $\hat{\epsilon}_t \sim BB(0, \sigma^2)$  où  $\hat{\epsilon}_t$  est l'estimateur de l'erreur  $\epsilon_t$ ).

### 2.3.1 Test de Student sur les coefficients

On vérifie d'abord que les racines des polynômes AR et MA ne sont pas égales à 1. On teste la significativité des coefficients par des tests de student. On calcule donc la statistique de Student du coefficients:

$$t_{\hat{\varphi}_j} = \frac{\hat{\varphi}_j}{\sqrt{\hat{V}[\hat{\varphi}_j]}}$$

que l'on compare à la valeur critique lue dans la table de la loi de Student. La règle de décision est alors :

- Si  $|t_{\hat{\varphi}_j}| < t_{1-\frac{\alpha}{2}}$ , on accepte l'hypothèse nulle  $H_0: \varphi_j = 0$ .
- Si  $|t_{\hat{\varphi}_j}| > t_{1-\frac{\alpha}{2}}$ , l'hypothèse nulle est rejetée :  $H_0: \varphi_j \neq 0$ .

Où  $t_{1-\frac{\alpha}{2}}$  est le quantile d'ordre  $(1-\frac{\alpha}{2})$  de loi de student à  $(T-h)$  degrés de liberté,  $h$  étant le nombre des paramètres estimés.

### 2.3.2 Test sur les résidus

Il s'agit de vérifier que les résidus du modèle estimé, noté  $\hat{\epsilon}_t$ , vérifient les propriétés requises pour que l'estimation soit valide c.à.d : qu'ils suivent un processus bruit blanc non autocorrélé et de même variance et qu'ils suivent une loi normale. D'abord il faut regarder le graphique des  $\hat{\epsilon}_t$  pour voir s'il apparaît des points aberrants, une tendance, une rupture, de l'autocorrélation etc, mais ceci n'est qu'indicatif.

## Test d'autocorrélation

**Analyse des FAC:** Regarder les AC simples et partielles, elles doivent être significativement égal à 0 si les résidus sont un BB.

$$\hat{\rho}_h = \frac{\hat{\gamma}_h}{\hat{\gamma}_0}, l'IC \text{ dans le cas d'un BBG : } \left[ \frac{\pm 1.96}{\sqrt{T}} \right]$$

Pour un BB il faut qu'aucune valeur du corrélogramme ne soit significativement différente de 0.

## Test de Ljung-Box(1978)

$$H_0 : \rho_1 = \rho_2 = \dots = \rho_h = 0 \quad VS \quad H_1 : \exists \text{ un } \rho_i \neq 0$$

La statistique du test est:  $Q_{LB}(h) = T(T+2) \sum_{h=1}^H \frac{\hat{\rho}_h^2}{T-h}$  Sous l'hypothèse nulle d'absence d'autocorrélation :

$$\hat{\rho}_1^2 = \hat{\rho}_2^2 = \dots = \hat{\rho}_h^2 = 0$$

La statistique  $Q_{LB}(h)$  suit une loi de Khi-Deux à  $(H-p-q)$  degrés de liberté.

**Tests d'hétéroscédasticité** Nous étudierons ici le test ARCH car il est très fréquemment employé en économétrie des séries temporelles financières.

**Test ARCH** Ce test consiste à effectuer une régression autorégressive des résidus carrés sur  $q$  retard :  $e_t^2 = \theta_0 + \sum_{j=1}^q \theta_j e_{t-j}^2$

Où  $e_t$  désigne le résidu à l'instant  $t$  issu de l'estimation des paramètres du processus ARMA( $p, q$ ). Pour déterminer le nombre de retards  $q$ , on étudie le corrélogramme des résidus au carré. Les hypothèses du test ARCH sont les suivantes :

- $H_0$ :homoscédasticité:  $\theta_1 = \dots = \theta_q = 0$ .
- $H_1$ :hétéroscédasticité:  $\exists \text{ un } \theta_i \neq 0$

La statistique du test est  $T \times R^2$  où  $T$  correspond au nombre d'observations de la série  $e_t$  et  $R^2$  représente le coefficient de détermination associé à la régression. Sous l'hypothèse  $H_0$  la statistique  $T \times R^2$  suit la loi du Khi-deux 0 q degrés de liberté. La règle de décision est alors :

- Si  $T \times R^2 \leq X_q^2$  où  $X_q^2$  on accepte  $H_0$  d'homoscédasticité.
- Si  $T \times R^2 > X_q^2$  où  $X_q^2$  on rejette  $H_0$  et on admet qu'il y a de l'hétéroscédasticité.

**Test de normalité** On utilise le test de Jarque et Bera (1984) basé sur la notion de skewness (coefficient d'asymétrie) et de Kurtosis (aplatissement-épaisseur des queues de distribution). Soit  $\mu_k$  le moment empirique d'ordre  $k$  du processus:

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

Les coefficients de la Skewness ( $S_k$ ) et de la Kurtosis ( $k_u$ ) est alors définie par:

$$(S_k)^{\frac{1}{2}} = \frac{\mu_3}{\mu_2^{\frac{3}{2}}} \xrightarrow{T \rightarrow \infty} \mathcal{N}(0, \sqrt{\frac{6}{T}})$$

$$K_u = \frac{\mu_4}{\mu_2^2} \xrightarrow{T \rightarrow \infty} \mathcal{N}(3, \sqrt{\frac{24}{T}})$$

On construit alors les statistiques centrées réduites correspondantes à  $(S_k)^{\frac{1}{2}}$  et  $K_u$ , qu'on compare aux seuils d'une loi normale centrée réduite. La statistique est calculée comme suit:

$$\frac{(S_k)^{\frac{1}{2}}}{\sqrt{\frac{6}{T}}} \xrightarrow{T \rightarrow \infty} \mathcal{N}(0, 1)$$

$$\frac{k_u - 3}{\frac{24}{T}} \xrightarrow{T \rightarrow \infty} \mathcal{N}(0, 1)$$

Si la statistique centrée réduite de  $(S_k)^{\frac{1}{2}}$  est inférieure au seuil 1,96 à 5%, on accepte l'hypothèse de symétrie et l'hypothèse de normalité. Si la statistique centrée réduite de  $K_u$  est inférieure au seuil 1,96 à 5%, on accepte l'hypothèse de queue de distributions plates et l'hypothèse de normalité.

Le test de Jarque et Bera regroupe ces deux tests en un seul test. On construit la statistique :

$$S = \frac{T}{6}S_k + \frac{T}{24}(K_u - 3)^2 \xrightarrow{T \rightarrow \infty} \chi_2^2 = 5.99$$

Donc si  $S \geq \chi_2^2(1 - \alpha)$  on rejette l'hypothèse  $H_0$  de normalité des résidus au seuil de  $\alpha\%$ .

### 2.3.3 Critères de choix des modèles

Après examen des coefficients et des résidus, certains modèles sont écartés. Pour départager les modèles restants, on fait appel aux critères standards et critères d'information.

#### Critères standards

L'erreur absolue moyenne (MAE):

$$MAE = \frac{1}{T} \sum_t |e_t|$$

Racine de l'erreur quadratique moyenne (RMSE)

$$RMSE = \sqrt{\frac{1}{T} \sum_t e_t^2}$$

Ecart absolu moyen en pourcentage (MAPE)

$$MAPE = 100 \frac{1}{T} \sum_t \left| \frac{e_t}{X_t} \right|$$

Plus la valeur de ces critères est faible, plus le modèle estimé est proche des observations.

### 2.3.4 Critère d'information

AIC (Akaike):

$$AIC(p, q) = \log \hat{\sigma}^2 + \frac{2(p + q)}{T}$$

**Le critère de Schwarz (1978):**

$$SC = (p, q) = \log(\sigma_{\hat{\varepsilon}_t}^2) + (p + q) \frac{\log T}{T}$$

*On choisit le modèle qui minimise les critères standards et les critères d'information. Le modèle sélectionné sera alors utilisé pour la prévision.*

## 2.4 Prévision

*Après validation du modèle on procède à des prévisions sur un horizon limité car la variance de l'erreur de prévision croît rapidement avec l'horizon. Si l'on désire avoir la valeur de  $X_{t+h}$ , tout en supposant qu'on est à l'instant  $t$ , on utilise l'estimateur:*

$$\widehat{X_t(h)} = E(X_{t+h}/X_t, X_{t-1}, \dots, X_1)$$

*avec  $t$  l'origine de prévision et  $h$  l'horizon de prévision.*

*Dans le cas d'un modèle ARMA, on aura la forme :  $\widehat{X_t(h)} = \varphi_1 E(X_{t+h-1}/X_t, X_{t-1}, \dots, X_1) + \dots + \varphi_p E(X_{t+h-p}/X_t, X_{t-1}, \dots, X_1) - \theta_1 E(\varepsilon_{t+h-1}/X_t, X_{t-1}, \dots, X_1) - \dots - \theta_q E(\varepsilon_{t+h-q}/X_t, X_{t-1}, \dots, X_1) + E(\varepsilon_{t+h}/X_t)$  avec:*

- $E(X_{t-j}/X_t, X_{t-1}, \dots, X_1) = X_{t-j}, \forall j \geq 0.$
- $E(X_{t+j}/X_t, X_{t-1}, \dots, X_1) = \widehat{X_t(j)}, \forall j \geq 1.$
- $E(\varepsilon_{t-j}/X_{t-1}, \dots, X_1) = \varepsilon_{t-j}, \forall j \geq 0.$
- $E(\varepsilon_{t+j}/X_{t-1}, \dots, X_1) = 0, \forall j \geq 1.$

### 2.4.1 Erreur de prévision

*On considère un processus ARMA( $p, q$ ) telle que :*

$$X_t = \delta + \varphi_1 X_{t-1} + \dots + \varphi_p X_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q}$$



avec  $(\varphi_p, \theta_q) \in \mathcal{R}^{2*}$  et  $\varepsilon_t \sim iid(0, \sigma_\varepsilon^2)$  Appliquons le th eor eme de Wold au processus  $(X_t, t \in \mathcal{Z})$  et consid erons la forme  $MA(\infty)$  correspondante :(avec  $\delta = 0$ )

$$X_t = \sum_{i=0}^{\infty} \psi_i \varepsilon_{t-i}; \quad \varepsilon_0 = 1$$

Il s'ensuit que la meilleur pr evision que l'on peut faire de  $X_{t+1}$  compte tenu de toute l'information disponible jusqu'a la date  $t$  not ee  $\widehat{X}_t(1)$  est donn ee par:

$$\widehat{X}_t(1) = E(X_{t+1}/X_t, X_{t-1}, \dots, X_1) = E(X_{t+1}/\varepsilon_t, \varepsilon_{t-1}, \dots, \varepsilon_1) = \sum_{i=1}^{\infty} \psi_i \varepsilon_{t+1-i}$$

L'erreur de pr evision " $\varepsilon_t$ " est appel e aussi l'innovation :

$$X_{t+1} - \widehat{X}_t(1) = \varepsilon_{t+1}$$

plus g en eralement pour une pr evision   un horizon  $k$  on a :

$$\widehat{X}_t(k) = \sum_{i=k}^{\infty} \psi_i \varepsilon_{t+k-i}$$

$$X_{t+k} - \widehat{X}_t(k) = \sum_{i=0}^{k-1} \psi_i \varepsilon_{t+k-i}$$

D eterminons un intervalle de confiance sur la pr evision  $\widehat{X}_t(k)$ , sous l'hypoth ese de normalit e des r esidus  $\varepsilon_t$ . On montre alors que:

$$E[X_{t+k} - \widehat{X}_t(k)]^2 = E\left[\sum_{i=0}^{k-1} \psi_i \varepsilon_{t+k-i}\right]^2 = \sum_{i=0}^{k-1} \psi_i^2 \sigma_\varepsilon^2$$

D'o u:

$$\frac{X_{t+k} - \widehat{X}_t(k)}{\delta_\varepsilon [\sum_{i=0}^{k-1} \psi_i^2]^{1/2}} \sim \mathcal{N}(0, 1)$$

On peut donc construire un intervalle de confiance sous la forme :

$$IC = [\widehat{X}_t(k) \pm t_{\frac{\alpha}{2}} \widehat{\sigma} [\sum_{i=0}^{k-1} \psi_i^2]^{1/2}].$$

---

# Chapitre 3

## Applications sous R

---

*Ce chapitre présente les notions de R spécifiques à la manipulation et l'analyse de base de séries temporelles. On considère les données de <https://fr.finance.yahoo.com/> contenant les valeurs de clôture journalière de la société/composante **Lenovo Group Limited** de l'indice boursier **Hang Seng** du 1er novembre 2021 au 7 juillet 2022.*

### Importer et lire les données de la série temporelle

*Avant toute analyse d'une série temporelle, il est nécessaire d'importer les données dans un format reconnu par R. Un format classique d'échange de données est le fichier texte. Comme nous l'avons vu, l'import en R de ce type de fichier se fait par les commandes R:*

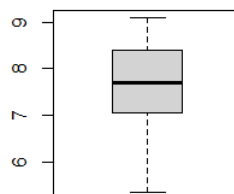
```
setwd("/Users/hp/Documents/dm_sc")
lenovo<-read.csv("lenovo.csv", header=TRUE, skip = 0)
str(lenovo)
```

### Statistiques de base d'une série temporelle avec R

*En plus des statistiques de base présentées ci-dessus, l'analyse d'une série temporelle passe bien souvent par une analyse descriptive dont nous présentons certains éléments. **Boxplot***

*La boxplot permet de représenter de manière synthétique les caractéristiques centrales et de dispersion des données.*

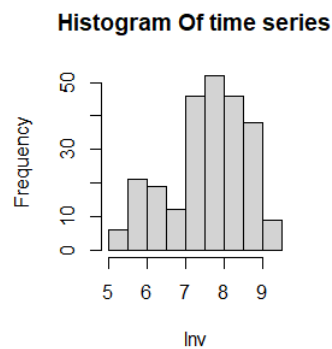
```
boxplot(lnv)
```



## ***Histogrammes***

*L'histogramme donne une estimation de la densité correspondant aux différentes réalisations dans le temps d'un processus.*

```
hist(lnv, main = 'Histogram Of time series')
```



*Le calcul des statistiques de base sur cette série s'effectue ainsi:*

```
mean(lnv)
sd(lnv)
summary(lnv)
acf(lnv)
acf(lnv, plot = FALSE)
pacf(lnv)
eacf(lnv)
```

```

> mean(lnv)
[1] 7.575422
> sd(lnv)
[1] 0.9969501
> summary(lnv)
      Close
Min.   :5.350
1st Qu.:7.060
Median :7.710
Mean   :7.575
3rd Qu.:8.400
Max.   :9.100
> acf(lnv)
> acf(lnv, plot = FALSE)

Autocorrelations of series 'lnv', by lag

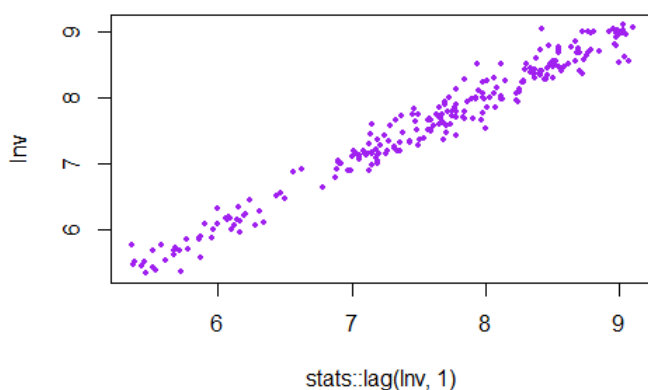
    1      2      3      4      5      6      7      8      9     10     11     12     13     14     15     16     17
0.978 0.956 0.941 0.926 0.909 0.890 0.872 0.851 0.831 0.814 0.798 0.781 0.767 0.755 0.745 0.732 0.718
    18     19     20     21     22     23
0.708 0.698 0.686 0.673 0.664 0.650

```

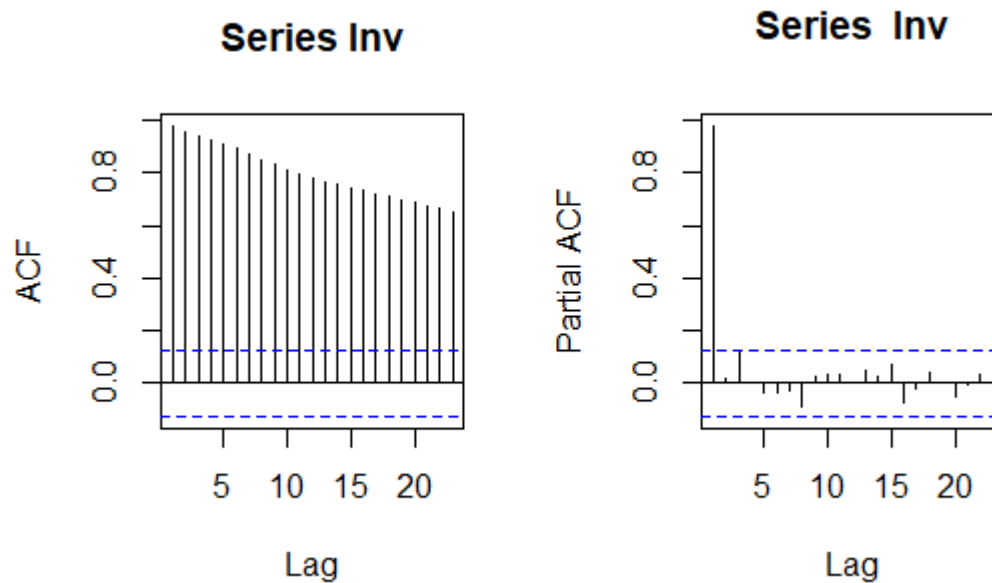
Le résultat de ces calculs s'affiche comme suit:

**Remark 3** On constate d'après les résultats affichés ci-dessus une forte corrélation d'ordre 1 (0.978) Voici le nuage de point de la série en fonction de la série retardée d'une heure le confirme

```
plot(stats::lag(lnv,1),lnv,pch=20,cex=0.8,col = "purple")
```



L'estimation de la fonction d'auto-corrélation et d'auto-corrélation partielle se fait via les fonction *Acf* et *Pacf*. Les graphique ci-dessous représentent la fonction d'auto-corrélation empirique (resp la fonction d'auto-corrélation partielle empirique) sont appelés les auto-corrélogrammes:



### La Manipulation de la série temporelle

- *En utilisant la commande "ts":* Voici l'affichage de la série temporelle dans le

```
lINV = ts(lenovo)
lINV
```

*console:*

```
Time Series:
Start = 1
End = 249
Frequency = 1
      Close
[1,]  8.50
[2,]  8.30
[3,]  8.06
[4,]  8.06
[5,]  7.98
[6,]  8.13
[7,]  8.25
[8,]  8.15
[9,]  8.08
[10,] 8.25
[11,] 8.01
[12,] 7.98
[13,] 7.93
```

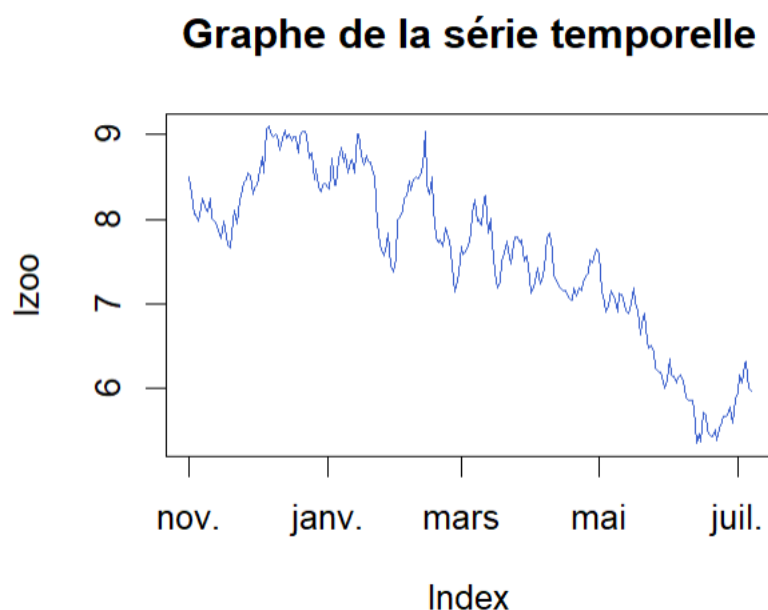
- *En utilisant la commande "zoo":*

```
#ajouter les dates
days <- seq(as.Date("2021-11-01"), as.Date("2022-07-07"), by = "days")
zoo(lenovo, days)
lzoo = zoo(lnv, days)
```

## Représentation graphique de la série temporelle

Une fois importée, il est intéressant de représenter graphiquement une série temporelle. Pour représenter un objet en R, la fonction de base est `plot()` qui se décline ensuite de différentes manières selon les objets qu'on représente.

```
plot(lzoo,col='royalblue3',main="Graphe de la série temporelle")
```



**Remark 4** D'après le graphe on déduit intuitivement qu'il s'agit d'une décomposition multiplicative car on constate une décroissance non constante, en effet, à mesure que la série temporelle décroît la variation saisonnière décroît également.

## Détection de la tendance

Pour détecter la tendance sous-jacente on utilise un lissage moyenne mobile sur la série temporelle, en indiquant que la période de la saisonnalité est égale à 12:

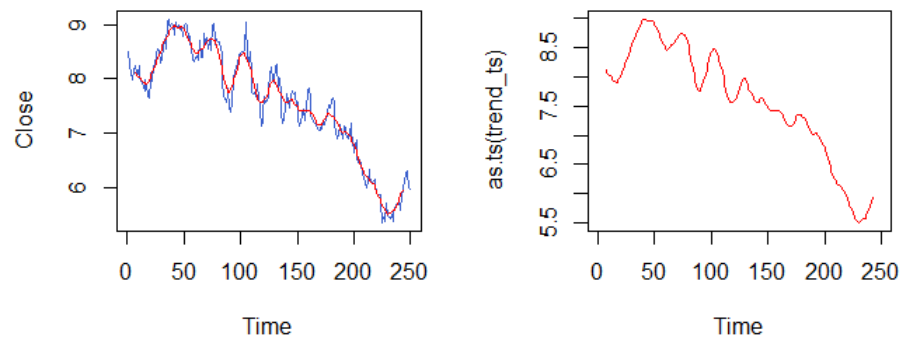


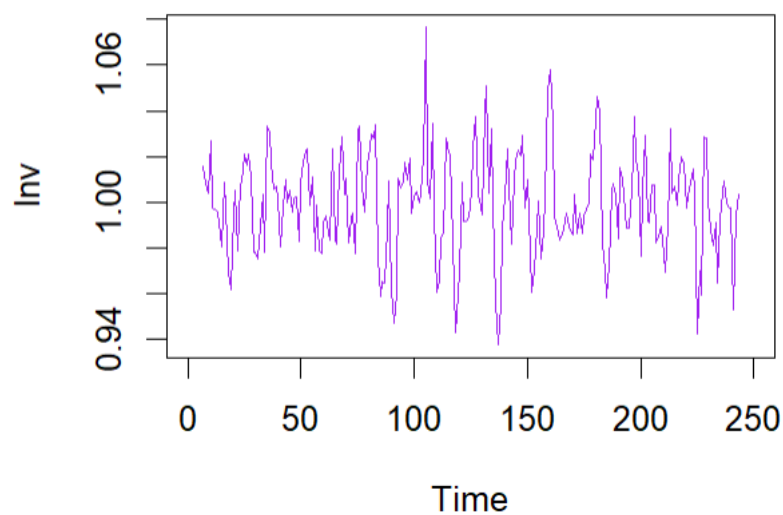
Figure 3.1: Modélisation de la tendance

### Éliminer la tendance(Detrend The time Series)

Puisque il s'agit d'une décomposition multiplicative on divise la série temporelle par la tendance.

```
detrend_ts = lnv / trend_ts
plot(as.ts(detrend_ts), col='purple')
```

### Série temporelle sans tendance



## Modélisation de la saisonnalité

*D'après l'élimination de la tendance il est facile de représenter la saisonnalité.*

```
m_ts = t(matrix(data = detrend_ts, nrow = 20))
seasonal_ts = colMeans(m_ts, na.rm = T)
plot(as.ts(rep(seasonal_ts, 12)), col='blue')
```

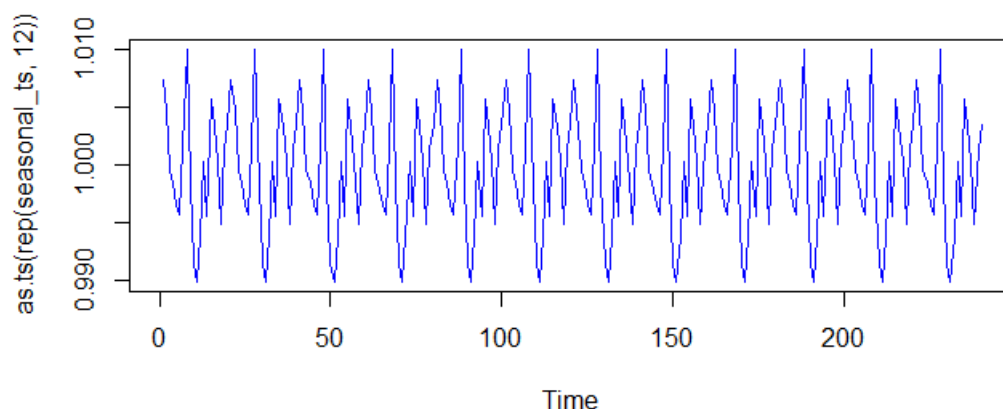


Figure 3.2: Modélisation de la saisonnalité

## Correction de la tendance par la méthode de différentiation

```
diff.lnv <- diff(lzoo, lag = 1, differences = 1)
acf(diff.lnv)
plot(lnv, col = "blue")
plot(diff.lnv, col = "orangered2")
```

## Test de la stationarité

*Il en résulte que: En appliquant deux test de la stationnarité(adf.test et adfTest, on obtient une p-value est superieur à 0.05,alors on conclut que la série temporelle est non stationnaire(sachant que le test adfTest "type C" donne le meilleur résultat).*



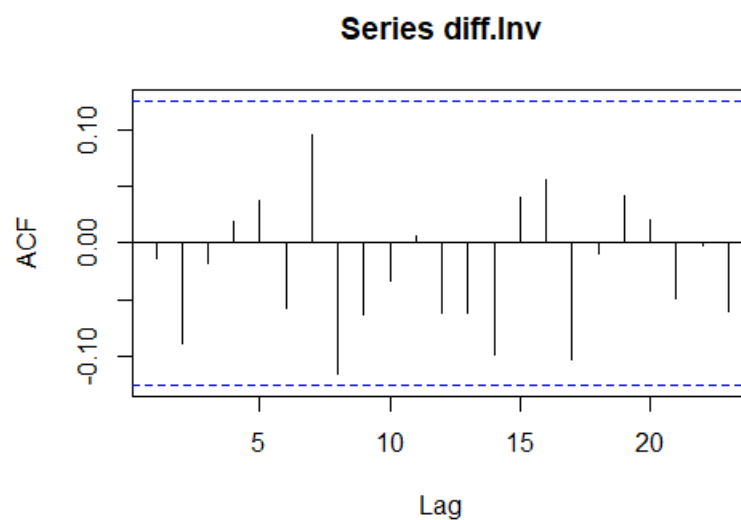


Figure 3.3: l'auto-corrélogramme après la correction de la tendance

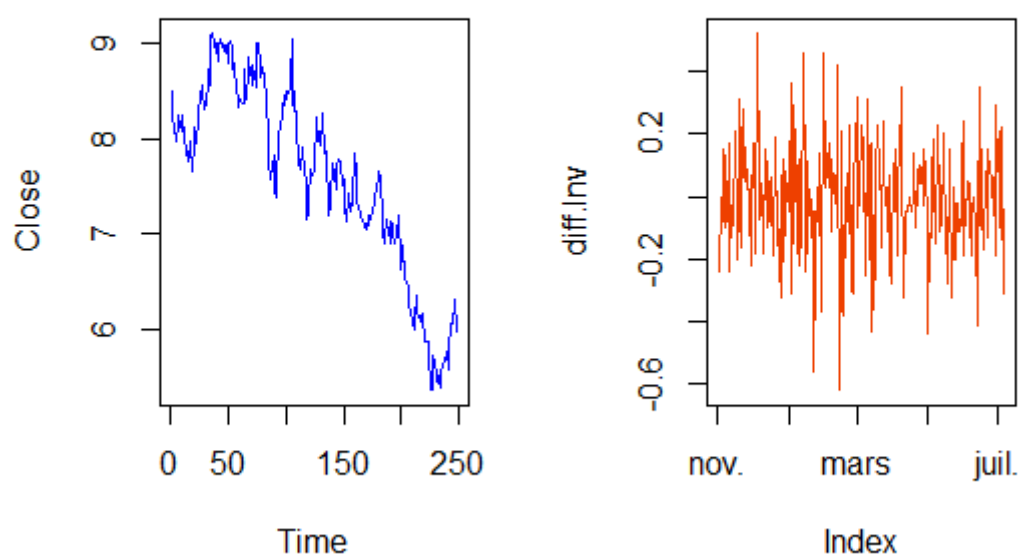


Figure 3.4: La série temporelle avant et après la correction de la tendance

### Test de la normalité

$H_0$ : la variable dont provient les données suit une loi normale

$H_1$ : la variable dont provient les données ne suit pas une loi normale. Etant donné que la  $p$ -

```
stat=ts(data=k[,1])
stat
adf.test(lnv,alternative="stationary")
adfTest(lnv,lags=1,type=c("c"),title=NULL,description=NULL)
```

Figure 3.5: La série temporelle avant et après la correction de la tendance

```
Lag Order: 1
STATISTIC:
  Dickey-Fuller: -0.9817
P VALUE:
  0.691

shapiro.test(lnv)

jarque.bera.test(lnv)

      Jarque Bera Test

data:  lnv
X-squared = 15.723, df = 2, p-value = 0.0003854

> shapiro.test(lnv)

      shapiro-wilk normality test

data:  lnv
W = 0.9467, p-value = 6.852e-08
```

value calculée est inférieure au niveau de signification  $\alpha=0.05$ , on doit rejeter l'hypothèse nulle  $H_0$ , et retenir l'hypothèse alternative  $H_1$  donc d'après les deux tests les données ne suivent pas une loi normale.

## Modélisation des modèles ARIMA

Dans les deux méthodes le meilleur modèle est celui avec le plus faible AIC.

### 1ere méthode du maximum vraisemblance

On modélise notre série temporelle avec deux modèles d'ARIMA, les voici:

```
## Modélisation (Estimation des modèles ARIMA)"méthode 1"==
fit=arima(diff.lnv,order=c(1,0,1),method = "ML")
fit
fit=arima(diff.lnv,order=c(1,1,1),method = "ML")
fit
```

```
Call:
arima(x = diff.lnv, order = c(1, 0, 1), method = "ML")
```

```
Coefficients:
          ar1          ma1  intercept
          0.9435         -1.0000         -0.0115
s.e.        0.0229         0.0151         0.0022
```

```
sigma^2 estimated as 0.02968:  log likelihood = 83.19,  aic = -160.39
> fit=arima(diff.lnv,order=c(1,1,1),method = "ML")
> fit
```

```
Call:
arima(x = diff.lnv, order = c(1, 1, 1), method = "ML")
```

```
Coefficients:
          ar1          ma1
          -0.0100         -1.0000
s.e.        0.0638         0.0119
```

```
sigma^2 estimated as 0.03069:  log likelihood = 77.01,  aic = -150.03
```

## 2eme méthode: Box-Jenkins

```
### Estimation des modèles ("méthode 2")
#ARIMA(1,1,0)
arima(diff.lnv,order=c(1,1,0))
#ARIMA(0,1,1)
arima(d.s,order=c(0,1,1))
#ARIMA(1,1,1)
arima(d.s,order=c(1,1,1))
#ARIMA(1,1,3)
arima(d.s,order=c(1,1,3))
#ARIMA(2,1,3)
arima(d.s,order=c(2,1,3))
```

*D'après les résultats affichés il est bien clair que le meilleur modèle est: ARIMA(0,1,1)*

```

> #ARIMA(1,1,0)
> arima(diff.lnv,order=c(1,1,0))

Call:
arima(x = diff.lnv, order = c(1, 1, 0))

Coefficients:
      ar1
    -0.4658
s.e.    0.0563

sigma^2 estimated as 0.04859:  log likelihood = 22.92,  aic = -43.84
> #ARIMA(0,1,1)
> arima(diff.lnv,order=c(0,1,1))

Call:
arima(x = diff.lnv, order = c(0, 1, 1))

Coefficients:
      ma1
    -1.0000
s.e.    0.0118

sigma^2 estimated as 0.03069:  log likelihood = 77,  aic = -152

> #ARIMA(1,1,1)
> arima(diff.lnv,order=c(1,1,1))

Call:
arima(x = diff.lnv, order = c(1, 1, 1))

Coefficients:
      ar1      ma1
    -0.0101    -1.0000
s.e.    0.0638    0.0119

sigma^2 estimated as 0.03069:  log likelihood = 77.01,  aic = -150.03
> #ARIMA(1,1,3)
> arima(diff.lnv,order=c(1,1,3))

Call:
arima(x = diff.lnv, order = c(1, 1, 3))

Coefficients:
      ar1      ma1      ma2      ma3
    -0.8528   -0.1469   -0.9261    0.0730
s.e.    0.0836    0.1083    0.0587    0.0739

sigma^2 estimated as 0.0301:  log likelihood = 79.3,  aic = -150.59

> #ARIMA(2,1,3)
> arima(diff.lnv,order=c(2,1,3))

Call:
arima(x = diff.lnv, order = c(2, 1, 3))

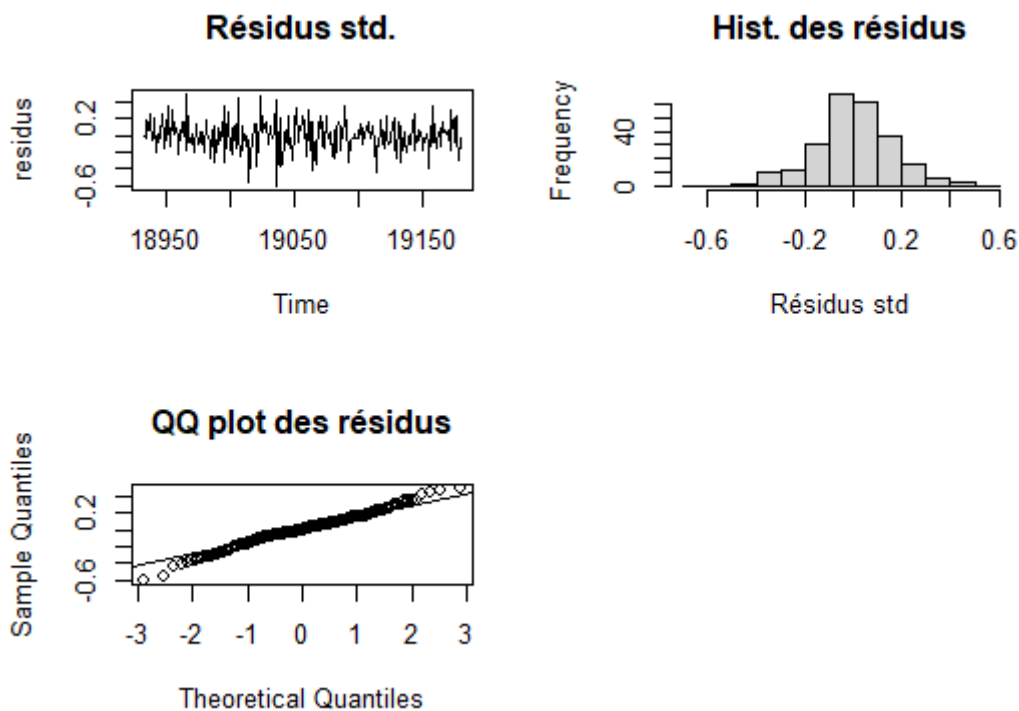
Coefficients:
      ar1      ar2      ma1      ma2      ma3
    0.0083    0.8005   -1.0083   -0.8985    0.9083
s.e.    0.0485    0.0473    0.0056    0.0772    0.0106

sigma^2 estimated as 0.02981:  log likelihood = 80.66,  aic = -151.31

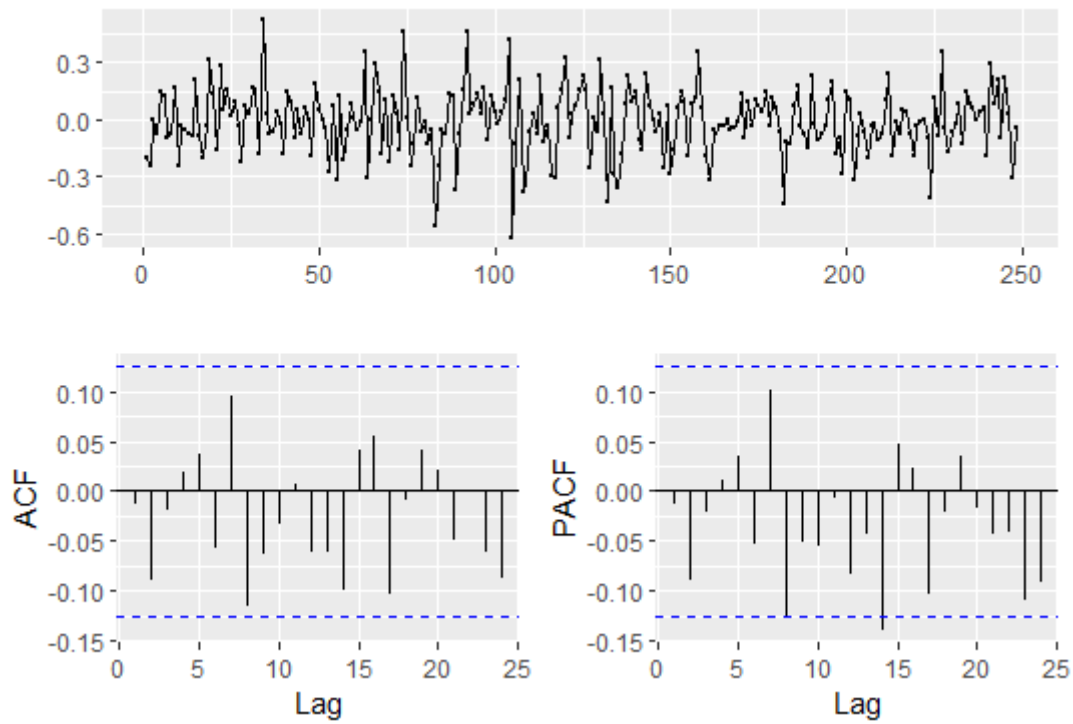
```

## Analyse des résiduels

```
##Analyse des résidus
residus=fit$residuals
par(mfrow=c(2,2))
plot.ts(residus,main="Résidus std.")
hist(residus,main="Hist. des résidus",xlab="Résidus std")
qqnorm(residus,main="QQ plot des résidus")
qqline(residus)
shapiro.test(residus)
```



### Exemple d'une Modélisation d'un modèles SARIMA



*Le Plot de l'acf des résidus de ce modèle SARIMA, montre que toutes les autocorrélations sont dans les limites du seuil, ce qui indique que les résidus se comportent comme un bruit blanc*

# Bibliography

---

- [1] **Yves Aargon**, *Séries temporelles avec R*, 2011, Springer Publishing, incorporated
- [2] **Régis Bourbonnais**, **Michel Terraza**, *Analyse des séries temporelles*, 3e édition, 2010, Donud, Paris.
- [3] **Terence C.Mills**, *Applied Time Series Analysis*, 2019, ACADEMIC PRESS, An imprint of Elsevier.
- [4] **Douglas C.Montgomery**, **Cheryl L.Jennings**, **Murat Kulahci**, *Time Series Analysis And Forecasting*, 2015, Published by John Wiley Sons, Inc., Hoboken, New Jersey..
- [5] **Hamilton, J. D**(1994): *Time Series Analysis*. Princeton University Press, Princeton, New Jersey.
- [6] <https://otexts.com/fpp2/>
- [7] <https://www.math.u-bordeaux.fr/~hzhang/m2/st/TS1.pdf>