

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

UNIVERSITÉ MOHAMED KHIDER, BISKRA

Faculté des Sciences Exactes et des Sciences de la Nature et de la Vie

DÉPARTEMENT DE MATHÉMATIQUES



Mémoire présenté en vue de l'obtention du Diplôme :

MASTER en Mathématiques

Option : Statistique

Présenté par : **MABROUKI ZINEB OUIAM**

THEME

Volatilité Stochastique et Applications

Membres du Comité d'Examen :

Pr.	BAHIMI Brahim	Université de Biskra	Président
Pr.	YAHIA Djabrane	Université de Biskra	Rapporteur
Dr.	ZOUAOUI Nour Elhouda	Université de Biskra	Examinatrice

Année Universitaire : 2022 – 2023

Dédicace

Je dédie ce modeste travail

À mes chers parents, à mes sœurs et mes frères, ainsi à toute ma famille.

A tous mes amis

A l'ensemble des étudiants et collègues de la promotion 2022/2023 - Master 2
Mathématiques.

REMERCIEMENTS

Je remercie dieu le tout puissant de m'avoir donné la santé et la volonté d'entamer et de terminer ce mémoire.

Tout d'abord, ce travail ne serait pas aussi riche et n'aurait pas pu avoir le jour sans l'aide, le suivi et l'encouragement de mon rapporteur Professeur YAHIA Djabrane. Je lui adresse mes remerciements, pour la qualité de son encadrement exceptionnel, pour sa patience, sa rigueur et sa disponibilité durant la préparation de ce mémoire.

Je tiens à remercier également les membres de Jury, Prof. BRAHIMI Brahim en qualité de Président et Mme. Zouaoui Nour Elhouda en qualité d'examinatrice. Merci d'avoir accepté de juger et évaluer ce travail.

Mes remerciements s'adressent également à tous mes professeurs, pour leurs générosités et pour leur aide, soutien moral et encouragements, dont ils ont su faire preuve. Particulièrement Prof. Khelfallah, qui ma bien aidé sur la partie stochastique du sujet.

Table des matières

Remerciements	ii
Table des matières	iii
Table des figures	v
Liste des tables	vi
Introduction	1
1 Sur la théorie des séries temporelles	5
1.1 Généralités	5
1.2 Analyse des séries temporelles	8
1.2.1 Modélisation d'une série temporelle	8
1.2.2 Modèle stationnaire	11
1.3 Modèles ARMA	11
1.4 Prévisions linéaires	14
2 Volatilité stochastique : ARCH	17
2.1 Construction du modèle	18
2.1.1 Volatilité et rendements	18
2.1.2 Aplatissement et asymétrie	20

2.2	Propriétés principales des séries de rendements	22
2.3	Processus ARCH	23
2.3.1	Rappels de probabilités	23
2.3.2	Notions et propriétés	24
2.3.3	Propriétés des processus ARCH	26
2.3.4	Processus GARCH et propriétés	27
2.4	Estimation et choix de modèle	30
2.5	Prévisions	34
3	Application	36
3.1	Présentation des données	36
3.2	Tests et validation du modèle	38
3.2.1	Test de l'hypothèse de normalité	38
3.2.2	Estimation et choix de modèle	39
3.2.3	Test de significativité des paramètres	41
3.3	Prévisions	42
	Conclusion	43
	Bibliographie	44
	Annexe : Abréviations et Notations	46

Table des figures

1	Graphique d'une série temporelle datée du X^e ou XI^e siècle. Source :[6]. . .	2
1.1	Décomposition de la série du nombre de naissances par mois à New York, de janvier 1946 à décembre 1958.	9
1.2	Série mensuelle des accidents mortelles en USA : 1973-1979 (Schéma additif).	9
1.3	Série mensuelle des passagères aériens en USA : 1949-1960 (Schéma multiplicatif).	10
1.4	Représentation d'un modèle $AR(3)$ stationnaire	12
1.5	Représentation d'un modèle $AR(2)$ non-stationnaire	12
1.6	Représentation d'un modèle $MA(1)$ stationnaire.	13
2.1	Série des rendements de l'indice boursier CAC 40 : France,1991-1998.	19
2.2	Queues de distributions Exponentielle, Normale et de Pareto.	20
2.3	Représentation graphique d'un modèle $ARCH(2)$, $n = 1000$	26
3.1	Prix mensuel du pétrole brut, (USD par baril) : 1986 - 2006.	37
3.2	Rendements de prix mensuel du pétrole brut, (USD par baril) : 1986 - 2006.	37
3.3	Autocorrélogramme des rendements et des erreurs.	38

Liste des tableaux

1.1	Principales caractéristiques des processus ARMA	14
3.1	Statistiques de la série des rendements mensuel du pétrole brut	37
3.2	Différentes valeurs AIC des modèles estimés	40
3.3	Prévisions des rendements et intervalle de confiance	42

Introduction

Les séries chronologiques, sont généralement constituées d'une suite d'observations aléatoires, ordonnées dans le temps, qui suit par exemple, l'évolution d'un stock (population, effectifs salariés, clientèle,...), un flux (ventes, nombre de diagnostics,...). Le temps joue ici le rôle de variable explicative. Dans ce type d'étude, la prise de décision repose sur le calcul des prévisions, qui lui même basé sur la bonne modélisation de la série étudiée. L'étude des séries temporelles, ou séries chronologiques, correspond à l'analyse statistique d'observations régulièrement espacées dans le temps. Elles ont été utilisées en astronomie (périodicité des taches solaires), en météorologie (Niveaux d'eau dans les mers, les lacs et les rivières),... etc. Citons [6], pour plus de détails sur les utilisations des séries temporelles.

Historiquement, ce sont les astronomes qui ont travaillé initialement sur des séries chronologiques. L'observation systématique du ciel remonte à l'Antiquité. Par exemple, les Romains savaient déjà que l'année dure approximativement 365 et 1/4 jours (calendrier Julien¹). C'est en particulier grâce à ce genre de donnée que Kepler² a pu énoncer ces lois sur le mouvement des planètes. La Figure Fig.1 suivante est tirée d'un manuscrit de *X^e* siècle représentant l'inclinaison des orbites des planètes en fonction du temps (voir, [6]). Ces visualisations graphiques ont permis, grâce aux différents outils mathématiques mis en place au *XVIII^e* et *XIX^e* siècle, de mettre en place les premières techniques d'étude des séries chronologiques, parmi lesquelles, l'analyse harmonique.

¹Le calendrier Julien est un calendrier solaire utilisé dans la Rome antique, introduit par Jules César en 46 av. J.-C.

²Johannes Kepler (1571 - 1630), est un astronome allemand célèbre pour avoir étudié les relations mathématiques des mouvements des planètes sur leur orbite.

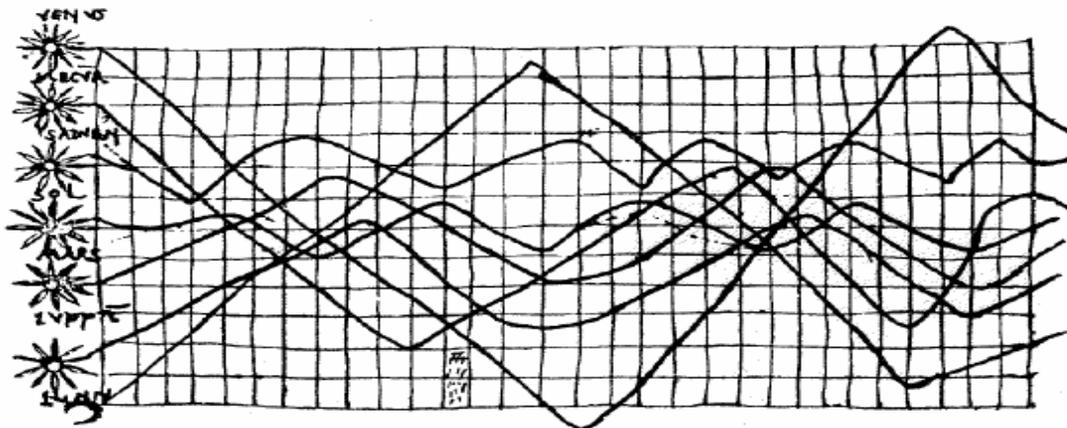


FIG. 1 – Graphique d’une série temporelle datée du X^e ou XI^e siècle. Source :[6].

Une des préoccupations majeures dans le monde financier, est le fort croisement du nombre d’actifs échangés, ainsi des risques fatals résultant de mouvements de marché. D’où l’apparition du terme de séries chronologiques financières. Dans les années 80, des développements ont été apportés dans l’étude de la non-linéarité de certaines séries, et sur leur modélisation. En 1982, *Engle* a introduit la classe des modèles *ARCH* (*autorégressifs conditionnellement hétéroscédastiques*). Ces modèles ont été introduits pour palier à une observation empirique qui n’était pas prise en compte par les modèles : la volatilité conditionnelle de Y_t n’a aucune raison d’être constante.

Un modèle économétrique est dit homosédastique si la variance des erreurs est constante quelque soit la période d’étude. Dans le cas contraire, on parlera d’hétéroscédasticité. Les modèles *ARCH* sont ainsi conditionnellement hétéroscédastique. *Engle* en 1982, a cherché un modèle dans lequel la variance conditionnelle de X_t sachant X_{t-1} dépendrait de X_{t-1} , et plus particulièrement,

$$V(X_t | X_{t-1}) = [\alpha + \beta X_{t-1}^2] \sigma^2.$$

Pour cela, il considéré les modèles de la forme :

$$X_t = \sigma_t \varepsilon_t, \text{ où } \sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2.$$

Cette classe de modèle, appelée $ARCH(1)$ a été généralisée sous la forme $ARCH(p)$,

$$X_t = \sigma_t \varepsilon_t, \text{ où } \sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + \dots + \alpha_p X_{t-p}^2,$$

ce qui a permis l'analogie entre les modèles *autorégressifs* (AR) et les modèles $ARCH$. De plus, cette classe de modèles $ARCH$ a été plus généralisée (par *Bollerslev*, 1986), en considérant des volatilités σ_t^2 de la forme :

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

De nombreuses méthodes conventionnelles de mesure du risque, associées aux actifs, se font par le biais d'études de la variance (*volatilité*) de l'actif. Cette mesure de la volatilité inconditionnelle ne prend pas compte qu'il pourrait y avoir un schéma prévisible de la volatilité des marchés boursiers. Dans la théorie des rendements financiers, une idéalisation de base est que les rendements suivent un modèle de série chronologique stationnaire avec une structure de volatilité stochastique. La présence de volatilité stochastique implique donc, que les rendements ne sont pas nécessairement indépendants dans le temps.

Ce mémoire est une compilation de travaux de recherche liés à la volatilité stochastique, ainsi qu'au diverses propriétés probabilistes et statistiques du modèle $ARCH$, en apportant des réponses à certaines questions notamment :

- Quel est l'objectif de l'étude de la volatilité et des modèles $ARCH$ et comment se fait la modélisation ?
- Comment choisir un modèle qui présente le mieux les données ?
- Quelles sont les conclusions à tirés de cette modélisation.

L'organisation de la thèse est la suivante :

Nous rappelons dans un premier chapitre les définitions principales, notations et résultats que nous verrons souvent dans la théorie des séries temporelles, et que l'on va utiliser par la

suite, notamment : les modèles *AR*, *MA*, *ARMA*.

Dans le deuxième chapitre, nous introduisons quelques modèles économétriques simples utiles pour analyser les séries chronologiques financières et appliquons ces modèles aux rendements des actifs. Ce chapitre est consacré donc à l'étude de la variance conditionnelle et la volatilité conditionnelle du processus des rendements, ainsi qu'aux modèles *hétéroscédastiques ARCH* et *GARCH*.

Nous présentons dans le dernier chapitre, une application sur des données réelles d'une série financière de l'évolution du prix de pétrole (Série mensuel du pétrole brut (\$ par baril), du janvier 1986 au janvier 2006).

Chapitre 1

Sur la théorie des séries temporelles

La théorie des séries temporelles est une combinaison de deux concepts, probabiliste et statistique, le probabiliste dont on étudie les caractéristiques des variables aléatoires X_t . Le problème statistique est de donner les caractéristiques des distributions de la série temporelle X_t , pour les observations X_1, X_2, \dots, X_n au temps $t = 1, 2, \dots, n$. Le modèle statistique résultant sert à la compréhension du système stochastique d'une part et la prédiction des valeurs futures (*i.e.*, X_{n+1}, X_{n+2}, \dots) d'autre part.

1.1 Généralités

En 1927, deux excellents articles ont ouvert une nouvelle voie d'étude et d'analyse de série temporelle : l'article de *Yule* ("On the methode of investigating periodicities in disturbated series with spacial referance to Wolfer's sunspot numbers'") et celui de *Slutsky* ("The summation of random causes as the source of cyclical processes'"). En effet, *Yule* a introduit des modèles autorégressifs, en considérant des modèles de la forme :

$$Y_t = \alpha Y_{t-1} + \beta Y_{t-2}.$$

Etant données deux valeurs initiales, cette suite présente un comportement saisonnier, fonction des paramètres α et β . *Yule* remarque qu'en fait, le comportement dépend des racines

(complexes) de l'équations $z^2 - \alpha z - \beta = 0$, et plus particulièrement de leur position par rapport au disque unité. Si leur module est l'inférieur 1, alors on observe un comportement sinusoïdal amorti. En fait, la forme générale des solutions sera :

$$Y_t = A\rho^t \cos(\omega t - \theta), \text{ lorsque } 0 < \rho < 1.$$

Le modèle autorégressif noté AR , proposé par *Yule* est le suivant :

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \varepsilon_t, \tag{1.1}$$

où ε_t correspond à un 'bruit blanc', i.e., un processus 'indépendant' (ou plus faiblement, non corrélé avec son passé). Néanmoins, des hypothèses plus forts doivent parfois être faites : on veut que ce bruit soit également indépendant du passé de la variable Y_t , i.e. ε_t indépendant de Y_{t-h} pour tout $h \geq 1$, et on parle alors d'innovation du processus Y_t .

Slutsky a introduit les moyennes mobiles notée MA , la même année que *Yule* a introduit les processus autorégressifs. Mais son article en 1927 en russe n'a été traduit qu'en 1937 en anglais. La forme générale était la suivante :

$$Y_t = \mu_0 \varepsilon_t + \mu_1 \varepsilon_{t-1} + \dots + \mu_q \varepsilon_{t-q}, \tag{1.2}$$

où ε_t est un bruit blanc. Cette écriture a suggéré d'élargir la relation (1.1) sous une forme proche de (1.2), à savoir :

$$\phi_0 Y_t + \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} = \varepsilon_t.$$

Les processus introduits par *Yule* sont les processus $AR(p)$ et ceux introduits par *Slutsky* sont les processus $MA(q)$.

L'étude des séries temporelles semble avoir atteint sa maturité au cours des années 1970, où des développements significatifs sont apparus. En 1970, *Box* et *Jenkins* ont publié leur

ouvrage "*Time series analysis, forecasting and control*", montrant que l'étude des séries temporelles à l'aide de processus de type *ARMA* (obtenus en associant les écritures eq (1.1) et eq (1.2) des processus *AR* et *MA*) pouvait s'appliquer à de nombreux domaines, et pouvait être facilement implémentée informatiquement.

Les modèles *ARMA* sont un mélange des modèles (1.1) et (1.2) proposés par *Yule* et *Slutsky*.

Un processus X_t est un processus *ARMA*(p, q) s'il existe un bruit blanc ε_t , tel que :

$$X_t = \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}, \text{ pour tout } t.$$

Sous certaines conditions, ces processus sont stationnaires. Comme nous le verrons par la suite, ces processus peuvent s'écrire sous la forme :

$$\Phi(L)X_t = \Theta(L)\varepsilon_t, \text{ où } \Phi(L) = \mathbf{I} - \phi_1 L - \dots - \phi_p L^p \text{ et } \Theta(L) = \mathbf{I} - \theta_1 L - \dots - \theta_q L^q,$$

L représentant l'**opérateur retard**, au sens où $LX_t = X_{t-1}$, et avec la convention

$$L^p = L \circ L^{p-1}, \quad \text{soit} \quad L^p X_t = X_{t-p}$$

la série Y_t telle que $Y_t = L^p X_t$ est alors la série X_t retardée de p périodes. Parallèlement, on dira qu'un processus non-stationnaire est **intégré** d'ordre 1, si en la différenciant une fois, on obtient un **processus stationnaire** : X_t (non-stationnaire) sera dit intégré d'ordre 1 si le processus Y_t définit $Y_t = \Delta X_t = X_t - X_{t-1} = (1 - L)X_t$ est stationnaire. On appellera alors de processus *ARIMA*(p, d, q) dont la forme est :

$$\Pi(L)X_t = \Phi(L)(1 - L)^d X_t = \Theta(L)\varepsilon_t, \text{ où } (\varepsilon_t) \text{ est un bruit blanc.}$$

Pour les données réelles, on notera que $d = 1, 2$ ou 3 (au maximum). Cela signifie que (Y_t) définit comme différence d'ordre d du processus (X_t), soit $Y_t = (1 - L)^d X_t$, "suit un processus *ARMA*(p, q)". On parlera d'ailleurs de présence de racine unité : $z = 1$ est alors racine du

polynôme autorégressif $\Pi(z)$. Par généralisation, on peut considérer le cas où $\exp(2\pi/s)$ est racine du polynôme autorégressif : c'est à dire que $\Pi(z) = (1 - L^s)\Phi(L)$. On dira alors que l'on est en présence d'une racine unité saisonnière, qui engendreront les modèles *SARIMA*. Les modèles intégrés sont très présents dans les séries économiques, par exemple les séries d'indices boursiers, d'indice de production, d'indice de prix... Les modèles *SARIMA* sont également très présents des lors que les séries sont très saisonnières (*avec une forte périodicité trimestrielle ou annuelle,...etc*).

1.2 Analyse des séries temporelles

Une série temporelle (ou série chronologique) est une suite finie $(X_t)_{1 \leq t \leq n}$, où t représente le temps (en minute, jour, année,...).

1.2.1 Modélisation d'une série temporelle

La modélisation des séries temporelles se fait à partir de la décomposition classique, "*décomposition de Persons*" en fonction des quatre éléments suivants :

1. Tendance (Z_t) : mouvement à long terme (*longue période*).
2. Saisonnière (S_t) : fonction périodique du temps (*période courte*).
3. Cycle (C_t) : cycle d'affaires, fluctuation périodique (*moyenne terme*).
4. Résidu (ε_t) : partie irrégulière, correspondante à la notion d'écart au modèle ou encore bruit blanc.

D'une manière générale, on peut proposer un modèle qui représente la série temporelle étudiée en combinaison les éléments précédents. La figure Fig.1.1 suivante, donne une ullustration de la décomposition en tandance (trend), saisonnalité (saisonal) et résidus (random) de la série (observed) des nombres de naissances par mois à New York, de janvier 1946 à décembre 1958.

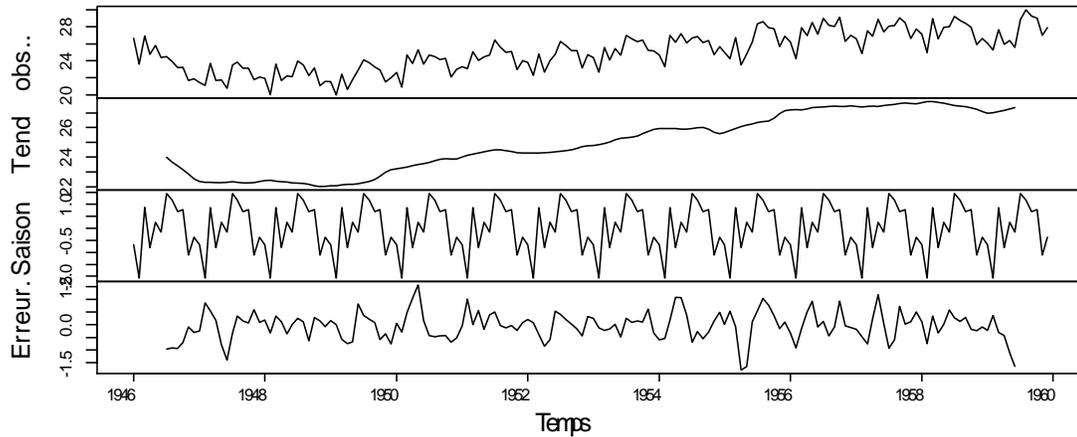


FIG. 1.1 – Décomposition de la série du nombre de naissances par mois à New York, de janvier 1946 à décembre 1958.

Pour cela, on a deux types de modèles d’ajustement de forme **additive** ou **multiplicative** comme suit :

$$X_t := Z_t + S_t + C_t + \varepsilon_t \quad \text{ou} \quad X_t := Z_t * S_t * C_t + \varepsilon_t.$$

Les deux figures Fig.1.2 et Fig.1.3 suivantes, présentent respectivement la série mensuelle des accidents mortelles en USA : 1973-1979 (Schéma additif) et la série mensuelle des passagers aériens en USA : 1949-1960 (Schéma multiplicatif).

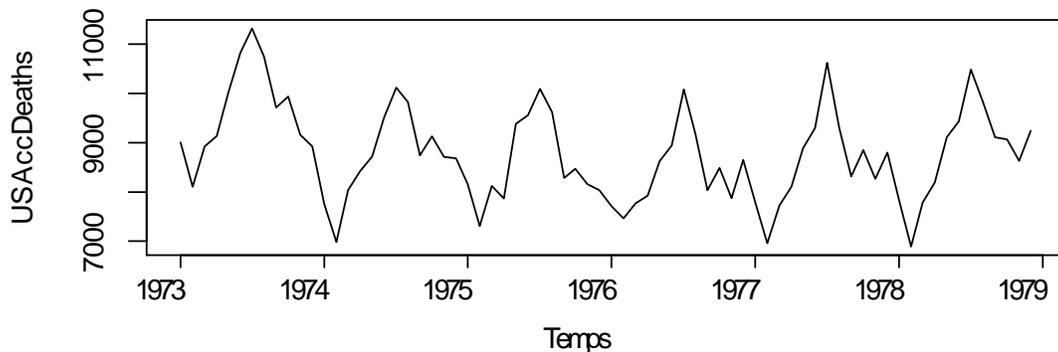


FIG. 1.2 – Série mensuelle des accidents mortelles en USA : 1973-1979 (Schéma additif).

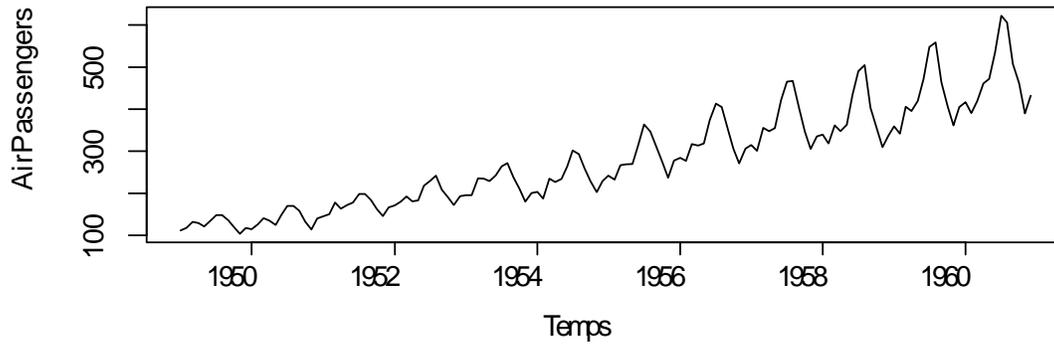


FIG. 1.3 – Série mensuelle des passagères aériens en USA : 1949-1960 (Schéma multiplicatif).

Le deuxième type est le modèle autoprojectif, dont on suppose que X_t est une fonction de ces valeurs passées et d'une perturbation aléatoire ε_t :

$$X_t := f(X_{t-1}, X_{t-2}, \dots; \varepsilon_t),$$

dans cette classe, on peut citer les modèles *AR*, *MA*, *ARIMA*, *ARMA*, ...

Le troisième type est le plus important, c'est le modèle explicatif. Dans cette catégorie de modèle, la variable aléatoire X_t est exprimée en fonction d'un vecteur aléatoire Y_t et d'une perturbation aléatoire ε_t :

$$X_t := f(Y_t, \varepsilon_t),$$

où Y_t est soit déterministe ou aléatoire, dans ce dernier cas les processus $(Y_t)_t$ et $(\varepsilon_t)_t$ ont certaines propriétés d'indépendance ou du non-corrélation. Ces modèles sont les modèles de base de l'économétrie et nous les considérons essentiellement pour faire le lien entre eux et les modèles autoprojectifs. On a ainsi deux cas particuliers de modèle explicatif, modèle explicatif statique où les variables Y_t ne contiennent pas de valeurs passées de X_t et les ε_t sont indépendantes entre eux. Le deuxième cas est le modèle explicatif dynamique, où les ε_t sont autocorrélés et les Y_t contiennent des valeurs passées de X_t .

1.2.2 Modèle stationnaire

Définition 1.2.1 Un processus $(X_t, t \in \mathbf{Z})$ est **stationnaire au second ordre** ou **stationnaire au sens faible** ou **stationnaire d'ordre deux** si les trois conditions suivantes sont satisfaites :

i) $\forall t \in \mathbf{Z}, E(X_t^2) < \infty,$

ii) $\forall t \in \mathbf{Z}, E(X_t) = m,$ constante indépendante de $t,$

iii) $\forall t, h \in \mathbf{Z}, cov(X_t, X_{t+h}) = E([X_t - m][X_{t+h} - m]) =: \gamma(h),$ indépendante de $t.$

Définition 1.2.2 Un processus X_t est **stationnaire au sens forte** si pour tous t_1, \dots, t_n et h on a l'égalité en loi :

$$(X_{t_1}, \dots, X_{t_n}) \stackrel{d}{=} (X_{t_1+h}, \dots, X_{t_n+h}).$$

Définition 1.2.3 La fonction $\gamma(\cdot)$ sera appelée **fonction d'autocovariance**. Il s'agit d'une fonction paire, au sens où $\gamma(h) = \gamma(-h)$ pour tout h et de variance $V(X_t) = \gamma(0)$ constante, indépendante de $t.$

Définition 1.2.4 On appelle **bruit blanc** (par fois appelé **bruit blanc faible**) un processus ε_t stationnaire dont les autocovariance sont tout nulles : $\gamma_\varepsilon(h) = 0$ pour $h \neq 0.$

1.3 Modèles ARMA

Généralement, les séries temporelles sont représentées par deux types de modèles, le premier est linéaire (modèles : *AR, MA,...*), le deuxième est non-linéaire (modèles : *ARCH, GARCH,...*).

Définition 1.3.1 On appelle **processus autoregressif d'ordre p** , noté $AR(p)$, un processus stationnaire X_t vérifiant une relation du type :

$$X_t - \sum_{i=1}^p \phi_i X_{t-i} = \varepsilon_t \text{ pour tout } t \in \mathbf{Z},$$

où les ϕ_i sont des réels et ε_t est un bruit blanc de variance σ^2 , qui est équivalent à :

$$\Phi(L)X_t = \varepsilon_t \quad \text{où} \quad \Phi(L) = \mathbf{I} - \Phi_1 L - \dots - \Phi_p L^p.$$

Dans les figures Fig.1.4 et Fig.1.5, nous donnons une représentation des modèles $AR(3)$ stationnaire et celle d'un $AR(2)$ non-stationnaire respectivement pour $t = 1, \dots, 200$.

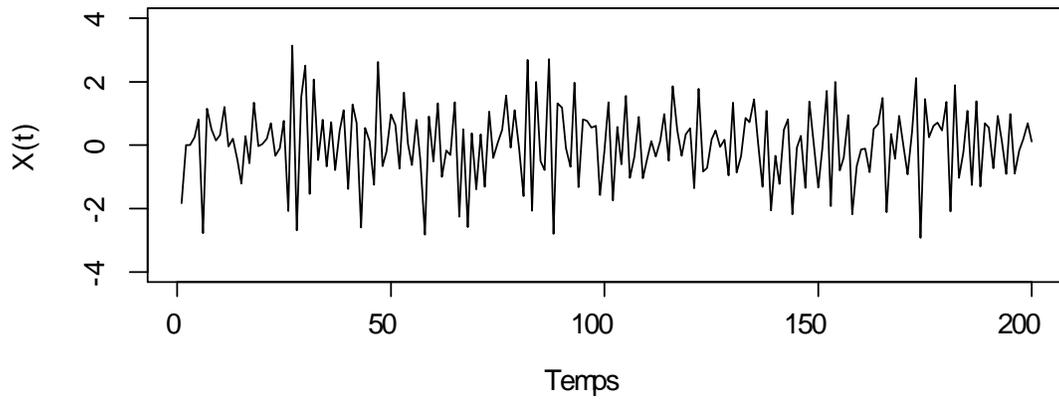


FIG. 1.4 – Représentation d'un modèle $AR(3)$ stationnaire

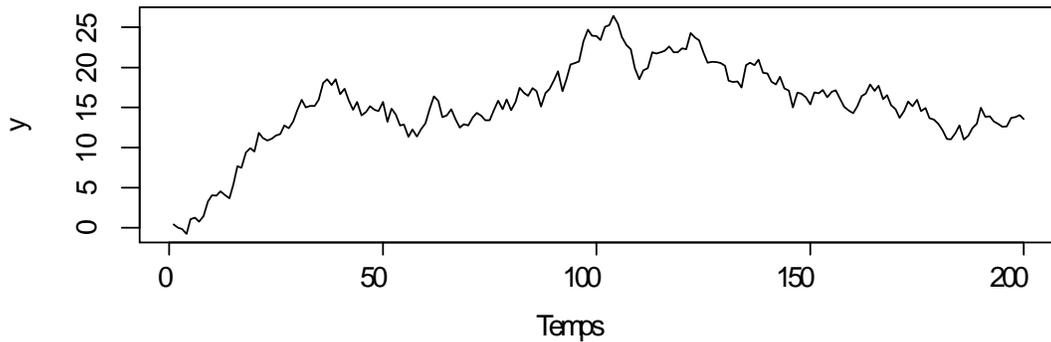


FIG. 1.5 – Représentation d'un modèle $AR(2)$ non-stationnaire

Définition 1.3.2 On appelle *moyenne mobile (Moving Average)* un processus de la forme :

$$X_t = \varepsilon_t + a_1\varepsilon_{t-1} + \dots + a_q\varepsilon_{t-q},$$

où les ε_t est un bruit blanc, centré de variance σ^2 , et où l'entier non nul q est l'ordre de la moyenne mobile notée $MA(q)$.

Remarque 1.3.1 Par construction, les modèles MA sont toujours stationnaires. Ils dépendent du ε_t seulement. La figure Fig.1.6 suivante est une représentation graphique d'un modèle $MA(1)$.

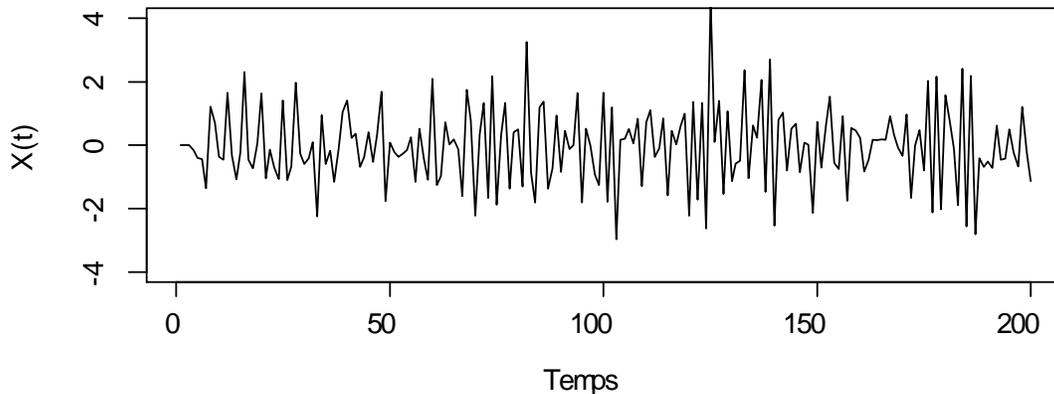


FIG. 1.6 – Représentation d'un modèle $MA(1)$ stationnaire.

Définition 1.3.3 Le nombre de paramètres d'un processus moyenne mobile ou d'un auto-régressif peut parfois être réduit en considérant un processus mixte : auto-régressif moyenne mobile (ARMA) d'ordres p et q :

$$X_t - \sum_{k=1}^p a_k X_{t-k} = \varepsilon_t - \sum_{j=1}^q b_j \varepsilon_{t-j}.$$

Remarque 1.3.2 La stationnarité d'un $ARMA_{p,q}$ est assurée lorsque toutes les racines du polynôme $A(z) = 1 + a_1z + \dots + a_pz^p$ sont de module strictement supérieur à 1. Ce polynôme

forme avec $B(z) = 1 + b_1z + \dots + b_qz^q$ les deux polynômes caractéristiques du processus. On supposera également que les polynômes A et B n'ont pas de racine commune, afin de s'assurer qu'il n'y a pas de représentation plus courte.

Remarque 1.3.3 On a

- i) Un processus $AR(p)$ est un $ARMA_{p,0}$ et un processus $MA(q)$ est un $ARMA_{0,q}$.
- ii) De plus, dans le cas où $p = \infty$ (resp. $q = \infty$), on dit que X_t admet est un modèle autoregressif infini (resp. moyenne mobile infini) noté $AR(\infty)$ (resp. $MA(\infty)$).

Proposition 1.3.1 Pour une large classe de suite d'auto-covariances $\gamma(h)$ et pour tout entier k positif, il existe un processus $ARMA$ tel que ses k premières auto-covariances coïncident avec $\gamma(1) \leq \dots \leq \gamma(k)$.

Dans le tableau Tab.1.1 suivant, nous regroupons les principales caractéristiques des procesus $ARMA$ en fonction des valeurs de $\gamma(h)$ et celles d'autocorrélation

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

Modèle	MA_q	AR_p	$ARMA_{p,q}$
$\gamma(h)$	$0, \forall h > q$	$\gamma \xrightarrow{h \rightarrow \infty} 0$	$\forall h > q, \gamma \xrightarrow{h \rightarrow \infty} 0$
$\rho(h)$	$0, \forall h > q$	$\rho \xrightarrow{h \rightarrow \infty} 0$	$\forall h > q, \rho \xrightarrow{h \rightarrow \infty} 0$

TAB. 1.1 – Principales caractéristiques des procesus $ARMA$

1.4 Prévisions linéaires

Soient X_1, X_2, \dots, X_n n -observations données de la série X_t :

$$X_t - \sum_{k=1}^p a_k X_{t-k} = \varepsilon_t - \sum_{j=1}^q b_j \varepsilon_{t-j} \iff A(L) X_t = B(L) \varepsilon_t.$$

La prévision d'une valeur future $X_{n+h} = X_n(h)$ d'horizon h de la série X_t est inconnue, mais elle est estimée par la quantité

$$\hat{X}_{n+h} = E(X_{n+h} | X_1, X_2, \dots, X_n),$$

Il s'agit de la projection linéaire de X_{n+h} sur l'ensemble engendrée par les n -observations données X_1, X_2, \dots, X_n .

Exemple 1.4.1 Si X_t est un modèle $AR(2)$: $X_t - a_1X_{t-1} - a_2X_{t-2} = \varepsilon_t$. Alors,

$$\begin{aligned} X_t &= a_1X_{t-1} + a_2X_{t-2} + \varepsilon_t \Rightarrow X_{n+h} = a_1X_{n+h-1} + a_2X_{n+h-2} + \varepsilon_{n+h} \quad (h \geq 1) \\ \Rightarrow \hat{X}_{n+h} &= E(X_{n+h} | X_1, X_2, \dots, X_n) = E(a_1X_{n+h-1} + a_2X_{n+h-2} + \varepsilon_{n+h} | X_1, X_2, \dots, X_n). \end{aligned}$$

D'où

$$\begin{aligned} \hat{X}_{n+1} &= E(a_1X_n + a_2X_{n-1} + \varepsilon_{n+1} | X_1, X_2, \dots, X_n) = a_1X_n + a_2X_{n-1} \\ \hat{X}_{n+2} &= E(a_1X_{n+1} + a_2X_n + \varepsilon_{n+2} | X_1, X_2, \dots, X_n) = a_1\hat{X}_{n+1} + a_2X_{n-1} \\ &= a_1(a_1X_n + a_2X_{n-1}) + a_2X_{n-1} = a_1^2X_n + a_2(1 + a_1)X_{n-1}. \end{aligned}$$

Exemple 1.4.2 De même, si X_t est un modèle $MA(1)$: $X_t = \varepsilon_t + b\varepsilon_{t-1}$. Alors,

$$\begin{aligned} X_{n+h} &= \varepsilon_{n+h} + b\varepsilon_{n+h-1} \quad (h \geq 1) \\ \Rightarrow \hat{X}_{n+h} &= E(X_{n+h} | X_1, X_2, \dots, X_n) = E(\varepsilon_{n+h} + b\varepsilon_{n+h-1} | X_1, X_2, \dots, X_n) \end{aligned}$$

On obtient donc,

$$\begin{aligned} \hat{X}_{n+1} &= E(\varepsilon_{n+1} + b\varepsilon_n | X_1, X_2, \dots, X_n) = b\varepsilon_n, \\ \hat{X}_{n+2} &= E(\varepsilon_{n+2} + b\varepsilon_{n+1} | X_1, X_2, \dots, X_n) = 0. \end{aligned}$$

Remarque 1.4.1 *Pour la prévision linéaire $X_t(k)$ des processus $ARMA_{p,d,q}$, on aura toujours besoin d'une estimation de $\varepsilon_{t-1}, \varepsilon_{t-2}, \dots$, i.e. du "bruit inobservable passé" du modèle. On peut aussi recourir à la représentation $AR(\infty)$, dans quel cas on aura besoin de X_{t-1}, X_{t-2}, \dots , qui sont aussi inobservables.*

Chapitre 2

Volatilité stochastique : ARCH

En 1982, *Engle* a proposé **un processus de volatilité** avec **une variance conditionnelle** variant dans le temps, le processus *ARCH*. Quatre ans après l'introduction par *Engle* du processus *ARCH*, *Bollerslev* en 1986, a proposé le modèle *ARCH* généralisé (noté *GARCH*) comme solution naturelle au problème des ordres *ARCH* élevés. Ce modèle est basé sur une spécification *ARCH* infinie et permet de réduire considérablement le nombre de paramètres estimés d'un nombre infini à quelques-uns.

Dans le modèle *GARCH* de *Bollerslev*, la variance conditionnelle est une fonction linéaire des innovations passées au carré et des variances conditionnelles calculées antérieurement.

Les séries précédemment étudiées étaient supposées stationnaires. Si besoin, tendances et saisonnalités étaient supprimées pour obtenir une série résiduelle stationnaire. Comme on peut le voir, la moyenne semble être constante alors que la variance change au cours du temps (on qualifie ce comportement d'« *hétéroscédastic* »). De plus, les moments de grande variabilité semblent regroupés. Les modèles de type *ARIMA* qui supposent un comportement « *homoscédastique* » (*variance constante*), ne sont pas adaptés à ce type de série à volatilité stochastique. Nous présentons dans ce chapitre des modèles adaptés à ce type de série : les processus *ARCH* introduits par *Engle* vers 1982, ainsi que leur généralisation, les processus *GARCH*.

2.1 Construction du modèle

2.1.1 Volatilité et rendements

L'un des modèles classiques qui fournit une bonne représentation des phénomènes financiers comme : les taux de change, prix d'actifs ou d'option, rendements d'indice boursier,... etc, en terme d'un processus $(P_t)_t$ qui désigne l'évolution de ces phénomènes est celui de *Black* et *Scholes*¹ :

$$P_t := P_0 \exp(\sigma_t W_t + \mu_t t),$$

où W_t est un mouvement brownien, σ_t et μ_t sont des processus stochastiques (voir [11], Définition 2.3. page 11). Ce modèle est basé sur l'équation de diffusion de type :

$$dP_t = \mu_t P_t dt + \sigma_t P_t dW_t,$$

ou encore sous forme d'intégrale :

$$\log\left(\frac{P_t}{P_0}\right) = \int_0^t \left(\mu_s - \frac{1}{2}\sigma_s^2\right) ds + \int_0^t \sigma_s^2 dW_s.$$

A l'origine de ce modèle, le processus σ_t est supposé constant et on dit qu'il représente la volatilité de P_t . On a deux cas : le premier est que σ_t et μ_t sont des processus déterministes les *log rendements* $(\log(\frac{P_t}{P_{t-1}}), t = 1, 2, \dots)$ dans une intervalle $]t-1, t]$ sont des variables aléatoires indépendantes, d'une distribution normale, de **moyenne** et **variance** données respectivement par :

$$\bar{\mu}_t = \int_{t-1}^t \left(\mu_s - \frac{1}{2}\sigma_s^2\right) ds \quad \text{et} \quad \bar{\sigma}_t^2 = \int_{t-1}^t \sigma_s^2 dW_s,$$

en d'autre terme, les variables $\varepsilon_t = \frac{1}{\bar{\sigma}_t} (\log(\frac{P_t}{P_{t-1}}) - \bar{\mu}_t)$, forment un échantillon *i.i.d.*, d'une **loi normale standard centrée et réduite**. Dans ce cas $\bar{\sigma}_t$ est appelée la "**volatilité moyenne**" sur l'intervalle $]t-1, t]$. Le deuxième cas, si σ_t et μ_t sont aléatoires, alors ε_t n'est pas né-

¹Le modèle Black-Scholes est un modèle mathématique du marché pour une action, dans lequel le prix de l'action est un processus stochastique.

cessairement gaussien. De plus, si l'unité du temps est petite, alors on peut dire que ε_t est d'une distribution approximativement normale, de même que les processus σ_t et μ_t sont approximativement constants, on peut les remplacer donc par $\bar{\sigma}_t$ et $\bar{\mu}_t$ respectivement. Pour simplifier les notations, on prend $\bar{\mu}_t = 0$, ce qui revient à définir un modèle de type (*ARCH*, *GARCH*,...), d'un processus à volatilité stochastique σ_t et un processus centré réduit $(\varepsilon_t)_t$. Un des apports de ces modèles **hétéroscédastiques**, était donc de mieux s'ajuster aux données, en particulier aux données financières. Pour cela, on pose :

$$X_t =: \log\left(\frac{P_t}{P_{t-1}}\right), \quad t = 1, 2, \dots$$

P_t désigne : les prix d'indice, prix d'actions,... etc. Notre étude sera donc basée sur les caractéristiques de la série temporelle $(X_t)_t$, dont le modèle souhaitable pour la représenter est de type hétéroscédastique : *ARCH*, *GARCH*,...

La figure Fig.2.1 présente la série des rendements des cours de clôture quotidiens de l'indice boursier CAC 40 (France). Les données sont échantillonnées pendant les heures ouvrables, c'est-à-dire que le fins de semaine et les jours fériés sont omis, entre 1991 et 1998.

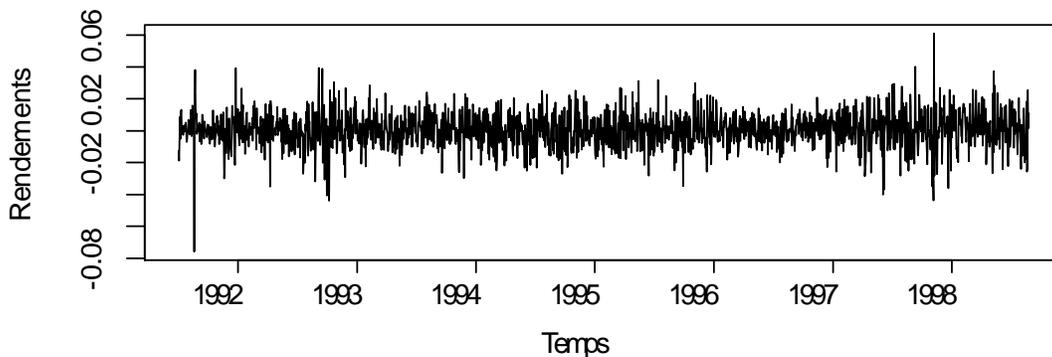


FIG. 2.1 – Série des rendements de l'indice boursier CAC 40 : France,1991-1998.

2.1.2 Aplatissement et asymetrie

Les observations de la série temporelle X_t ont une distribution, qui est souvent supposée normale (**Gaussienne**). Cependant, les études empiriques (voir : [13] pages 9–12), de pratiquement toutes les séries financières montrent que ce n'est pas tout à fait correct. Une façon de quantifier cette propriété est de regarder l'aplatissement des distributions. **L'aplatissement** est une mesure dans laquelle les données observées se situent près du centre d'une distribution ou dans les queues :

- Une distribution platykurtique a une valeur d'aplatissement inférieure à celle d'une distribution normale standard.
- Une distribution leptokurtique a une valeur d'aplatissement supérieure à celle d'une distribution normale standard. Une distribution est leptokurtique si elle a une queue lourde ou de type-Pareto.

Cette dernière caractéristique est courante dans les données de séries chronologiques sur les prix, les taux et les rendements observés. Cela implique qu'il y a dans les séries financières une probabilité plus élevée d'événements extrêmes que dans les données normalement distribuées. La figure Fig.2.2 suivante présente une comparaison entre les queues de distributions : Exponentielle, Normale et de Pareto.

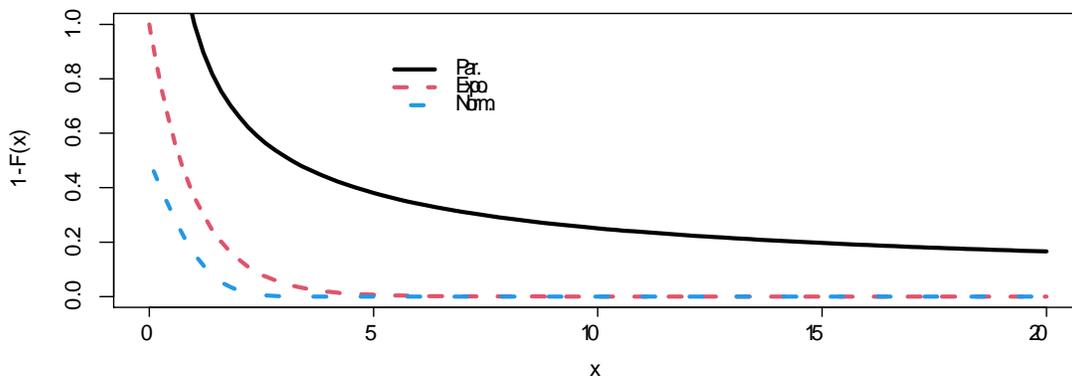


FIG. 2.2 – Queues de distributions Exponentielle, Normale et de Pareto.

Définition 2.1.1 (Kurtosis) La *kurtosis* notée κ de Fisher est défini comme suit (pour la série temporelle $X_t, t = 1, \dots, n$) :

$$\kappa = \frac{\mu_4}{\mu_2^2} - 3 = \frac{\mu_4}{\sigma^4} - 3, \quad \text{où } \mu_k = E \left[(X - \mu)^k \right], \quad \mu = E(X).$$

Cette mesure d'aplatissement n'existe que si le quatrième moment existe et fini. L'aplatissement est une forme normalisée du quatrième moment central d'une distribution. Dans l'aplatissement de Fisher, le chiffre trois est soustrait pour donner à la distribution normale le kurtosis zéro (la distribution est mésokurtique).

Remarque 2.1.1 En supposant que $X_t \sim \mathcal{N}(0, 1)$, les quatrième et deuxième moments centraux sont donnés par :

$$\mu_4 = E \left[(X - \mu)^4 \right] = E(X^4) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} t^4 e^{-t^2/2} dt = 3$$

$$\mu_2 = E \left[(X - \mu)^2 \right] = V(X) = 1,$$

ce qui montre que l'aplatissement de Fisher pour les distributions normales est nul.

Remarque 2.1.2 Le moment central de degré k est estimé par :

$$\mu_k = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^k.$$

Dans les cas ci-dessus, nous avons les définitions suivantes :

- ▷ distributions platykurtiques : $\kappa < 0$
- ▷ distributions leptokurtiques : $\kappa > 0$ (*aplatissement excessif*)
- ▷ distributions mésokurtiques : $\kappa = 0$

Définition 2.1.2 (Skewness) Les observations de la distribution empirique de X_t montrent souvent que la distribution est **leptokurtique**. Une autre propriété qui s'écarte de la distribution gaussienne si souvent supposée est que la distribution empirique n'est pas symétrique.

L'asymétrie de Fisher (le type d'asymétrie le plus courant, généralement appelé simplement asymétrie) est définie par :

$$s = \text{skewness} = \text{coefficient d'asymétrie} = \frac{\mu_3}{\mu_2^{3/2}},$$

où μ_3 est le troisième moment central et $\mu_2^{1/2}$ est l'écart type.

Remarque 2.1.3 Une valeur d'asymétrie négative indique que les données ont une distribution asymétrique vers la gauche. Cela signifie que la queue gauche est plus lourde que la queue droite dans la distribution. Respectivement, une valeur d'asymétrie positive indique une distribution asymétrique à droite avec une queue droite plus lourde que sa queue gauche.

2.2 Propriétés principales des séries de rendements

- "Stationnarité" : Les processus stochastiques P_t associés aux prix d'actif sont généralement non stationnaires au sens de la stationnarité du second ordre, tandis que les processus associés aux rendements sont compatibles avec la propriété de stationnarité au second ordre.
- "Autocorrélations des carrés des variations de prix" : La série r_t^2 associée aux carrés des rendements présente généralement de fortes auto-corrélations tandis que les auto-corrélations de la série $r_t = \log\left(\frac{P_t}{P_{t-1}}\right)$ sont souvent très faibles (hypothèse de bruit blanc). L'absence d'auto-corrélation des rendements renvoie à la notion d'efficience (voir, [7] pour plus de détails).
- "Queues de distribution épaisses" : L'hypothèse de normalité des rendements est généralement rejetée. Les queues des distributions empiriques des rendements sont généralement plus épaisses que celles d'une loi gaussienne. On parle alors de distribution **leptokurtique** dont le coefficient d'aplatissement (kurtosis) est plus grand que 3, celle de la distribution normale.
- "Clusters de Volatilité" : On observe empiriquement (voir : [11] page 8), que de fortes

variations des rendements sont généralement suivies de fortes variations. On assiste ainsi à un regroupement des extrêmes en cluster ou paquets de volatilités.

- "Queues épaisses conditionnelles" : Même une fois corrigée de la volatilité clustering (*ex : avec ARCH*), la distribution des résidus demeure **leptokurtique** même si la kurtosis est plus faible que dans le cas non conditionnelle.
- "Effet de levier" : Il existe une asymétrie entre l'effet des valeurs passées négatives et l'effet des valeurs passées positives sur la volatilité des cours ou de rendements. Les baisses de cours tendent à engendrer une augmentation de la volatilité supérieure à celle induite par une hausse des cours de même ampleur (voir [11] pages 9-10).
- "Saisonnalité" : Les rendements présentent de nombreux phénomènes de saisonnalité (effets week end, effet janvier etc..).
- "Asymétrie perte/gain" : La distribution des cours est généralement **asymétrique** : il y a plus de mouvements forts à la baisse qu'à la hausse.

2.3 Processus ARCH

2.3.1 Rappels de probabilités

Soit un couple de variables aléatoires continues (X, Y) , de densité $f(., .)$. Soit $f_Y(.)$ la densité marginale de Y . La densité conditionnelle de X sachant que $Y = y$ est définie par :

$$f_{X|Y=y}(x, y) = \frac{f(x, y)}{f_Y(y)} \quad \text{si } f_Y(y) > 0$$

L'espérance conditionnelle est donc l'espérance par rapport à cette loi :

$$E[X | Y = y] = \int_{\mathbf{R}} x f_{X|Y=y}(x, y) dx$$

La variance conditionnelle est :

$$\sigma_t^2 = V(X_t | \mathcal{F}_{t-1}) = E[X_t^2 | \mathcal{F}_{t-1}] - E[X_t | \mathcal{F}_{t-1}]^2$$

Soit \mathcal{F}_{t-1} l'ensemble engendré par les observations $\{X_{t-1}, X_{t-2}, \dots\}$ précédant l'instant t (Filtration naturelle). On a :

$$E[X_t | \mathcal{F}_s] = X_t, \quad \forall t \leq s$$

$$E[X_t | \mathcal{F}_r] = E[E(X_t | \mathcal{F}_s) | \mathcal{F}_r], \quad \forall r \leq s$$

et en particulier $E[X_t] = E[E(X_t | \mathcal{F}_s)]$.

2.3.2 Notions et propriétés

Définition 2.3.1 Un modèle X_t est dit *autoregressif conditionnellement hétéroscédastique* d'ordre 1 qu'on note par $ARCH(1)$, s'il admet l'écriture suivante :

$$X_t := \sigma_t \varepsilon_t, \quad \sigma_t^2 = E[X_t^2 | \mathcal{F}_{t-1}] = \alpha_0 + \alpha_1 X_{t-1}^2,$$

pour tout $t \in \mathbb{Z}$ et $\alpha_0 > 0$, $\alpha_1 > 0$, des constantes données et ε_t un bruit blanc centrée et réduite.

Remarque 2.3.1 On suppose souvent que les variables ε_t sont indépendantes de la filtration \mathcal{F}_{t-1} et que σ_t dépend de X_t .

Propriété 2.3.1 Le modèle $ARCH(1)$ admet les propriétés suivantes :

1) X_t est hétéroscédastique conditionnellement à son passé :

$$\text{var}(X_t | \mathcal{F}_{t-1}) = \alpha_0 + \alpha_1 X_{t-1}^2.$$

2) De variance constante : $\text{var}(X_t) = \alpha_0 + \alpha_1 \text{var}(X_{t-1})$.

Preuve. Notons que les moments conditionnels d'ordre 2 et 4 du processus $X_t \sim \mathcal{N}(0, 1)$ sont liés :

$$E[X_t^4 | \mathcal{F}_{t-1}] = 3(E[X_t^2 | \mathcal{F}_{t-1}])^2$$

Il suffit d'appliquer la décomposition de la **variance** et la **variance conditionnelle** comme suit : pour tout $X_t \in \mathbf{L}^2$:

$$\begin{aligned} \text{var}(X_t | \mathcal{F}_{t-1}) &= E[X_t^2 | \mathcal{F}_{t-1}] - E[X_t | \mathcal{F}_{t-1}]^2. \\ \text{var}(X_t) &= \text{var}(E[X_t | \mathcal{F}_{t-1}]) + E[\text{var}(X_t | \mathcal{F}_{t-1})]. \end{aligned} \quad (2.1)$$

Pour la preuve de (1), utilisons (2.1), on a

$$\begin{aligned} \text{var}(X_t) &= \text{var}(E[X_t | \mathcal{F}_{t-1}]) + E[\text{var}(X_t | \mathcal{F}_{t-1})] \\ &= E(E[X_t | \mathcal{F}_{t-1}] - EE[X_t | \mathcal{F}_{t-1}])^2 + E[E[X_t^2 | \mathcal{F}_{t-1}] - E[X_t | \mathcal{F}_{t-1}]^2] \\ &= E(E[X_t | \mathcal{F}_{t-1}]^2) - E[X_t]^2 + E[X_t^2] - E[E[X_t | \mathcal{F}_{t-1}]^2] \\ &= E(E[X_t | \mathcal{F}_{t-1}]^2) + E[X_t]^2 - 2E[X_t | \mathcal{F}_{t-1}]E[X_t] + E[X_t^2] - E[E[X_t | \mathcal{F}_{t-1}]^2] \\ &= E[X_t]^2 - 2E[X_t | \mathcal{F}_{t-1}]E[X_t] + E[X_t^2] \\ &= E[X_t^2] - E[X_t]^2. \end{aligned}$$

$$\begin{aligned} \text{var}(X_t | \mathcal{F}_{t-1}) &= E[\varepsilon_t^2(\alpha_0 + \alpha_1 X_{t-1}^2) | \mathcal{F}_{t-1}] - E[\varepsilon_t]^2 E[\sigma_t | \mathcal{F}_{t-1}]^2 \\ &= E[\varepsilon_t^2] E[\alpha_0 + \alpha_1 X_{t-1}^2 | \mathcal{F}_{t-1}] - 0 = \alpha_0 + \alpha_1 X_{t-1}^2 = \sigma_t^2. \end{aligned}$$

Pour le deuxième point, on a

$$\begin{aligned} \text{var}(X_t) &= \text{var}(E[X_t | \mathcal{F}_{t-1}]) + E[\text{var}(X_t | \mathcal{F}_{t-1})] \\ &= \text{var}(0) + E[\alpha_0 + \alpha_1 X_{t-1}^2] = \alpha_0 + \alpha_1 \text{var}(X_{t-1}). \end{aligned}$$

■

Remarque 2.3.2 Sous-hypothèse de *stationnarité* (i.e. $\text{var}(X_t) = \text{var}(X_{t-1})$), on a :

$$\text{var}(X_t) = \frac{\alpha_0}{1 - \alpha_1},$$

pour que cette variance non-conditionnelle existe, il suffit donc que : $\alpha_1 \in]0, 1[$.

Définition 2.3.2 Un processus $ARCH_p$ est défini par :

$$X_t = \sigma_t \varepsilon_t \quad \text{avec } \sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + \dots + \alpha_p X_{t-p}^2 \quad \text{où } \varepsilon_t \mid \mathcal{F}_{t-1} \sim \mathcal{N}(0, \sigma_t^2).$$

Remarque 2.3.3 La variance conditionnelle dépend du temps : si les valeurs précédentes sont grandes (en valeur absolue), la variance sera grande, et inversement. Ainsi, si on observe un choc dans la série (valeur anormalement grande), elle sera suivie d'une période de haute volatilité, dont la durée dépend de l'ordre p du modèle ARCH. La figure Fig.2.3 suivante est une représentation graphique d'un modèle $ARCH(2)$.

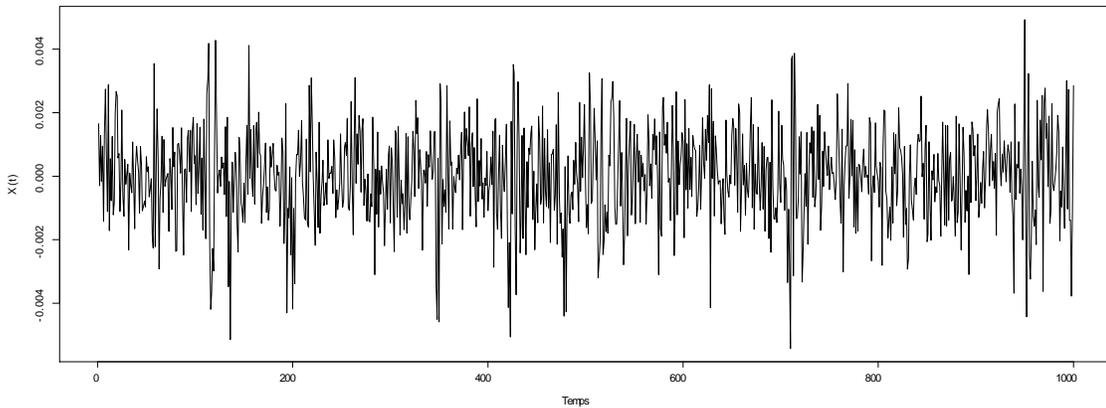


FIG. 2.3 – Représentation graphique d'un modèle $ARCH(2)$, $n = 1000$.

2.3.3 Propriétés des processus ARCH

Si X_t est un processus $ARCH$, alors :

$$E[X_t] = 0 \quad \text{et} \quad E[X_t \mid \mathcal{F}_{t-1}] = 0$$

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

$$V(X_t | \mathcal{F}_{t-1}) = E[X_t^2 | \mathcal{F}_{t-1}] = \alpha_0 + \alpha_1 X_{t-1}^2 + \dots + \alpha_p X_{t-p}^2 = \sigma_t^2.$$

Propriété 2.3.2 La condition suffisante de stationnarité d'un ARCH (p) est $\sum_{i=1}^p \alpha_i < 1$.

Proposition 2.3.1 On peut montrer également, ce qui peut être intéressant pour détecter un ARCH en pratique, que la distribution d'un processus ARCH a :

- i) un skewness nul (moment centré d'ordre 3) : la distribution est donc **symétrique**,
- ii) un kurtosis (moment centré d'ordre 4) **supérieur à 3** : la distribution est donc plus aplatie qu'une gaussienne.

2.3.4 Processus GARCH et propriétés

Définition 2.3.3 On dit qu'un modèle X_t est **autorégressif conditionnellement hétéroscédastique généralisé** (GARCH(1,1)), s'il admet l'écriture suivante :

$$X_t := \sigma_t \varepsilon_t, \quad \sigma_t^2 := \alpha_0 + \alpha X_{t-1}^2 + \beta \sigma_{t-1}^2,$$

pour tout $t \in \mathbf{Z}$ et α_0, α et β des constantes positives données avec σ_t^2 est \mathcal{F}_t -mesurable, $(\varepsilon_t)_t$ est un **bruit blanc** telle que

$$E[\varepsilon_t | \mathcal{F}_{t-1}] = 0, \quad E[\varepsilon_t^2 | \mathcal{F}_{t-1}] = 1.$$

Propriété 2.3.3 Le modèle GARCH(1,1) à les propriétés suivantes :

- 1) X_t est **hétéroscédastique conditionnellement** à X_{t-1} :

$$\text{var}(X_t | \mathcal{F}_{t-1}) = \alpha_0 + \alpha X_{t-1}^2 + \beta \sigma_{t-1}^2 = \sigma_t^2,$$

- 2) $\text{var}(X_t) = \alpha_0 + (\alpha + \beta)\text{var}(X_{t-1})$.

Preuve. 1) On a

$$\begin{aligned} V(X_t | \mathcal{F}_{t-1}) &= E[X_t^2 | \mathcal{F}_{t-1}] - E[X_t | \mathcal{F}_{t-1}]^2 \\ &= E[\sigma_t^2 \varepsilon_t^2 | \mathcal{F}_{t-1}] - E[\sigma_t \varepsilon_t | \mathcal{F}_{t-1}]^2 = \alpha_0 + \alpha X_{t-1}^2 + \beta \sigma_{t-1}^2 = \sigma_t^2. \end{aligned}$$

2) De plus,

$$\begin{aligned} \text{var}(X_t) &= \text{var}(E[X_t | \mathcal{F}_{t-1}]) + E[\text{var}(X_t | \mathcal{F}_{t-1})] \\ &= \text{var}(0) + E[\alpha_0 + \alpha X_{t-1}^2 + \beta \sigma_{t-1}^2 | \mathcal{F}_{t-1}]. \end{aligned}$$

Comme,

$$E[X_{t-1}^2 | \mathcal{F}_{t-1}] = E[\sigma_{t-1}^2 \varepsilon_{t-1}^2 | \mathcal{F}_{t-1}] = E[\sigma_{t-1}^2 | \mathcal{F}_{t-1}] = \text{var}(X_{t-1}),$$

donc,

$$\text{var}(X_t) = \alpha_0 + (\alpha + \beta)\text{var}(X_{t-1}).$$

Sous hypothèse de stationnarité, (*i.e.* $\text{var}(X_t) = \text{var}(X_{t-1})$), ce qui implique que :

$$\text{var}(X_t) = \frac{\alpha_0}{1 - (\alpha + \beta)},$$

pour que cette variance existe, il faut donc que : $\alpha + \beta < 1$. ■

Définition 2.3.4 Un processus $GARCH_{p,q}$ est défini par :

$$X_t = \sigma_t \varepsilon_t; \quad \text{avec} \quad \sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + \dots + \alpha_p X_{t-p}^2 + \beta_1 \sigma_{t-1}^2 + \dots + \beta_q \sigma_{t-q}^2$$

où

$$\varepsilon_t | \mathcal{F}_{t-1} = X_{t-1}, X_{t-2}, \dots \sim \mathcal{N}(0, \sigma_t^2)$$

avec $\alpha_0 > 0$, $\alpha_i \geq 0$ pour $i = 1, \dots, p$ et $\beta_j \geq 0$ pour $j = 1, \dots, q$.

Remarque 2.3.4 *Un $GARCH_{p,0}$ est un $ARCH_p$. De plus, un processus $GARCH$ peut être vu comme un processus $ARCH$ d'ordre infini, et peut ainsi représenter formellement de façon plus parcimonieuse un processus $ARCH$ comprenant un nombre élevé de paramètres.*

Si X_t est un processus $GARCH_{p,q}$, alors :

$$E[X_t] = 0, \quad E[X_t | \mathcal{F}_{t-1}] = 0$$

$$Cov(X_t, X_{t+h}) = 0, \quad \forall h > 0 \quad \text{et} \quad Cov(X_t, X_{t+h} | \mathcal{F}_{t-1}) = 0.$$

Remarque 2.3.5 *Le modèle $ARCH(p)$ est une forme restreinte du modèle $GARCH(p, q)$, avec comme restriction $q = 0$. On réécrit l'équation (σ_t^2) avec les notations :*

$$\phi(L) = \Phi(L) + \Psi(L) \quad \text{et} \quad Z_t = X_t^2 - \sigma_t^2,$$

on trouve

$$\sigma_t^2 = \alpha_0 + \Phi(L)X_t^2 + \Psi(L)\sigma_t^2 = \alpha_0 + \Phi(L)X_t^2 - \Psi(L)Z_t,$$

d'où

$$(1 - \phi(L))Y_t = \alpha + (1 - \Psi(L))Z_t,$$

ce qui implique que $Y_t = X_t^2$ admet la représentation d'un modèle $ARMA$, voir eq. (1.3.3).

Propriétés 2.3.1 *Soit X_t un processus $GARCH_{p,q}$, et soit $m = \max(p, q)$.*

- *Le processus X_t^2 admet une représentation $ARMA(m, q)$.*
- *Pour identifier un $GARCH_{p,q}$, on identifiera tout d'abord le processus $ARMA(m, q)$ qui modélise X_t^2 .*
- *Pour identifier p dans le cas où $m = q (p \leq q)$, il faut effectuer des tests de significativité des coefficients $\alpha_q, \dots, \alpha_1$ du processus $ARMA(m, q)$.*
- *Pour les conditions de **stationnarité**, on peut citer à titre d'exemple [3] et [14].*

Proposition 2.3.2 *Si le processus X_t satisfait une représentation $GARCH(p, q)$ conditionnellement gaussienne, telle que :*

$$X_t = \sigma_t \varepsilon_t, \quad \text{avec } V(X_t | \mathcal{F}_{t-1}) = \alpha + \sum_{i=1}^q \alpha_i X_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2$$

où ε_t désigne bruit blanc faible gaussien. Alors,

(i) *la loi marginale de X_t a des queues plus épaisses qu'une loi normale (distribution leptokurtique) :*

$$E[X_t^4] \geq 3 [E(X_t^2)]^2$$

(ii) *son coefficient d'excès de **kurtosis** peut s'exprimer sous la forme suivante :*

$$\kappa = \frac{E[X_t^4]}{E(X_t^2)^2} - 3 = 3 \frac{V[E(X_t^2 | \mathcal{F}_{t-1})]}{E(X_t^2)^2}.$$

2.4 Estimation et choix de modèle

Quand on est on face du problème d'ajustement d'un modèle hétéroscédastiques aux données, la première étape est de choisir l'ordre (p, q) et la classe-type du modèle à utiliser. Généralement, pour le choix du modèle, l'approche la plus utile est d'estimer un modèle, faire une validation et passer à d'autres modèles,... Ainsi de suite jusqu'à ce que nous arrivions à la classe préférable. Une démarche inspirée (voir :[9]) de celle de Box et Jenkins, pour les modèles *ARIMA*, qui s'appuie sur l'identification a priori des paramètres p et q en utilisant des méthodes qui ne nécessitent pas l'estimation des autres paramètres $(\alpha_0, \alpha$ et $\beta)$.

L'estimation des paramètres de ces modèles, est généralement fondée sur la méthode du maximum de vraisemblance notée *MV*. L'optimisation est effectuée en supposant que la suite ε_t est un bruit blanc faible et gaussien, centré et de variance unité.

Définition 2.4.1 La fonction de log-vraisemblance associée à un échantillon de T observations (y_1, y_2, \dots, y_T) de Y_t sous l'hypothèse de normalité de la loi conditionnelle de Y_t sachant Y_{t-1} s'écrit :

$$\log L(\theta) = -\frac{T}{2} \log(2\pi) - \frac{1}{2} \sum_{t=1}^T \log(\sigma_t^2(\theta)) - \frac{1}{2} \sum_{t=1}^T \frac{[y_t - m_t(\theta)]^2}{\sigma_t^2(\theta)}$$

dans ces notations la fonction $\sigma_t^2(\theta)$ désigne la variance conditionnelle et donc il n'y a pas lieu de l'élever au carré.

Définition 2.4.2 Les estimateurs du MV sous l'hypothèse de normalité, notés $\hat{\theta}$, des paramètres $\theta \in \mathbb{R}^k$, satisfont un système non linéaire à k équations :

$$\frac{\partial \log L(\theta)}{\partial \theta} \Big|_{\theta=\hat{\theta}} = 0 \quad \text{et} \quad \frac{\partial^2 \log L(\theta)}{\partial \theta^2} \Big|_{\theta=\hat{\theta}} < 0 \quad (2.2)$$

avec

$$\begin{aligned} \frac{\partial \log L(\theta)}{\partial \theta} \Big|_{\theta=\hat{\theta}} = & -\frac{1}{2} \sum_{t=1}^T \frac{1}{\sigma_t^2(\hat{\theta})} \frac{\partial \sigma_t^2(\theta)}{\partial \theta} \Big|_{\theta=\hat{\theta}} + \frac{1}{2} \sum_{t=1}^T \frac{[y_t - m_t(\hat{\theta})]^2}{\sigma_t^4(\hat{\theta})} \frac{\partial \sigma_t^2(\theta)}{\partial \theta} \Big|_{\theta=\hat{\theta}} \\ & + \frac{1}{2} \sum_{t=1}^T \left[\frac{y_t - m_t(\hat{\theta})}{\sigma_t^2(\hat{\theta})} \right] \frac{\partial m_t(\theta)}{\partial \theta} \Big|_{\theta=\hat{\theta}} \end{aligned}$$

Remarque 2.4.1 On peut montrer que ce système peut se décomposer en deux sous systèmes lorsque les paramètres θ interviennent de façon séparée dans l'écriture de l'espérance et de la variance conditionnelle. Ainsi, si l'on a $\theta = (\alpha\beta)$ où α n'apparaît que dans l'espérance conditionnelle et β dans la variance conditionnelle, on peut décomposer ce système en deux sous système puisque :

$$\frac{\partial \log L(\alpha)}{\partial \alpha} \Big|_{\theta=\hat{\theta}} = \sum_{t=1}^T \left[\frac{y_t - m_t(\hat{\alpha})}{\sigma_t^2(\hat{\beta})} \right] \frac{\partial m_t(\alpha)}{\partial \theta} \Big|_{\alpha=\hat{\alpha}}$$

$$\frac{\partial \log L(\beta)}{\partial \beta} \Big|_{\theta=\hat{\theta}} = -\frac{1}{2} \sum_{t=1}^T \frac{1}{\sigma_t^2(\hat{\beta})} \frac{\partial \sigma_t^2(\beta)}{\partial \beta} \Big|_{\beta=\hat{\beta}} + \frac{1}{2} \sum_{t=1}^T \frac{[y_t - m_t(\hat{\alpha})]^2}{\sigma_t^4(\hat{\alpha})} \frac{\partial \sigma_t^2(\beta)}{\partial \beta} \Big|_{\beta=\hat{\beta}}$$

Revenant maintenant au modèle $GARCH(1, 1)$, comme nous ne connaissons pas la vraie distribution de ce modèle, on doit utiliser les faits que la distribution conditionnelle est normale

et que le bruit ε_t suit une loi gaussienne centré réduite, à savoir :

$$X_t | \mathcal{F}_{t-1} \sim \mathcal{N}(0, \sigma_t^2) \quad , \quad \frac{X_t}{\sigma_t} \sim \mathcal{N}(0, 1).$$

L'estimateur du maximum vraisemblance des paramètres d'un modèle $GARCH(1, 1)$, est non linéaire et peut être obtenu en résolvant les conditions de premier ordre (2.2). Le log-vraisemblance du modèle se calcul généralement à partir de la distribution conditionnelle. Pour T observations (X_1, X_2, \dots, X_T) , le log vraisemblance d'un modèle $GARCH(1, 1)$ du paramètre $\theta = (\alpha_0, \alpha, \beta)^T$ se calcul donc, à partir du produit de toutes les densités conditionnelles de X_t :

$$L_T(\theta) = \sum_{t=1}^T l_t(\theta) \quad , \quad l_t(\theta) = \log f_\varepsilon\left(\frac{X_t}{\sigma_t}\right),$$

f_ε : désigne la densité de ε_t , qui vaut dans le cas gaussien $\mathcal{N}(0, 1)$:

$$f_\varepsilon\left(\frac{X_t}{\sigma_t}\right) = \frac{1}{\sqrt{2\pi\sigma_t^2}} \exp\left\{-\frac{1}{2} \left(\frac{X_t}{\sigma_t}\right)^2\right\}.$$

D'où :

$$l_t(\theta) = \frac{-1}{2} \log(2\pi) - \frac{1}{2} \log(\sigma_t^2) - \frac{1}{2} \left(\frac{X_t}{\sigma_t}\right)^2.$$

Le log-vraisemblance par rapport à θ vaut donc :

$$\begin{aligned} L_T(\theta) &= \sum_{t=1}^T \left[\frac{-1}{2} \log(2\pi) - \frac{1}{2} \log(\sigma_t^2) - \frac{1}{2} \left(\frac{X_t}{\sigma_t}\right)^2 \right] \\ &= \frac{-T}{2} \log(2\pi) - \frac{1}{2} \sum_{t=1}^T \log(\alpha_0 + \alpha X_{t-1}^2 + \beta \sigma_{t-1}^2) - \frac{1}{2} \sum_{t=1}^T X_t^2 (\alpha_0 + \alpha X_{t-1}^2 + \beta \sigma_{t-1}^2)^{-1}. \end{aligned}$$

Remarque 2.4.2 *La méthode de maximum vraisemblance pose un problème concernant la valeur initiale du processus de volatilité σ_t , quand on veut maximiser $L_T(\theta)$. Le processus σ_t dépend à la fois de X_{t-1} et de σ_{t-1} . Or, si la première observation a lieu au temps $t = 1$, les valeurs X_0 et σ_0 vont être spécifiées, donc par récurrence σ_t est calculée.*

Cette méthode donne des estimateurs consistants, la matrice des variances covariances égale

à l'inverse de la matrice d'information. Par définition, la matrice d'information notée $I_T(\theta)$ est donnée (voir, [10, page44]) par :

$$I_T(\theta) = E \left[\frac{\partial L_T(\theta)}{\partial \theta} \frac{\partial L_T(\theta)}{\partial \theta^T} \right] = -E \left[\frac{\partial^2 L_T(\theta)}{\partial \theta \partial \theta^T} \right] \quad (2.3)$$

Dans le cas d'un modèle $GARCH(1, 1)$ gaussien,

$$\frac{\partial L_T(\theta)}{\partial \theta} = \sum_{t=1}^T \frac{\partial l_t(\theta)}{\partial \theta},$$

avec

$$\begin{aligned} \frac{\partial l_t(\theta)}{\partial \theta} &= \frac{\partial}{\partial \theta} \left(\frac{-1}{2} \log(2\pi) - \frac{1}{2} \log(\sigma_t^2) - \frac{1}{2} \left(\frac{X_t}{\sigma_t} \right)^2 \right) \\ &= \frac{-1}{2\sigma_t^2} \frac{\partial \sigma_t^2}{\partial \theta} + \frac{1}{2\sigma_t^2} \frac{X_t^2}{\sigma_t^2} \frac{\partial \sigma_t^2}{\partial \theta} = \frac{1}{2\sigma_t^2} \left(\frac{X_t^2}{\sigma_t^2} - 1 \right) \frac{\partial \sigma_t^2}{\partial \theta}. \end{aligned}$$

Pour calculer la matrice d'information sous (2.3), les valeurs passées de σ_t^2 et celles de $\frac{\partial \sigma_t^2}{\partial \theta}$ doivent être spécifiées :

$$\left(i.e. \frac{\partial \sigma_{t-1}^2}{\partial \theta}, \frac{\partial \sigma_{t-2}^2}{\partial \theta}, \dots, \frac{\partial \sigma_0^2}{\partial \theta} \right).$$

Une approche qui est largement utilisé est d'estimer la matrice d'information $I_T(\theta)$ par le Hessian, à savoir :

$$I_T(\theta) \simeq H := -\frac{\partial^2 L_T(\theta)}{\partial \theta \partial \theta^T},$$

la valeur de $\frac{\partial L_T(\theta)}{\partial \theta}$ est nulle quand θ maximise $L_T(\cdot)$, ce qui indique que l'approximation de l'espérance dans (2.3) par les valeurs observées des dérivées d'ordre deux est toujours vérifiée, si aucune des contraintes sur les paramètres ne paraît :

$$\alpha > 0, \alpha \geq 0, \beta \geq 0 \text{ et } \alpha + \beta < 1.$$

2.5 Prévisions

L'un des applications importantes de la modélisation *GARCH* de la volatilité est le calcul des prévisions. Soient X_1, X_2, \dots, X_n n -observations données de la série X_t :

$$X_t = \sigma_t \varepsilon_t; \quad \text{avec} \quad \sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$

où

$$\varepsilon_t \mid \mathcal{F}_{t-1} \sim \mathcal{N}(0, 1)$$

avec $\alpha_0 > 0$, $\alpha_1 \geq 0$ et $\beta_1 \geq 0$.

En supposant que les données jusqu'au temps n sont disponibles. La prévision d'une valeur future de la volatilité $\sigma_{n+h} = \sigma_n(h)$ d'horizon h pour la série X_t est inconnue, mais elle est estimée par la quantité

$$\hat{\sigma}_{n+h}^2 = E(\sigma_{n+h}^2 \mid \mathcal{F}_{t-1}) = \alpha_0 + \alpha_1 \hat{X}_{n+h-1}^2 + \beta_1 \hat{\sigma}_{n+h-1}^2,$$

Pour l'horizon $h = 1$:

$$\hat{\sigma}_{n+1} = E(\sigma_{n+1} \mid \mathcal{F}_{t-1}) = \alpha_0 + \alpha_1 \hat{X}_n^2 + \beta_1 \hat{\sigma}_n^2 = \alpha_0 + \alpha_1 X_n^2 + \beta_1 \sigma_n^2$$

car dans ce cas $\hat{X}_n^2, \hat{\sigma}_n^2$ sont connues est elle present respectivement les valeurs : X_n^2, σ_n^2 .

Pour h quelconque, on a $X_t = \sigma_t \varepsilon_t$, donc

$$X_t^2 = \sigma_t^2 \varepsilon_t^2,$$

cela implique,

$$\begin{aligned}
 \sigma_{n+h}^2 &= \alpha_0 + \alpha_1 (\sigma_{n+h-1}^2 \hat{\varepsilon}_{n+h-1}^2) + \beta_1 \sigma_{n+h-1}^2 \\
 &= \alpha_0 + \alpha_1 (\sigma_{n+h-1}^2 \hat{\varepsilon}_{n+h-1}^2) + \beta_1 \sigma_{n+h-1}^2 + \alpha_1 \sigma_{n+h-1}^2 - \alpha_1 \sigma_{n+h-1}^2 \\
 &= \alpha_0 + (\alpha_1 + \beta_1) \sigma_{n+h-1}^2 + \alpha_1 (\hat{\varepsilon}_{n+h-1}^2 - 1) \sigma_{n+h-1}^2.
 \end{aligned}$$

Maintenant, sachant que $E(\varepsilon_t^2 | \mathcal{F}_{t-1}) = 1$, alors

$$\hat{\sigma}_{n+h}^2 = E(\sigma_{n+h}^2 | \mathcal{F}_{t-1}) = \alpha_0 + (\alpha_1 + \beta_1) \hat{\sigma}_{n+h-1}^2,$$

puisque $E(\hat{\varepsilon}_{n+h-1}^2 - 1 | \mathcal{F}_{t-1}) = 0$.

Chapitre 3

Application

Nous présentons et étudions dans ce dernier chapitre, une application sur des données réelles d'une série financière de l'évolution du prix de pétrole (Série mensuel du pétrole brut (\$ par baril), du janvier 1986 au janvier 2006). Cette information est tirée de : `data(oil.price)` au format : séries temporelles avec 241 observations de 1986 à 2006 : 22, 915, 412, 612, 815, 4..., du package TSA en logiciel **R** [4]. Les fonctions R utilisées dans cette étude pratiques sont extraites des packages **R** suivants : `tseries`, `TSA`, `fGarch`, `cvar` et `datasets`.

3.1 Présentation des données

Les séries de prix du pétrole sont constituées d'un ensemble de prix déterminés par les opérateurs du marché pétrolier, concernant principalement sa valeur propre, mais aussi sa valeur spéculée à différents termes. Le graphique Fig.3.1, présente le prix au comptant mensuel du pétrole brut (USD par baril), 01/1986–01/2006.

La figure Fig.3.2 représente l'évolution de la série des rendements de le prix au comptant mensuel du pétrole brut, de la période allant du 01/1986 au 01/2006. L'axe des abscisses présente le temps et l'axe des ordonnées présente le rendement de l'indice journalier.

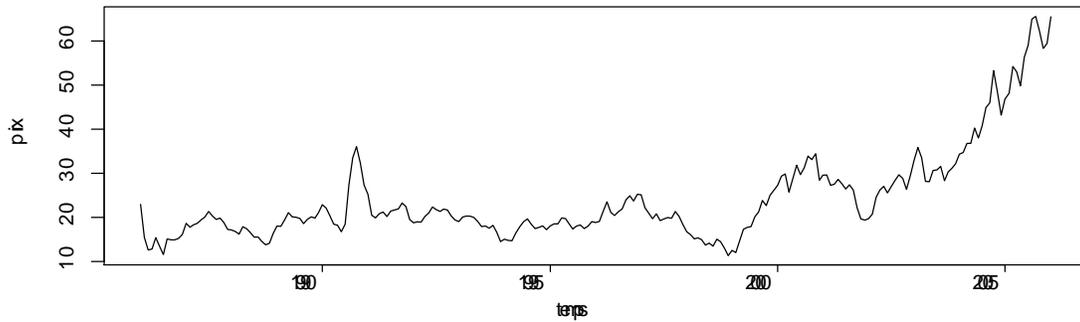


FIG. 3.1 – Prix mensuel du pétrole brut, (USD par baril) : 1986 - 2006.

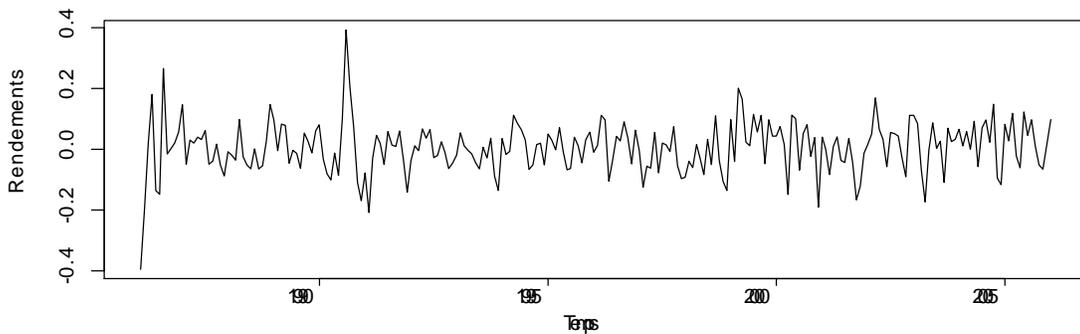


FIG. 3.2 – Rendements de prix mensuel du pétrole brut, (USD par baril) : 1986 - 2006.

Les différentes statistiques de la série étudiée sont présentés dans la table Tab.3.1 suivante. Nous remarquons par exemple que la $\kappa = 3.2281 > 0$, la distribution des rendements est donc leptokurtiques.

Statistique	Valeur
Minimum	-2.7417
Moyenne	0.0044
Maximum	4.8450
kurtosis(κ)	3.2281
skewness(s)	-0.0552

TAB. 3.1 – Statistiques de la série des rendements mensuel du pétrole brut

3.2 Tests et validation du modèle

3.2.1 Test de l'hypothèse de normalité

Les premières études sur la loi de probabilité des **rendements** ont été basées sur la loi normale. L'objet de cette section est donc de tester cette hypothèse de normalité soit qu'on l'accepte ou on la rejette. Notons par

$$X_t = \log\left(\frac{P_t}{P_{t-1}}\right), \quad t = 1, \dots, n = 241$$

la série des rendements, avec P_t est le prix mensuel du pétrole brut.

L'hypothèse de normalité à tester est la suivante :

$$H_0 : X_t \sim N(\mu, \sigma^2).$$

Graphiquement, la figure Fig.3.3 illustre que l'ACF (fonction d'autocorrélation empirique) de la série de rendements est différent de celle des erreurs dont la loi est supposée normale, ce qui signifie que l'hypothèse nulle de normalité est rejetée, donc la série des rendements n'est pas un **bruit blanc**.

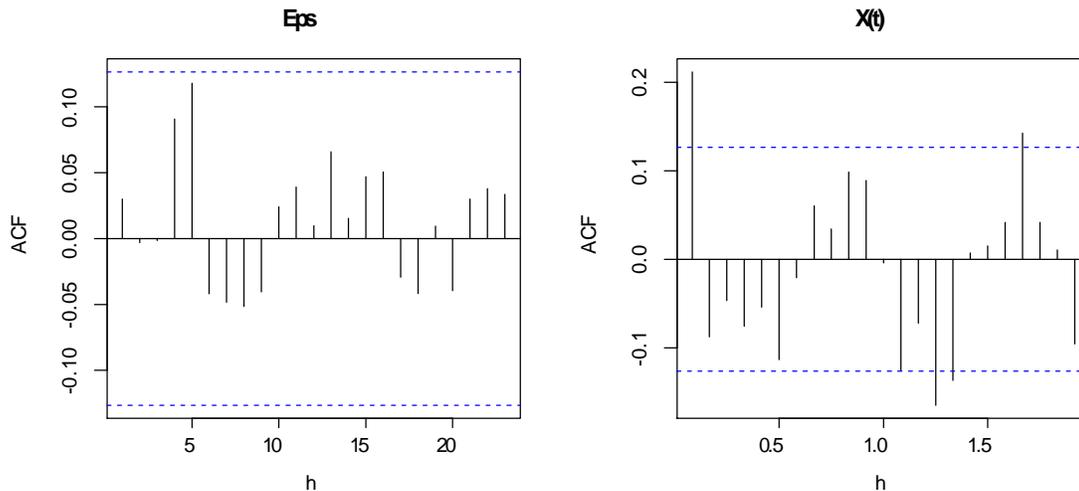


FIG. 3.3 – Autocorrélogramme des rendements et des erreurs.

Théoriquement, pour tester l'hypothèse de normalité de la série $(X_t)_t$, on utilise le test de Jarque-Bera (1984) (Voir :[15] page 95). Ce test statistique permet de vérifier la normalité de la distribution d'un échantillon $(X_t, t = 1, \dots, N)$, il est basé sur les estimateurs empiriques des coefficients d'aplatissement et d'asymétrie. Le test de **Jarque-Bera** (voir [12]) confirme donc, que si la distribution de l'échantillon est normale, alors la statistique **JB** :

$$JB = \frac{n}{6} \left(s^2 + \frac{1}{4} (\kappa - 3)^2 \right)$$

suit asymptotiquement une loi de khi deux à 2 degré de liberté. Ainsi, si $JB \geq \chi_{1-\alpha}^2(2)$, on rejette l'hypothèse H_0 de normalité au seuil α .

Dans notre cas, la kurtosis vaut 3.2281 et le skewness -0.0552. La statistique de **Jarque & Bera** vaut alors 26.633, qui est supérieur au quantile du χ^2 à deux degré de liberté au seuil $\alpha = 0.05$, soit 5.991. L'hypothèse de normalité est donc fortement rejetée au seuil $\alpha = 0.05$. La $p - value = 1.65 \times 10^{-06}$ de test de **Jarque-Bera** inférieur a 0.05, on rejette fortement donc l'hypothèse H_0 de normalité.

3.2.2 Estimation et choix de modèle

Nous suivant la procedure décrite dans le deuxième chapitre, la série de rendements est de type *ARCH* : $X_t := \sigma_t \varepsilon_t$.

L'estimation des paramètres des modèles, est fondée sur la méthode du maximum de vraisemblance *MV*. L'optimisation est effectuée en supposant que la suite des erreurs ε_t est de loi normale : $\varepsilon_t \sim N(0, 1)$. Pour l'identification du modèle, il suffit de comparer le critère d'information d'Akaike (notée AIC) des modèles :

$$ARCH(p), \quad GARCH(p, q) \quad \text{pour } p = 1, 2, 3$$

Le tableau Tab.3.2 donne les différentes valeurs AIC des modèles estimés, on choisit alors le modèle avec l'AIC le plus faible.

/	<i>ARCH</i> (1)	<i>ARCH</i> (2)	<i>ARCH</i> (3)	<i>GARCH</i> (1, 1)	<i>GARCH</i> (2, 1)	<i>GARCH</i> (3, 1)
<i>AIC</i>	-534.30	-534.19	-532.27	-539.56	-536.23	-529.19

TAB. 3.2 – Différentes valeurs AIC des modèles estimés

Dans notre cas, on remarque que le modèle qui réalise l'AIC le plus petit est le modèle *GARCH*(1, 1), dont on a $AIC = -\mathbf{539.56}$. On conclut, que le modèle retenu est le modèle *GARCH*(1, 1), qui admet l'écriture suivante :

$$X_t := \sigma_t \varepsilon_t \quad , \quad \sigma_t^2 := \alpha_0 + \alpha X_{t-1}^2 + \beta \sigma_{t-1}^2,$$

α_0 , α et β , des constantes vérifiant les contraintes suivantes :

$$\alpha_0 > 0 \quad , \quad \alpha \geq 0 \quad , \quad \beta \geq 0 \quad \text{et} \quad \alpha + \beta < 1, \tag{3.1}$$

avec σ_t est \mathcal{F}_{t-1} -mesurable, ε_t est un bruit blanc.

Grâce à la puissance de calcul des ordinateurs actuels, il est possible de minimiser effectivement le **log-vraisemblance** de façon numérique. Pour l'instant, le calcul des estimateurs du modèle sous les contraintes (3.1) sera effectué en utilisant le **logiciel R**, qui réalise ce type de calcul de manière automatique. Le modèle retenu est donc un *GARCH*(1, 1) de paramètres :

$$\alpha_0 = 0.0013 \quad , \quad \alpha = 0.1655 \quad \text{et} \quad \beta = 0.6291$$

On remarque que la somme $\alpha + \beta = 0.7946 < 1$, donc la condition de la stationnarité du modèle est vérifiée.

La volatilité (variance conditionnelle) du modèle est estimée par :

$$\sigma_t^2 = 0.0013 + 0.1655 X_{t-1}^2 + 0.6291 \sigma_{t-1}^2.$$

3.2.3 Test de significativité des paramètres

Pour tester la signification des paramètres du modèle, nous utilisons le test de student, comme suit :

Les hypothèses à tester sont :

$$\begin{cases} H_0 : b = 0 \\ H_1 : b \neq 0 \end{cases}$$

La statistique de test Student est

$$T = \frac{|\hat{b} - b|}{\hat{\sigma}_{\hat{b}}}$$

b est le paramètre à tester, \hat{b} sont estimateur, $\hat{\sigma}_{\hat{b}}$ est l'écart-type de \hat{b} . On accepte H_0 si $|T| < t_{1-\frac{\alpha}{2}}(n-3)$, avec $t_{1-\frac{\alpha}{2}}(n-3)$ le quantile de loi de Student à $(n-3)$ degrés de libertés.

On a les statistiques de Student des paramètres α_0 , α , β respectivement sont :

$$T_{\alpha_0} = 1.81, \quad T_{\alpha} = 2.45, \quad T_{\beta} = 4.24$$

au seuil $\alpha = 0.05$ on a

$$t_{1-\frac{\alpha}{2}}(238) = 1.97.$$

On rejette donc l'hypothèse H_0 pour les deux paramètres α , β du modèle :

$$H_1 : \alpha \neq 0 \quad \text{et} \quad H_1 : \beta \neq 0$$

ce qui signifie leurs signification. Pour le paramètre α_0 , on a la $p - value = 0.06 > 0.05$. Le paramètre α_0 est faiblement significatif (à 93% seulement).

3.3 Prévisions

Soient X_1, X_2, \dots, X_n les n -observations ($n = 241$) de la série des rendements étudiée X_t dont le modèle est

$$X_t := \hat{\sigma}_t \varepsilon_t \quad \text{avec} \quad \hat{\sigma}_t^2 = 0.0013 + 0.1655X_{t-1}^2 + 0.6291\hat{\sigma}_{t-1}^2.$$

La prévision d'une valeur future $X_{n+h} = X_n(h)$ d'horizon h de la série X_t est inconnue, mais elle est estimée par la quantité

$$\hat{X}_{n+h} = \varepsilon_t \sqrt{0.0013 + 0.1655X_{t-1}^2 + 0.6291\hat{\sigma}_{t-1}^2},$$

avec l'intervalle de confiance à $(1 - \alpha)\%$:

$$X_{n+h} = \hat{X}_{n+h} \pm \hat{\sigma}_h t_{1-\frac{\alpha}{2}}(n-3).$$

Dans le tableau Tab.3.3, nous donnons les valeurs prédites de la série des rendements X_t avec l'intervalle de confiance (*I.C*) à 95% des 12 mois de l'année 2007 :

h	\hat{X}_{n+h}	<i>I.C</i>
1	-0.014	[-0.145, 0.148]
2	0.076	[-0.149, 0.141]
3	0.044	[-0.146, 0.150]
4	-0.154	[-0.159, 0.148]
5	0.022	[-0.149, 0.157]
6	0.140	[-0.149, 0.155]
7	0.011	[-0.148, 0.147]
8	0.131	[-0.152, 0.148]
9	0.048	[-0.138, 0.140]
10	-0.062	[-0.149, 0.156]
11	-0.116	[-0.153, 0.153]
12	0.039	[-0.153, 0.153]

TAB. 3.3 – Prévisions des rendements et intervalle de confiance

Conclusion

Ce mémoire donne une idée générale des séries chronologiques, en particulier les modèles ARCH qui modalisent les séries financières présentant une volatilité stochastique.

En conclusion, les séries chronologiques financières présentent souvent des caractéristiques bien connues. Les modèles ARCH capturent avec succès ces propriétés, cela a conduit à l'utilisation de distributions non normales et des modèles non linéaires pour une bonne modélisation de la volatilité stochastique.

Depuis que Bollerslev a introduit le modèle GARCH, de nouvelles extensions ont été proposés, par exemple le GARCH exponentiel (noté EGARCH), le GARCH avec seuil (noté TGARCH) et autres, avec différentes caractéristiques, avantages et inconvénients. Il sera donc intéressant d'étudier ces extensions dans des travaux futurs.

Bibliographie

- [1] Akaike, H. (1973) Information theory and an extension of the maximum like lihood principal. In : Petrov BN, Csake F (eds) 2nd Int Symp Inform Theory. Akademia Kiado, Budapest, pp 267–281.
- [2] Bollerslev, T. (1986) Generalized autoregressive conditional heteroscedasticity. J Econ 31 :307–327.
- [3] Bougerol, P., Picard, N. (1992). Stationarity of *GARCH* processes and of some nonnegative time series. Journal of econometrics, 52(1-2), 115-127.
- [4] Boulmelh, Z., Farour, M. (2017). Estimation des paramètres pour un modèle *GARCH*(p, q). Doctoral dissertation, Universite de Jijel.
- [5] Box, G., Jenkins G. (1970) Time series analysis, forecasting, and control. Holden-Day, Oakland.
- [6] Charpentier, A. (2002). Séries temporelles : théorie et applications. Polycopié de Cours, Université Paris IX Dauphine.
- [7] Cuthbertson, K. (2000). Economie financiere quantitative. Actions, Obligations et taux dechange, De Boeck Université.
- [8] Engle, R. (1982) Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. Econometrica 50 :987–1007.
- [9] Gouriéroux, C., Monfort, A. (1989). Statistique et modeles econometriques, 2 vols. Paris : Economica.
- [10] Gouriéroux, C. (1997). *ARCH* models and financial applications. Springer Science & Business Media.

- [11] Hurlin, C. (2007). Master Econometrie Et Statistique Appliquée (ESA). Université d'Orléans www.univ-orleans.fr/deg/masters/ESA.
- [12] Jarque, C.M., Bera, A.K. (1987). A test for normality of observations and regression residuals". *International Statistical Review*. 55 (2) : 163–172.
- [13] Lars, K. M. (2002). *GARCH*–Modelling : Theoretical Survey, Model Implementation and Robustness Analysis. Doctoral dissertation, Masters Thesis, Royal Institute of Technology (KTH), Stockholm.
- [14] Nelson, D. B. (1990). *ARCH* models as diffusion approximations. *Journal of econometrics*, 45(1-2), 7-38.
- [15] Yahia, D. (2005). Séries temporelles et test d'adéquation pour un modèle *GARCH*(1, 1). Mémoire de magister, Université de Biskra.

- [16] <https://CRAN.R-project.org/package=tseries>.
- [17] <https://CRAN.R-project.org/package=TSA>.
- [18] <https://CRAN.R-project.org/package=fGarch>
- [19] <https://CRAN.R-project.org/package=cvar>
- [20] <https://CRAN.R-project.org/package=datsets>

Annexe : Abréviations et Notations

Les différentes abréviations et notations utilisées tout au long de ce mémoire sont expliquées ci-dessous :

$E(\cdot)$: <i>Espérance mathématique (Moyenne).</i>
$V(\cdot)$ ou $var(\cdot)$: <i>Variance mathématique.</i>
$Cov(X, Y)$: <i>Covariance mathématique du couple (X, Y).</i>
$\gamma_X(\cdot)$: <i>Auto – covariance de X_t.</i>
$\rho_X(\cdot)$: <i>Auto – corrélation de X_t.</i>
\mathcal{F}_t	: <i>Filtration.</i>
$\stackrel{d}{=}$: <i>Par définition.</i>
$N(0, 1)$: <i>Loi normale centrée et réduite.</i>
$(X_t)_t$: <i>Série temporelle.</i>
$E[\cdot \mathcal{F}_t]$: <i>Espérance conditionnelle par rapport à \mathcal{F}_t.</i>
$V[\cdot \mathcal{F}_t]$: <i>Variance conditionnelle par rapport à \mathcal{F}_t.</i>

$I.C$: <i>Intervalle de confiance</i>
AR	: <i>Modèle autoregressif</i>
MA	: <i>Modèle moyenne mobile</i>
$ARMA$: <i>Modèle auto – régressif moyenne mobile</i>
$ARIMA$: <i>Modèle ARMA – Intégré</i>
$ARCH$: <i>Modèle Autorégressifs conditionnellement hétéroscédastiques</i>
$GARCH$: <i>Modèle Autorégressifs conditionnellement hétéroscédastiques Généralisé</i>
κ	: <i>Kurtosis.</i>
s	: <i>skewness.</i>
AIC	: <i>Critère d'information d'Akaike.</i>
JB	: <i>Statistique de Jarque – Bera</i>
\hat{X}_{t+h}	: <i>Modèle de prévision</i>
P_t	: <i>Prix d'une action</i>
R_t	: <i>Rendement</i>

ملخص

نقدم في هذه المذكرة لمحة عامة حول التقلب العشوائي والنماذج ARCH، والتي تستخدم على نطاق واسع لنمذجة وتنبؤ التقلب المشروط للسلاسل الزمنية المالية. حيث يجد هذا التقلب العشوائي تطبيقات عديدة في نمذجة العوائد المالية، تقييم الخيارات، إدارة المخاطر والتنبؤ بالتقلبات. كما نقدم أيضا دراسة تطبيقية باستخدام معطيات حقيقية للسلسلة المالية لتطور أسعار النفط.

Résumé

Dans ce mémoire, nous donnons un aperçu sur la volatilité stochastique et les modèles ARCH, qui sont largement utilisés pour modéliser et prédire la volatilité conditionnelle des séries chronologiques financières. Cette volatilité stochastique trouve de nombreuses applications dans la modélisation des rendements financiers, l'évaluation des options, la gestion des risques et la prévision de la volatilité. Nous présentons également une étude empirique utilisant des données réelles de la série financière de l'évolution des prix du pétrole.

Abstract

In this memory, we provide an overview on stochastic volatility and ARCH models, which are widely used to model and predict the conditional volatility of financial time series. This stochastic volatility finds many applications in modeling financial returns, option pricing, risk management, and volatility forecasting. We also present an empirical study using real data from the financial time series of the evolution of oil prices.