

الجمهورية الجزائرية الديمقراطية الشعبية  
République Algérienne Démocratique et Populaire

وزارة التعليم العالي والبحث العلمي  
Ministère de l'Enseignement Supérieur et de la Recherche  
Scientifique  
Centre Universitaire Salhi Ahmed- Naama  
Institut des sciences et technologies  
Département de Mathématiques et Informatique



## *Mémoire de fin d'études*

En vue de l'obtention du diplôme de Master  
En : Mathématiques

Spécialité : Probabilités, Statistique et Application

## Intitulé

---

**Les modèles GARCH et application dans les prévisions**

---

**Présenté par :**

AMARA Mohammed Ridha

**Soutenu :** Septembre 2022

**Devant le jury composé de :**

|                           |     |              |              |
|---------------------------|-----|--------------|--------------|
| Dr.DAOUIDI Hemza          | MCB | C-Univ Naâma | Président    |
| Dr.BELGUERNA Abderrahmane | MCA | C-Univ Naâma | Encadreur    |
| Dr.LAALA Zeyneb           | MCB | C-Univ Naâma | Examinatrice |

**Année universitaire 2021/2022**

# Remerciement

Avant tout, je remercie **ALLAH** qui m'a donné le courage et la volonté d'aller jusqu'au bout dans ce modeste travail.

Je tiens à exprimer toute ma reconnaissance à mon encadrant de mémoire Monsieur **BELGUERNA Abderrahmane**, je le remercie de m'avoir encadré, orienté, aidé et conseillé durant toute la période du projet.

J'adresse mes sincères remerciements à tous les professeurs, qui ont acceptés d'être membres du jury de juger et d'examiner ce travail madame **LAALA Zeyneb** et monsieur **DAOUIDI Hemza**.

Un grand merci d'adresse particulièrement à ma famille qui ma soutenus pour réaliser mes études et qui ma toujours encouragés dans chaque pas vers l'avant.

Et enfin je remercie tous mes collègues de la promo Pour leur soutien, et à tous ceux qui ont contribué dans ma formation de prés et de loin.

***Merci a tous.***

# Table des matières

|   |           |
|---|-----------|
| <b>Introduction</b>   | <b>5</b>  |
| <b>1 Concepts des séries chronologiques</b>                                 | <b>7</b>  |
| 1.1 Série chronologique . . . . .   | 7         |
| 1.1.1 Définition . . . . .  | 7         |
| 1.1.2 Domaines d'application . . . . .                                      | 7         |
| 1.1.3 L'objectif de l'analyse des séries chronologique . . . . .            | 8         |
| 1.2 Analyse classique des séries chronologiques . . . . .                   | 9         |
| 1.2.1 Indices descriptifs d'une série chronologie . . . . .                 | 9         |
| 1.2.2 Composantes et schéma d'une série chronologique . . . . .             | 10        |
| 1.3 Processus aléatoires stationnaires . . . . .                            | 12        |
| 1.3.1 Processus stochastique . . . . .                                      | 12        |
| 1.3.2 La stationnarité au sens strict (forte) . . . . .                     | 13        |
| 1.3.3 La stationnarité d'ordre deux des processus : la stationnarité faible | 13        |
| 1.3.4 Le processus Bruit Blanc (White Noise) . . . . .                      | 14        |
| 1.3.5 Auto-covariance, auto-corrélation simple, auto-corrélation partielle  | 14        |
| 1.3.6 Les opérateurs . . . . .  | 16        |
| 1.3.7 Processus auto-régressifs . . . . .                                   | 19        |
| 1.3.8 Processus moyennes mobiles . . . . .                                  | 20        |
| 1.3.9 Les processus mixtes (ARMA) . . . . .                                 | 21        |
| 1.3.10 La méthodologie de Box-Jenkins . . . . .                             | 22        |
| 1.4 Processus aléatoires non stationnaires . . . . .                        | 23        |
| 1.4.1 Les processus TS et DS . . . . .                                      | 23        |
| 1.4.2 Les processus ARIMA . . . . .   | 25        |
| 1.4.3 Les processus SARIMA . . . . .  | 26        |
| <b>2 Processus Conditionnellement Hétéroscédastiques</b>                    | <b>27</b> |
| 2.1 Le modèle ARCH . . . . .  | 27        |
| 2.1.1 Présentation du processus ARCH . . . . .                              | 27        |
| 2.1.2 Propriétés du processus ARCH . . . . .                                | 28        |
| 2.1.3 Condition de stationnarité . . . . .                                  | 30        |

|          |  |           |
|----------|--|-----------|
| 2.1.4    | Le cas d'un ARCH(1) . . . . .                                      | 30        |
| 2.2      | Le modèle GARCH . . . . .  | 30        |
| 2.2.1    | Présentation du processus GARCH . . . . .                          | 30        |
| 2.2.2    | Propriétés du processus GARCH . . . . .                            | 31        |
| 2.2.3    | Condition de stationnarité . . . . .                               | 32        |
| 2.3      | Extentions des modèles ARCH/GARCH, et les modèles asymétriques . . | 32        |
| 2.4      | Modélisation des modèles GARCH . . . . .                           | 33        |
| 2.4.1    | Identification . . . . .   | 33        |
| 2.4.2    | Estimation . . . . .   | 33        |
| 2.4.2.1  | Estimation des paramètres pour un modèle ARCH . . .                | 34        |
| 2.4.2.2  | Estimation des paramètres pour un modèle GARCH . .                 | 34        |
| 2.4.3    | Prévision (dans le cas d'un ARMA-GARCH) . . . . .                  | 35        |
| <b>3</b> | <b>Application</b>   | <b>36</b> |
| 3.1      | Installation et Libération (packages) . . . . .                    | 36        |
| 3.1.1    | Installation . . . . .   | 36        |
| 3.1.2    | libération . . . . .   | 37        |
| 3.2      | Lecture et Analyse . . . . .                                       | 37        |
| 3.2.1    | Représentation graphique . . . . .                                 | 37        |
| 3.2.1.1  | Lecture et Positionnement . . . . .                                | 37        |
| 3.2.2    | Stationnarité . . . . .  | 37        |
| 3.2.3    | Tests . . . . .  | 39        |
| 3.2.3.1  | KPSS . . . . .   | 39        |
| 3.2.3.2  | ADF . . . . .  | 39        |
| 3.3      | Modélisation ARIMA-GARCH . . . . .                                 | 39        |
| 3.3.1    | La partie ARIMA . . . . .  | 40        |
| 3.3.1.1  | Ajustement vers ARIMA avec le moins d'AIC . . . . .                | 40        |
| 3.3.1.2  | Analyse résiduelle du ARIMA trouvé . . . . .                       | 40        |
| 3.3.1.3  | Test de Box-Ljung . . . . .  | 41        |
| 3.3.1.4  | Effet ARCH sur les résidus . . . . .                               | 41        |
| 3.3.2    | La partie GARCH . . . . .  | 42        |
| 3.4      | Prévisions . . . . .   | 43        |
| <b>4</b> | <b>Conclusion</b>  | <b>46</b> |
|          | <b>Bibliographie</b>   | <b>46</b> |

# Table des figures

|     |  |    |
|-----|--|----|
| 1.1 | prix d'un produit et leur prévision . . . . .                          | 8  |
| 1.2 | Représentation d'une à une tendance . . . . .                          | 10 |
| 1.3 | Illustration de la composante saisonnière dans une chronique . . . . . | 11 |
| 1.4 | Illustration de la stationnarité dans une chronique . . . . .          | 11 |
| 1.5 | Les trois graphes de l'application . . . . .                           | 18 |
| 3.1 | Représentation graphique du série initiale . . . . .                   | 38 |
| 3.2 | Représentation graphique du série après différenciation . . . . .      | 38 |
| 3.3 | Graphismes de l'analyse . . . . .                                      | 40 |
| 3.4 | l'ACF et la PACF . . . . .   | 42 |

# Introduction

Connaître le futur, ou du moins avoir une idée sur le futur est l'un des soucis de l'homme depuis toujours. De nos jours aussi, les raisons socioéconomiques poussent à anticiper l'avenir. Une question importante est de savoir sur quoi nous appuyer pour prédire l'avenir. Il est donc primordial d'arriver à prévoir le mieux possible le futur en s'appuyant sur le passé.

D'une façon mathématique, on peut formuler le problème de la prévision, en utilisant la discipline statistique (particulièrement, analyse des séries chronologique).

L'analyse des séries chronologiques (séries temporelles), se réfère à la branche de statistique où les observations sont régulièrement espacées dans le temps. Le thème d'analyse des séries ont l'un des sujets les plus statistiques que les statisticiens ont traités en raison de son importance pour expliquer les phénomènes qui se produisent pendant une période donnée.

Cette analyse vise à obtenir une description précise de la série et à construire un modèle approprié pour expliquer son comportement et utiliser les résultats pour prédire le comportement de la série chronologique dans le futur.

Dans l'analyse traditionnelle de la prévision, la construction des valeurs futures prévues est fondée sur la moyenne conditionnelle de la série utilisée. Les méthodes classiques de prévision basées sur les processus *ARMA* supposent des séries temporelles à volatilité constante. Cette modélisation n'est pas toujours conforme à la réalité, elle néglige l'information contenue dans la partie non expliquée du processus dévolution des séries temporelles. Ces modèles linéaires des séries temporelles n'étaient finalement fondés que sur des combinaisons linéaires de valeurs présentes et passés de chocs.

Une autre approche qui a muni une nouvelle modélisation des séries chronologiques (qui ont une volatilité stochastique), c'est les modèles *ARCH* et *GARCH*. Les travaux sur les processus conditionnellement hétéroscédastiques d'*ARCH* introduits par Engle en 1982 et leur extension *GARCH* (*ARCH* généralisés) sont dus à Bollerslev (1986). Leurs caractérisations repose essentiellement sur le concept de variance conditionnelle. Dans ces modèles, celle-ci s'écrit comme une fonction  $\alpha$  ne des valeurs passées du carré de la série. Ces modèles ont permis de modéliser la majeure partie de la dynamique non

linéaire des séries ont la volatilité stochastique.

L'objectif général de ce travail est de tenter montrer théoriquement la méthode utilisée pour faire la prédiction par un processus qui suit le modèle *GARCH*, ainsi de comment faire l'application sur des données réels.

Ce mémoire est partagé en trois chapitres : deux théoriques, et une application. Le premier est consacré pour les concepts des séries chronologique (notions, méthodes, caractéristiques, processus stationnaires, modèles classiques). Le deuxième c'était pour les Processus Conditionnellement Hétéroscédastiques (présentation, propriétés, méthode de modélisation). En fin une application sur un modèle particulier qui est défini sur la partie théorique.

# Chapitre 1

## Concepts des séries chronologiques

### 1.1 Série chronologique

#### 1.1.1 Définition

Une suite d'observations  $\{X_t, t \in \mathbb{T}\}$  indexées par un ensemble  $T$  (qui exprime le temps, avec  $t = 1, \dots, n$ ), est une série chronologique de grandeur  $n$ . L'indice  $T$  ( $T \in \mathbb{Z}, \mathbb{N}$ ) être selon le cas (l'année, le mois, le jour, l'heure etc....), et il peut aussi être un ensemble des réels ( $T \in \mathbb{R}$ ).

*Elle s'appelle aussi chronique ou série temporelle.*

#### 1.1.2 Domaines d'application

- Finance et économétrie : évolution des indices boursiers, des prix, des données économiques des entreprises, des ventes et achats de biens, des productions agricoles ou industrielles .
- Energie : prévision de la consommation d'électricité, pétrole .
- Assurance : analyse des sinistres .
- Medecine / Biologie : suivi des évolutions des pathologies, analyse d'électro-encéphalogrammes et d'électrocardiogrammes .
- Science de la terre et de l'espace : indices de marées, variations des phénomènes physiques (météorologie), évolution des taches solaires ,phénomènes d'avalanche .
- Traitement de signal : signaux de communications, de radars, de sonars, analyse de la parole.
- Traitement des données : mesures successives de position ou de direction d'un objet mobile (trajectographie) .

### 1.1.3 L'objectif de l'analyse des séries chronologique

Le principal objectif de l'analyse des séries temporelles est de prédire l'avenir d'un certain processus sur la base de ce qui s'est passé dans le passé. Par exemple, sur le marché financier, les traders veulent prédire le prix d'une certaine action dans les jours à venir sur la base des prix précédents. Par exemple :

la figure 1.1 illustre la prédiction d'une série chronologique financière ( prix d'un produit durant une période)

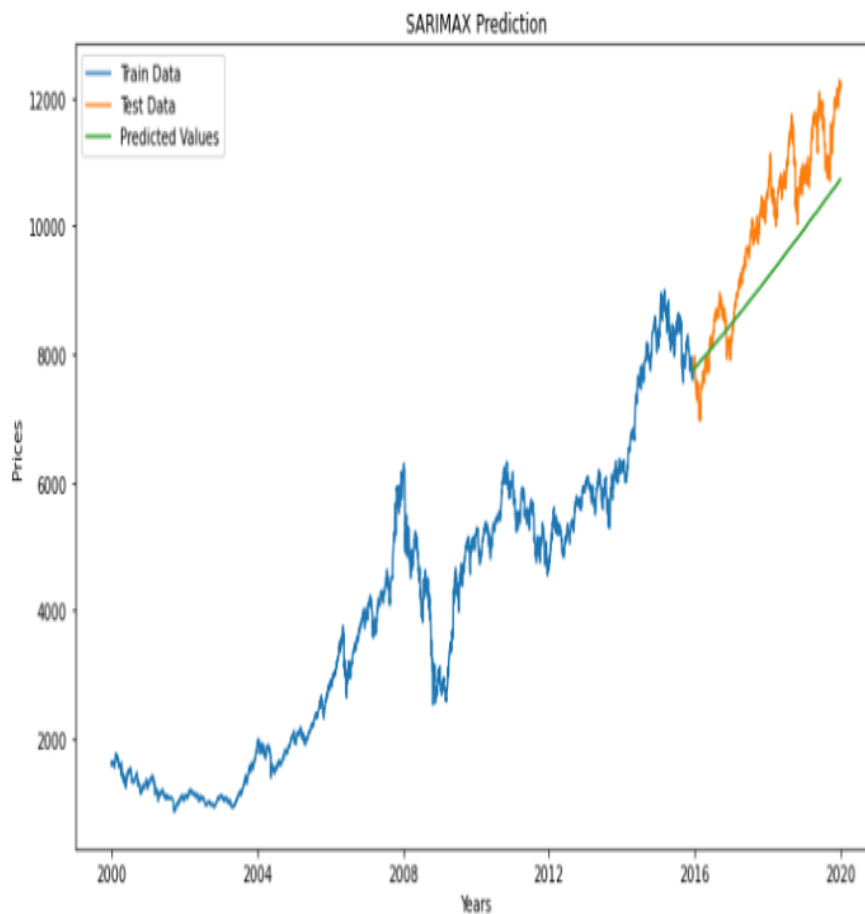


FIGURE 1.1 – prix d'un produit et leur prévision

le résultat du prévision est illustré par la couleur orange (la prévision était après l'année 2016).

## 1.2 Analyse classique des séries chronologiques

Considérons  $n$  observations  $\{x_1, \dots, x_n\}$  d'une série chronologique  $\{X_t, t \in \mathbb{T}\}$ .

### 1.2.1 Indices descriptifs d'une série chronologie

#### La moyenne :

Elle indique le centre de la distribution.

$$\bar{x}_n = \frac{1}{n} \sum_{i=1}^n x_t$$

#### La variance :

La variance empirique est utilisée comme indicateur de dispersion de la série.

$$\hat{\sigma}_n = \frac{1}{n} \sum_{i=1}^n (x_t - \bar{x}_n)^2$$

#### Indice de dépendance

##### - Autocovariance :

La fonction d'auto-covariance empirique est donnée par :

$$\hat{\sigma}_n(h) = \frac{1}{n-h} \sum_{t=h+1}^n (x_t - \bar{x}_n)(x_{t-h} - \bar{x}_n)$$

Pour trouver la dépendance entre deux données successives, on cherche l'auto-covariance empirique d'ordre 1 :

$$\hat{\sigma}_n(1) = \frac{1}{n-1} \sum_{t=2}^n (x_t - \bar{x}_n)(x_{t-1} - \bar{x}_n)$$

Pour trouver la dépendance entre deux données écartées par deux pas de temps, on cherche l'auto-covariance empirique d'ordre 2 :

$$\hat{\sigma}_n(2) = \frac{1}{n-2} \sum_{t=3}^n (x_t - \bar{x}_n)(x_{t-2} - \bar{x}_n)$$

##### - Auto-corrélation :

Les quotients des covariances empiriques par la variance empirique, donnent Les auto-corrélations empiriques.

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

On utilise ces indices pour caractériser la dépendance entre les variables.

### 1.2.2 Composantes et schéma d'une série chronologique

Une technique traditionnelle de traitement des séries chronologiques est la décomposition de la chronique, et puis la reconstitution pour faire les prévisions.

Cette approche décompose la série en quatre composantes principales :

**1. La tendance (trend)  $T$  :** Représente l'évolution à long terme de la série étudiée. Elle traduit le comportement "moyen" de la série. Par exemple la figure 1.4 représente une tendance qui augmente d'une façon linéaire :

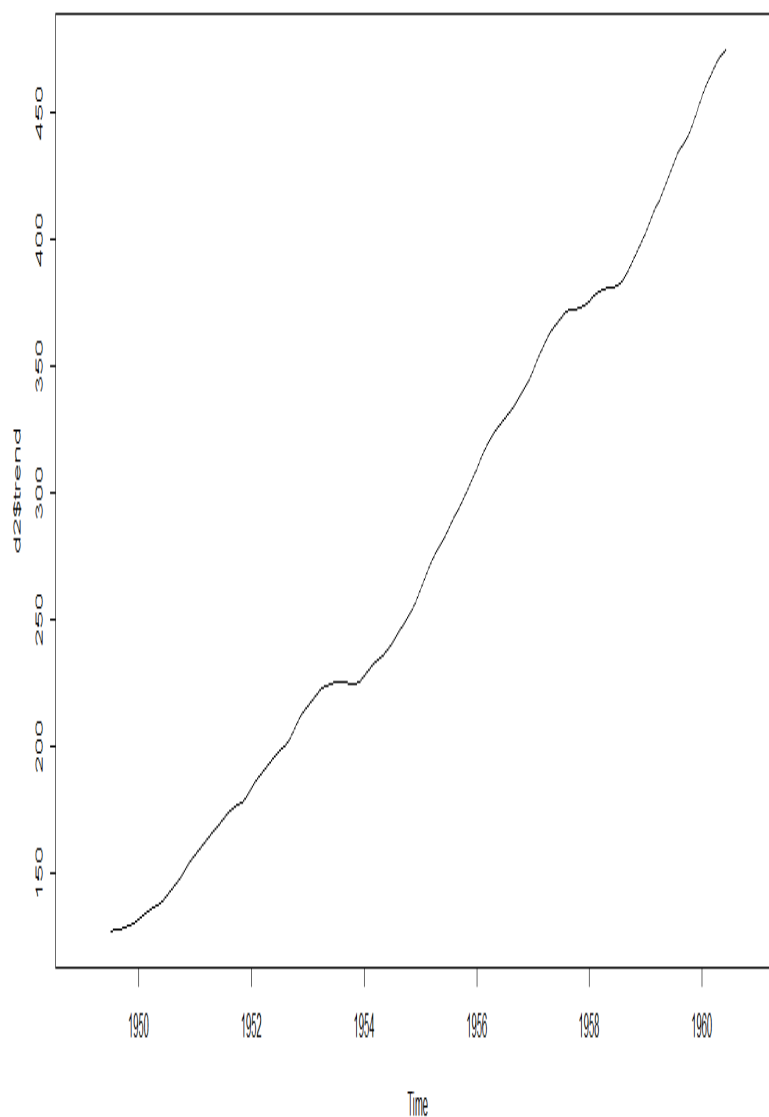


FIGURE 1.2 – Représentation d'une à une tendance

2. **La composante cyclique  $C_t$**  : elle est représentée par le cycle de Kitchin d'une période de 3 à 5 ans. Dans la plupart des travaux sur les séries chronologiques, la tendance et le cycle sont regroupés en une seule composante appelée **l'extra-saisonnier  $E_t$** .

3. **La composante saisonnière (ou saisonnalité)  $S_t$**  : Correspond à un phénomène qui se répète à intervalles de temps réguliers (périodique). En général, c'est un phénomène saisonnier d'où le terme de variations saisonnières. Par exemple la figure 1.5 représente des cycles réguliers au cours du temps et de même amplitude :

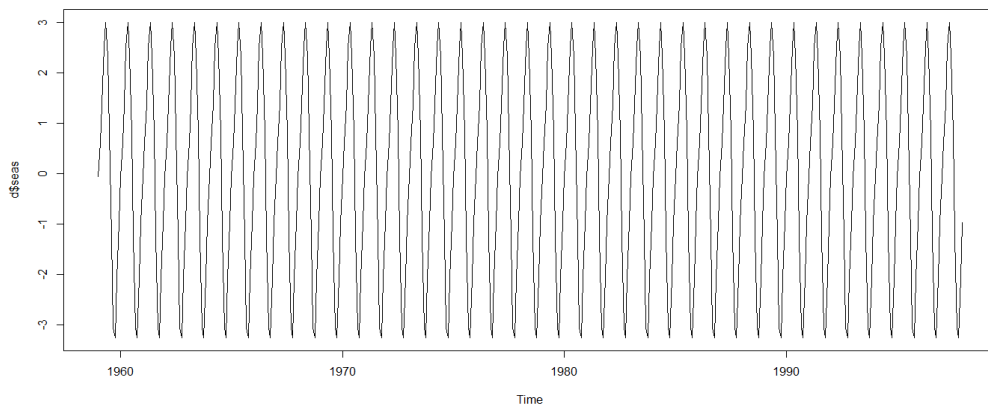


FIGURE 1.3 – Illustration de la composante saisonnière dans une chronique

4. **La composante résiduelle (bruit)  $R_t$**  : Correspond à des fluctuations irrégulières, en général de faible intensité mais de nature aléatoire. Par exemple la figure 1.6 représente des fluctuations irrégulières :

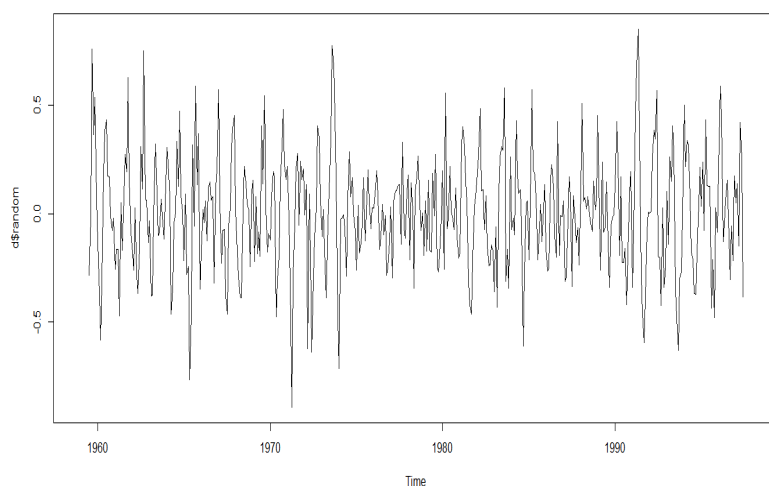


FIGURE 1.4 – Illustration de la stationnarité dans une chronique

**schémas de décomposition :****1. Le schéma additif :**

Qui suppose qu'il y a une indépendance entre les différentes composantes de la chronique. Il s'écrit :

$$X_t = E_t + S_t + R_t$$

**2. Le schéma multiplicatif :** Dans ce schéma, la composante saisonnière est liée à l'extra-saisonnier. Il s'écrit :

$$X_t = E_t \times S_t + R_t$$

**3. Le schéma multiplicatif complet :**

Qui peut lier les trois composantes  $(E, S, R)$ . Il s'écrit :

$$X_t = E_t \times S_t \times R_t$$

## 1.3 Processus aléatoires stationnaires

On va définir dans cette section les processus stochastiques, les notions de stationnarité. Ainsi on va donner des modèles (stationnaires, non stationnaires) usuels.

### 1.3.1 Processus stochastique

**Définition 1.3.1**

Soit  $X$  une application qui associe au couple  $(\omega, t)$  la quantité  $X_t(\omega)$ . Donc  $\forall t \in T$  fixé,  $X_t$  est une variable aléatoire définie sur un espace probabilisé.

Une famille de variables aléatoires indicées par  $t$  ( $t \in T$ ) représente un processus stochastique  $(X_t)_{t \in T}$  (dit aussi processus aléatoire).

$T$  est l'espace des indices, il exprime le temps.

$t$  est l'instant d'observation de la variable pour un individu  $\omega$

$\omega$  un est un individu ( $\omega \in \Omega$ , ou  $\Omega$  est l'espace fondamental)

On dit que le processus est discret ou bien continu, selon le cas de  $\mathbb{T}$  :

— Si  $T = \mathbb{R}$  ou un intervalle de  $\mathbb{R}$ . Alors  $X_t$  est un processus aléatoire "continu".

— Si  $T = \mathbb{Z}$  ou  $T = \mathbb{N}$ . Alors  $X_t$  est un processus aléatoire "discret".

On considère dans la suite que  $T = \mathbb{Z}$

### 1.3.2 La stationnarité au sens strict (forte)

#### Définition 1.3.2

Soit  $(X_t)_{t \in \mathbb{Z}}$  un processus aléatoire.

Pour un  $n$ -uplet du temps  $t_1 < t_2 < \dots < t_n$  avec  $t_i \in \mathbb{Z}$

Et pour tout  $h \in \mathbb{Z}$  tel que  $t_i + h \in \mathbb{Z}$ ,  $\forall i$  ( $i = 1, \dots, n$ )

Si les deux suites  $(X_{t_1+h}, \dots, X_{t_n+h})$  et  $(X_{t_1}, \dots, X_{t_n})$  ont la même loi de probabilité, on dit que le processus stochastique  $(X_t)_{t \in \mathbb{Z}}$  est strictement stationnaire (ou fortement stationnaire).

La loi de probabilité de la suite  $(X_{t_1}, \dots, X_{t_n})$  peut être caractérisée par sa fonction de répartition. Donc on peut définir une condition équivalente de la condition précédente :

$$\forall (x_1, \dots, x_n), \forall (t_1, \dots, t_n), \text{ et } \forall h :$$

$$P[x_{t_1} < x_1, \dots, x_{t_n} < x_n] = P[x_{t_1+h} < x_1, \dots, x_{t_n+h} < x_n]$$

#### conclusion 1.3.1

On remarque, d'après la dernière condition, que la loi du processus est toujours inchangée pour toute translation régulière dans le temps. Donc on conclure que la loi d'un processus fortement stationnaire ne dépend pas du temps. Ainsi on peut définir la propriété suivante :

#### Proposition 1.3.1

Un processus est fortement stationnaire si tout ses moments ne sont pas variées pour tout changement dans le temps.

### 1.3.3 La stationnarité d'ordre deux des processus : la stationnarité faible

#### Définition 1.3.3

Si  $(X_t)_{t \in \mathbb{Z}}$  un processus aléatoire vérifie les propriétés suivantes :

1.  $\forall t \in \mathbb{Z}, E(x_t) = m$
2.  $\forall t \in \mathbb{Z}, V(x_t) = \sigma^2$  (finie)
3.  $\forall t \in \mathbb{Z}, \forall h \in \mathbb{Z}, Cov(x_t, x_{t+h}) = \gamma(h)$  (la covariance dépend de la différence du temps ( $h$ ) seule et non du temps)  
 $\gamma(h)$  est l'auto-covariance d'ordre  $h$  de  $X_t$ .

Alors il est faiblement stationnaire (stationnaire au second ordre)

#### Proposition 1.3.2

Si les moments d'ordre 1 et d'ordre 2 (les deux seulement) d'un processus  $(X_t)_{t \in \mathbb{Z}}$  sont stationnaires, Alors il est faiblement stationnaire. (si par exemple  $E(x_t^3)$  n'est pas stationnaire, c'est-à-dire dépend de temps, alors dans ce cas le processus est faiblement stationnaire)

**Propriété 1.3.1**

Tout processus stationnaire au sens strict est un processus faiblement stationnaire.  
La réciproque est généralement fausse.

**1.3.4 Le processus Bruit Blanc (White Noise)****Définition 1.3.4**

On dit que le processus  $(\varepsilon_t)_{t \in \mathbb{Z}}$  est un **bruit blanc faible** s'il possède les propriétés suivantes :

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

En d'autres termes les variables aléatoires  $\varepsilon_t$  sont de moyenne nulle, de variance constante et non corrélées.

On dit que le processus  $(\varepsilon_t)_{t \in \mathbb{Z}}$  est un **bruit blanc fort**, s'il est bruit blanc faible et que les variables  $\varepsilon_t$  sont indépendantes (dans ce cas, elles sont indépendantes et identiquement distribuées (iid)).

On dit que le processus  $(\varepsilon_t)_{t \in \mathbb{Z}}$  est un **bruit blanc gaussien**, s'il est bruit blanc fort et si  $\forall t, \varepsilon_t \in \mathcal{N}(0, \sigma^2)$ .

**Notation :**

- Si  $(\varepsilon_t)_{t \in \mathbb{Z}}$  un bruit blanc faible, on le note par :  $\varepsilon_t \sim WN(0, \sigma_\varepsilon^2)$
- Si  $(\varepsilon_t)_{t \in \mathbb{Z}}$  un bruit blanc fort, on le note par :  $\varepsilon_t \sim IID(0, \sigma_\varepsilon^2)$
- Si  $(\varepsilon_t)_{t \in \mathbb{Z}}$  un bruit blanc gaussien, on le note par :  $\varepsilon_t \sim \mathcal{N}(0, \sigma^2)$

**Remarque 1.3.1**

Le bruit blanc est un processus stationnaire.

**1.3.5 Auto-covariance, auto-corrélation simple, auto-corrélation partielle****Définition 1.3.5 (La fonction d'auto-covariance)**

La covariance entre les valeurs d'un processus  $(X_t)_{t \in \mathbb{Z}}$  à deux dates décalées par une distance  $h$ , est mesurée par la fonction d'auto-covariance. Elle est définie par :

$$\begin{aligned} \gamma : \mathbb{Z} &\rightarrow \mathbb{R} \\ h &\mapsto \gamma(h) = \text{Cov}(X_t, X_{t-h}) \end{aligned}$$

**Proposition 1.3.3**

1. La fonction  $\gamma$  est une fonction paire :  $\gamma(-h) = \gamma(h) \forall h$

2. La fonction  $\gamma$  est de type positif :  $\forall n \in \mathbb{N}, \forall (\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n, \forall (t_1, \dots, t_n) \in \mathbb{Z}^n :$

$$\sum_{1 \leq i, j \leq n} \alpha_i \alpha_j \gamma(t_i - t_j) > 0$$

**Définition 1.3.6 (La fonction d'auto-corrélation simple)**

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

$$\rho(h) = \text{Corr}(X_t, X_{t-h}) = \frac{\text{Cov}(X_t, X_{t-h})}{\sqrt{\text{Var}(X_t)\text{Var}(X_{t-h})}} = \frac{\gamma(h)}{\gamma(0)} \quad \forall h \in \mathbb{Z}$$

On appelle coefficient d'auto-corrélation d'ordre 1 (respectivement d'ordre  $k$ ) le coefficient de corrélation linéaire  $\rho(1)$  (resp.  $\rho(k)$ ) calculé entre la série et cette série décalée d'une période (respectivement  $k$  périodes).

**Proposition 1.3.4**

1.  $\rho(0) = 1$
2.  $|\rho(h)| \leq 1 \quad \forall h$
3.  $\rho(-h) = \rho(h) \quad \forall h$

**Définition 1.3.7 (La Fonction d'auto-corrélation partielle (PACF))**

Par définition, l'auto-corrélation partielle d'ordre  $h$  d'un processus stationnaire  $(X_t)_{t \in \mathbb{Z}}$  est donnée par :

$$r(h) = \text{Corr}(X_t - E_L(X_t | X_{t-1}, \dots, X_{t-h+1}); X_{t-h} - E_L(X_{t-h} | X_{t-1}, \dots, X_{t-h+1})) \quad \forall h \in \mathbb{Z}$$

où  $E_L$  désigne l'espérance conditionnelle linéaire.

Elle désigne la corrélation entre  $X_t$  et  $X_{t-h}$  obtenue lorsque l'influence des variables  $X_{t-h-i}$  (avec  $0 < i < h$ ) a été retirée.

**Calcul :**

On note  $\rho_h$  et  $\rho_{hh}$  les fonctions respectivement l'auto-corrélation et l'auto-corrélation partielle de  $X_t$  au retard  $h$ . Soit  $R_h$  la matrice symétrique formée de  $(h - 1)$  premières auto-corrélations de  $X_t$  :

$$R_h = \begin{pmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{h-1} \\ \rho_1 & 1 & \cdots & \cdots & \rho_{h-2} \\ \rho_2 & \rho_1 & 1 & \cdots & \rho_{h-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_{h-1} & \rho_{h-2} & \cdots & \cdots & 1 \end{pmatrix}$$

La PACF est la succession des  $\rho_{hh} = \frac{|R_h^*|}{|R_h|}$  avec :

$|R_h^*|$  = déterminant de la matrice  $R_h$  dans laquelle on remplace la dernière colonne par le vecteur  $[\rho_1, \dots, \rho_h]$

$$R_h^* = \begin{pmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_1 \\ \rho_1 & 1 & \cdots & \cdots & \rho_2 \\ \rho_2 & \rho_1 & 1 & \cdots & \rho_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_{h-1} & \rho_{h-2} & \cdots & \cdots & \rho_h \end{pmatrix}$$

Calculons  $\rho_{11}$ ,  $\rho_{22}$  et  $\rho_{33}$  par exemple :

$$\rho_{11} = \frac{|R_1^*|}{|R_1|} \quad \text{avec } R_1 = [1] \quad R_1^* = [\rho_1]$$

$$\rho_{11} = \frac{\rho_1}{1} = \rho_1$$

On constate que  $\rho_{11} = \rho_1$

$$\rho_{22} = \frac{|R_2^*|}{|R_2|} \quad \text{avec } R_2 = \begin{pmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{pmatrix} \quad R_2^* = \begin{pmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{pmatrix}$$

$$\rho_{22} = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2}$$

$$\rho_{33} = \frac{|R_3^*|}{|R_3|} \quad \text{avec } R_3 = \begin{pmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{pmatrix} \quad R_3^* = \begin{pmatrix} 1 & \rho_1 & \rho_1 \\ \rho_1 & 1 & \rho_2 \\ \rho_2 & \rho_1 & \rho_2 \end{pmatrix}$$

$$\rho_{33} = \frac{\rho_1^3 - \rho_1\rho_2(2 - \rho_2) + \rho_3(1 - \rho_1^2)}{1 - \rho_2^2 - 2\rho_1^2(1 - \rho_2)}$$

Pour simplifier les calculs, on utilise l'écriture récurrente de  $\rho_{ii}$  :

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

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

### 1.3.6 Les opérateurs

#### Définition 1.3.8 (Opérateur de retard)

Soit  $(X_t)_{t \in \mathbb{Z}}$  un processus stochastique. L'opérateur qui peut décaler le processus d'une unité de temps vers le passé est s'appelle l'opérateur retard  $B$  :

$$B(X_t) = X_{t-1}$$

Pour décaler le processus de  $h$  unité de temps, on doit faire l'application précédente  $h$  fois :

$$\underbrace{B(B(\dots B(X_t)\dots))}_{h \text{ fois}} = B^h(X_t) = X_{t-h}$$

**Définition 1.3.9 (Opérateur de différence)**

Pour faire la différence entre le processus et sa version d'écalée d'une unité de temps, on utilise l'opérateur différence  $\Delta$ . Il se construit en utilisant l'opérateur retard :

$$\Delta(X_t) = X_t - X_{t-1} = X_t - B(X_t) = (1 - B)X_t \quad \forall t \geq 2$$

On peut définir l'opérateur différence d'ordre  $k$  par la formule de récurrence suivante :

$$\Delta^{(k)}X_t = \Delta(\Delta^{(k-1)}X_t) \quad \forall k \geq 2$$

**Application de l'opérateur de différence pour éliminer la tendance et de la saisonnalité**

Cette méthode permet de supprimer les tendances et saisonnalité d'une série temporelle sans les estimer.

Soit  $\Delta_t$  l'opérateur qui associe  $(X_t - X_{t-T})$  à  $(X_t)$

$$\Delta_t = X_t - X_{t-T}$$

On note  $\Delta$  l'opérateur  $\Delta_1$ , et  $\Delta_t^k$  l'opérateur  $\underbrace{\Delta_t \circ \dots \circ \Delta_t}_{k \text{ fois}}$

**Proposition 1.3.5**

Soit un processus admettant une tendance polynomiale d'ordre  $k$  :

$$X_t = \underbrace{\sum_{j=0}^k a_j t^j}_{m_t} + \epsilon_t$$

Le processus  $\Delta X_t$  admet une tendance polynomiale d'ordre  $k - 1$ .

Ainsi, en appliquant  $k$  fois  $\Delta$ , on élimine la tendance.

**Exemple :**

Pour une série a une tendance linéaire et une saisonnalité (de période 12)

La figure 1.6 présente :

- La série initiale sur le premier graphe.
- La série après l'élimination de la tendance
- La série après l'élimination de saisonnalité (ainsi que la tendance qui est déjà éliminé)

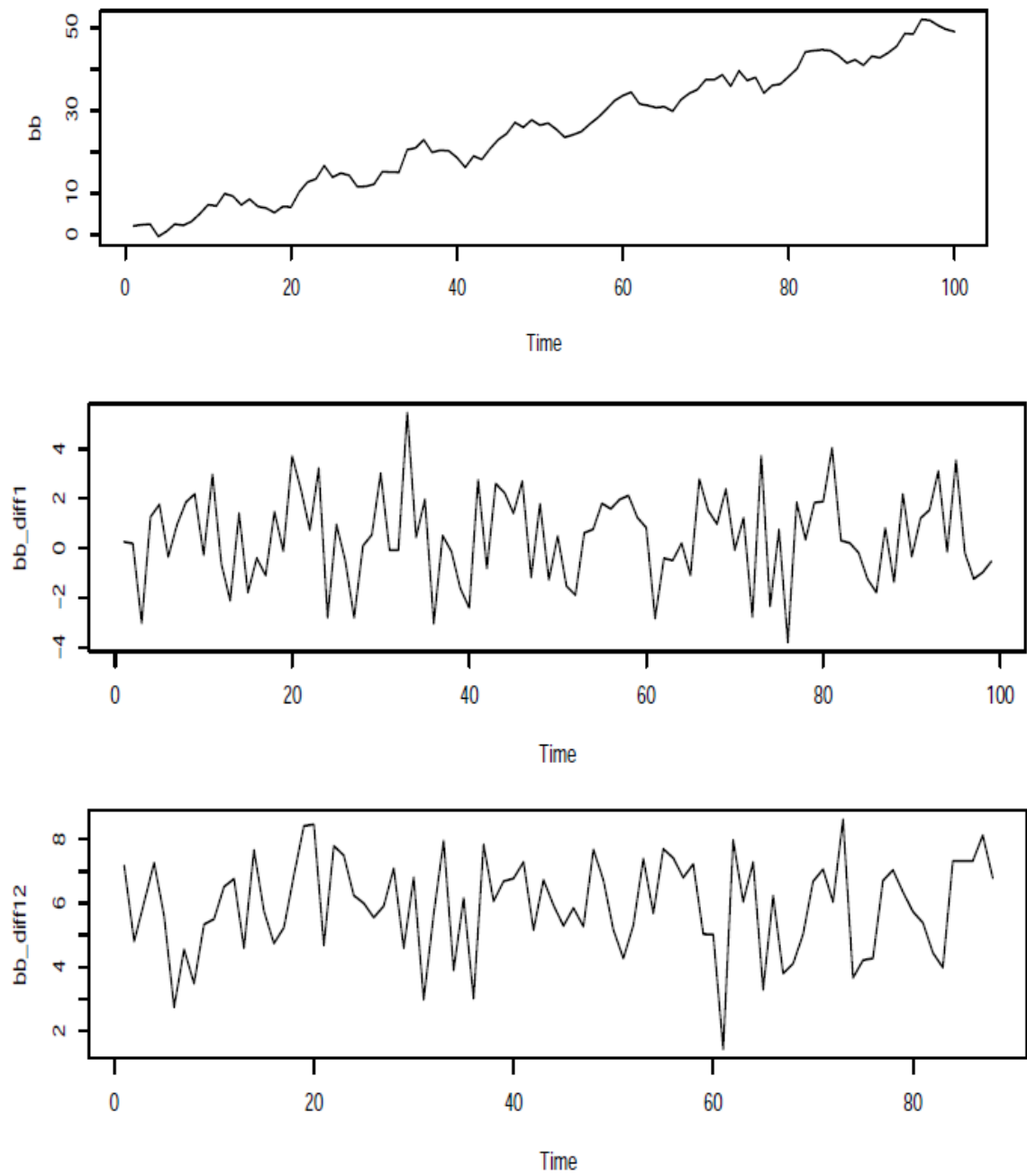


FIGURE 1.5 – Les trois graphes de l'application

### 1.3.7 Processus auto-régressifs

Dans les modèles auto-régressifs, la valeur du processus à un instant  $t$  est expliquée par la somme d'une combinaison linéaire des valeurs précédentes et d'un terme de bruit (le processus dépend linéairement des valeurs passées).

#### Définition 1.3.10

Les processus auto-régressifs (Auto-Regressive) d'ordre  $p$  (notés  $AR(p)$ ) sont définis par la relation suivante :

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

où  $(\varepsilon_t)_{t \in \mathbb{Z}}$  est un bruit blanc  $WN(0, \sigma_\varepsilon^2)$  et  $\phi_1, \dots, \phi_p$  sont des coefficients réels, avec  $\phi_p \neq 0$ .

On applique l'opérateur retard, la relation précédente peut s'écrire :

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

avec

$$\Phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$$

#### La stationnarité :

Un processus  $AR(p)$  est stationnaire si tous les racines (réelles ou complexes) du polynôme caractéristique  $\Phi(z)$  ont des modules strictement supérieur à 1.

où

$$\Phi(z) = 1 - \phi_1 z - \phi_2 z^2 - \dots - \phi_p z^p$$

Dans ce cas on peut inverser ce polynôme et écrire  $AR(p)$  sous la forme  $MA(\infty)$  :

$$X_t = \Phi^{-1}(B)\varepsilon_t = \sum_{i=0}^{+\infty} \varphi_i \varepsilon_{t-i}$$

avec  $\varphi_0 = 1$  et  $\sum_{i=0}^{+\infty} |\varphi_i| < \infty$

#### l'inversibilité :

Les processus auto-régressifs sont inversibles par définition.

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

#### Exemples :

$$X_t = 0.05X_{t-1} + \varepsilon_t \sim AR(1)$$

$$X_t = -0.1X_{t-1} + 0.02X_{t-2} + \varepsilon_t \sim AR(2)$$

### 1.3.8 Processus moyennes mobiles

#### Définition 1.3.11

Les processus qui portent le nom moyenne mobile (Moving Average) d'ordre  $q$  (notés  $MA(q)$ ) sont des processus stationnaires définis par la relation suivante :

$$X_t = \varepsilon_t - \theta_1\varepsilon_{t-1} - \theta_2\varepsilon_{t-2} - \dots - \theta_q\varepsilon_{t-q}$$

où  $(\varepsilon_t)_{t \in \mathbb{Z}}$  est un bruit blanc  $WN(0, \sigma_\varepsilon^2)$  et  $\theta_1, \dots, \theta_q$  sont des coefficients réels, avec  $\theta_q \neq 0$ .

On applique l'opérateur retard, la relation précédente peut s'écrire :

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

avec

$$\Theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q$$

#### La stationnarité :

Par définition, ce type des processus est automatiquement stationnaire.

#### l'inversibilité :

Un processus  $MA(q)$  est stationnaire si tous les racines (réelles ou complexes) du polynôme caractéristique  $\Theta(z)$  ont des modules strictement supérieur à 1.

où

$$\Theta(z) = 1 - \theta_1 z - \theta_2 z^2 - \dots - \theta_q z^q$$

Dans ce cas on peut inverser ce polynôme et écrire  $MA(q)$  sous la forme  $AR(\infty)$  :

$$\varepsilon_t = \Theta^{-1}(B)X_t = \sum_{i=0}^{+\infty} \vartheta_i X_{t-i}$$

avec  $\vartheta_0 = 1$  et  $\sum_{i=0}^{+\infty} |\vartheta_i| < \infty$

#### Exemples :

$$X_t = \varepsilon_t + 0.01\varepsilon_{t-1} \sim MA(1)$$

$$X_t = \varepsilon_t + 0.04\varepsilon_{t-1} + 0.02\varepsilon_{t-2} \sim MA(2)$$

### 1.3.9 Les processus mixtes (ARMA)

La combinaison des deux modèles précédents ( $AR$  et  $MA$ ) introduire un nouveau modèle mixte, qui s'appelle auto-régressif moyenne mobile (Auto-Regressive Moving Average) noté  $ARMA$  (une combinaison des erreurs passées et des valeurs passées).

#### Définition 1.3.12

Les processus de type auto-régressif moyenne mobile d'ordres  $p$  et  $q$  (notés  $ARMA(p, q)$ ) sont définis par la relation suivante :

$$X_t - \phi_1 X_{t-1} - \phi_2 X_{t-2} - \dots - \phi_p X_{t-p} = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q}$$

où

les coefficients  $\phi_i$  ( $i = 1, \dots, p$ ) et  $\theta_j$  ( $j = 1, \dots, q$ ) sont des réels, avec  $\phi_p \neq 0$  et  $\theta_q \neq 0$ .

et  $(\varepsilon_t)_{t \in \mathbb{Z}}$  est un bruit blanc  $WN(0, \sigma_\varepsilon^2)$ .

$p$  est l'ordre de la partie auto-régressive.

$q$  est l'ordre de la partie moyenne mobile

On applique l'opérateur retard, la relation précédente s'écrit :

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

avec

$$\Phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p \quad \text{et} \quad \Theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q$$

#### La stationnarité :

Un processus  $ARMA(p, q)$  est stationnaire si tous les racines (réelles ou complexes) du polynôme caractéristique  $\Phi(z)$  ont des modules strictement supérieurs à 1, où

$$\Phi(z) = 1 - \phi_1 z - \phi_2 z^2 - \dots - \phi_p z^p$$

Dans ce cas on peut inverser ce polynôme et écrire le processus  $ARMA$  sous forme d'un  $MA(\infty)$  :

$$X_t = \frac{\Theta(B)}{\Phi(B)} \varepsilon_t = \sum_{i=0}^{+\infty} \varphi_i \varepsilon_{t-i}$$

$$\text{avec } \varphi_0 = 1 \text{ et } \sum_{i=0}^{+\infty} |\varphi_i| < \infty$$

#### L'inversibilité :

Un processus  $ARMA(p, q)$  est inversible si tous les racines (réelles ou complexes) du polynôme caractéristique  $\Theta(z)$  ont des modules strictement supérieurs à 1, où

$$\Theta(z) = 1 - \theta_1 z - \theta_2 z^2 - \dots - \theta_q z^q$$

Dans ce cas on peut inverser ce polynôme et écrire le processus *ARMA* sous forme d'un *AR*( $\infty$ ) :

$$\varepsilon_t = \frac{\Phi(B)}{\Theta(B)} X_t = \sum_{i=0}^{+\infty} \vartheta_i X_{t-i}$$

avec  $\vartheta_0 = 1$  et  $\sum_{i=0}^{+\infty} |\vartheta_i| < \infty$

### Remarque 1.3.2

1. Si  $p = q = 0$ , on a  $X_t = \varepsilon_t \sim WN(0, \sigma^2)$
2. Si  $p = 0$  et  $q \neq 0$ , on a  $\{X_t\} \sim MA(q)$
3. Si  $p \neq 0$  et  $q = 0$ , on a  $\{X_t\} \sim AR(p)$

Exemples :

$$X_t - 2X_{t-1} = \varepsilon_t + 0.5\varepsilon_{t-1} \sim ARMA(1, 1)$$

$$X_t - X_{t-1} + 0.25X_{t-2} = \varepsilon_t - \varepsilon_{t-1} \sim ARMA(2, 1)$$

### 1.3.10 La méthodologie de Box-Jenkins

La méthode de *Box-Jenkins* (1970) est un outil systématique qui permet de :

- Déterminer le meilleur modèle de type *ARMA* décrivant le processus stochastique d'une série observée ou d'une transformation stationnaire de celle-ci ;
- Estimer ce modèle
- L'utiliser pour extrapoler les valeurs de la série.

Cette méthodologie comporte essentiellement cinq étapes :

1. Transformation des données afin de stabiliser la variance, et différenciation des données pour les stationnariser.
2. Visualisation de ACF et de PACF empiriques pour identifier les paramètres  $p$  et  $q$  appropriés.
3. Estimation des paramètres du(des) modèle(s) sélectionné(s).
4. Diagnostique et tests adéquatation du modèle.
5. Prévision des valeurs futures à travers le modèle retenu.

## 1.4 Processus aléatoires non stationnaires

### 1.4.1 Les processus TS et DS

#### Les processus TS :

Les processus *TS* (Trend Stationary) sont caractérisés par une non stationnarité de nature déterministe.

#### Définition 1.4.1

Les processus non stationnaire de type *TS* sont définis par la relation suivante :

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

où

$f(t)$  est une fonction du temps.

$Z_t$  est un processus stochastique stationnaire.

#### La non stationnarité du processus TS :

Le processus *TS* s'écrit comme la somme d'une composante stochastique stationnaire, éventuellement de type *ARMA* et d'une fonction déterministe du temps.

Il est évident que l'espérance de ce processus est dépend du temps,

$E(X_t) = f(t) + E(Z_t)$  et alors le processus *TS* ne satisfait plus la première condition de la stationnarité faible. Donc par conséquent le *TS* n'est pas stationnaire.

#### Proposition 1.4.1

On prend le cas le plus simple d'un processus *TS* : une tendance linéaire perturbée par un bruit blanc,

$$X_t = \alpha + \beta t + \varepsilon_t \quad \text{où } (\alpha, \beta) \in \mathbb{R}^2 \quad \text{et } \varepsilon_t \sim WN(0, \sigma_\varepsilon^2)$$

on a alors les propriétés suivantes :

$$E(X_t) = E(\alpha + \beta t + \varepsilon_t) = \alpha + \beta t$$

$$V(X_t) = V(\alpha + \beta t + \varepsilon_t) = V(\varepsilon_t) = \sigma_\varepsilon^2$$

$$Cov(X_t, X_s) = E[(X_t - E(X_t))(X_s - E(X_s))] = E(\varepsilon_s \varepsilon_t) = 0 \quad , \quad \forall t \neq s$$

#### conclusion 1.4.1

- L'espérance d'un *TS* exhibe une tendance déterministe.
- La variance d'un *TS* est constante dans le temps, donc il est stationnaire en variance.

#### Les processus DS :

Les processus *DS* (Difference Stationary) est caractérisés par une non stationnarité de nature stochastique.

**Définition 1.4.2**

Soit  $(X_t)_{t \in \mathbb{Z}}$ . Si le processus filtré qui est défini par  $(1 - L)^d X_t$  (où  $d$  désigne l'ordre d'intégration (différentiation)) est stationnaire, alors  $(X_t)_{t \in \mathbb{Z}}$  représente un processus non stationnaire de type DS d'ordre  $d$  (il dit aussi un processus intégré d'ordre  $d$ , noté  $I(d)$ ).

Pour un processus non stationnaire DS d'ordre un (intégré d'ordre un,  $I(1)$ ), la différence première  $\Delta X_t = X_t - X_{t-1}$  est stationnaire. Ainsi pour un DS d'ordre deux ( $I(2)$ ), la différence deuxième  $(1 - L)^2 X_t = (1 - L)\Delta X_t = X_t - X_{t-1} - X_{t-2}$  est stationnaire.

On va définir un type de processus qui appartient à la classe générale des processus DS. Dans la classe générale des processus DS, un type de processus apparaît de façon régulière. Ce processus est s'appelle : *la marche aléatoire*.

**Définition 1.4.3 (Marche aléatoire)**

Un processus  $AR(1)$  intégré d'ordre un, représente une marche aléatoire (Random Walk)

$$\Delta X_t = (1 - L)X_t = \beta + \varepsilon_t \iff X_t = \beta + X_{t-1} + \varepsilon_t$$

$$\text{où } \varepsilon_t \sim WN(0, \sigma_\varepsilon^2)$$

**Marche aléatoire avec dérive ( $\beta \neq 0$ ) :**

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

Qui représente une non stationnarité de nature stochastique qu'on peut démontrer par récurrence :

$$X_1 = X_0 + \beta + \varepsilon_1$$

$$X_2 = X_1 + \beta + \varepsilon_2 = X_0 + \beta + \varepsilon_1 + \beta + \varepsilon_2$$

$$X_3 = X_2 + \beta + \varepsilon_3 = X_1 + \beta + \varepsilon_2 + \beta + \varepsilon_3 = X_0 + \beta + \varepsilon_1 + \beta + \varepsilon_2 + \beta + \varepsilon_3$$

⋮

$$X_t = X_0 + t\beta + \sum_{i=1}^t \varepsilon_i$$

$$\text{où } \varepsilon_t \sim WN(0, \sigma_\varepsilon^2)$$

donc :

$$E(X_t) = X_0 + t\beta$$

$$V(X_t) = t\sigma_\varepsilon^2$$

$$Cov(X_t, X_s) = \min(t, s)\sigma_\varepsilon^2 \quad \text{si } t \neq s$$

Alors : le processus (marche aléatoire avec dérive) est non stationnaire ni en moyenne ni en variance.

**Marche aléatoire sans dérive ( $\beta = 0$ ) :**

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

Que nous montrons sa non stationnarité par récurrence :

$$X_t = X_0 + \varepsilon_1$$

$$X_t = X_1 + \varepsilon_2 = X_0 + \varepsilon_1 + \varepsilon_2$$

$$X_t = X_2 + \varepsilon_3 = X_1 + \varepsilon_2 + \varepsilon_3 = X_0 + \varepsilon_1 + \varepsilon_2 + \varepsilon_3$$

⋮

$$X_t = X_0 + \sum_{i=1}^t \varepsilon_i$$

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

donc :

$$E(X_t) = X_0$$

$$V(X_t) = t\sigma_\varepsilon^2$$

$$Cov(X_t, X_s) = \min(t, s)\sigma_\varepsilon^2 \quad \text{si } t \neq s$$

Alors : le processus (marche aléatoire sans dérive) est stationnaire en moyenne mais pas en variance.

## 1.4.2 Les processus ARIMA

Les modèles *ARMA* sont utilisés pour modéliser les processus stationnaire. Généralement dans la pratique, les séries chronologiques sont non stationnaires, et il est nécessaire de supprimer les tendances et les saisonnalités.

Une généralisation des processus *ARMA* aux cas non stationnaires avec tendance polynômiale, nous donne un autre type non stationnaire. c'est le *ARIMA* (Processus Autorégressif moyenne mobile intégré / AutoRegressive Integrated Moving Average).

### Définition 1.4.4

Les processus *ARIMA*( $p, d, q$ ) sont définis par la relation suivante :

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

tel que :

$$\nabla^d = (1 - B)^d$$

$$\Phi_p(B) = 1 - \phi_1 B - \phi_2 B^2 \dots - \phi_p B^p$$

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

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

Les coefficients  $\phi_1, \dots, \phi_p; \theta_1, \dots, \theta_q$  sont des réels et tels que  $\phi_p \neq 0$  et  $\theta_q \neq 0$ .  
 $d$  un entier positif appelé paramètre d'intégration ou de différenciation.

### 1.4.3 Les processus SARIMA

Une autre généralisation des modèles *ARMA*, c'est le modèle *SARIMA* (*ARMA* intégré saisonnier). Il est utilisé pour la modélisation des séries qui présentent une saisonnalité et qui ont une tendance polynômiale.

#### Définition 1.4.5

Un processus  $(X_t)_{t \in \mathbb{Z}}$  est un *SARIMA* $(p, d, q) * (P, D, Q)_s$ , si le processus différencié :

*contenu...*

est un processus *ARMA*, on peut le définir par la relation suivante :

$$\phi_p(L)\phi_P(L^s)(1-L)^d(1-L^s)^D X_t = \theta_q(L)\theta_Q(L^s)\varepsilon_t$$

tel que :  $\phi_p, \theta_q, \phi_P, \theta_Q$  sont des polynômes de degrés respectifs  $p, q, P, Q$  :

$$\begin{aligned}\phi_p(z) &= 1 - \phi_1 z - \dots - \phi_p z^p \\ \phi_P(z^s) &= 1 - \phi_{1s} z^s - \dots - \phi_{Ps} z^{Ps} \\ \theta_q(z) &= 1 - \theta_1 z - \dots - \theta_q z^q \\ \theta_Q(z^s) &= 1 - \theta_{1s} z^s - \dots - \theta_{Qs} z^{Qs}\end{aligned}$$

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

$s$  : exprime la saisonnalité .

# Chapitre 2

## Processus Conditionnellement Hétéroscédastiques

### 2.1 Le modèle ARCH

#### 2.1.1 Présentation du processus ARCH

Les modèles autorégressifs conditionnellement hétéroscédastiques (Auto-regressive Conditionally Heteroscedastic) introduits par Engle (1982). ils reposent sur une paramétrisation endogène de la variance conditionnelle et permettent la prise en compte de ce type de propriétés.

##### **Définition 2.1.1**

On dit que  $(X_t)_{t \in \mathbb{Z}}$  est un processus auto-régressif conditionnellement hétéroscédastique d'ordre  $p$  noté  $ARCH(p)$ , s'il vérifie :

$$X_t = h_t \varepsilon_t$$

avec

$$h_t^2 = \alpha_0 + \sum_{i=1}^p \alpha_i X_{t-i}^2$$

où :

les  $\alpha_i$  sont des réels ;  $\alpha_0 > 0$ ,  $\alpha_i \geq 0$  pour  $i = 1, \dots, p$

$\varepsilon_t$  est un bruit blanc gaussien  $\varepsilon_t \sim \mathcal{N}(0, 1)$

$h_t^2$  représente la variance conditionnelle du processus  $X_t$

(est un processus qui conditionnellement à l'ensemble d'information des valeurs passées de  $X_t$   $\{X_{t-1}, X_{t-2}, \dots\}$ , il est déterministe et positif.)

##### **Remarque 2.1.1**

On note par  $\mathcal{F}_{t-1}$  l'ensemble des valeurs passées de  $X_t$  :

$$\mathcal{F}_{t-1} = \{X_{t-1}, X_{t-2}, \dots\}$$

## 2.1.2 Propriétés du processus ARCH

### Propriété 2.1.1

1. L'espérance conditionnelle d'un processus ARCH :

$$E(X_t | \mathcal{F}_{t-1}) = 0$$

2. L'espérance non conditionnelle :

$$E(X_t) = 0$$

**Preuve :**

$$\begin{aligned}
 1. \quad E(X_t | \mathcal{F}_{t-1}) &= E(h_t \varepsilon_t | \mathcal{F}_{t-1}) \\
 &= h_t E(\varepsilon_t | \mathcal{F}_{t-1}) && \text{car } h_t \text{ est mesurable par rapport à} \\
 & && \text{la tribu } \mathcal{F}_{t-1} \\
 &= h_t E(\varepsilon_t) && \text{car } \varepsilon_t \text{ est indépendante de } \mathcal{F}_{t-1} \\
 &= 0 && \text{car } \varepsilon_t \text{ est de moyenne nulle}
 \end{aligned}$$

$$\begin{aligned}
 2. \quad E(X_t) &= E(E(X_t | \mathcal{F}_{t-1})) \\
 &= E(0 | \mathcal{F}_{t-1}) \\
 &= 0
 \end{aligned}$$

### Propriété 2.1.2

1. la variance conditionnelle d'un processus ARCH est varié dans le temps :

$$V(X_t | \mathcal{F}_{t-1}) = E(X_t^2 | \mathcal{F}_{t-1}) = \alpha_0 + \sum_{i=1}^p \alpha_i X_{t-i}^2 = h_t^2$$

2. La variance non conditionnelle d'un processus ARCH est constante dans le temps, elle est exprimée par :

$$V(X_t) = \frac{\alpha_0}{1 - \sum_{i=1}^p \alpha_i} \quad \text{si} \quad \sum_{i=1}^p \alpha_i < 1$$

**Preuve :**

$$\begin{aligned}
 1. \quad V(X_t | \mathcal{F}_{t-1}) &= E(X_t^2 | \mathcal{F}_{t-1}) \\
 &= E(h_t^2 \varepsilon_t^2 | \mathcal{F}_{t-1}) \\
 &= h_t^2 E(\varepsilon_t^2 | \mathcal{F}_{t-1}) \\
 &= h_t^2 && \text{car } E(\varepsilon_t^2 | \mathcal{F}_{t-1}) = E(\varepsilon_t^2) = 1 \\
 &= \alpha_0 + \sum_{i=1}^p \alpha_i X_{t-i}^2
 \end{aligned}$$

2. Avec la formule de décomposition de la variance du processus  $X_t$ , on trouve :

$$\begin{aligned} V(X_t) &= E(V(X_t|\mathcal{F}_{t-1})) + V(E(X_t|\mathcal{F}_{t-1})) \\ &= 0 + E(h_t) + 0 \end{aligned}$$

et pour un ARCH(1), on a :

$$\begin{aligned} V(X_t) &= E(\alpha_0 + \alpha_1 X_{t-1}^2) \\ &= \alpha_0 + \alpha_1 V(X_{t-1}) \end{aligned}$$

Sous-hypothèse de stationnarité (i.e :  $V(X_t) = V(X_{t-1})$ ), on obtient :

$$\begin{aligned} V(X_t) &= \alpha_0 + \alpha_1 V(X_t) \iff V(X_t) - \alpha_1 V(X_t) = \alpha_0 \\ &\iff (1 - \alpha_1)V(X_t) = \alpha_0 \\ &\iff V(X_t) = \frac{\alpha_0}{1 - \alpha_1} \end{aligned}$$

on généralise sur ARCH( $q$ ), on trouve :

$$V(X_t) = \frac{\alpha_0}{1 - \sum_{i=1}^q \alpha_i}$$

### Propriété 2.1.3

pour tout  $h > 0$  :

1. La fonction d'autocovariance conditionnelle :

$$Cov(X_t, X_{t+h}|\mathcal{F}_{t-1}) = 0$$

2. La fonction d'autocovariance non conditionnelle :

$$Cov(X_t, X_{t+h}) = 0$$

**Preuve :**

$$\begin{aligned} 1. \quad Cov(X_t, X_{t+h}|\mathcal{F}_{t-1}) &= E(X_t X_{t+h}|\mathcal{F}_{t-1}) - E(X_t|\mathcal{F}_{t-1})E(X_{t+h}|\mathcal{F}_{t-1}) \\ &= E(X_t X_{t+h}|\mathcal{F}_{t-1}) \end{aligned}$$

on utilise la loi des espérances itérées :

$$\begin{aligned} Cov(X_t, X_{t+h}|\mathcal{F}_{t-1}) &= E(E(X_t X_{t+h}|\mathcal{F}_{t+h-1})|\mathcal{F}_{t-1}) \\ &= E(X_t E(X_{t+h}|\mathcal{F}_{t+h-1})|\mathcal{F}_{t-1}) \quad \text{car } X_t \mathcal{F}_{t+h-1} \\ &= E(X_t * 0|\mathcal{F}_{t-1}) \\ &= 0 \end{aligned}$$

$$\begin{aligned}
2. \quad \text{Cov}(X_t, X_{t+h}) &= E(X_t X_{t+h}) - E(X_t)E(X_{t+h}) \\
&= E(X_t X_{t+h}) \\
&= E(E(X_t X_{t+h} | \mathcal{F}_{t+h-1})) \\
&= E(X_t E(X_{t+h} | \mathcal{F}_{t+h-1})) \\
&= E(0 * E(X_{t+h} | \mathcal{F}_{t+h-1})) \\
&= 0
\end{aligned}$$

### 2.1.3 Condition de stationnarité

D'après les propriétés précédentes, on déduit les conditions concernant les paramètres  $\alpha_0, \dots, \alpha_p$  qui vérifient la stationnarité du processus *ARCH*. En effet, la moyenne du processus est constante (nulle), ainsi la covariance ne dépend que du temps (nulle pour tout décalage). Dans ce cas, si la variance existe (c'est-à-dire, si elle est finie) On dit que une le processus est faiblement stationnaire. Alors la condition suffisante de la stationnarité du modèle *ARCH* est :

$$\sum_{i=1}^p \alpha_i < 1$$

### 2.1.4 Le cas d'un ARCH(1)

soit  $(X_t)_{t \in \mathbb{Z}}$  un *ARCH*(1) :

$$X_t = h_t \varepsilon_t$$

avec

$$h_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2$$

où :

$\alpha_0, \alpha_1 > 0$  des réels

$\varepsilon_t$  est un bruit blanc gaussien  $\varepsilon_t \sim \mathcal{N}(0, 1)$

$h_t^2$  représente la variance conditionnelle du processus  $X_t$

## 2.2 Le modèle GARCH

### 2.2.1 Présentation du processus GARCH

Bollerslev (1986) a proposé une généralisation du modèle *ARCH*, qui s'appelle le modèle *GARCH* (Hétéroscédasticité conditionnelle autorégressive généralisée / Generalized Auto-regressive Conditionally Heteroscedastic).

**Définition 2.2.1**

On dit que  $(X_t)_{t \in \mathbb{Z}}$  est un processus autorégressif conditionnellement hétéroscédastique généralisé d'ordre  $p$  et  $q$  noté  $GARCH(p, q)$ , s'il vérifie :

$$X_t = h_t \varepsilon_t$$

avec

$$h_t^2 = \alpha_0 + \sum_{i=1}^p \alpha_i X_{t-i}^2 + \sum_{j=1}^q \beta_j h_{t-j}^2$$

où :

les  $\alpha_i$  sont des réels ;  $\alpha_0 > 0$ ,  $\alpha_i \geq 0$  pour  $i = 1, \dots, p$

$\varepsilon_t$  est un bruit blanc gaussien  $\varepsilon_t \sim \mathcal{N}(0, 1)$

$h_t$  représente la variance conditionnelle du processus  $X_t$

**2.2.2 Propriétés du processus GARCH**

Les propriétés théoriques des processus  $GARCH$  se déduisent de la même façon que nous avons développé les propriétés des processus  $ARCH$ .

**Propriété 2.2.1**

Comme le processus  $ARCH$ , les espérances conditionnelle et non conditionnelle d'un processus  $GARCH$  sont nulles :

$$E(X_t | X_{\mathcal{F}^{-\infty}}) = 0 \quad \text{et} \quad E(X_t) = 0$$

**Propriété 2.2.2**

Les variances conditionnelle et non conditionnelle d'un processus  $GARCH$  sont aussi présentées comme le processus  $ARCH$ .

1. La variance conditionnelle est variée dans le temps, elle est exprimée par :

$$V(X_t | X_{\mathcal{F}^{-\infty}}) = \alpha_0 + \sum_{i=1}^p \alpha_i X_{t-i}^2 + \sum_{j=1}^q \beta_j h_{t-j}^2 = h_t$$

2. La variance non conditionnelle est constante dans le temps, elle est exprimée par :

$$V(X_t) = \frac{\alpha_0}{1 - \left( \sum_{i=1}^p \alpha_i + \sum_{j=1}^q \beta_j \right)}$$

**Propriété 2.2.3**

Comme pour le modèle  $ARCH(q)$ , les auto-covariances conditionnelles et non conditionnelles du processus  $GARCH(p, q)$  sont nulles.

### 2.2.3 Condition de stationnarité

Dans le cas d'un modèle *GARCH*, généralement on ne peut pas calculer la variance directement, et il faut vérifier une telle condition sur les paramètres du processus. Bollerslev (1986) a montré que la variance du processus *GARCH* existe (finie) si la somme des paramètres  $\alpha_1 + \dots + \alpha_p + \beta_1 + \dots + \beta_q$  est plus petite que 1.

#### **Théorème 2.2.1**

Un processus *GARCH*( $p, q$ ) est faiblement stationnaire, si :

$$\sum_{i=1}^p \alpha_i + \sum_{j=1}^q \beta_j < 1$$

#### **Remarque 2.2.1**

Dans le cas où :  $\sum_{i=1}^p \alpha_i + \sum_{j=1}^q \beta_j = 1$  on dit que le processus *GARCH* est intégré, nommé processus *IGARCH*.

## 2.3 Extentions des modèles ARCH/GARCH, et les modèles asymétriques

#### **Modèle GARCH-M :**

Le modèle *GARCH-M* (General Autoregressive Conditional Heteroskedasticity in **Mean**) a été introduit par Engle-Lilien-Robbins (1987). Dans ces modèles, la variance conditionnelle est une variable explicative de la moyenne conditionnelle.

#### **Modèle IGARCH :**

Le modèle *IGARCH* (**I**ntegrated Gènèral Autoregréssive Conditional Heteroskedasticity) correspond au cas d'une racine unitaire dans la variance conditionnelle.

#### **Modèle EGARCH :**

Le modèle *EGARCH* (**E**xponential Gènèral Autoregréssive Conditional Heteroskedasticity) a été introduit par Nelson (1991). La spécification porte sur le logarithme de la variance conditionnelle.

#### **Modèle APARCH :**

Introduit par Ding, Granger et Engle (1993)

#### **Modèles TGARCH :**

*TARCH* (Zakoian, 1991)

*TGARCH* (Zakoian, 1994)

## 2.4 Modélisation des modèles GARCH

### 2.4.1 Identification

Pour faire l'identification d'un processus  $GARCH(p, q)$ , on doit passer à la représentation  $ARMA$  du carré de processus  $GARCH$  :

Pour un processus  $GARCH(p, q)$  vs :

$$\begin{aligned} X_t &= \varepsilon_t h_t \\ h_t^2 &= \alpha_0 + \sum_{i=1}^p \alpha_i X_{t-i}^2 + \sum_{j=1}^q \beta_j h_{t-j}^2, \end{aligned}$$

On définit le processus d'innovation  $\eta_t$  du carré de  $X_t$  :

$$\eta_t = X_t^2 - h_t^2$$

Donc le carré du processus  $X_t$  s'écrit :

$$\begin{aligned} X_t^2 &= h_t^2 + \eta_t \\ X_t^2 &= h_t^2 = \alpha_0 + \sum_{i=1}^p \alpha_i X_{t-i}^2 + \sum_{j=1}^q \beta_j h_{t-j}^2 + \eta_t \end{aligned}$$

Si on note  $m = \max(p, q)$  et avec les contraintes ( $\alpha_i = 0$  si  $i > p$  ;  $\beta_j = 0$  si  $j > q$ ), alors on peut développer la dernière expression par :

$$\begin{aligned} X_t^2 &= \alpha_0 + \sum_{i=1}^m (\alpha_i + \beta_i - \beta_i) X_{t-i}^2 + \sum_{j=1}^q \beta_j h_{t-j}^2 + \eta_t \\ &= \alpha_0 + \sum_{i=1}^m (\alpha_i + \beta_i) X_{t-i}^2 + \sum_{j=1}^q \beta_j (h_{t-j}^2 - X_{t-j}^2) + \eta_t \end{aligned}$$

Finalement on trouve une représentation  $ARMA(m, q)$  pour le processus  $X_t^2$  :

$$X_t^2 = \alpha_0 + \sum_{i=1}^m (\alpha_i + \beta_i) X_{t-i}^2 + \sum_{j=1}^q \beta_j \eta_{t-j} + \eta_t$$

Puisque  $E(\eta_t | \mathcal{F}_{t-1}) = 0$ , et par la propriété des espérances itérées, on déduit que  $\eta_t$  est de moyenne nulle et est non corrélé. Le processus  $\{\eta\}$  joue le rôle du processus des innovations dans cette représentation.

Maintenant pour identifier les nombres  $p$  et  $q$  du processus  $GARCH$ , on peut utiliser les fonctions d'auto-corrélation et d'auto-corrélation partielle du processus  $X_t^2$  suivant la même procédure utilisée pour trouver le nombre de paramètres d'un processus  $ARMA$ .

### 2.4.2 Estimation

La méthode la plus utilisée pour estimer les modèles Conditionnellement Hétéroscédastiques est la méthode du maximum de vraisemblance. Il y a aussi la méthode des moindres carrés, et la méthode des deux phases.

Nous utilisons la première méthode (maximum de vraisemblance) pour les deux modèles *ARCH* et *GARCH*.

### 2.4.2.1 Estimation des paramètres pour un modèle ARCH

Soit  $X_t \sim ARCH(p)$

$$\begin{aligned} X_t &= h_t \varepsilon_t \\ h_t^2 &= \alpha_0 + \sum_{i=1}^p \alpha_i X_{t-i}^2 \end{aligned}$$

Par hypothèse, la loi conditionnelle du processus  $X_t$  est gaussienne (i.e  $X_t | \mathcal{F}_{t-1} \sim \mathcal{N}(0, h_t^2)$ ).

Donc on peut définir la vraisemblance associée à  $X_t$  conditionnellement au passé  $\mathcal{F}_{t-1}$  :

$$\mathcal{L}(x_t | \mathcal{F}_{t-1}; \theta) = \frac{1}{\sqrt{2\pi}h_t} \exp\left(-\frac{x_t^2}{2h_t^2}\right)$$

qui dépend de  $\theta$ , où  $\theta$  est le vecteur qui contient les paramètres du processus  $X_t$   
 $\theta = (\alpha_0, \alpha_1, \dots, \alpha_p)$  à travers  $h_t$

La fonction de vraisemblance de  $N$  observations de  $X_t$  ( $x_1, \dots, x_N$ ) est par conséquent :

$$\mathcal{L}_N(x_1, \dots, x_N; \theta) = \prod_{t=1}^N \mathcal{L}(x_t | \mathcal{F}_{t-1}; \theta)$$

Par la maximisation du logarithme de cette dernière fonction, on trouve l'estimateur de/du maximum de vraisemblance  $\hat{\theta} = (\hat{\alpha}_0, \hat{\alpha}_1, \dots, \hat{\alpha}_p)$

$$\hat{\theta} = \arg \max \log \mathcal{L}_N(x_1, \dots, x_N; \theta)$$

### 2.4.2.2 Estimation des paramètres pour un modèle GARCH

Soit  $X_t \sim GARCH(p, q)$

$$\begin{aligned} X_t &= h_t \varepsilon_t \\ h_t^2 &= \alpha_0 + \sum_{i=1}^p \alpha_i X_{t-i}^2 + \sum_{j=1}^q \beta_j h_{t-j}^2 \end{aligned}$$

La loi conditionnelle du processus  $X_t$  est Par hypothèse gaussienne (i.e  $X_t | \mathcal{F}_{t-1} \sim \mathcal{N}(0, h_t^2)$ ).

Donc on peut définir la vraisemblance associée à  $X_t$  conditionnellement au passé  $\mathcal{F}_{t-1}$  :

$$\mathcal{L}(x_t | \mathcal{F}_{t-1}; \theta) = \frac{1}{\sqrt{2\pi}h_t} \exp\left(-\frac{x_t^2}{2h_t^2}\right)$$

mais dans le cas d'un *GARCH*, la variance  $h_t^2$  dépend des valeurs passées de la variance conditionnelle  $h_{t-1}^2, h_{t-2}^2, \dots, h_{t-q}^2$ . On pratique on ne pas observ(er) ces valeurs, et il est impossible de maximiser la fonction de vraisemblance directement.

avant de calculer la vraisemblance, on (fait) une estimation pour les valeurs de  $\sigma_1^2, \sigma_2^2, \dots, \sigma_{t-1}^2$ .

Ainsi pour un vecteur  $\theta^\circ = (\alpha_0^\circ, \dots, \alpha_p^\circ, \beta_1^\circ, \dots, \beta_q^\circ)$  fixé de paramètres, on calcule récursivement

$$\hat{h}_s^2 = \alpha_0^\circ + \sum_{i=1}^p \alpha_i^\circ X_{s-i}^2 + \sum_{j=1}^q \beta_j^\circ \hat{h}_{s-j}^2$$

avec la convention  $X_i = 0$  et  $h_i^2 = 0$  si  $i \leq 0$ .

On remplace donc la fonction de vraisemblance par :

$$\mathcal{L}(x_t | \mathcal{F}_{t-1}; \theta^\circ) = \frac{1}{\sqrt{2\pi\hat{h}_t}} \exp\left(-\frac{x_t^2}{2\hat{h}_t^2}\right)$$

La fonction de vraisemblance totale est :

$$\mathcal{L}_N(x_1, \dots, x_N; \theta^\circ) = \prod_{t=1}^N \mathcal{L}(x_t | \mathcal{F}_{t-1}; \theta^\circ)$$

Par la maximisation du logarithme de cette dernière fonction, on trouve l'estimateur de/du maximum de vraisemblance  $\theta^\circ = (\alpha_0^\circ, \dots, \alpha_p^\circ, \beta_1^\circ, \dots, \beta_q^\circ)$

### 2.4.3 Prévision (dans le cas d'un ARMA-GARCH)

#### Définition 2.4.1

Weiss (1984) a appliqué la théorie des modèles GARCH sur les erreurs d'une régression linéaire, et sur le processus d'innovation d'un modèle ARMA.

— Soit  $X_t$  une régression linéaire :

$$X_t = aY_t + \varepsilon_t$$

Si  $\varepsilon_t \sim GARCH(p, q)$ , alors on dit que  $X_t$  est un **modèle de régression avec erreurs GARCH**.

— Soit  $X_t$  un processus ARMA :

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

Si  $\varepsilon_t \sim GARCH(p, q)$ , alors on dit que  $X_t$  est un **modèle ARMA-GARCH**.

**La méthode :** Soit  $X_t$  un processus ARMA( $r, s$ )

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

où  $\Phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_r B^r$  et  $\Theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_s B^s$  et  $\varepsilon_t \sim GARCH(p, q)$

On suppose que cette s'est observée jusqu'au temps  $t$ .  $X_{t+h}$  est l'observation à l'instant  $t+h$  :

$$X_{t+h} = \sum_{i=1}^r \phi_i X_{t+h-i} + \sum_{j=1}^s \theta_j \varepsilon_{t+h-j} + \varepsilon_{t+h}$$

Le prédicteur optimal  $\hat{X}_t(h)$  est donné par :

$$\begin{aligned} \hat{X}_t(h) &= E(X_{t+h} | \mathcal{F}_t) = \sum_{i=1}^r \phi_i E(X_{t+h-i} | \mathcal{F}_t) + \sum_{j=1}^s \theta_j E(\varepsilon_{t+h-j} | \mathcal{F}_t) + E(\varepsilon_{t+h}) \\ &= \sum_{i=1}^r \phi_i E(X_{t+h-i} | \mathcal{F}_t) + \sum_{j=1}^s \theta_j E(\varepsilon_{t+h-j} | \mathcal{F}_t) \end{aligned}$$

# Chapitre 3

## Application

On va faire une petite application (logiciel *R*), sur une série chronologique financière.

Cette application montre pratiquement la méthodologie qui peut être suivie pour prédire une telle série à une volatilité hétéroscédastique, à travers le modèle **ARIMA-GARCH**.

**Les données utilisées :**

Données journalières (de : 20 octobre 1997, jusqu'au 20 Mai 2017)

### 3.1 Installation et Libération (packages)

#### 3.1.1 Installation

Avant de faire aucune opération, il faut nécessairement de installer les packages (de séries temporelles).

La commande sur *R* : *install.packages("nom du package")* La liste des packages :

1. aTSA
2. tseries
3. e1071
4. data.table
5. plotly
6. fGarch
7. tidyquant
8. fp33
9. ggplotify
10. parallel

11. mclust
12. rugarch
13. rmgarch
14. qrmtools

### 3.1.2 libération

Après l'installation, il suffit de libérer ces packages, à l'aide de la fonction *library*, par exemple :

```
library(nom-pack)
```

## 3.2 Lecture et Analyse

### 3.2.1 Représentation graphique

#### 3.2.1.1 Lecture et Positionnement

```
D <- fread("C :/GD.data.csv")

RTS <- D[, .(Date, Close = as.numeric(Close))]
```

#### 3.2.2 Stationnarité

```
RTS <- RTS %>%
  tsibble(index = Date) %>%
  mutate(Return = difference(log(RTS$Close)) * 100)
```

```
RTS <- RTS %>%
  na.omit() %>%
  mutate(time = row_number()) %>%
  update_tsibble(index = time)
head(RTS)
```

```
RTS %>%
  ggplot(aes(x = Date, y = Close)) +
  geom_line() +
  theme_tq() +
  labs(title = " l'Indice RTS $ (cloture) $ ")
```

```
RT %>%  
ggplot(aes(x = Date, y = Return)) +  
geom_line() +  
theme_tq() +  
labs(title = "l'Indice RTS *(Return)* ")
```

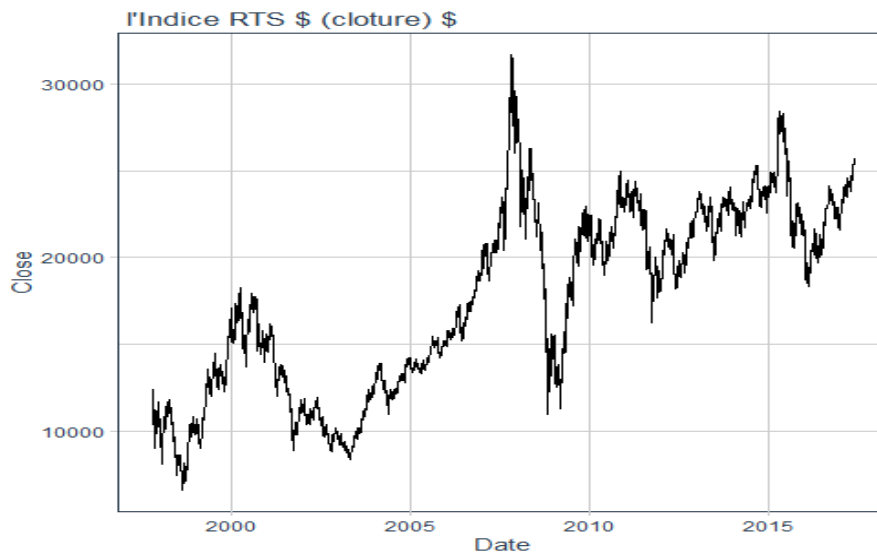


FIGURE 3.1 – Représentation graphique du série initiale

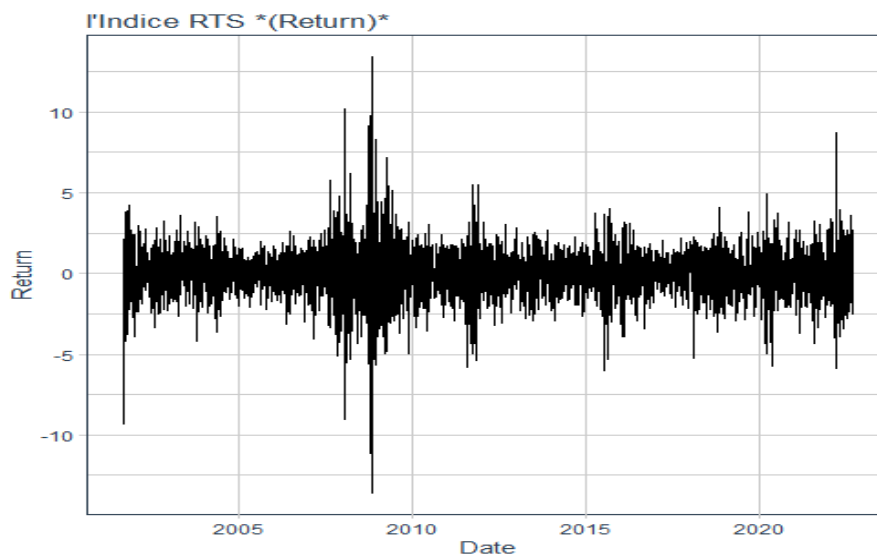


FIGURE 3.2 – Représentation graphique du série après différenciation

### 3.2.3 Tests

#### 3.2.3.1 KPSS

$H_0$  : la donnée est stationnaire autour d'une tendance déterministe.

**Code :**

```
RTS %>%
  features(Return, unitroot_kpss)
```

**Résultats :**

```
kpss_stat kpss_pvalue
<dbl>      <dbl>
1      0.0876      0.1
```

**Interprétation :**  $p - value = 0,1 > 0,05$ , donc l'hypothèse nulle est acceptée.

#### 3.2.3.2 ADF

$H_0$  : une racine unitaire est présente.

$H_1$  : la série temporelle est stationnaire.

**Code :**

```
adf.test(RTS$Return, alternative = "stationary")
```

**Résultats :**

```
data: RTS$Return
Dickey-Fuller = -16.153,
Lag order = 16, p-value = 0.01
alternative hypothesis: stationary
```

**Interprétation :**  $p - value = 0,01 < 0,05$ , donc l'hypothèse nulle est rejetée.

En conclusion, la série temporelle est stationnaire pour un modèle *ARIMA*.

## 3.3 Modélisation ARIMA-GARCH

On essaye d'ajuster cette série stationnaire vers un *ARIMA - GARCH*.

### 3.3.1 La partie ARIMA

#### 3.3.1.1 Ajustement vers ARIMA avec le moins d'AIC

```
arima.mod <- RTS %>%  
model(ARIMA(Return, stepwise = FALSE, approximation = FALSE))  
  
<ARIMA(2, 0, 4) >
```

#### 3.3.1.2 Analyse résiduelle du ARIMA trouvé

##### Graphiquement :

```
arima.mod%>%  
gg_tsresiduals()
```

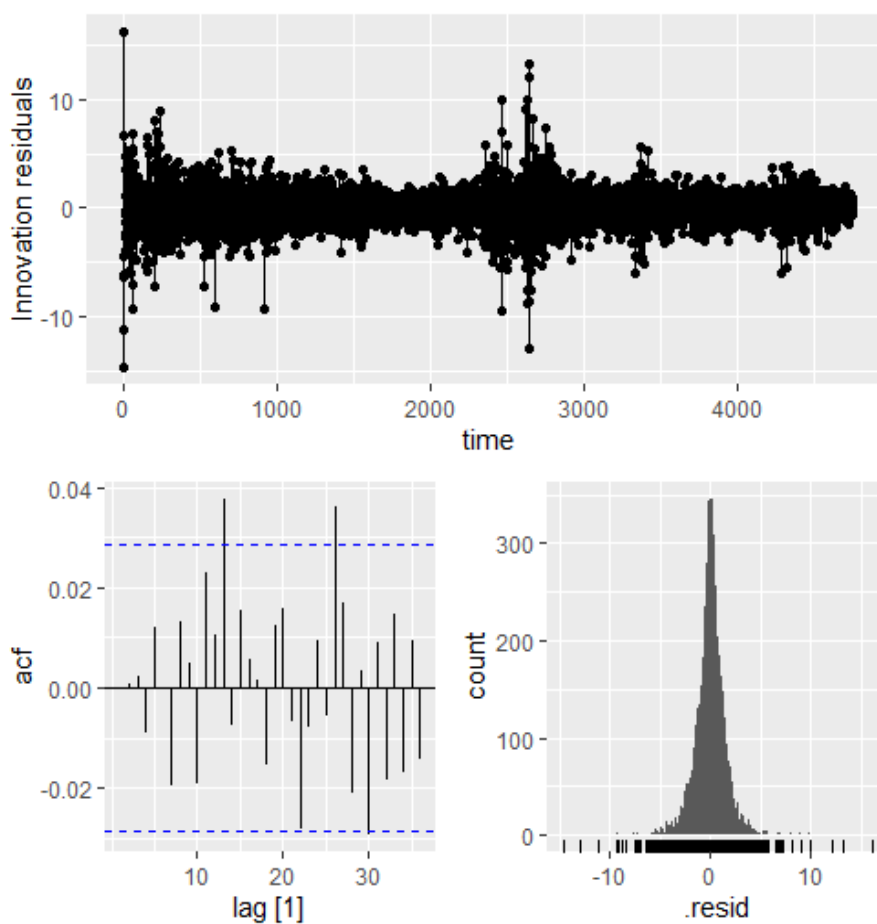


FIGURE 3.3 – Graphismes de l'analyse

### 3.3.1.3 Test de Box-Ljung

```
augment(arima.mod) %>%
features(.innov, ljung_box)

.model          lb_stat lb_pvalue
<chr>          <dbl>   <dbl>
1 ARIMA()      0.000829  0.977
```

**Interprétation :** p-value > 0,05, les résidus du modèle ne peuvent pas être distingués du bruit blanc. Il est rationnel d'accepter *ARIMA*(2, 0, 4)

### 3.3.1.4 Effet ARCH sur les résidus

L'Effet ARCH existe, si les acf et pacf des résidus au carré sont significatifs.

```
Acorr <- arima.mod %>%
augment() %>%
select(.innov) %>%
ACF(.innov ^ 2) %>%
autoplot() +
labs(y = "resid^2 acf")
```

```
AcorrP <- arima.mod %>%
augment() %>%
select(.innov) %>%
PACF(.innov ^ 2) %>%
autoplot() +
labs(y = "resid^2 pacf")
```

```
gridExtra::grid.arrange(Acorr, AcorrP)
```

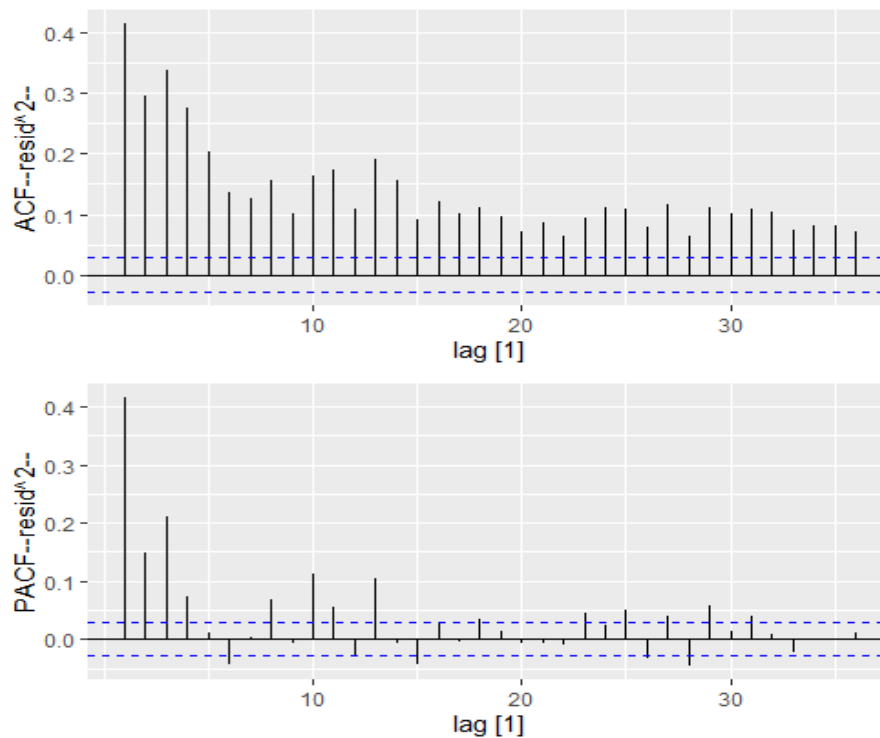


FIGURE 3.4 – l'ACF et la PACF

l'Effet *ARCH* sur les résidus est significatif, donc il est possible d'ajuster un modèle *GARCH* sur les résidus.

### 3.3.2 La partie GARCH

#### ARIMA(2,0,4)+GARCH(1,1)

```
spec <-
ugarchspec(
variance.model = list(
model = "sGARCH",
garchOrder = c(1, 1),
submodel = NULL,
external.regressors = NULL,
variance.targeting = FALSE
),
mean.model = list(armaOrder = c(2, 4),
include.mean = TRUE),
distribution.model = "sged"
)
```

```
fit <- ugarchfit(spec, RTS$Return,
solver = "hybrid")
```

```
print(fit)
```

|           | Estimate  | Std. Error | t value   | Pr(> t ) |
|-----------|-----------|------------|-----------|----------|
| ## mu     | 0.025299  | 0.018994   | 1.33197   | 0.182871 |
| ## ar1    | -1.312107 | 0.015853   | -82.76469 | 0.000000 |
| ## ar2    | -0.440748 | 0.020674   | -21.31875 | 0.000000 |
| ## ma1    | 1.311944  | 0.002183   | 601.05156 | 0.000000 |
| ## ma2    | 0.429959  | 0.029283   | 14.68293  | 0.000000 |
| ## ma3    | 0.006806  | 0.011885   | 0.57264   | 0.566887 |
| ## omega  | 0.020438  | 0.007478   | 2.73295   | 0.006277 |
| ## alpha1 | 0.051404  | 0.009379   | 5.48086   | 0.000000 |
| ## beta1  | 0.933010  | 0.012840   | 72.66518  | 0.000000 |
| ## skew   | 0.916142  | 0.021129   | 43.36049  | 0.000000 |
| ## shape  | 1.383882  | 0.053582   | 25.82747  | 0.000000 |

Maintenant, on peut donc faire la prévision.

### 3.4 Prévisions

```
fore<- function(input, p, q, n) {
spec <-
ugarchspec(
variance.model = list(
model = "sGARCH",
garchOrder = c(1, 1),
submodel = NULL,
external.regressors = NULL,
variance.targeting = FALSE
),
mean.model = list(armaOrder = c(p, q),
include.mean = TRUE),
distribution.model = "sged"
)
fit <- ugarchfit(spec, input$Return,
```

```
solver = "hybrid")
fore <- ugarchforecast(fit, n.ahead = n)
pred <- fore@forecast$seriesFor
pred <- as.data.table(pred)
setnames(pred, "Return")
pred[, time := seq.int(from = 1, to = n, by = 1)][]

}

pred <- fore(RTS, 2, 4, 10)

pred

f.Cltr <- function(input, p, q, n) {
spec <-
ugarchspec(
variance.model = list(
model = "sGARCH",
garchOrder = c(1, 1),
submodel = NULL,
external.regressors = NULL,
variance.targeting = FALSE
),
mean.model = list(armaOrder = c(p, q),
include.mean = TRUE),
distribution.model = "sged"
)
fit <- ugarchfit(spec, input$Close,
solver = "hybrid")
f <- ugarchforecast(fit, n.ahead = n)
pred <- f@forecast$seriesFor
pred <- as.data.table(pred)
setnames(pred, "Return")
pred[, time := seq.int(from = 1, to = n, by = 1)][]

}
```

```
pred <- f.Cltr(RTS, 2, 4, 10)
```

**Résultat :**

```
> pred <- fore(RTS, 2, 4, 10)
```

```
> pred
```

```
Return time
```

|     |             |    |
|-----|-------------|----|
| 1:  | 0.035924894 | 1  |
| 2:  | 0.008326899 | 2  |
| 3:  | 0.022540847 | 3  |
| 4:  | 0.021691669 | 4  |
| 5:  | 0.027899268 | 5  |
| 6:  | 0.032694806 | 6  |
| 7:  | 0.033879211 | 7  |
| 8:  | 0.032973809 | 8  |
| 9:  | 0.031884934 | 9  |
| 10: | 0.031430902 | 10 |

```
Return time
```

|     |          |    |
|-----|----------|----|
| 1:  | 25712.02 | 1  |
| 2:  | 25714.27 | 2  |
| 3:  | 25724.26 | 3  |
| 4:  | 25730.46 | 4  |
| 5:  | 25738.74 | 5  |
| 6:  | 25746.47 | 6  |
| 7:  | 25754.36 | 7  |
| 8:  | 25762.21 | 8  |
| 9:  | 25770.09 | 9  |
| 10: | 25777.96 | 10 |

# Chapitre 4

## Conclusion

Le plus important dans les études des séries temporelles est de savoir comment faire l'ajustement de la série à un modèle usuel. Dans notre travail nous avons vu la modélisation *ARCH/GARCH* pour les séries non linéaires (ont une volatilité hétéroscésastique), plus précisément la modélisation *ARIMA – GARCH*, et leur techniques pour prédire ce type des séries. Puisque ce domaine est très vaste, les résultats de l'application sont restés ouvertes, pour aller plus loin par des recherches approfondies.

# Bibliographie

- [1] Régis Bourbonnais, Michel Terraza, (2004), Analyse des séries Temporelles 3e édition, Dunod. Paris.
- [2] Bourbonnais. R, (2005), Econométrie, 6 e Edition. Dunod. Paris
- [3] Charpentier. A, (2004), Cours des séries temporelles. Théorie et Applications, Université de Paris.
- [4] Corinne. P, (2005), Séries chronologiques : Quelques éléments du cours. Paris.
- [5] Brockwell P.J. et Davis R.A, (2001), Introduction to Time Series and Forecasting, Springer.
- [6] Olivier Cappé, Maurice Charbit, Eric Moulines, (2007), Introduction aux séries temporelles.
- [7] J.J. Daudin, C. Duby ,S. Robin, P.Trécourt, (1996), Analyse de Séries Chronologiques, INA-PG, Mathématiques
- [8] Agnès Lagnoux, (2010), Séries Chronologiques, lagnoux@univ-tlse2.fr
- [9] Engle, R.F. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation, *Econometrica* 50, 987-1007, 1982
- [10] Bollerslev, T.Generalized autoregressive conditional heteroskedasticity, *J. Econometrics* 31, 307-327, 1986.
- [11] Rainer, Von Sachs Sébastien, Van-Bellegem, Séries Chronologiques.
- [12] Christophe Hurlin : Econométrie pour la finance, Rue de Blois BP 6739, Orléans, 2006-2007.
- [13] Brockwell P.J. et Davis R.A, Introduction to Time Series and Forecasting, Springer, 2001.
- [14] Arthur Charpentier, Cours de séries temporelles théorie et applications, volume1, Paris Dauphine.
- [15] Bensaber A. et Bleuse-Trillon B,(1989), Pratique des chroniques et de la prévision à court terme, Masson, Paris.
- [16] GOURIROUX Christian, (1992), Modèle ARCH et Applications financières , Composé par ECONOMICA, Paris.
- [17] Aragon, Y. Séries temporelles avec R. Méthodes et cas., Springer, Paris, 2011.