

Mémoire longue des rendements journaliers du S&P500

Ronna Fitoussi & Paul Boquant
Un mémoire présenté pour le diplôme de
M1 MBFA : ARM



**FACULTÉ
DE SCIENCES ÉCONOMIQUES
ET DE GESTION**

*Département
d'économie*

Sommaire

1	Remerciements	2
2	Résumé	2
3	Partie théorique	3
4	Logiciels	7
5	Introduction	8
6	Les données	8
7	Analyse de l'autocorrélation du rendement	14
8	Sensibilité de la structure d'autocorrélation	22
9	Conclusion	26
10	Annexes	27
11	Sources (à rajouter)	30

Remerciements

Notre binôme tient à remercier Monsieur François Legendre, Professeur et Directeur du Département d'économie à l'Université Paris-Est Créteil. Son intense volonté de transmission et sa capacité à partager curiosité et connaissances ont été les moteurs de notre projet. Sans son aide constante en cette période compliquée nous n'aurions pu mener à bien ce mémoire.

Résumé

La dépendance à long terme (LRD), également appelée mémoire longue, est un phénomène qui peut survenir lors de l'analyse de données spatiales ou de séries chronologiques. Il se rapporte au taux de décroissance de la dépendance statistique de deux points avec un intervalle de temps croissant, voici quelques phénomènes illustrant sa présence sur les marchés financiers :

Excès de volatilité : de nombreuses études empiriques soulignent l'importante difficulté de justification des variabilités observées des rendements des actifs financiers par des variations des variables économiques "fondamentales". L'apparition de rendements importants ne s'explique pas toujours par l'arrivée de nouvelles informations sur le marché.

Les queues lourdes (ou épaisses) : la distribution des rendements présente une queue lourde avec un excès de kurtosis positif.

Absence d'autocorrélation dans des rendements : les autocorrélations (linéaires) des rendements des actifs financiers sont souvent insignifiantes, à l'exception de très petites échelles de temps intra journalières (20 minutes) où les effets de microstructure entrent en jeu.

Les clusters de volatilité: Comme le note Mandelbrot : "Les grands changements ont tendance à être suivis de grands changements, les petits changements ont tendance à être suivis de petits changements" . Bien que les rendements eux-mêmes ne soient pas corrélés, les rendements absolus $|rt|$ ou leurs carrés affichent une fonction d'autocorrélation positive, significative, et qui se dégrade lentement.

Nous allons étudier la propriété de "mémoire longue" des rendements boursiers. Nous constatons que la corrélation des rendements absolus est plus importante que celle des rendements. La transformation puissance des ren-

dements $|rt|^d$ présente également une autocorrélation assez élevée pour de longs lags (retards). Il est possible de définir $|rt|^d$ comme étant de "mémoire longue", cette propriété est renforcée à mesure que d est autour de 1. Ces résultats semblent aller à l'encontre des spécifications du type ARCH, basés sur des rendements au carré.

Partie théorique

Notre étude traitant de la mémoire longue des rendements boursiers journaliers, nous allons au sein de cette partie rappeler et préciser les différentes théories de la littérature financière classique. Le point de départ de ce projet est l'Hypothèse des marchés financiers efficients d'Eugène Fama.

"Un marché dans lequel les prix reflètent totalement et constamment toute l'information disponible est appelé **marché efficient**".

Ainsi, dès qu'une information concernant un actif est dévoilée, son prix est modifié à la hausse ou à la baisse selon la teneur de cette information. Nous pouvons définir l'efficience comme : l'aptitude qu'a le marché à rétablir l'équilibre, sur la base de prix fiables permettant aux investisseurs de prendre des décisions en fonction de l'information disponible en temps réel. Fama décrit 3 formes d'efficience du marché efficient :

- Faible efficience : dépend de l'historique des cours.
- Efficience semi-forte : informations publiques entourant l'entreprise qui enrichissent l'ensemble des informations.
- Efficience forte : la totalité des informations précédentes ainsi que les informations "privées".

La théorie des marchés efficients fait l'objet de critiques, la finance comportementale, application de la psychologie à la finance vient souligner la présence de nombreux biais chez l'agent supposé rationnel. Ces biais peuvent être assimilés à des "anomalies" :

- Biais cognitifs, liés à la compréhension, à la mémoire.
- Emotionnels (peurs, envies, admirations, répulsions, fierté).
- Individuels ou collectifs (mimétismes de groupe ou de foule).

- Prophéties auto-réalisatrices (un investisseur entend dire que le cours d'une action va augmenter : il achète et d'autres en font autant si bien que le cours augmente effectivement, la prophétie s'est réalisée).

Le comportement mimétique

Le "mimétisme" dans le cadre des marchés financiers peut-être défini comme un ensemble de comportements individuels présentant des corrélations. Dans le processus d'investissement, les gérants de portefeuilles sont encouragés à suivre les "convictions" de leurs pairs. En effet les gérants doivent avoir des "convictions" (arguments justifiant l'orientation des portefeuilles). Lorsque le gestionnaire ne dégage pas de conviction, il peut alors soit rester neutre sur le marché (dans ce cas son rôle de gérant et les frais inhérents à la gestion deviennent illégitimes aux yeux du détenteur, puisque celui-ci pourrait souscrire à un ETF pour répliquer l'indice tout en limitant les frais), la deuxième option serait de suivre l'avis général de ses collègues (il est plus facile de se tromper en groupe que seul). Le comportement du gérant est également biaisé par la structure de rémunération et les modalités de gestion. Une grande part fixe et un rejet du risque des détenteurs n'invitent ni à une prise de risque, ni un comportement solitaire face au marché. A l'inverse un gérant intéressé à la performance de ses fonds tentera de dégager de l'alpha et n'hésitera pas à aller contre le marché.

Alpha : est un indicateur essentiel du couple rendement/risque des placements, il permet de calculer le surplus de performance dégagé par un gérant en fonction de son niveau de prise de risque. Ce qui le rend par exemple très utile au moment de choisir un fonds.

$$\alpha_p = \bar{r}_p - [r_f + \beta_p(\bar{r}_m - r_f)]$$

- α_p : l'alpha de Jensen
- \bar{r}_p : la rentabilité espérée du portefeuille
- r_f : le taux sans risque
- β_p : le Beta du portefeuille
- \bar{r}_m : la rentabilité espérée du marché de l'actif

La définition d'Eugène Fama suppose une information accessible rapidement et facilement, or il n'en est rien, chaque professionnel sur les marchés financiers connaît la place, le rôle des données. Celles-ci sont difficilement

accessibles, souvent chers, et non-traitées, parfois bloquées sur le logiciel du *data provider* (fournisseur des données). La finance de marché moderne connaît une utilisation massive des données, en résulte une plus grande volatilité du fait de la vitesse d'accès à l'information.

Une anomalie de marché consiste en une variation du cours sur le marché boursier qui diffère du cours prévisionnel. Elles peuvent être rapidement intégrées par le marché et donc disparaître promptement ou demeurer.

L'arbitrage existera tant que l'opportunité d'arbitrage ne sera visible d'aucun agent. De nombreux effets sont connus des professionnels, nous allons citer brièvement quelques-uns :

- L'effet lundi : tendance des cours du marché à clôturer en baisse le lundi, par rapport au vendredi précédent.
- Tournant du mois : tendance haussière suivie par les cours lors du dernier jour de chaque mois
- L'effet janvier : tendance haussière des volumes boursiers suivie d'une augmentation des cours des actions durant la dernière semaine de décembre et les premières semaines de janvier.
- L'effet vacances : anomalie liée au calendrier qui décrit la tendance du marché boursier à générer un gain le jour de négociation précédant un jour férié.
- L'effet Momentum : basé sur une analyse technique, les "gagnants" récents du marché boursier ont de grandes chances de continuer à s'imposer face aux "perdants".

Nous avons vu avec l'hypothèse des marchés financiers efficients que le prix d'un actif intègre toute l'information privée comme publique. Seulement la présence de mémoire longue vient faire dévier notre compréhension de la volatilité boursière. Désormais nous devons raisonner avec la mémoire de la volatilité passée. En effet, la volatilité des chocs demeure et persiste dans l'information qui détermine les prix boursiers.

Les queues épaisses : Les marchés financiers fluctuent de manière imprédictible. En s'intéressant à la distribution des rendements journaliers nous pouvons voir qu'en grande majorité, la fluctuation est comprise entre -2% et 2%. L'histogramme de r_t peut apparaître à nos yeux comme une distribution gaussienne. Il n'en est rien, la distribution gaussienne est une distribution

qui connaît très peu d'évènements extrêmes. Le seuil 3% est présent 653 fois, celui de 5% l'est 160 fois. Il y a donc eu 653 jours sur 23 195 soit presque 3% de toute la distribution. Si les fluctuations étaient gaussiennes nous n'aurions jamais de fluctuations de l'ordre de 5% or on en recense 1.6 par an.

Forte volatilité : la théorie de l'efficiencia des marchés financiers (EMH) propose que le prix d'une action n'est autre que la valeur actuelle des dividendes futurs anticipés. Cependant les cours boursiers et la valeur fondamentale n'évoluent pas toujours dans le même sens (Tesla, Nikola Motor). Nicolas Motor est un constructeur qui n'a jamais livré de camion, sa capitalisation boursière est aujourd'hui plus élevée que celle de Ford. LeRoy et Porter (1981) démontrent que les cours du S&P 500 sont excessivement volatiles en comparaison à la fluctuation de leur valeur fondamentale.

La volatilité mesure la dispersion d'une variable autour de sa moyenne. En statistique et en théorie des probabilités, la variance est une mesure de la dispersion des valeurs d'un échantillon ou d'une distribution de probabilité.

$$\sigma^2 = \frac{\sum_{i=1}^n (x_i - \mu)^2}{n}$$

Une volatilité faible suppose des écarts de cours réduits. Il existe de nombreuses explications comportementalistes de la volatilité nous n'en citerons que quelques-unes.

- L'heuristique de jugement : les investisseurs surestiment la capacité des petits échantillons à refléter les propriétés de la population mère
- Biais d'excès de confiance : permet aussi d'expliquer l'écart des prix à la valeur fondamentale, nous pensons évidemment à Tesla avec son fondateur charismatique Elon Musk. L'agent sur-confiant surestime ses connaissances. Lorsque le marché lui rappelle qu'il peut se tromper il réagit dans le meilleur des cas rapidement, mais souvent de manière excessive, cela peut expliquer une partie des effets de masse sur les marchés.

La théorie des anticipations rationnelles (TAR) (John F. Muth) est utilisée pour décrire de nombreuses situations économiques dans lesquelles le résultat dépend en partie de ce que les personnes s'attendent à obtenir. La TAR ne faisant que prolonger l'hypothèse du consommateur rationnel. L'existence de biais comportementaux soulevés par la finance comportementale sont irrationnels au regard de la théorie économique standard. Certains auteurs

de la finance comportementale remettent en cause la thèse d'efficience des marchés. Pour terminer, certains comportements irrationnels peuvent, dans le cadre développé par Fama devenir rationnels (exemple : mimétisme). La finance comportementale affina les hypothèses classiques pour développer des méthodes de compréhension des comportements financiers. En cela elle n'est pas en rupture avec la littérature classique. Nous pouvons l'envisager comme une réécriture du modèle de rationalité.

Logiciels

Avec ce mémoire nous cherchons à exposer notre capacité à manier l'outil informatique dans ce qu'il a de plus performant à nous offrir ; Ainsi nous ne nous sommes pas limités à un seul logiciel statistique ou langage. L'utilisation de SAS et R résulte des modalités d'évaluation de ce projet cependant, Python, Matlab et Excel furent eux aussi utilisés. Python est un incontournable du data wrangling (traitement de données), le module Matplotlib permet de réaliser de remarquables graphiques, seulement un problème de stabilité lié au poids conséquent des données survint et remit en question l'utilisation du langage comme outil de visualisation. Cette même difficulté se manifesta sur Excel, les 23 000 lignes du S&P500 imposaient un standard plus stable. C'est ainsi que Matlab entra en jeu, forte d'une solide réputation, l'entreprise MathWorks offrit la possibilité aux étudiants d'utiliser gratuitement son environnement pendant la période de confinement. L'interface étant claire et moderne, l'importation des données y est rapide et facile c'est donc sur ce logiciel que nous avons construit nos graphiques.

Introduction

Si r_t est le rendement d'un actif spéculatif tel qu'une obligation ou une action, nous considérerons les propriétés temporelles des fonctions $|r_t|^d$ pour des valeurs positives de d . Il est normal que les rendements eux-mêmes contiennent une légère corrélation en série, en accord avec l'hypothèse des marchés efficients. Toutefois, Taylor (1986) a constaté que $|r_t|$ présente une corrélation positive significative sur de longs retards. Cette propriété est étudiée sur de longues séries de cours boursiers journaliers. Il est possible de définir $|r_t|^d$ comme étant de "mémoire longue" avec une assez forte autocorrélation pour les longs retards. Autre fait empirique, cette propriété est plus forte pour $d = 1$ ou proche de 1 par rapport à des valeurs positives plus grandes ou plus petites que d . Ces résultats semblent s'opposer aux spécifications du type ARCH basées sur des rendements au carré. Le papier vérifie si les différentes classes de modèles sont compatibles avec cette observation. Une nouvelle classe générale de modèle est alors proposée, qui permet à la puissance δ de l'équation d'hétéroscédasticité d'être estimée à partir des données.

Le reste du document est organisé de la manière suivante : Dans la section 2, nous donnerons une brève description des données que nous utilisons. Dans la section 3, nous effectuerons l'autocorrélation et l'analyse de corrélation croisée. Le modèle spécial de l'autocorrélogramme et corrélogramme croisé des rendements boursiers seront exploités et examinés.

Les données

Avant de présenter brièvement les données de notre projets nous devons mettre en lumière une des problématiques du mémoire. Ce papier publié en 1993 dans le Journal of Empirical Finance a été réalisé par Zhuanxin Ding, Clive W.J. Granger et Robert F. Engle. Les données ont été transférées aux auteurs par William Schwert, professeur de Finance et de Statistiques de l'Université de Rochester dans l'État de New-York, aucune information sur la provenance des données ni les méthodes utilisées pour le data wrangling ne sont accessibles. De même les auteurs ne transmettent aucune information sur leurs propres méthodologies statistiques, logiciels utilisés. Le graphique 3.8 illustre assez bien ce décalage entre nos données et celles proposés par Docteur Schwert à l'équipe de recherche. Notre graphique semble effroyablement brut à côté de celui du papier. Dans ce cas précis nous pensons que le graphique a été filtré. Pris par le temps, si nous avions disposé de plus de

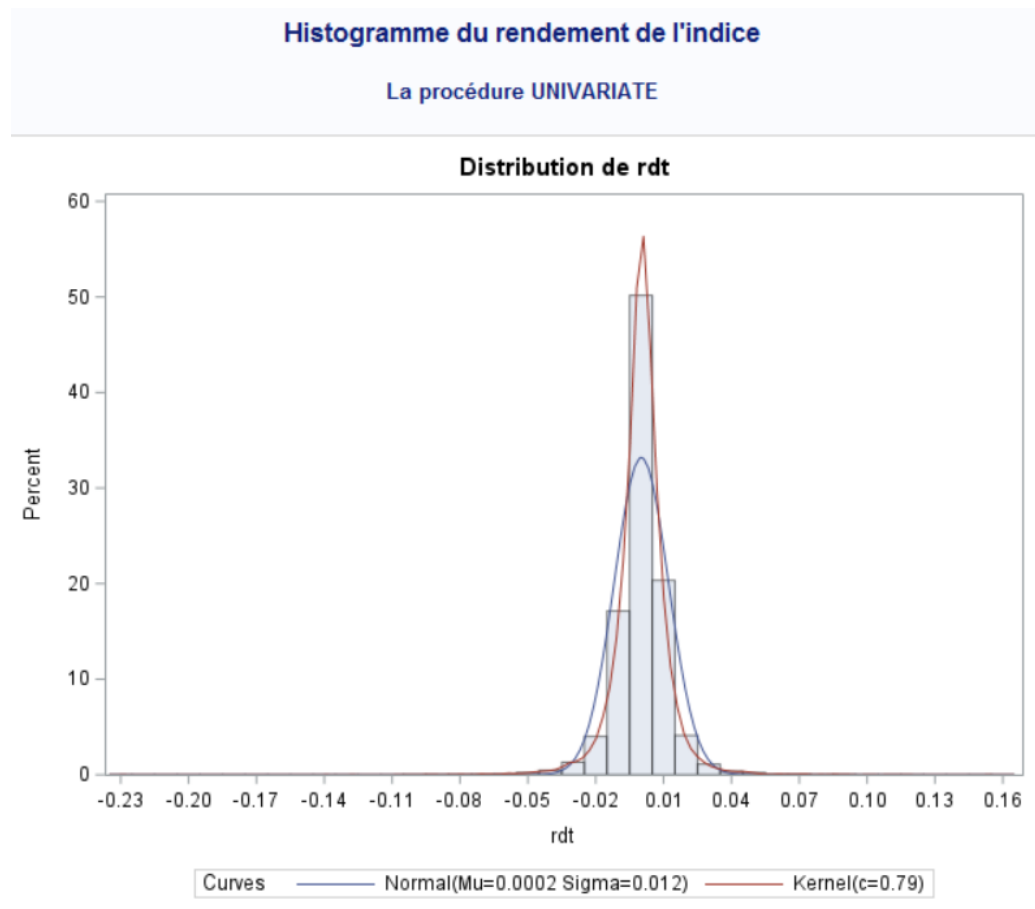
temps nous aurions fait davantage de recherches pour essayer d'approcher cette méthode de filtre sur graphique, nous pensons qu'un filtre de moyenne mobile permettrait de lisser notre graphique, l'approchant ainsi visuellement de celui de Monsieur Schwert.

L'échantillon de données est constitué des prix de clôture du Standard & Poor 500 (S&P500). Il y a en tout 23 195 observations, du 31 décembre 1927 au 5 Mai 2020. Notons p_t comme le prix de l'indice S&P500 au temps t ($t = 0, \dots, 23\ 195$).

$$r_t = \ln p_t - \ln p_{t-1}$$

Soit le rendement composé de l'indice des prix S&P500 au temps t ($t = 0, \dots, 23\ 195$).

Le tableau 2.1 présente les statistiques pour r_t . Nous pouvons voir depuis la table des statistiques descriptives que le kurtosis de r_t est de 19.23, bien loin d'une distribution normale de 3.



Le kurtosis, coefficient d'aplatissement, exprime la caractéristique "queue épaisse" de distribution.

Les graphiques 2.1, 2.2 donnent les figures de p_t , r_t , $|r_t|$. Nous pouvons voir dans les graphiques le mouvement long-terme des p_t , r_t , $|r_t|$ journaliers au cours des 92 dernières années. Il existe une tendance haussière pour p_t , r_t demeure plutôt stable autour de la moyenne $\mu = 0.00022$. Nous pouvons clairement voir l'observation de Mandelbrot 1963 et de Fama (1965) au sein de la série $|r_t|$ soit que les hauts rendements absolus ont une probabilité plus élevée que les petits rendements absolus d'être suivis par un grand rendement absolu.

Concernant la normalité des résidus nous nous intéressons au test de Kolmogorov-Smirnov qui est le test à privilégier lorsque nous avons un nombre d'observations > 2000 donc, $T > 2000$;

Ce test a pour hypothèses :

- H_0 : normalité
- H_A : non-normalité

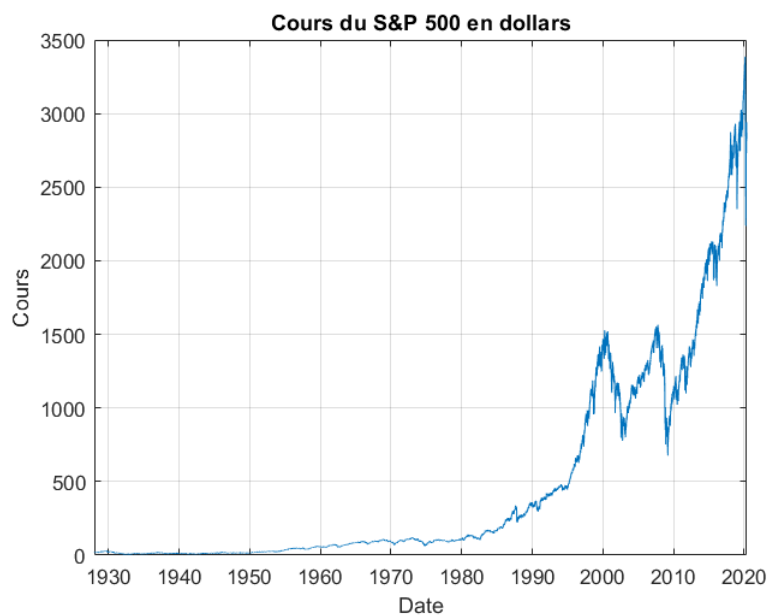
Comme nous pouvons le voir, la p-value $< 0,0100$ donc au seuil de 5% on rejette H_0 et nous pouvons conclure d'une non-normalité.

Test de normalité				
Test	Statistique		p-value	
Kolmogorov-Smirnov	D	0.099	Pr > D	< 0.01
Cramer-von Mises	W-Sq	104.79	Pr > W-Sq	< 0.005
Anderson-Darling	A-Sq	594.831	Pr > A-Sq	< 0.005

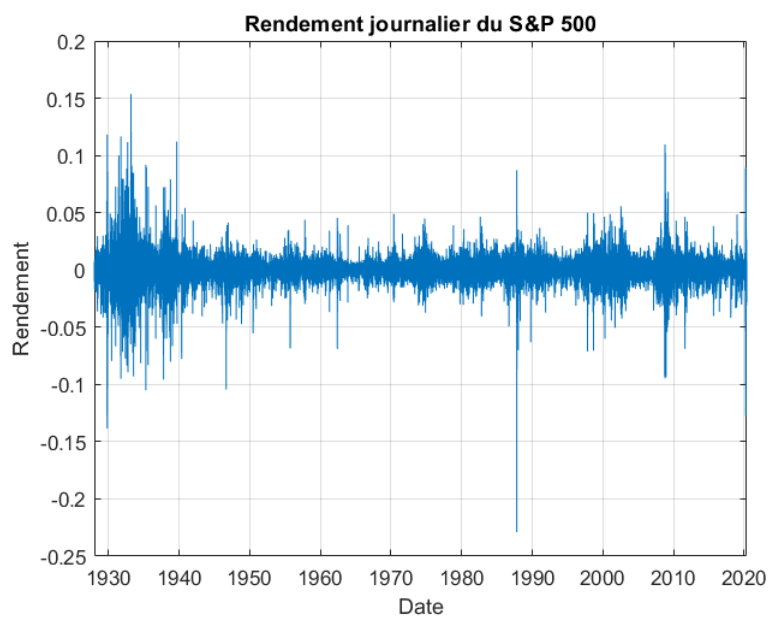
Tableau 2.1 : Statistiques descriptives

Moments			
N	23194	Somme des poids	23194
Moyenne	0.0002	Somme des observations	5.09
Ecart-type	0.012	Variance	0
Skewness	-0.47	Kurtosis	19.22
Somme des carrés non corrigée	3.34	Somme des carrés corrigée	3.34
Coeff Variation	5474.59	Std Error Mean	0

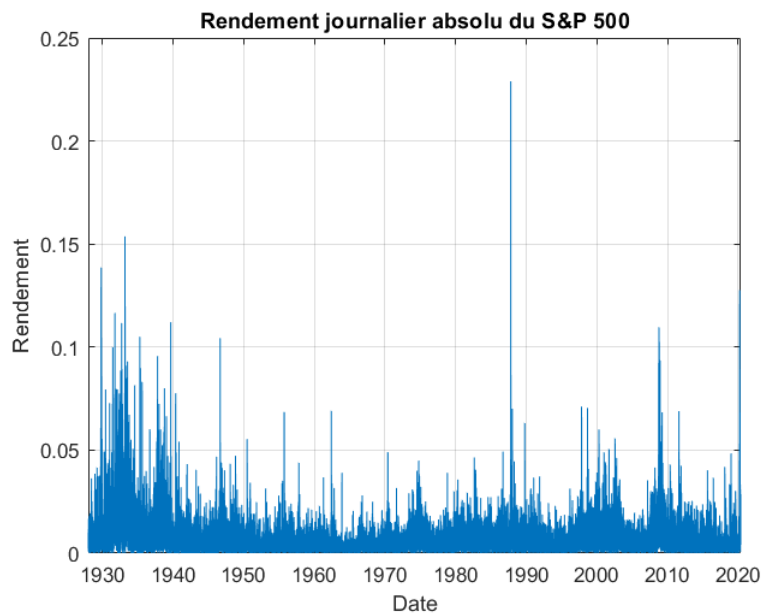
Graphique 2.1



Graphique 2.2



Graphique 2.3



La volatilité est changeante au fil du temps, ce qui suggère l'utilisation d'un modèle adapté à une volatilité fluctuante dans le temps tel que le modèle ARCH. Pendant la Grande Dépression de 1929 ainsi qu'au début des années 1930, la volatilité était très élevée, une pareille volatilité ne se vérifiera que pendant le black Monday de 1987.

L'expression "lundi noir", traduction de l'anglais Black Monday, utilisée parfois pour désigner le 19 octobre 1987, constitue une référence indirecte au Black Thursday, ou jeudi noir, du 24 octobre 1929, première journée du long krach de 1929 de la bourse de New York, qui fit entrer les États-Unis dans la Grande Dépression.

L'origine du krach est avant tout à chercher dans :

- Le doublement de l'endettement global depuis la fin des années 1970, endettement matérialisé par l'afflux énorme de liquidité. L'indice du Dow Jones triple en dix ans.
- Les importantes et brutales fluctuations du dollar au milieu des années 1980 dans le système de changes flottants (adopté depuis mars 1973) et entériné avec les accords de la Jamaïque (1976)

- Un mécanisme d'immunisation des portefeuilles, dit "portfolio insurance" en anglais, très en vogue à l'époque, destiné à réduire les risques mais que sa généralisation transforma en système vicieux
- Les systèmes automatiques d'achats et de ventes d'actions (program trading) relativement récents à l'époque, qui ont amplifié et accéléré le plongeon des cours. Après le krach, Wall Street se dote de coupe-circuits qui se déclenchent automatiquement au-delà d'une certaine baisse de l'indice Dow Jones.

Notons que la crise des Subprimes mais aussi l'épisode récent de covid-19 sont visibles sur le graphique. La Grande Dépression, à l'inverse du black Monday entraîna une longue période de haute volatilité. En dehors de ces périodes spécifiques de crises le S&P500 est relativement stable.

Nous allons faire un test de GARCH, plus spécialement un GARCH (1,1)

Estimer un modèle GARCH (1,1) équivaut à estimer un modèle ARCH infini. Avec pour contrainte que les paramètres sur le modèle ARCH infini sont décroissants. A l'aide de la procédure autoreg nous obtenons les estimations suivantes ;

Paramètres estimés					
Variable	DDL	Estimation	Erreur type	Valeur du test t	Approx $P_r > t $
Intercept	1	0.0004	0.0000591	8.38	< 0.0001
$ARCH_0$	1	1.06E-06	1.52E-07	6.98	< 0.0001
$ARCH_1$	1	0.0987	0.007691	12.83	< 0.0001
$GARCH_1$	1	0.8974	0.006944	129.24	< 0.0001

Résultats :

- $ARCH_0 = 1.0587E - 6$
- $ARCH_1 = 0.0987$
- $GARCH_1 = 0.8974$

$ARCH_0$ représente la constante dans notre écriture de modèle, Garch représente le δ .

Quand on veut interpréter un paramètre avec de l'incertitude, on fait la dériver. Pour trouver γ_1 ($= ARCH_1$) on fait la dérivée de $\frac{h_t}{u_{t-1}^2}$.

Plus γ_1 sera élevé, plus la volatilité sera capable de décrocher quand on aura des news (le prix varie, ce qui correspond à des personnes qui achètent et qui vendent) entraîne une variation de prix. Ces paramètres que nous avons estimé sont tous significatifs au seuil de 5% car $p\text{-value} < 5\%$ on rejette donc l'hypothèse nulle.

Sous l'hypothèse de normalité des résidus :

- Le risque est que les paramètres estimés (au niveau des écarts types) soient mal estimés.

COVEST = HESSIAN | QML, permet de spécifier la forme de la matrice de variance covariance asymptotique des paramètres. Sous l'hypothèse de normalité, cette matrice est construite de façon standard par inversion de la matrice hessienne (estimateur du MV). En revanche dans le cas du PMV, l'option QML (quasi maximum likelihood) permet de construire cette matrice conformément aux formulations du programme utilisé.

Lorsque l'on rajoute "covest" la seule chose qui change sera la façon de calculer les écarts types.

Analyse de l'autocorrélation du rendement

Il est communément accepté que les rendements boursiers contiennent eux-mêmes un peu d'autocorrélation [Fama (1970), Taylor (1986)], cette caractéristique est en accord avec l'hypothèse des marchés efficients. Cependant ce fait empirique n'implique pas nécessairement que les rendements soient indépendamment distribués de manière identique comme le supposent de nombreux modèles théoriques. Il est possible que la série ne soit pas autocorrélée mais qu'elle soit dépendante. Les données boursières le sont d'autant plus si le marché suit l'hypothèse des marchés efficients ainsi le cours de l'action devrait changer avec l'arrivée d'information. Si l'information arrive par lots, la répartition du prochain rendement dépendra des précédents bien qu'ils ne soient pas corrélés. En 1986 Taylor étudia les corrélations des rendements transformés pour 40 séries et conclut que le processus se caractérise par une corrélation sensiblement plus élevée des rendements carrés et absolus relativement aux rendements eux-mêmes.

Kariya et al (1990) ont obtenu un résultat similaire en étudiant les cours des actions japonaises. Pour renforcer ce développement nous examinerons

l'autocorrélation de r_t et $|r_t|^d$ pour un d positif dans cette section, où r_t est le rendement du S&P500.

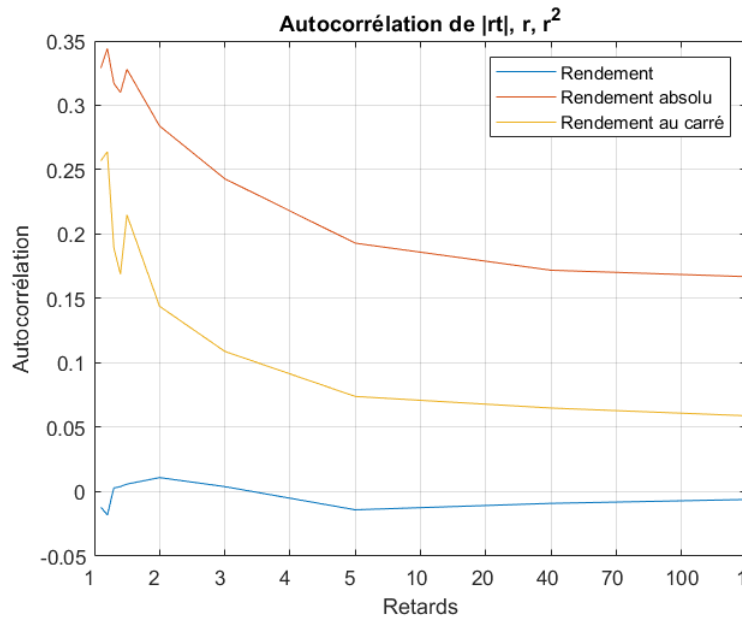
Le graphique 3.1 donne les autocorrélations de l'échantillon de r_t , $|r_t|$ et r_t^2 pour les retards de 1 à 5 mais aussi 10, 20, 40, 70, 100. Nous traçons les autocorrélogrammes de r_t , $|r_t|$ et r_t^2 pour les retards de 1 à 100 au sein du graphique 3.1. Les lignes pointillées montrent les $\pm \frac{1.96}{\sqrt{T}} = 0.0128$. Barlett (1946) prouva que si r_t est un indépendant et identiquement distribué alors l'autocorrélation de l'échantillon ρ est approximativement $N(0, 1/T)$. Dans le graphique 3.1 environ un quart des autocorrélations de l'échantillon avec un retard de 100 se situent en dehors de l'intervalle de confiance de 95% pour un processus indépendant et identiquement distribué. La première autocorrélation du retard est de -0.012.

De nombreux autres chercheurs [voir Fama (1976), Taylor (1986), Hamao et al. (1990)] ont également constaté que la plupart des séries de rendements boursiers présentent une très faible autocorrélation positive du premier ordre.

Tableau 3.1 Autocorrélation de r_t

Variables	Lag1	2	3	4	5	10	20	40	70	100
r_t	-0,012	-0,018	0,003	0,004	0,006	0,011	0,004	-0,014	-0,009	-0,006
$ r_t $	0,329	0,344	0,317	0,310	0,328	0,284	0,243	0,193	0,172	0,167
Rt^2	0,257	0,264	0,189	0,169	0,215	0,144	0,109	0,074	0,065	0,059

Graphique 3.1



L'autocorrélation négative du premier et deuxième ordre suggèrent qu'un rendement positif en t augmente la probabilité d'avoir un rendement négatif pour le prochain retard. Une autocorrélation négative change le sens de l'influence de t sur $t+1$. L'autocorrélation du deuxième retard (-0,018) est significativement négative, ce qui soutient le comportement dit de "retour à la moyenne" des rendements boursiers. Cela soutient l'idée que la série de rendements boursiers de S&P500 n'est pas une réalisation d'un processus d'i.i.d (indépendant et identiquement distribué).

En outre, si r_t est un processus indépendant et identiquement distribué, alors toute transformation du rendement est aussi un processus indépendant et identiquement distribué, même chose pour $|r_t|$ et r_t^2 . L'erreur type de l'autocorrélation de l'échantillon $|r_t|$ sera $\frac{1}{\sqrt{T}} = 0.0066$.

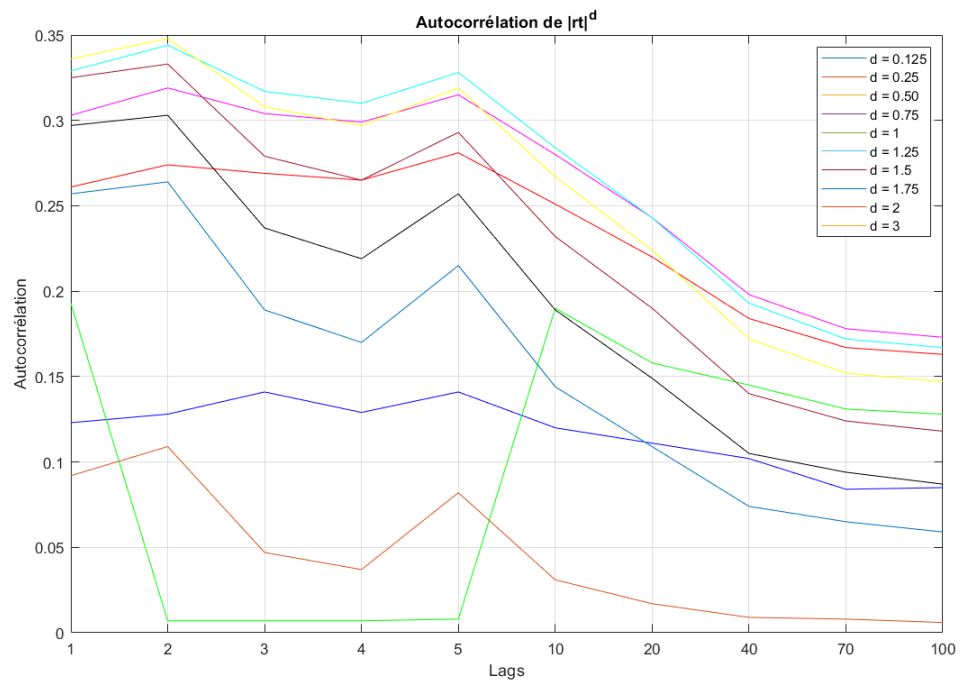
D'après la figure 3.1, les autocorrélations pour l'échantillon des rendements absolus sont majoritairement supérieurs aux autocorrélations pour les rendements au carré à chaque retard jusqu'au centième. Il est clair que le processus de rendement journalier du S&P500 n'est pas un processus indépendant et identiquement distribué.

MANQUE paragraphe (sur les bases)

3.2 Autocorrélations de $|r_t|^d$ cf. annexes pour SAS et R

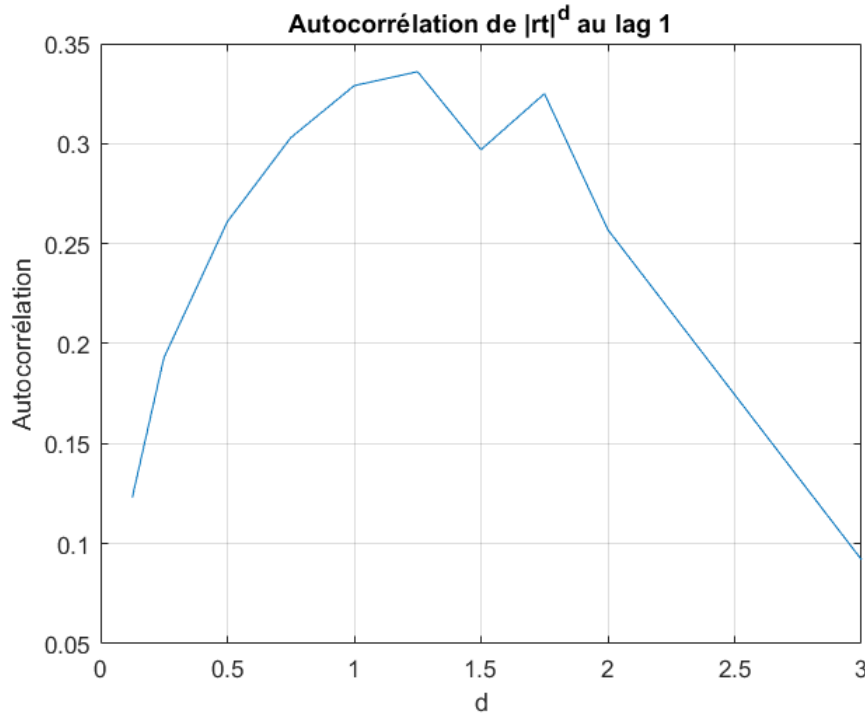
d	Lag1	2	3	4	5	10	20	40	70	100
0.125	0.123	0.128	0.141	0.129	0.141	0.120	0.111	0.102	0.084	0.085
0.25	0.193	0.017	0.017	0.017	0.209	0.190	0.168	0.145	0.131	0.128
0.5	0.261	0.274	0.269	0.265	0.21	0.251	0.220	0.184	0.167	0.163
0.75	0.303	0.319	0.304	0.299	0.315	0.270	0.243	0.198	0.178	0.173
1	0.329	0.334	0.317	0.300	0.328	0.284	0.243	0.199	0.176	0.177
1.25	0.336	0.348	0.308	0.297	0.319	0.267	0.224	0.177	0.157	0.157
1.5	0.309	0.323	0.271	0.259	0.277	0.219	0.189	0.150	0.129	0.128
1.75	0.280	0.289	0.229	0.211	0.243	0.181	0.150	0.110	0.101	0.091
2	0.247	0.254	0.189	0.160	0.205	0.134	0.109	0.076	0.065	0.059
3	0.082	0.109	0.047	0.037	0.082	0.031	0.017	0.009	0.008	0.006

Graphique 3.2/3.3



Graphique 3.4

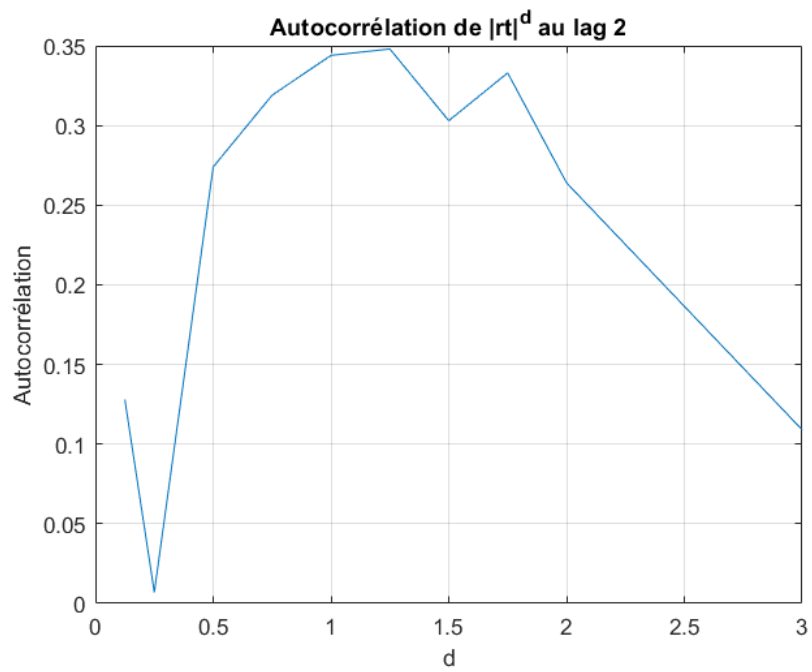
Erratum : les graphiques 3.4/3.5/3.6/3.7 possèdent des mauvaises informations au niveau de l'axe x, il s'agit bien de (0,1,2,3,4,5), nous améliorerons ces graphiques dans le pdf final.



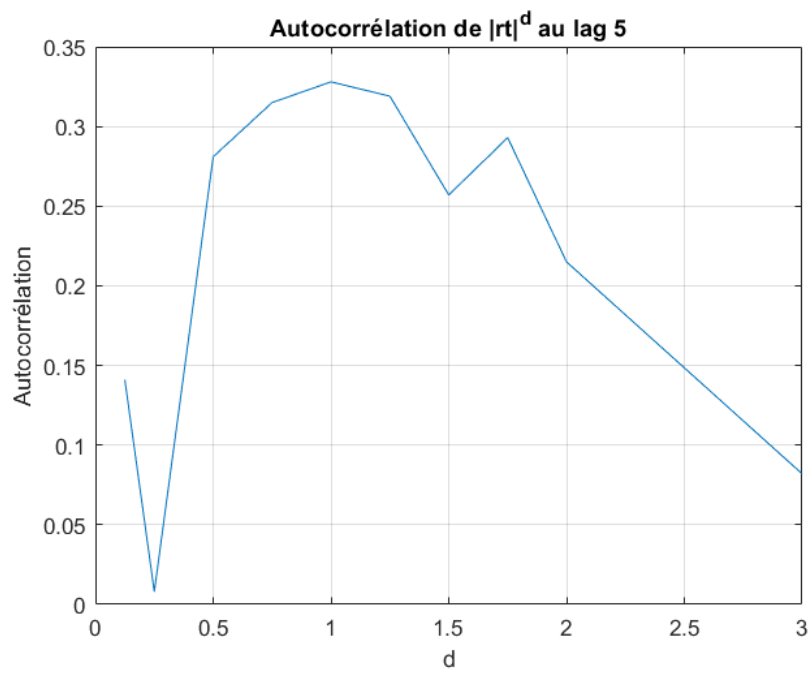
La conclusion la plus intéressante de l'autocorrélogramme est que $|r_t|^d$ présente la plus grande autocorrélation au moins jusqu'à un retard de 100 lorsque $d = 1$ ou est proche de 1. L'autocorrélation se réduit de façon presque monotone lorsque d s'éloigne de 1.

Pour illustrer cela plus clairement, nous calculons les autocorrélations de l'échantillon $\rho(d)$ en fonction de $d, d > 0$ pour un $\tau = 1, 2, 5, 10$ et en prenant $d = 0,125, 0,130, \dots, 1,745, 1,750, 2, 2,25, \dots, 4,75, 5$. Les graphiques 3.4, 3.5, 3.6 et 3.7 donnent les tracés de $\rho_t(d)$ calculé au $\tau = 1, 2, 5, 10$. Il ressort clairement de ces chiffres que l'autocorrélation $\rho(d)$ est une fonction lisse de d . Il existe un point \tilde{d} entre 2 et 3 tel que lorsque $d < \tilde{d}$, $\rho_t(d)$ est une fonction concave et lorsque $d > \tilde{d}$, $\rho_t(d)$ est une fonction convexe de d . Il existe un point unique d^* autour de 1 tel que $\rho_t(d)$ atteint son maximum en ce point, $\rho_t(d^*) > \rho_t(d)$ pour $d \neq d^*$.

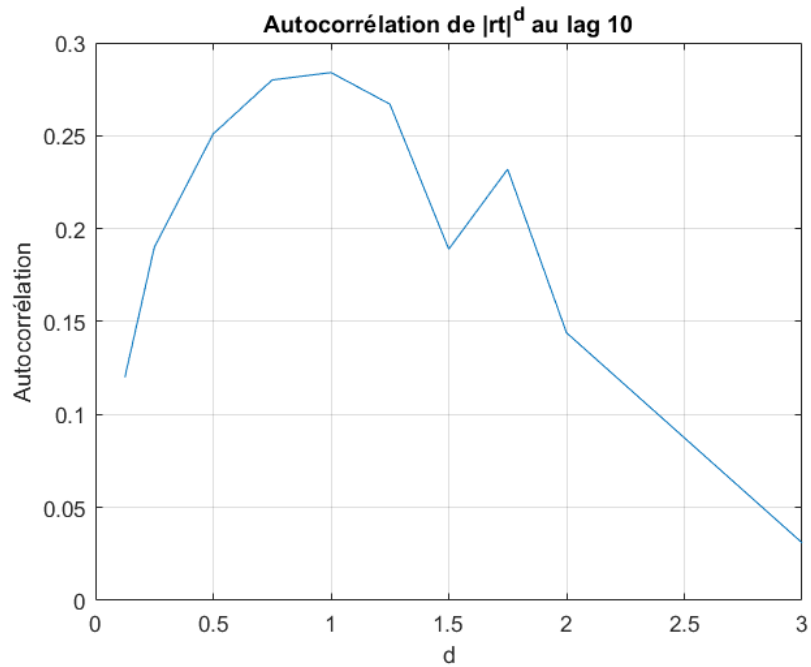
Graphique 3.5



Graphique 3.6



Graphique 3.7

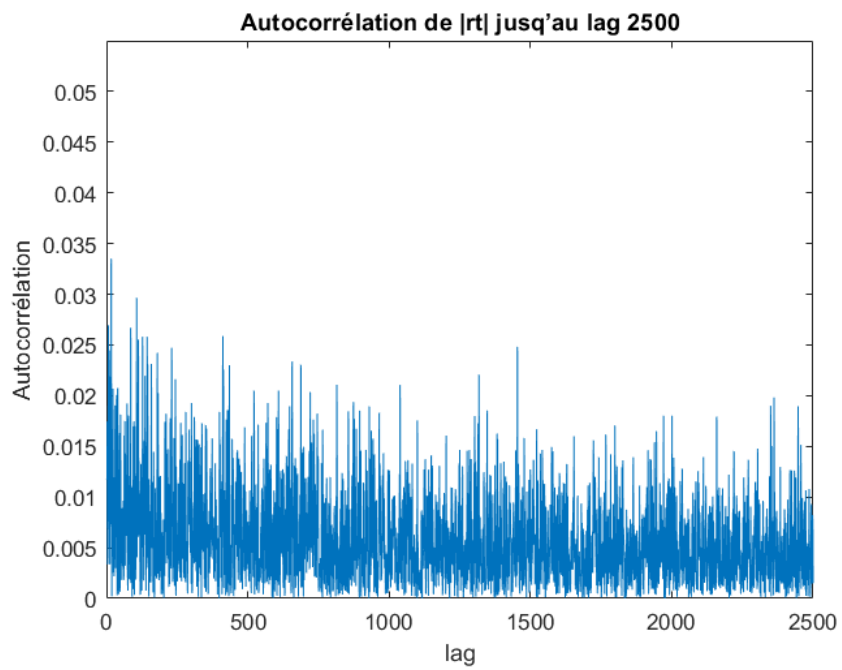


$|r_t|^d$ a une autocorrélation positive sur un délai beaucoup plus long que 100. Le tableau 3.3 indique que les retards (τ^*), retards auxquels la première autocorrélation négative de $|r_t|^d$ se produit pour divers d . On peut voir dans le tableau que dans la plupart des cas, $|r_t|^d$ a des autocorrélations positives sur plus de 2500 retards. Comme il y a environ 250 jours de travail chaque année, les résultats empiriques suggèrent que $|r_t|^d$ présente des autocorrélations positives pendant plus de 10 ans !

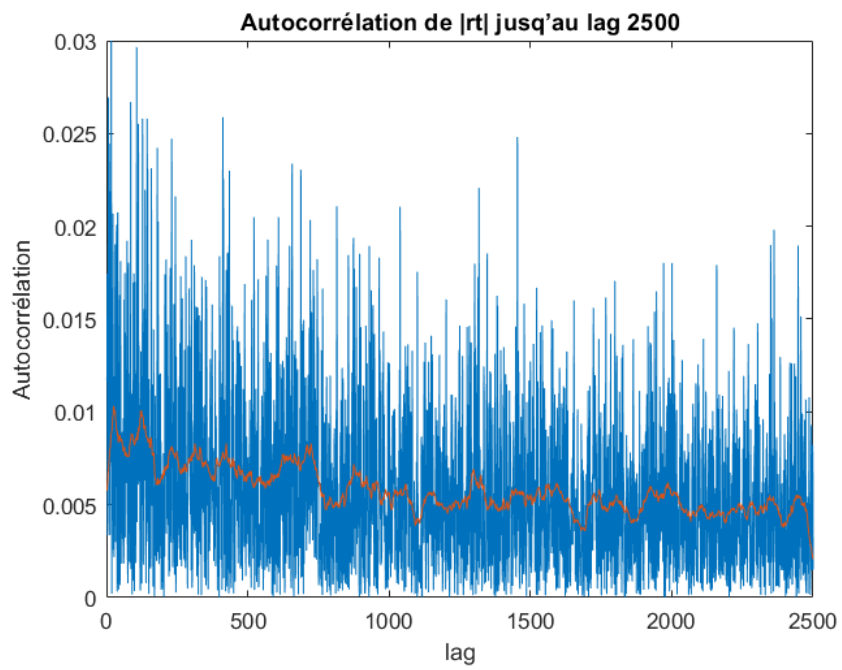
Tableau 3.3 : Retard à partir duquel apparaît la première autocorrélation négative de $|r_t|^d$

d	0.125	0.25	0.5	0.75	1	1.25	1.5	1.75	2	3
r^*	1893	2249	2429	2664	2588	2429	2315	2313	2049	788

Graphique 3.8



Graphique 3.8.1



Comme nous avons pu l'évoquer au sein de la sous-partie données, l'apparence du graphique 3.8 apparaît terriblement brute. Nous avons, pour rapprocher l'apparence de ce graphique à celui des auteurs, créé un *moving average filter* (filtre de moyenne mobile) sous Matlab. L'idée étant d'afficher un trend lisible de l'autocorrélation de $|r_t|$ de 1928-1934. Le code sera au sein de notre dossier en format .txt. Le résultat de ce filtre est bien évidemment la courbe orange du graphique 3.8.1.

Nous choisissons ici $|r_t|$ comme transformation typique de la série du rendement et traçons ses échantillons d'autocorrélations jusqu'au retard 2500 sur la figure 3.8.

Sensibilité de la structure d'autocorrélation

Nous allons étudier plus en détail l'effet de l'agrégation temporelle sur la structure d'autocorrélation. Le tableau 3.4 nous donne les autocorrélations de $|r_{t,5}|^d$ où $|r_{t,5}|^d$ est la moyenne sur 5 jours du rendement.

Où $t = 1, 2, \dots, 3410$ et $t' = 5(t - 1)$. On peut voir que l'agrégation de la moyenne temporelle ne modifie pas la propriété de mémoire longue de la série du rendement absolu. $\rho_t(|r_{t,5}|^d)$ atteint toujours un maximum unique lorsque d est autour de 1 ou 1,25 pour différents retards τ .

Par rapport à la série journalière originale, l'autocorrélation de premier ordre pour $|r_{t,5}|^d$ est bien plus élevée que la deuxième. Bien que la série de rendements agrégée ne soit pas exactement la même que la série hebdomadaire, nous nous attendons à ce qu'un résultat similaire soit obtenu pour les données hebdomadaires.

Il convient également de noter, d'après le graphique 2.2, que la structure de la volatilité diffère considérablement entre l'avant-guerre et l'après-guerre. La période d'avant-guerre (1928-1945) est beaucoup plus volatile que la période d'après-guerre (1946-1986). Il est intéressant d'examiner la structure de la mémoire pour ces deux périodes. [ATTENTION EXAMINER PERIODE 2020]

Tableau 3.4 cf. annexes R

Tableau 3.5 - 1928/1945 cf. annexes SAS & R

d	Lag1	2	3	4	5	10	20	40	70	100
0.125	0.113	0.131	0.160	0.143	0.137	0.111	0.119	0.107	0.097	0.087
0.25	0.203	0.224	0.228	0.217	0.229	0.207	0.186	0.159	0.162	0.143
0.5	0.276	0.294	0.282	0.273	0.291	0.268	0.228	0.187	0.197	0.174
0.75	0.307	0.317	0.298	0.286	0.303	0.276	0.232	0.185	0.195	0.173
1	0.320	0.317	0.294	0.279	0.293	0.262	0.218	0.169	0.177	0.158
1.25	0.322	0.305	0.279	0.259	0.272	0.238	0.195	0.146	0.154	0.134
1.5	0.318	0.287	0.256	0.231	0.246	0.209	0.168	0.121	0.128	0.108
1.75	0.309	0.265	0.231	0.201	0.219	0.179	0.139	0.096	0.102	0.082
2	0.297	0.244	0.205	0.172	0.195	0.152	0.112	0.074	0.079	0.060
3	0.233	0.175	0.118	0.085	0.126	0.073	0.038	0.021	0.020	0.011

Tableau 3.6 – 1946/1986 cf. annexes SAS

d	Lag1	2	3	4	5	10	20	40	70	100
0.125	0.082	0.059	0.055	0.052	0.076	0.059	0.054	0.052	0.028	0.039
0.25	0.128	0.093	0.089	0.096	0.120	0.088	0.084	0.064	0.044	0.055
0.5	0.165	0.125	0.122	0.135	0.155	0.112	0.109	0.075	0.054	0.064
0.75	0.185	0.146	0.142	0.158	0.171	0.124	0.119	0.081	0.055	0.064
1	0.196	0.161	0.152	0.176	0.178	0.129	0.121	0.081	0.049	0.057
1.25	0.198	0.169	0.153	0.190	0.174	0.127	0.114	0.076	0.041	0.048
1.5	0.191	0.169	0.144	0.200	0.159	0.117	0.098	0.065	0.031	0.037
1.75	0.173	0.159	0.126	0.206	0.137	0.100	0.078	0.052	0.020	0.025
2	0.148	0.141	0.102	0.204	0.109	0.079	0.056	0.038	0.012	0.016
3	0.051	0.052	0.024	0.142	0.028	0.016	0.007	0.004	0.000	0.000

Tableau 3.7 - 1987/2020 cf. annexes SAS

d	Lag1	2	3	4	5	10	20	40	70	100
0.125	0.134	0.169	0.189	0.168	0.184	0.162	0.126	0.109	0.090	0.089
0.25	0.162	0.201	0.219	0.203	0.214	0.192	0.150	0.125	0.099	0.097
0.5	0.211	0.265	0.263	0.252	0.264	0.235	0.185	0.143	0.103	0.103
0.75	0.256	0.316	0.295	0.286	0.305	0.261	0.205	0.148	0.096	0.100
1	0.289	0.355	0.309	0.298	0.331	0.267	0.206	0.138	0.080	0.087
1.25	0.303	0.372	0.298	0.283	0.334	0.245	0.186	0.115	0.059	0.068
1.5	0.286	0.358	0.258	0.241	0.309	0.202	0.149	0.083	0.038	0.047
1.75	0.244	0.316	0.199	0.183	0.262	0.149	0.104	0.053	0.020	0.028
2	0.189	0.259	0.140	0.126	0.208	0.099	0.066	0.030	0.010	0.015
3	0.044	0.084	0.021	0.017	0.065	0.014	0.006	0.002	0.000	0.000

Le tableau 3.5 montre les autocorrélations de $|r_t|^d$ pour la période d'avant-guerre (1928-1945).

On constate que l'ampleur de l'autocorrélation pour $|r_t|^d$ est à peu près la même que celle du tableau 3.2 $|r_t|$ présente la plus grande autocorrélation pour les deux premiers retards, puis cette caractéristique devient la plus forte pour $|r_t|^{.75}$ ou $|r_t|^{.5}$.

Le tableau 3.6 donne les autocorrélations de $|r_t|^d$ pour la période d'après-guerre (1946-1986). Il ressort clairement du tableau que cette période est moins volatile, le marché a une mémoire à la fois plus petite et plus courte, en ce sens que les autocorrélations sont plus petites et diminuent plus rapidement. Les autocorrélations ne représentent que deux tiers environ de celles de la période d'avant-guerre.

En comparant les tableaux 3.2, 3.5 et 3.6, nous pouvons probablement dire que la propriété de mémoire longue qui a été trouvée dans l'ensemble de la période de l'échantillon peut être principalement attribuée à la période d'avant-guerre. Le marché a une mémoire forte et longue des grands événements comme la grande dépression de 1929 et le début des années 1930 quand la volatilité était très élevée.

Il convient de noter qu'un processus peut avoir des autocorrélations nulles mais des autocorrélations de carrés supérieures à celles des modules.

Nous voyons donc que les résultats du tableau 3.2 ne se produisent pas nécessairement. Une des explications possibles de la grande autocorrélation positive entre $|r_t|$ et $|r_t + \tau|$ ou $|r_t|^d$ et $|r_t + \tau|^d$ est l'hétéroscédasticité des données, c'est-à-dire que la variance ou la variance conditionnelle évolue dans le temps.

Une famille de modèles de séries chronologiques non linéaires capables de saisir certains aspects de la structure de la volatilité variant dans le temps est le modèle ARCH (AutoRegressive Conditional Heteroskedasticity) de Engle [Engle (1982)]. Dans son cadre original, le modèle ARCH est défini comme un processus de génération de données pour une variable aléatoire qui a une distribution normale conditionnelle, une fonction linéaire de carrés des résidus retardés.

De nombreuses modifications et extensions du modèle ARCH original sont également apparues dans la littérature. Par exemple, afin de saisir la propriété de mémoire longue du processus de variance conditionnelle, Bollerslev (1986) a introduit le modèle GARCH(p, q), qui définit l'équation de variance conditionnelle.

Il est facile de montrer que les r_t , ne sont pas autocorrélés entre eux mais que $|r_t|^d$ l'est. Par conséquent, la distribution de r_t dépend de $r_t - i, i > 0$. Depuis son introduction par Engle (1982), le modèle ARCH a été largement utilisé pour modéliser la volatilité variable dans le temps et la persistance des chocs de volatilité. De nombreux travaux ont également été réalisés, tant sur le plan théorique qu'empirique.

Taylor (1986) a modélisé la fonction d'écart-type conditionnel au lieu de la variance conditionnelle. Schwert (1989), suivant l'argument de Davidian et Carroll (1987), a modélisé l'écart-type conditionnel comme une fonction linéaire des des résidus absolus retardés.

Nous pourrions, à première vue, penser qu'il serait préférable d'utiliser le modèle de Taylor/Schwert plutôt que le GARCH de Bollerslev, puisque le modèle est exprimé en termes de rendements absolus plutôt que de rendements au carré. Mais cette idée n'est pas nécessairement vraie lorsque le modèle est non linéaire.

Conclusion

Au sein de notre mémoire, une propriété "mémoire longue" de la série des rendements boursiers du S&P 500 a fait l'objet d'une analyse. Nous avons constaté que non seulement il existe une corrélation qui est nettement plus importante au niveau des rendements absolus que les rendements eux-mêmes, mais que la transformation du pouvoir des "absolute return" présente également une autocorrélation assez élevée pour les longs délais. Pour faire cette conclusion des rendements, nous avons utilisé deux logiciels statistiques qui sont R et SAS ces logiciels nous ont permis d'obtenir des résultats allant dans le sens de nos hypothèses. Par ailleurs le fait que les rendements absolus aient une corrélation plus importante que les rendements, cela va aux antipodes de la théorie des modèles ARCH qui sous entendent eux, que ce sont les rendements aux carrés qui doivent être étudiés et non les rendements absolus. Notre mémoire a fait l'objet d'une analyse statistique mais nous a également permis une analyse théorique notamment avec la théorie du mimétisme, ou encore celle des anticipations rationnelles qui sont assez présentes dans le domaine de la finance et qui nous ont permis d'avoir un regard plus affuté sur les corrélations possibles concernant nos rendements.

Annexes

Tableaux 3.1 SAS & R (mettre titre)

Variables	Lag1	Lag2	Lag3	Lag4	Lag5	Lag10	Lag20	Lag40	Lag70	Lag100
rt	-0,012	-0,018	0,003	0,004	0,006	0,011	0,004	-0,014	-0,009	-0,006
rt_abs	0,329	0,344	0,317	0,310	0,328	0,284	0,243	0,193	0,172	0,167
rt_2	0,257	0,264	0,189	0,169	0,215	0,144	0,109	0,074	0,065	0,059

	lag 1	2	3	4	5	10	20	40	70	100
rt	-0.01	-0.02	0.00	0.00	0.01	0.01	0.00	-0.01	-0.01	-0.01
rt	0.33	0.34	0.32	0.31	0.33	0.28	0.24	0.19	0.17	0.17
rt ²	0.26	0.26	0.19	0.17	0.22	0.14	0.11	0.07	0.07	0.06

Tableaux 3.2 SAS & R (mettre titre)

	lag 1	2	3	4	5	10	20	40	70	100
0.125	0.12	0.13	0.14	0.13	0.14	0.12	0.11	0.10	0.08	0.09
0.25	0.19	0.20	0.21	0.20	0.21	0.19	0.17	0.15	0.13	0.13
0.5	0.26	0.27	0.27	0.27	0.28	0.25	0.22	0.18	0.17	0.16
0.75	0.30	0.32	0.30	0.30	0.32	0.28	0.24	0.20	0.18	0.17
1	0.33	0.34	0.32	0.31	0.33	0.28	0.24	0.19	0.17	0.17
1.25	0.34	0.35	0.31	0.30	0.32	0.27	0.22	0.17	0.15	0.15
1.5	0.32	0.33	0.28	0.26	0.29	0.23	0.19	0.14	0.12	0.12
1.75	0.30	0.30	0.24	0.22	0.26	0.19	0.15	0.11	0.09	0.09
2	0.26	0.26	0.19	0.17	0.22	0.14	0.11	0.07	0.07	0.06
3	0.09	0.11	0.05	0.04	0.08	0.03	0.02	0.01	0.01	0.01

d	Lag1	Lag2	Lag3	Lag4	Lag5	Lag10	Lag20	Lag40	Lag70	Lag100
0.125	0.123	0.128	0.141	0.129	0.141	0.120	0.111	0.102	0.084	0.085
0.25	0.193	0.017	0.017	0.017	0.209	0.190	0.168	0.145	0.131	0.128
0.5	0.261	0.274	0.269	0.265	0.21	0.251	0.220	0.184	0.167	0.163
0.75	0.303	0.319	0.304	0.299	0.315	0.270	0.243	0.198	0.178	0.173
1	0.329	0.334	0.317	0.300	0.328	0.284	0.243	0.199	0.176	0.177
1.25	0.336	0.348	0.308	0.297	0.319	0.267	0.224	0.177	0.157	0.157
1.5	0.309	0.323	0.271	0.259	0.277	0.219	0.189	0.150	0.129	0.128
1.75	0.280	0.289	0.229	0.211	0.243	0.181	0.150	0.110	0.101	0.091
2	0.247	0.254	0.189	0.160	0.205	0.134	0.109	0.076	0.065	0.059
3	0.082	0.109	0.047	0.037	0.082	0.031	0.017	0.009	0.008	0.006

Tableau 3.4 R

	lag 1	2	3	4	5	10	20	40	70	100
0.125	0.92	0.85	0.78	0.70	0.63	0.59	0.53	0.47	0.42	0.41
0.25	0.93	0.86	0.79	0.72	0.64	0.60	0.54	0.48	0.42	0.41
0.5	0.94	0.87	0.80	0.73	0.66	0.62	0.55	0.48	0.42	0.41
0.75	0.94	0.88	0.81	0.74	0.67	0.61	0.54	0.46	0.40	0.39
1	0.94	0.88	0.81	0.74	0.67	0.59	0.51	0.43	0.37	0.36
1.25	0.94	0.87	0.80	0.72	0.65	0.55	0.46	0.38	0.32	0.31
1.5	0.94	0.86	0.77	0.69	0.61	0.50	0.40	0.33	0.27	0.26
1.75	0.93	0.85	0.74	0.65	0.56	0.44	0.34	0.27	0.22	0.21
2	0.93	0.83	0.71	0.60	0.51	0.38	0.27	0.21	0.17	0.16
3	0.90	0.74	0.55	0.40	0.29	0.18	0.09	0.06	0.04	0.04

Tableau 3.5 SAS

d	Lag1	Lag2	Lag3	Lag4	Lag5	Lag10	Lag20	Lag40	Lag70	Lag100
0.125	0.113	0.131	0.160	0.143	0.137	0.111	0.119	0.107	0.097	0.087
0.25	0.203	0.224	0.228	0.217	0.229	0.207	0.186	0.159	0.162	0.143
0.5	0.276	0.294	0.282	0.273	0.291	0.268	0.228	0.187	0.197	0.174
0.75	0.307	0.317	0.298	0.286	0.303	0.276	0.232	0.185	0.195	0.173
1	0.320	0.317	0.294	0.279	0.293	0.262	0.218	0.169	0.177	0.158
1.25	0.322	0.305	0.279	0.259	0.272	0.238	0.195	0.146	0.154	0.134
1.5	0.318	0.287	0.256	0.231	0.246	0.209	0.168	0.121	0.128	0.108
1.75	0.309	0.265	0.231	0.201	0.219	0.179	0.139	0.096	0.102	0.082
2	0.297	0.244	0.205	0.172	0.195	0.152	0.112	0.074	0.079	0.060
3	0.233	0.175	0.118	0.085	0.126	0.073	0.038	0.021	0.020	0.011

Tableau 3.6 SAS

d	Lag1	Lag2	Lag3	Lag4	Lag5	Lag10	Lag20	Lag40	Lag70	Lag100
0.125	0.082	0.059	0.055	0.052	0.076	0.059	0.054	0.052	0.028	0.039
0.25	0.128	0.093	0.089	0.096	0.120	0.088	0.084	0.064	0.044	0.055
0.5	0.165	0.125	0.122	0.135	0.155	0.112	0.109	0.075	0.054	0.064
0.75	0.185	0.146	0.142	0.158	0.171	0.124	0.119	0.081	0.055	0.064
1	0.196	0.161	0.152	0.176	0.178	0.129	0.121	0.081	0.049	0.057
1.25	0.198	0.169	0.153	0.190	0.174	0.127	0.114	0.076	0.041	0.048
1.5	0.191	0.169	0.144	0.200	0.159	0.117	0.098	0.065	0.031	0.037
1.75	0.173	0.159	0.126	0.206	0.137	0.100	0.078	0.052	0.020	0.025
2	0.148	0.141	0.102	0.204	0.109	0.079	0.056	0.038	0.012	0.016
3	0.051	0.052	0.024	0.142	0.028	0.016	0.007	0.004	0.000	0.000

Tableau 3.7 SAS

d	Lag1	Lag2	Lag3	Lag4	Lag5	Lag10	Lag20	Lag40	Lag70	Lag100
0.125	0.134	0.169	0.189	0.168	0.184	0.162	0.126	0.109	0.090	0.089
0.25	0.162	0.201	0.219	0.203	0.214	0.192	0.150	0.125	0.099	0.097
0.5	0.211	0.265	0.263	0.252	0.264	0.235	0.185	0.143	0.103	0.103
0.75	0.256	0.316	0.295	0.286	0.305	0.261	0.205	0.148	0.096	0.100
1	0.289	0.355	0.309	0.298	0.331	0.267	0.206	0.138	0.080	0.087
1.25	0.303	0.372	0.298	0.283	0.334	0.245	0.186	0.115	0.059	0.068
1.5	0.286	0.358	0.258	0.241	0.309	0.202	0.149	0.083	0.038	0.047
1.75	0.244	0.316	0.199	0.183	0.262	0.149	0.104	0.053	0.020	0.028
2	0.189	0.259	0.140	0.126	0.208	0.099	0.066	0.030	0.010	0.015
3	0.044	0.084	0.021	0.017	0.065	0.014	0.006	0.002	0.000	0.000

Sources (à rajouter)