

Correction du biais d'estimation sur les mesures de risque en finance

Jean-Cyprien Héam

▶ **To cite this version:**

Jean-Cyprien Héam. Correction du biais d'estimation sur les mesures de risque en finance. Autre. 2009. dumas-00516027

HAL Id: dumas-00516027

https://dumas.ccsd.cnrs.fr/dumas-00516027

Submitted on 8 Sep 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.





École Centrale de Lyon
TFE 2009

Centre de Recherche en Économie et Statistique
Malakoff

Rapport Final de Travail de Fin d'Etudes

Correction du biais d'estimation sur les mesures de risque en finance

Jean-Cyprien Héam

Tuteurs :
ECL :
Christian De Peretti (tuteur pédagogique)
Elisabeth Mironescu (responsable filière)
Entreprise :
Jean-Michel Zakoïan

Option : Mathématiques et Décision

Filière : Ingénierie Mathématique

Métier : Ingénieur d'Affaire

Remerciements :

Je tiens tout d'abord à remercier le Professeur Jean-Michel Zakoïan pour m'avoir permis de conduire ce travail de fin d'étude au Centre de Recherche en Économie et Statistiques. Sa disponibilité m'a été précieuse pour rapidement comprendre les éléments mis en jeu pour cette étude. La pertinence des conseils du Professeur Christian Gouriéroux m'a permis d'étoffer mon étude ; je l'en remercie sincèrement.

Mes remerciements vont enfin à l'ensemble des membres du laboratoire de finance/assurance et de macro-économie qui ont fait preuve d'une grande sympathie en m'accueillant : Kenza Benhima, Pierre Cahuc, Laura Hering, Guy Laroque, Etienne Lehmann, Céline Letrémy, Julien Martin, Isabelle Méjean, Gulten Mero, Sophie Osotimehin, Linas Tarasonis, Fanda Traore et Bertrand Villeneuve.

Résumé

Dans le cadre des procédures de backtesting de la Value-at-Risk (VaR), nous proposons une étude de la qualité de la correction de l'effet d'estimation du test de Kupiec (*Journal of Derivatives*, 1995) fondée sur les travaux de Escansiano et Olmo (2008). Cette étude conduit à apporter une correction de l'effet d'estimation de la VaR elle-même dans le cadre d'un modèle paramétrique. En intégrant les contraintes des praticiens, notamment la faible taille des échantillons, l'étude montre que les corrections de l'effet de l'estimation tant au niveau de la VaR qu'au niveau de la variance asymptotique du test de Kupiec sont inefficaces : le gain est minime voire inexistant tandis que le temps de calcul est sensiblement allongé.

Mot clefs : mesure de risque, Value-at-Risk (VaR), biais d'estimation, Estimation Paramétrique, Effet de Paramètre.

Abstract

Focusing on the value-at-risk (VaR) backtesting framework, we present a survey on the accuracy of the correction from estimation effect of the test of Kupiec (Journal of Derivatives, 1995) based on Escanciano and Olmo's papers (2008). This survey leads to build, in a parametric approach, a correction from the estimation effect for the VaR itself. Regarding to down-to-earth considerations, namely the small sizes of samples, the survey shows that the corrections of the estimation effect concerning the VaR and the asymptotical variance of Kupiec's test are inefficient : the upgrading is minimal, even null, while computing time increases.

Key-words : Risk Measure, Value-at-Risk (VaR), Estimation Biasis, Parametric Estimation, Parameter Effect

Sommaire

I Introduction : l'importance de la VaR comme mesure de risque 5

II Procédures de backtesting de la VaR 13

1 Modélisation 17

2 Tests statistiques 27

III Effet d'estimation 39

3 Éléments empiriques 43

4 Correction de la VaR 47

5 Variance asymptotique 53

6 Conclusion III 59

IV Conclusion 61

V Références 65

VI Table des figures 69

VII Annexes 73

A Éléments techniques 75

B Code Matlab 97

Première partie

Introduction : l'importance de la VaR comme mesure de risque

« Les mathématiciens financiers ne sont pas dans le projet Manhattan en train de fabriquer la bombe d'Hiroshima » déclarait Nicolas Bouleau [1] dans le Figaro (29/10/2008) pour répondre aux attaques contre la mathématisation extrême de la sphère financière.
La mathématisation du monde est un phénomène constaté depuis plusieurs siècles [12] : par exemple, mesurer le temps c'est se libérer des contraintes naturelles : jour/nuit, saison, localisation géographique (mesure de la longitude),... La mathématisation est utilisée comme un outil de domestication du monde L'économie n'a pas échappé à cette formalisation. Si les réflexions sur la monnaie d'Aristote ou de St-Thomas d'Aquin [19] influencent toujours les théories économiques, on appelle « révolution marginaliste » l'introduction du formalisme mathématique dans l'économie : la révolution marginaliste, portée par Walrass, fait reposer la valeur d'un bien sur la dérivée de l'utilité qu'il procure. Cette mathématisation se poursuit encore aujourd'hui ([6]).
Connexe à la science économique, la finance est devenue un lieu d'application des mathématiques à tel point qu'un site d'information américain titre à propos de la crise financière « La formule qui a tué Wall Street » [2]. Ne peut-on pas s'affranchir du risque financier comme l'on s'est affranchi du temps par une mathématisation ?

Le risque financier regroupe en réalité un ensemble de risques qui différent par leur nature : le risque opérationnel correspond aux aléas humains ou physiques (mauvaise saisie informatique, incendie,...) ; le risque de marché renvoie à la notion même de dépréciation soudaine de la valeur d'un actif ; le risque de liquidité traduit la difficulté à vendre ou à acheter tel actif (et donc à l'estimer)... Si aucun de ces risques ne peut être écarté, les institutions financières (banques, assureurs, fonds spéculatifs,...) sont particulièrement sensibles au risque de marché.
En 1993, la société JP Morgan propose une méthode d'évaluation du risque appelée RiskMetrics. Cette évaluation repose sur une grandeur appelée Value-at-Risk (VaR). Pour un seuil de confiance α, on définit la VaR d'un actif comme l'α-quantile de la fonction de pertes et profits (PnL) dans h périodes de l'actif conditionnelle à l'information disponible aujourd'hui.

$$VaR(t+h,t,\alpha) = F^{-1}_{PnL_{t+h}|\mathcal{F}_t}(\alpha) \tag{1}$$

Par exemple, une VaR au seuil de confiance 1% à horizon 1 ($h = 1$) de 10 000$ signifie qu'une fois sur 100 (1%) l'institution perd au moins 10 000$. La simplicité de sa signification a fait le succès de cette méthode. Son utilisation est devenue obligatoire en 2004 avec les accords dit de Bâle II. Ces accords internationaux sont des normes de gestion imposées aux institutions financières.
La VaR n'est pas exemptE de défaut [3] mais elle a l'avantage d'être synthétique et de ne pas afficher la stratégie des banques : un seul scalaire rend compte de l'ensemble des positions occupées sans qu'il puisse être possible de savoir quelles sont ces positions.

One-day, 99% VaR

in / end of period	Interest rate and credit spread	Foreign exchange	Commodity	Equity	Diversification benefit	**Total VaR**
2008 (CHF million)						
Average	149	22	36	65	(94)	178
Minimum	103	7	18	29	– [1]	123
Maximum	232	57	60	126	– [1]	272
End of period	107	16	18	39	(57)	123

Figure 1 – Exemple de publications - Rapport annuel du Crédit Suisse 2008

Si la mesure est imposée, le comité de Bâle II laisse aux institutions financières la liberté quant à la modélisation du risque. L'aspect technique de la détermination de l'α-quantile de la fonction de perte et

1. professeur de mathématiques financières à l'ENPC
2. « Recipe for Disaster : The Formula That Killed Wall Street », 23-02-09, *Wired Magazine*, `http://www.wired.com/techbiz/it/magazine/17-03/wp_quant?currentPage=all` (consulté en Juin 2009) expose l'utilisation abusive de la formule de Li (copule gaussienne) alors que la faiblesse de la modélisation est évidente dans la mesure où l'ensemble des interdépendances ne sont présentes que par un unique paramètre scalaire.
3. en particulier, elle n'incite pas à la diversificaton : la VaR de deux porte-feuilles fusionnés n'est pas nécessairement plus petite que la somme des VaR de ces deux portefeuilles.

profit est du ressort des banques. Comme cette étape est fondamentale, le législateur a accompagné la liberté d'une obligation d'évaluation de la qualité du modèle utilisé (voir par exemple, fig.1 p7 et fig.2 p8).
L'évaluation est faite a posteriori en comparant les résultats prédits et les résultats réalisés. On regroupe l'ensemble de cette démarche sous le terme de procédure de backtesting. C'est sur ce point que les mathématiques interviennent pour une troisième fois (après la modélisation, et le choix de la mesure). La théorie statistique fournit des outils, appelés tests statistiques, qui permettent de décider selon un critère précis si la procédure d'estimation de la VaR est correcte ou non. Cependant, le calcul de la VaR ainsi que les tests sont issus de procédures dont les propriétés statistiques ne sont vraies que lorsque l'estimation portent sur un échantillon de taille infinie. Le biais introduit correspond à l'effet d'estimation. De plus, les tests statistiques reposent sur des propriétés dites asymptotiques : elles ne sont vraies qu'à « l'infini »ie lorsque la taille de l'échantillon utilisée pour le test est infinie.

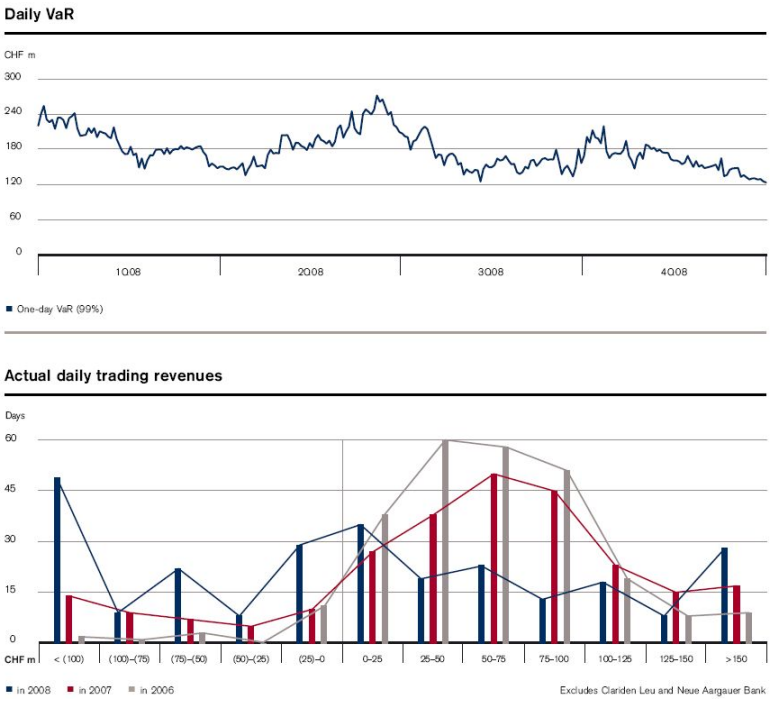


Figure 2 – Exemple de publications - Rapport annuel du Crédit Suisse 2008

Ce mémoire a pour but d'étudier et de quantifier le risque dû à l'effet de paramètre. La problématique retenue est donc de savoir si la non-prise en compte de ce biais (on parlera alors de risque) fausse les résultats des procédures de backtesting.

Utilisons une analogie pour expliquer ce qui se joue. On choisit de représenter le risque réel par une forme polygonale « en étoile ». C'est le risque économique. Un mesure de risque est une manière de quantifier la taille du risque : par exemple, le diamètre[4] renseigne sur la « taille » du risque : voir fig.3 (p9). Si la forme est compacte, c'est une information pertinente.
L'étape de modélisation (fig.4 p9) consiste à choisir une forme approchée du risque pour mener les calculs. A gauche, on a choisi un modèle de type circulaire tandis qu'on a opté pour une ellipse d'axe principal oblique sur la droite. Il est intéressant de voir que la proximité modèle/réel ne sera abordée que via la mesure de risque choisie. C'est ce qu'illustre la figure 5 (p9). Si on reporte ces distances parallèlement (Fig. 6 p10), on observe une « erreur ». C'est cette grandeur qui est utilisée pour procéder au backtesting : l'idée consiste à regarder si cette erreur est statistiquement importante.

C'est à l'étape de modélisation qu'intervient l'effet de paramètre. Tout comme il impossible de tracer un cercle parfait, le modèle n'est que purement théorique. Ainsi, on estime le cercle par un polygone

4. Le diamètre d'une surface est la plus grande distance entre deux points de la surface.

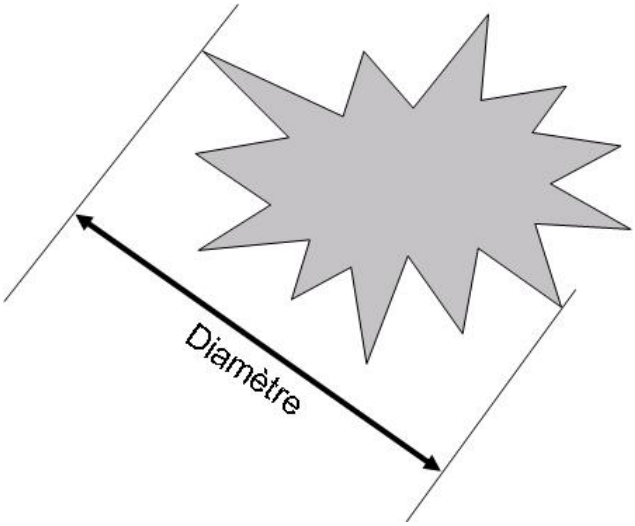


Figure 3 – Illustration de la notion de mesure de risque

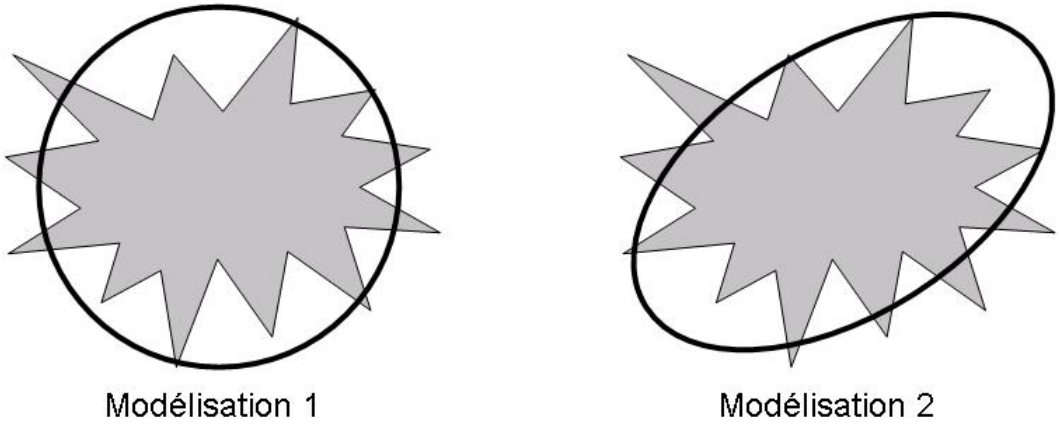


Figure 4 – Illustration de la notion de modélisation

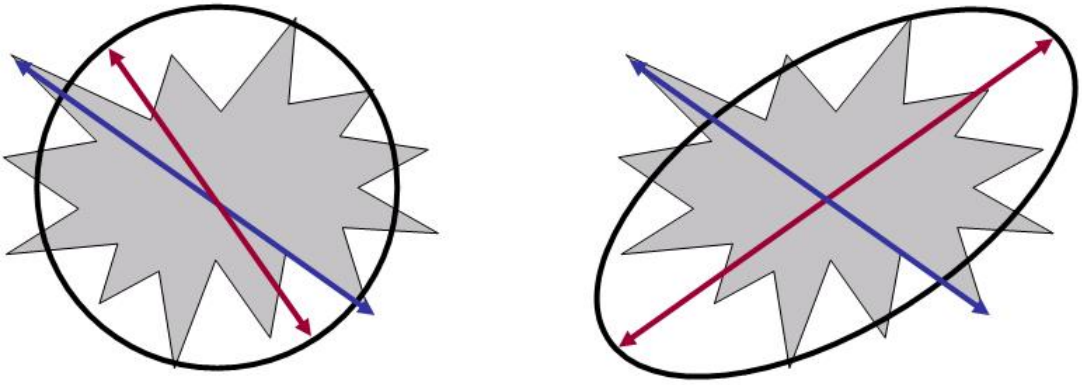


Figure 5 – Illustration des notions de modélisation et de mesure de risque

régulier (fig.7 p10). Le cercle est la limite des polygones réguliers lorsque l'on augmente le nombre de côtés. Cet effet d'approximation joue sur les mesures de risques comme l'illustre la figure 8 (p10). L'illustration présente montre un biais de paramètre qui accroît la mesure de risque mais on pourrait très bien imaginer un effet inverse (approximation du cercle par polygones inscrit). L'important est que seule l'erreur dans

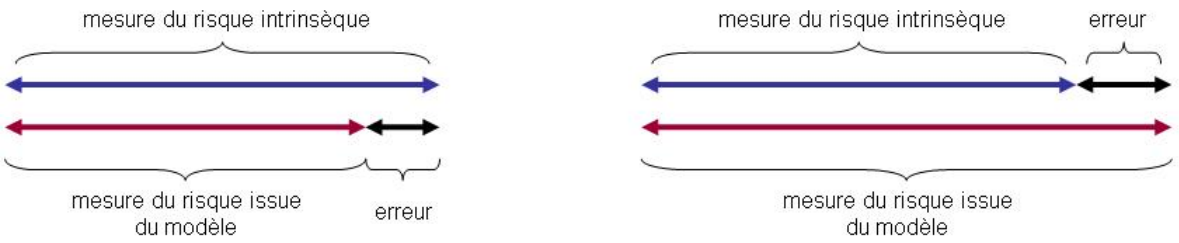


Figure 6 – Illustration des différences entres les mesures

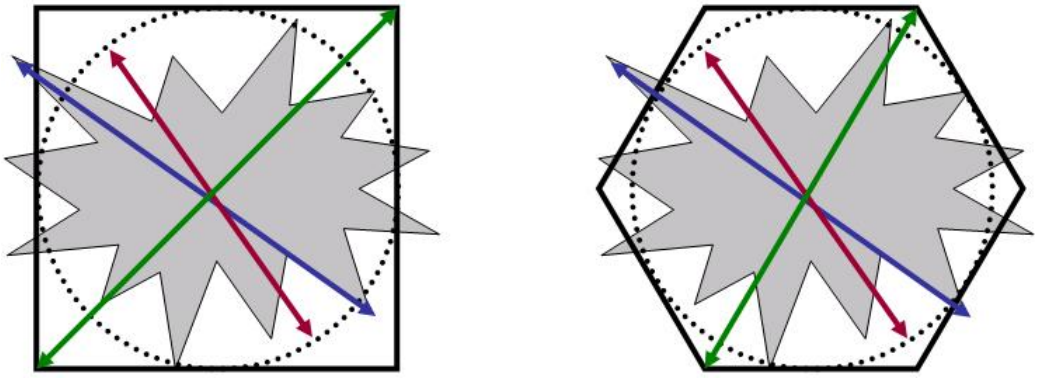


Figure 7 – Illustration de l'effet de paramètre

son ensemble est observable. Mais la procédure de backtesting est calibrée pour la différence entre la

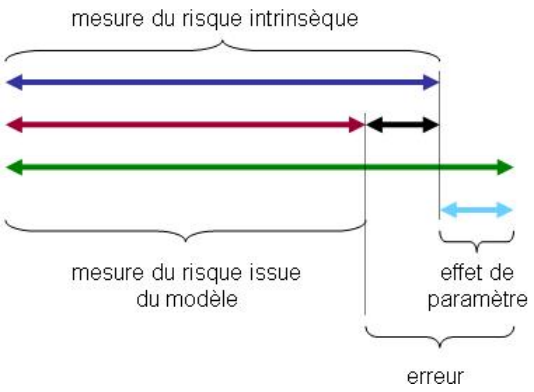


Figure 8 – Illustration de l'effet de paramètre sur les mesures de risques

mesure du risque intrinsèque et la mesure du risque issue du modèle. On peut voir la procédure de backtesting comme un entonnoir qui détermine la qualité du modèle : plus on est bas, plus le modèle est correct.
Comme, cette différence est parasitée par l'effet de paramètre il se *peut*, comme l'illustre la figure 9 p11, que cet effet fausse les résultats de backtesting : un modèle correct sera considéré comme faussé ou un modèle incorrect sera considéré comme juste.

Le mémoire se compose de deux grandes parties. La première se concentre sur les procédures de backtesting de la VaR afin de donner des éléments théoriques et pratiques. Ce processus complexe est

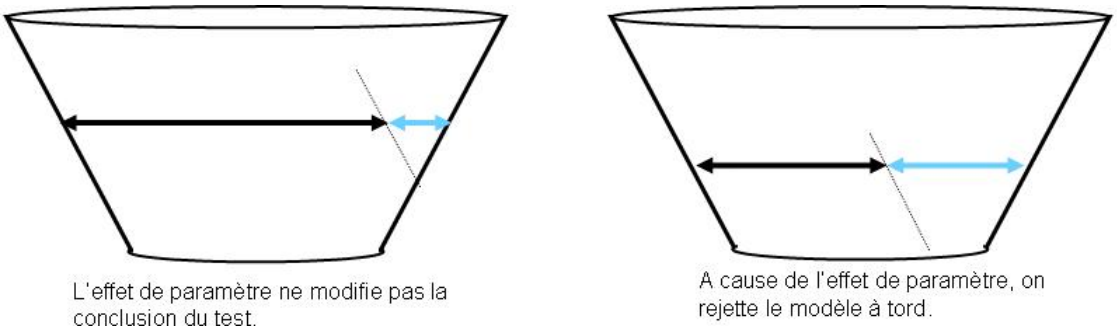


Figure 9 – Illustration de l'effet de paramètre sur le backtesting

découpé en deux sections portant respectivement sur l'étape de modélisation et sur l'étape de test. La seconde grande partie est le coeur de l'étude : une analyse des corrections pour le calcul de la VaR et pour le calcul de la variance asymptotique du test de Kupiec.
Le lecteur trouvera en annexe les passages techniques ainsi que les codes sources.

Notations :
Les notations introduites au fur et à mesure restent autant que possible homogènes sur l'ensemble du mémoire. Cette page a pour but de faciliter la lecture.

- PnL : fonction de pertes et profits (Profits & Loss)
- VaR : Value-at-Risk
- γ : seuil de confiance en la VaR
- GARCH(1,1)$-\mathcal{N}$: modélisation par un processus GARCH(1,1) à aléas gaussiens
- ω : premier paramètre dans un modèle GARCH(1,1)$-\mathcal{N}$
- α : deuxième paramètre dans un modèle GARCH(1,1)$-\mathcal{N}$
- β : troisième paramètre dans un modèle GARCH(1,1)$-\mathcal{N}$
- σ : volatilité du processus dans un modèle GARCH
- η : aléa (ou innovation)
- ε : rendement logarithmique d'un actif financier
- $\tilde{\cdot}$: grandeur calculée à partir d'estimation
- $\hat{\cdot}$: grandeur estimée
- $F_X^{\leftarrow}$: fonction de répartition (effectif cumulé croissant) de la variable aléatoire X
- Φ : fonction de répartition de la loi gaussienne
- Φ^{-1} : fonction quantile de la loi gaussienne
- $\mathcal{F}_t$: information disponible à la date t
- $\mathcal{N}(\mu, \sigma^2)$: loi normale d'escmpérance μ et de variance σ^2
- $\mathcal{G}(p)$: loi géométrique de paramètre p
- $\mathcal{B}(p)$: loi binomiale de paramètre p
- $\mathbb{P}(\cdot)$: probabilité
- $\mathbb{P}(\cdot|\cdot)$: probabilité conditionnelle
- $\mathbb{E}(\cdot)$: espérance
- $\mathbb{E}(\cdot|\cdot)$: espérance conditionnelle
- $\mathbb{V}(\cdot)$: variance
- $\mathbb{V}(\cdot|\cdot)$: variance conditionnelle
- $\perp$: relation d'indépendance entre variables aléatoires
- $\bowtie$: relie un processus markovien stationnaire à sa matrice de transition
- $\rightsquigarrow$: convergence en loi
- $\dot{\rightsquigarrow}$: convergence en loi sous hypothèse
- $\overset{iid}{\sim}$: suite de variables indépendantes et identiquement distribuées
- $\mathcal{L}$: loi de probabilité quelconque

Deuxième partie

Procédures de backtesting de la VaR

Les procédures de backtesting sont en aval de la procédure de gestion du risque. Elles succèdent à la modélisation et au choix de la mesure du risque. Il convient donc d'analyser ces étapes−modélisation et mesure de risque− pour pouvoir identifier précisément l'influence du biais de paramètre.

Une première partie s'attache à étudier la modélisation des cours boursiers. Il s'agit de formaliser mathématiquement les données financières. L'objectif est de trouver une famille d'objets mathématiques, appelés processus [5], qui rendent compte correctement du réel.

On attend que ces processus vérifient deux propriétés. D'abord, ils doivent pourvoir être correctement calibrés sur des données financières réelles. Ensuite, ces processus doivent posséder des propriétés mathématiques établies et utiles au calcul de la VaR. Cette condition se traduit par le fait que l'on doit pouvoir connaître facilement la loi de probabilité conditionnelle des pertes.

Une seconde partie détaille plusieurs tests couramment utilisés dans le milieu financier dans les procédures de backtesting. Le but est de comprendre à la fois l'objectif de chaque test et les moyens mis en oeuvre. On montrera que les tests ne reposent pas directement sur des propriétés de la VaR mais se fondent sur des conséquences des propriétés de la VaR.

5. Mathématiquement, un processus est une suite de variables aléatoires indexées par une grandeur représentant le temps.

Chapitre 1

Modélisation

Ce chapitre se concentre sur la construction de modèle d'évolution des cours boursiers. Les cours boursiers évoluent en continu : pendant les heures d'ouverture des marchés financiers, la dynamique de l'offre et de la demande entraîne une évolution constante des cours.

Une modélisation naturelle est alors de chercher à rendre compte de cette continuité : mathématiquement, on parle de processus en temps continu. Cette approche, illustrée par le fameux modèle de Black-Scholes ([10]), n'est pas adaptée à notre étude car on cherche à modéliser un processus journalier : la VaR réglementaire se calcule jour par jour. Par conséquent, le but de la modélisation conduit à considérer une approche en temps discret.

La première section expose la manière dont on mène une modélisation à partir d'un exemple concret fondé sur un indice boursier. Ensuite, on présente la manière dont on passe de la modélisation de la dynamique financière à la mesure du risque.

1.1 Modélisation financière

Dans cette section, on considère que l'ensemble des positions de la banque correspond au cours d'un seul actif financier. Une première partie donne un exemple de série financière avec l'indice S&P500 [1]. Cet exemple permet d'introduire, en deuxième partie, la modélisation de type « GARCH » [2].

1.1.1 Données empiriques

La figure 1 (p18) représente les séries brutes de l'indice S&P500 sur les 10 dernières années. La crise financière de 2008 est clairement visible en fin de série.
Mathématiquement, on définit la notion de stationnarité ainsi :

Définition 1.1 (Processus stationnaire)**.** *Le processus* (X_t) *est dit strictement stationnaire si les vecteurs* $(X_1, ...X_k)$ *et* $(X_{1+h}, ..., X_{k+h})$ *ont même loi jointe pour tout entier* k *et tout entier relatif* h*.*

Concrètement, la série du S&P500 n'est pas stationnaire : il y a notamment des périodes de fortes hausses. Ce caractère non-stable est une difficulté mathématique car si une série ne semble pas bornée, on ne peut exclure qu'elle puisse diverger.
Pour construire une série stationnaire à partir d'une série non-stationnaire, il existe plusieurs transformations standards à essayer. On peut par exemple considérer la série des accroissements (différence entre deux termes consécutifs), calculer la série des rendements,... Les figures suivantes (fig.2 p18, fig.3 p19 et fig.4 p19) présentent le résultats de plusieurs de ces transformations. A ce stade, les trois séries transformées semblent viables.

1.1.2 GARCH

Une fois que le choix de la série de travail est arrêté, il faut choisir un modèle. On retient le plus courant : le modèle GARCH(1,1).

1. Le S&P500 est un indice boursier fondé sur les 500 plus grandes sociétés cotées sur les bourses américaines ; il est disponible gratuitement sur la section finance de Yahoo `http://fr.finance.yahoo.com/`.
2. « GARCH »est l'acronyme de Generalized AutoRegressive Conditional Heteroscedasticity.

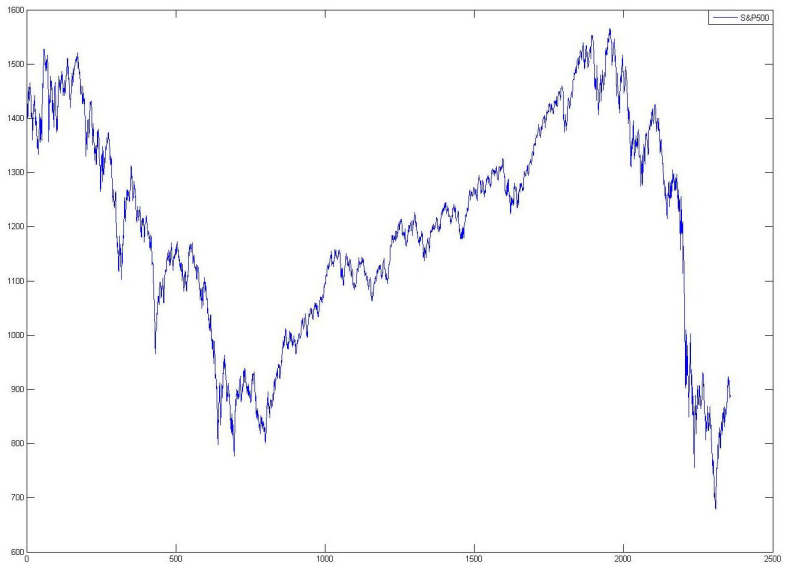


FIGURE 1 – Graphique de l'indice S&P500 entre mai 1999 et mai 2009

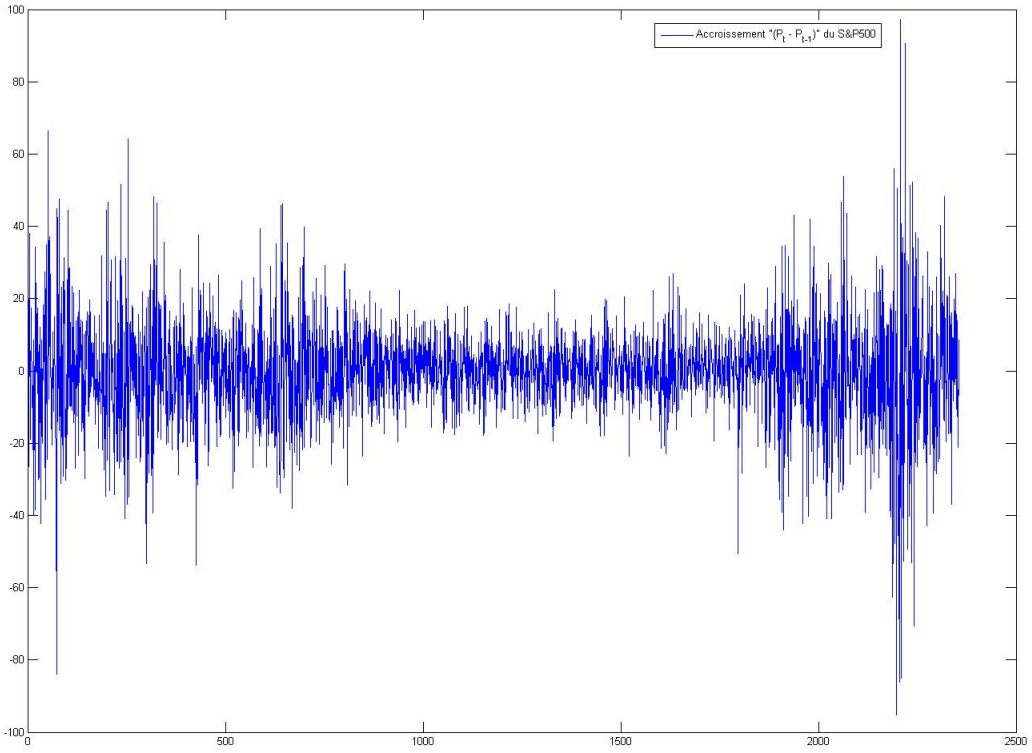


FIGURE 2 – Graphique des accroissements journaliers de l'indice S&P500 entre mai 1999 et mai 2009

L'intérêt d'un modèle GARCH est de transcrire une propriété observée sur les marchés. La variabilité des cours, appelée volatilité, n'est pas constante : des phases de grande variabilité alternent avec des phases de variabilité faible. On parle de « paquets de volatilité ».

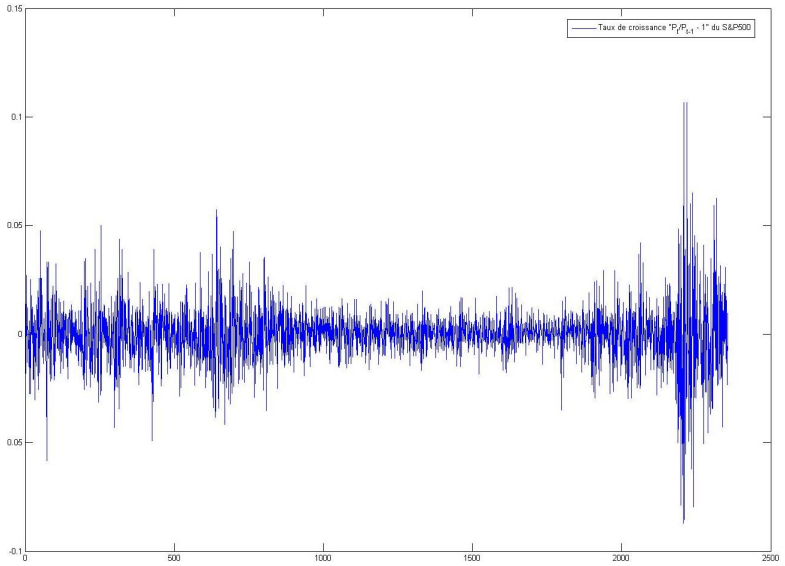


FIGURE 3 – Graphique des taux de croissance journaliers de l'indice S&P500 entre mai 1999 et mai 2009

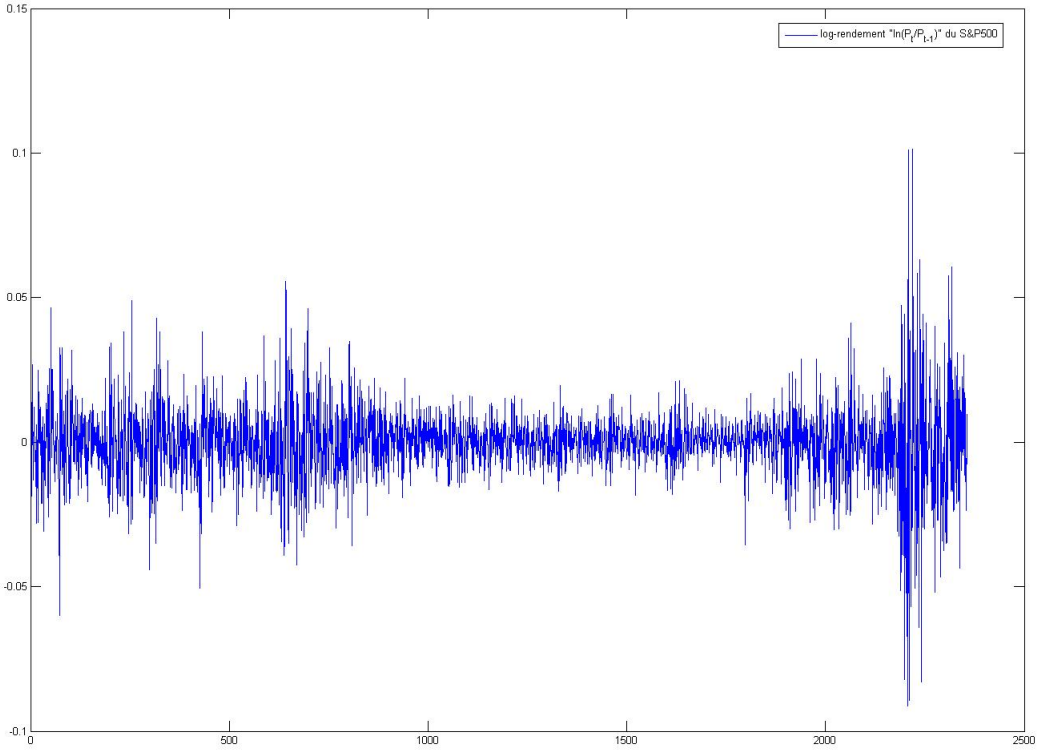


FIGURE 4 – Graphique des log-rendements journaliers de l'indice S&P500 entre mai 1999 et mai 2009

Éléments théoriques

Pour transcrire ce fait, il faut que la volatilité présente une certaine persistance. Mathématiquement, cela signifie que la volatilité est un phénomène à mémoire. Les processus GARCH font explicitement dépendre la volatilité présente des volatilités passées.

Définition 1.2 (Processus GARCH). *Le processus $(X_t)_{t\in[\![0;T]\!]}$ suit un processus GARCH de paramètres p et q, noté GARCH(p,q), si*

- *les deux premiers moments de X_t existent pour tout $t \in]\!]0;T[\![$*
- *$\forall t \in [\![0;T]\!], \mathbb{E}(X_t|X_{t-1}) = 0$*
- *il existe des constantes ω, α_i, $i \in \mathbb{N}_q$ et β_j, $j \in \mathbb{N}_p$ telles que*

$$\sigma_t^2 = V(X_t|X_{t-s}, s \in \mathbb{N}^*) = \omega + \sum_{i=1}^{q} \alpha_i X_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2 \tag{1.1}$$

◇

Cette définition très générale laisse notamment le choix de p et de q. En pratique, il faut garder à l'esprit qu'en règle générale, un modèle sera d'autant plus précis que le nombre de paramètres, $p+q$, est grand mais qu'il sera d'autant plus difficile d'estimer un modèle que le nombre de paramètres est important. Il y a donc un compromis [3] entre une qualité théorique (grand nombre de paramètres) et la qualité atteignable en pratique (faible nombre de paramètres).
Un compromis courant est de choisir $p = q = 1$:

$$\begin{cases} X_t = \sigma_t \eta_t \\ \sigma_t^2 = \omega + \alpha X_{t-1}^2 + \beta \sigma_{t-1}^2 \\ \omega \geq 0 \ ; \ \alpha \geq 0 \ ; \ \beta \geq 0 \end{cases} \tag{1.2}$$

Le dernier problème concerne la stationnarité : tout modèle GARCH(1,1) n'est pas stationnaire. Néanmoins, il existe une condition suffisante assez simple assurant la stationnarité :

Proposition 1.1 (Stationnarité - [25]). *Soit $(X_t)_t$ un processus GARCH(1,1). Une condition suffisante pour assurer la stationnarité au second ordre est :*

$$\omega > 0 \quad et \quad \alpha + \beta < 1 \tag{1.3}$$

◇

L'ensemble de ces considérations amène donc une modélisation par un GARCH(1,1) dont les principales caractéristiques sont le respect du caractère volatile des cours et la qualité pratique à être calibrée facilement :

$$\begin{cases} X_t = \sigma_t \eta_t \\ \sigma_t^2 = \omega + \alpha X_{t-1}^2 + \beta \sigma_{t-1}^2 \\ \eta_t \overset{iid}{\sim} \mathcal{N}(0,1) \ ; \ \omega > 0 \ ; \ \alpha \geq 0 \ ; \ \beta \geq 0 \ ; \ \alpha + \beta < 1 \end{cases} \tag{1.4}$$

Principe d'estimation

L'estimation d'un modèle consiste à déterminer les valeurs de ces paramètres à partir des données empiriques. La difficulté dans l'estimation réside dans le fait que le praticien ne peut observer qu'une réalisation de l'expérience où apparaissent des variables inobservables. Par exemple, dans notre cas, le financier observe X_t mais ne peut pas observer σ_t ni η_t.
Il existe plusieurs méthodes d'estimation. La principale dans le cas de modèles paramétriques est le principe de maximisation de vraisemblance. Concrètement, il s'agit de déterminer les valeurs des paramètres qui maximisent la probabilité de réalisation de l'échantillon. Pour cela, on suppose que les résidus (η_t) suivent une loi normale; comme cela n'est pas rigoureusement le cas, certains emploient le terme de « quasi-vraisemblance ». Formellement, dans le cas d'un modèle GARCH(1,1) à aléa gaussien [4] :

$$\begin{cases} (\omega, \alpha, \beta) = argMax\Big(\displaystyle\prod_{t=1}^{T} \frac{1}{\sqrt{2\pi}\sigma_t} e^{-\frac{X_t^2}{2\sigma_t^2}} \Big) \\ \text{sous contrainte : } \omega > 0 \ ; \ \alpha \geq 0 \ ; \ \beta \geq 0 \ ; \ \alpha + \beta < 1 \ ; \ \sigma_0 \geq 0 \ ; \ X_0 \in \mathbb{R} \end{cases} \tag{1.5}$$

3. On parle parfois de critère de parcimonie.
4. On notera par la suite GARCH(1,1)$-\mathcal{N}$.

Choix de la série

On a cherché à estimer un modèle GARCH(1,1) sur les séries construites à partir du S&P500 dans le but de les stationnariser :

- La série des accroissement est la simple différence entre deux dates consécutives : $Y_t := X_t - X_{t-1}$
- La série des taux de croissance revient à rapport la séries des accroissements à la valeur courante de manière à s'affranchir des unités : $Y_t := \dfrac{X_t - X_{t-1}}{X_{t-1}}$
- La séries des log-rendements (parfois appelés rendements géométriques) est la série des logarithmes du rapport des valeurs entre deux dates consécutives : $Y_t := ln\dfrac{X_t}{X_{t-1}}$

Les résultats complets sont présents en annexes (fig.11 p88, fig.12 p89 et fig.13 p90). Pour le choix d'un modèle, on analyse la valeur des paramètres estimés et l'allure de la distribution des $\tilde{\eta}_t$ correspondante par rapport à une distribution normale centrée réduite.
Les trois modèles GARCH estimés sont fortement stationnaires mais avec une très faible marge ($\alpha + \beta \lesssim 1$).

Séries	ω	α	β	$\alpha+\beta$
Accroissement	1.2242	0.0607	0.9332	$0.9939 < 1$
Taux de croissance	$1.0973E^{-6}$	0.0760	0.9176	$0.9936 < 1$
Log-rendement	$1.1008E^{-6}$	0.0754	0.9181	$0.9935 < 1$

FIGURE 5 – Estimation des paramètres à partir du S&P500

Le très fort poids du coefficient ω pour la série des accroissements laisse dubitatif. Le terme ω est un facteur d'échelle et d'habitude les séries financières sont estimées avec un facteur d'échelle très petit devant l'unité.
Les modèles sur les taux de croissance et les log-rendements sont très similaires, tant sur les paramètres que sur les distributions des η_t. Le choix se porte sur la modélisation des log-rendements car l'introduction d'un logarithme permet de combiner facilement ces grandeurs (c.f. A.8.1, p 92).

Estimation

Il est simple de comprendre que si la taille de la fenêtre d'estimation (T) augmente, la qualité de l'estimation s'accroît en général. Pour étudier la nature de la convergence des paramètres estimés $(\omega, \alpha, \beta) = f(T)$ à partir des réalisations $(X_t)_{t\in[\![1;T]\!]}$, des simulations sous Matlab ont été menées. Les détails de l'expérience sont mis-en-annexe [5] (c.f. A.1, p 76) tandis que seuls les résultats sont présentés ici.
Le principe consiste à générer des trajectoires selon un modèle GARCH(1,1) connu et à pratiquer la méthode d'estimation. Le grand nombre de trajectoires permet d'obtenir à la fois la convergence moyenne et d'estimer l'intervalle de confiance de cette convergence (principe de Monte Carlo).

T	ω ($\times 10^{-4}$)	α	β	Simulations
100	3.26	0.113	0.437	10 000
200	2.94	0.100	0.519	10 000
300	2.33	0.094	0.614	10 000
400	1.98	0.091	0.661	10 000
500	1.76	0.088	0.693	10 000
600	1.68	0.087	0.704	10 000
700	1.62	0.086	0.714	10 000
800	1.54	0.086	0.724	10 000
1000	1.43	0.085	0.736	10 000
5000	1.02	0.082	0.793	10 000
∞	1.00	0.080	0.80	

FIGURE 6 – Tableau de convergence des paramètres estimés en moyenne

5. De manière générale, c'est un modèle GARCH(1,1) de paramètres ($\omega = 0.001$, $\alpha = 0.08$, $\beta = 0.8$) qui est utilisé dans ce mémoire.

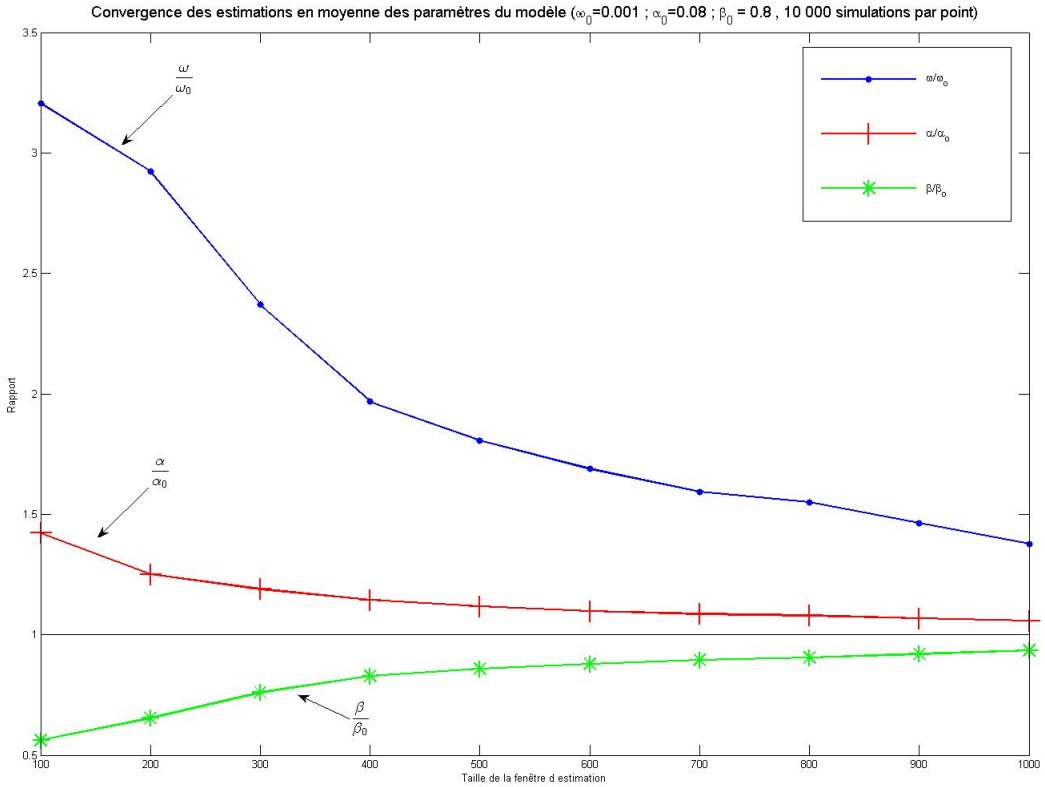


FIGURE 7 – Graphique de convergence du rapport moyen paramètres estimés-paramètres théoriques

Ce résultat (fig.6 p21 et fig.7 p22) permet de constater que la convergence en moyenne est assez rapide pour des données réelles : 300 ou 400 dates, soit de 15 à 20 mois de cotations environ suffisent [6]. Le retard de convergence du rapport $\frac{\omega}{\omega_0}$ est à nuancer par le fait que ω est en valeur bien plus petit que α ou β donc l'imprécision joue moins.
Si un résultat en moyenne permet de valider la démarche générale, une étude de la répartition des estimations, représentées sous forme de boîte-à-moustache [7], permet de vérifier que l'approche « en moyenne » est fiable (fig.8 p23, fig.9 p23 et fig.10 p24).
Il y a effectivement convergence de la moyenne et de la médiane. On observe aussi une réduction significative de la variance (taille du rectangle). Pour ω, la convergence de la médiane est progressive et par excès et s'accompagne d'une réduction de l'écart-type (qui transcrit l'étalement de la distribution). Par contre, les points extrêmes (5% de la population) ne se rapprochent pas de la valeur limite moyenne. Pour α, la convergence en moyenne est par excès mais on observe que la médiane est quasiment stable à la valeur limite. Le plus manifeste est le rétrécissement complet de la distribution (variance et points extrêmes). Pour β, la convergence en moyenne est par défaut. Cette convergence s'accompagne d'une convergence de la médiane et d'une réduction modérée de la variance. De même, la taille (95% de la population) semble se réduire bien moins rapidement que pour les deux autres paramètres (uniquement à partir de 300 dates).

1.1.3 Conclusion

La modélisation des cours ou positions des institutions financières se fondent sur des processus mathématiquement formalisés.

6. On s'accorde sur le fait qu'une année d'exercice/cotation correspond à 250 dates.
7. La représentation sous forme de boîte-à-moustache, « boxplot » en anglais, permet de visualiser la distribution. Le rectangle représente les quartiles centraux (50% de la population) ; il est coupé par la médiane. Les traits en pointillé qui s'en écartent représentent la taille de la distribution (95% de la population). Les points « aberrants »(5% extrêmes de la population) sont marqués par des croix.

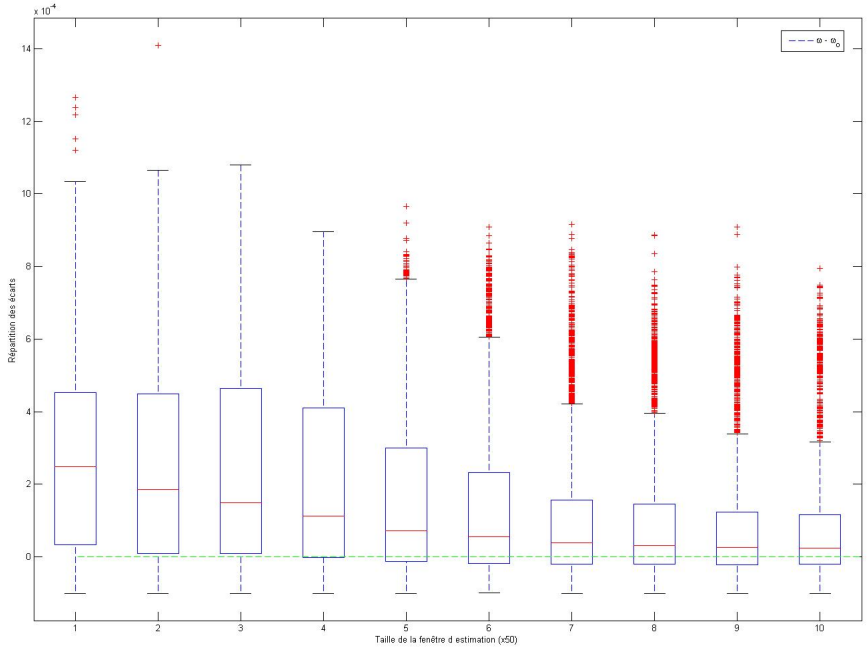


FIGURE 8 – Boxplot de la répartition des erreurs entre ω et ω_0 selon la taille de la fenêtre d'estimation

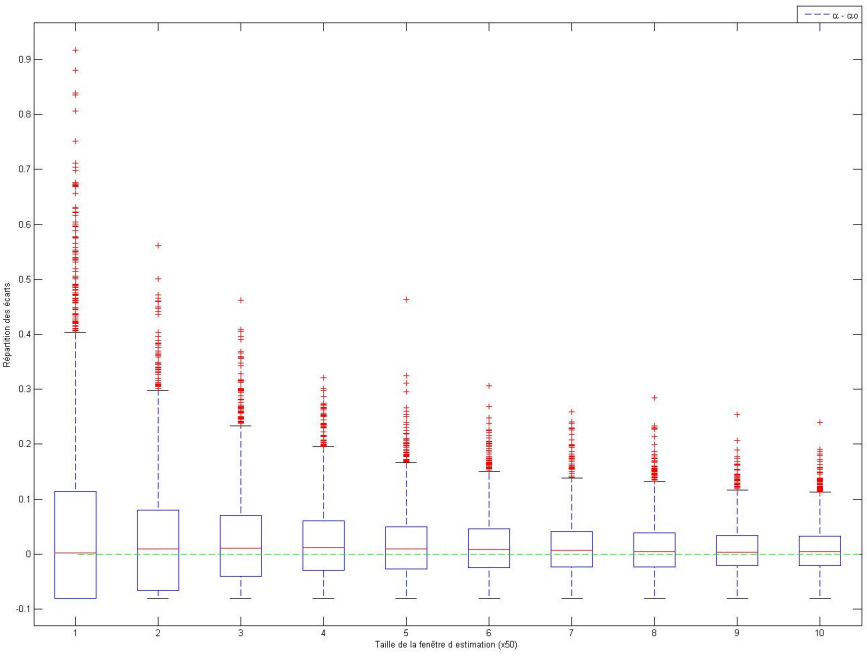


FIGURE 9 – Boxplot de la répartition des erreurs entre α et α_0 selon la taille de la fenêtre d'estimation

L'estimation repose sur le principe de maximisation de la vraisemblance. La vraisemblance fait intervenir directement la fonction de densité de probabilité en chaque point. La facilité de mise en oeuvre de la maximisation dépend donc de la facilité à exprimer en tout point cette fonction de densité.

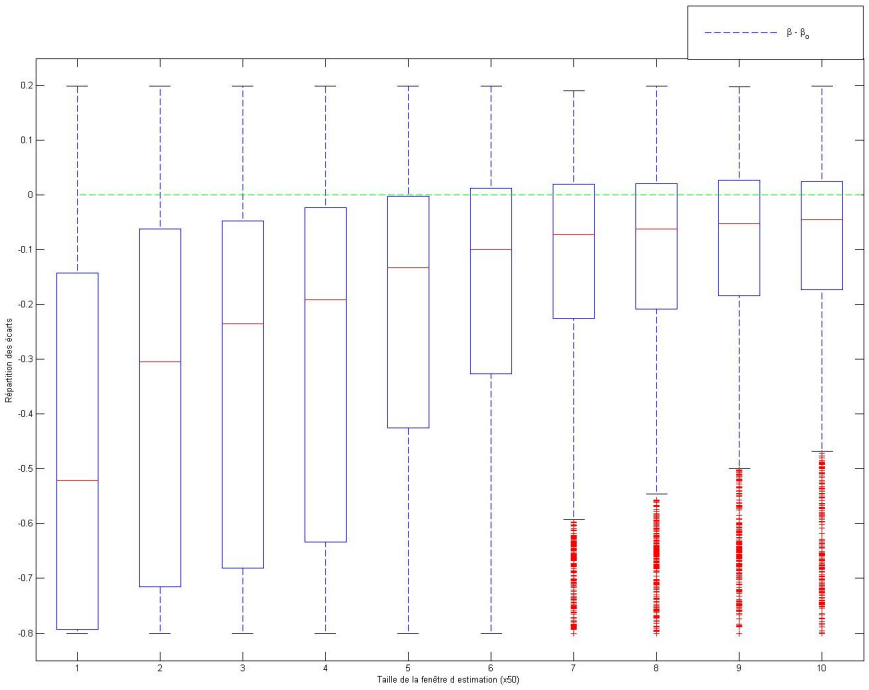


FIGURE 10 – Boxplot de la répartition des erreurs entre β et β_0 selon la taille de la fenêtre d'estimation

Dans le cas d'un modèle GARCH(1,1), quelques centaines de points suffisent en général à une estimation viable.

1.2 VaR : théorie et estimation

La VaR est une mesure de risque : elle a pour objet de quantifier le risque encouru par une institution financière. Elle repose sur l'idée de coupler le montant des pertes (quantification brute) et la fréquence des pertes (prise en compte du risque). Après avoir défini avec précision la VaR, on verra quelques développements de cette mesure, avant d'analyser sa mise en oeuvre.

1.2.1 Définition

Définition 1.3 (VaR). *On appelle VaR de niveau γ et d'horizon h, le quantile d'ordre $1-\gamma$ de la distribution conditionnelle des pertes $(-PnL_t)$:*

$$VaR_{t+h|t}(\gamma) = inf\Big\{x \in \mathbb{R} \ / \ \mathbb{P}\Big(-PnL_{t+h,t} \leq x|\mathcal{F}_t\Big) \ \geq \ 1-\gamma\Big\} \tag{1.6}$$

lorsque ce quantile est positif. Par convention, $VaR_{t+h|t}(\gamma) = 0$ sinon.

◇

La réglementation impose aux banques le calcul journalier de la VaR à une horizon de 10 jours pour un seuil de 1% ([13], [5]). Afin de garder une certaine mémoire, le capital réglementaire (K_t) est le maximum entre la VaR calculée et la moyenne mobile des VaR passées affectée d'un coefficient de sécurité S_t (dépendant de la qualité du modèle utilisé) :

$$K_t = max\Big(VaR_{t+10|t}(1\%); S_t \times \frac{1}{60}\sum_{i=0}^{59} VaR_{t+10-i|t-i}(1\%)\Big) \tag{1.7}$$

Le coefficient S_t est une fonction du nombre de dépassements (N) par les pertes de la VaR calculés durant les 250 derniers jours.

$$S_t = \begin{cases} 3 & \text{si } N \leq 4 \\ 3 + 0.2(N-4) & \text{si } 5 \leq N \leq 9 \\ 4 & \text{sinon} \end{cases} \tag{1.8}$$

La réglementation se fonde clairement sur la VaR comme indicateur principal tout en prenant en compte une dimension « qualité du modèle » via le coefficient S_t.

1.2.2 Développement - approximation

Cette formulation est très générale. Il est possible de reformuler la VaR lorsque la distribution conditionnelle des pertes admet ses deux premiers moments.
Si la distribution conditionnelle admet une espérance et une variance, on peut alors écrire, en introduisant PnL_h la variables des pertes et profits à horizon h :

$$PnL_{t+h,t} = \mathbb{E}(PnL_{t+h,t}|\mathcal{F}_t) + \sigma_{t+h,t} PnL_h \tag{1.9}$$

On réinjecte cette expression dans la formulation de la VaR [8] :

$$VaR_{t+h|t}(\gamma) = -\mathbb{E}(PnL_{t+h,t}|\mathcal{F}_t) - \sigma_{t+h,t} F_h^{\leftarrow}(1-\gamma) \tag{1.10}$$

Dans le cas où les pertes suivent conditionnellement des lois gaussiennes, on peut montrer la propriété suivante :

$$\begin{cases} PnL_{t+h,t} \overset{iid}{\sim} \mathcal{N}(\mathbb{E}(PnL_t)h, \mathbb{V}(PnL_t)h) \\ VaR_{t+h|t}(\gamma) = \mathbb{E}(PnL_t)h + \sqrt{\mathbb{V}(PnL_t)}\sqrt{h}\Phi^{-1}(1-\gamma)\} \end{cases} \tag{1.11}$$

Très souvent, l'espérance de perte est nulle et la formule se réduit à :

$$VaR_{t+h|t}(\gamma) = \underbrace{\sqrt{h}}_{\text{horizon}} \underbrace{\underbrace{\sqrt{V(L_t)}}_{\text{écart-type}} \underbrace{\Phi^{-1}(1-\gamma)}_{\text{quantile}}\}}_{VaR_{t+1,t}(\gamma)} \tag{1.12}$$

Même si cette formule n'est vraie que dans le cas gaussien, l'approximation suivante est couramment utilisée du fait de sa simplicité :

$$VaR_{t,h}(\gamma) \approx \sqrt{h} \times VaR_{t+1,t}(\gamma) \tag{1.13}$$

VaR dans le cas d'un modèle GARCH(1,1)

Appliquons ces éléments dans le cas pratique retenu. On suppose donc que les log-rendements de la fonction de pertes et profits suivent un processus GARCH(1,1)–$\mathcal{N}$:

$$\begin{cases} \varepsilon_t = ln(\frac{P_t}{P_{t-1}}) \\ \varepsilon_t = \sigma_t \eta_t \\ \sigma_t^2 = \omega + \alpha \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2 \\ \eta_t \overset{iid}{\sim} \mathcal{N}(0,1) \end{cases} \tag{1.14}$$

On peut alors démontrer que la VaR conditionnelle à 1 jour au seuil γ est donnée par l'expression suivante :

Proposition 1.2 (VaR journalière dans le cas d'un GARCH(1,1)–$\mathcal{N}$)**.**

$$\mathbb{P}\Big(PnL_t < -\underbrace{VaR_{t+1|t}(\gamma)}_{VaR(\gamma)} |\mathcal{F}_{t-1}\Big) = \gamma \Leftrightarrow VaR(\gamma) = P_{t-1}(1 - e^{\sigma_t \Phi^{-1}(\gamma)}) \tag{1.15}$$

◇ *Démonstration p92*

Cette expression est explicite. On montre par ailleurs qu'elle vérifie les deux propriétés de la VaR : les violations (dépassement) doivent être indépendantes et arriver en moyenne dans γ % des cas (démonstration p93).

8. La notation $F^{\leftarrow}$ signifie fonction réciproque de la fonction de répartition de la variable aléatoire réelle continue à gauche.

1.2.3 Estimation

Cette partie consiste à présenter l'agencement des différents éléments que sont les choix de la mesure et du modèle, l'estimation du modèle et le calcul de la mesure de risque.
L'objectif de la procédure est d'obtenir une méthode complète de calcul de la VaR.

Choix de la mesure :

La mesure du risque est imposée par la réglementation : on utilisera la VaR(1%). Pour l'horizon, même si la réglementation stipule « 10 jours », on préférera se fixer à 1 jour quitte à appliquer l'approximation présentée ci-dessus (1.2.2). D'un point de vue théorique, il est d'ailleurs nettement plus facile d'obtenir des résultats conditionnels à la date précédente qu'à une horizon supérieure à 1 jour.

Choix du modèle :

On considère que la fonction de pertes et profits d'une banque est issue des variations d'un unique cours P_t (ou plus vraisemblablement un indice). On choisit de modéliser les log-rendements par un processus GARCH(1,1) :

$$\begin{cases} \varepsilon_t := ln(\frac{P_t}{P_{t-1}}) \\ \varepsilon_t = \sigma_t \eta_t \\ \sigma_t^2 = \omega_0 + \alpha_0 \varepsilon_{t-1}^2 + \beta_0 \sigma_{t-1}^2 \\ \eta_t \overset{iid}{\sim} \mathcal{N}(0,1) \end{cases} \tag{1.16}$$

Estimation du modèle :

Sur une période donnée $[\![1;n]\!]$, on estime les paramètres ω_0, α_0 et β_0 par maximisation de la vraisemblance. On note ces paramètres estimés : $\hat{\omega}_n$, $\hat{\alpha}_n$ et $\hat{\beta}_n$. Ces paramètres permettent de construire une série $(\tilde{\sigma}_t)_t$ à partir des observations $(\varepsilon_t)_t$.

Estimation de la VaR :

Une fois que l'on a estimé le modèle, on peut construire, à partir des observations (ε_t) pour toute date $t > n$, une estimation de la volatilité $\tilde{\sigma}_t$ par la formule suivante :

$$\tilde{\sigma}_t = \sqrt{\hat{\omega}_n + \hat{\alpha}_n \varepsilon_{t-1}^2 + \hat{\beta}_n \tilde{\sigma}_{t-1}^2} \tag{1.17}$$

On estime alors simplement la VaR ainsi :

$$\widehat{VaR}_{t|t-1}(\gamma) = P_{t-1}(1 - e^{\tilde{\sigma}_t \Phi^{-1}(\gamma)}) \tag{1.18}$$

1.2.4 Conclusion

La Value-at-Rik (VaR) est une manière simple de mesurer le risque. Elle s'articule facilement avec un modèle GARCH(1,1) à aléas gaussiens, puisqu'on obtient une formule explicite qui se programme aisément.

1.3 Conclusion sur le Modélisation

L'objectif final de l'étape de modélisation est d'obtenir une méthode de prédiction de la VaR. Pour cela, les outils mathématiques utilisés sont des outils robustes ayant fait leurs preuves.

Le choix d'un modèle de type GARCH a l'avantage de rendre compte efficacement des évolutions des marchés financiers et d'être facilement calibré sur des données existantes. La prédiction de la VaR se déduit facilement de la VaR théorique par une simple substitution dans l'écriture.

Si l'on possède dorénavant une méthode de prédiction de la VaR, dans l'optique du praticien il est nécessaire de tester la qualité de la prédiction. Cette étape est celle des tests statistiques de backtesting.

Chapitre 2

Tests statistiques

P. Jorion donne du backtesting la définition suivante ([3]) :

« Le backtesting est un ensemble de procedures statistiques dont le but est de vérifier que les pertes réelles observées ex-post sont en adéquation avec pertes prévues. Cela implique de comparer systématiquement l'historique des prévisions de la Value-at-Risk aux rendements observés du portefeuille. »

L'étude a posteriori de la qualité de la VaR n'est pas évidente dans la mesure où les propriétés de la VaR sont probabilistes. Les mathématiciens contournent cet obstacle en utilisant des tests statistiques. Un test statistique est une procédure cohérente qui fournit une règle de décision : il répond à la question « la prédiction de la VaR est-elle correcte ? »

Les différents tests peuvent à première vue paraître complexes. En réalité, ils reposent sur des idées intuitives. La complexité vient de l'arsenal statistique qu'il est nécessaire d'utiliser pour atteindre les objectifs. Un test statistique n'est qu'un outil.

Une lecture à plusieurs niveaux est possible selon que l'on recherche plus ou moins de précisions sur la mise-en-oeuvre statistique. Le lecteur est donc invité à moduler son attention à ce chapitre selon ses goûts.

La première section présente les différents principes communs aux tests utilisés pour le backtesting de la VaR. Ensuite, 5 sections s'attachent à détailler les idées et la mise-en-oeuvre pour 5 principaux tests. Il s'en suit une section qui dresse un rapide tour d'horizon des autres tests proposés dans la littérature.

2.1 Principes des tests

2.1.1 Test statistique

Un test statistique correspond à une procédure de décision. Il est constitué d'une hypothèse , d'une statistique et d'un seuil de confiance (dit risque de première espèce).

L'hypothèse (dite hypothèse nulle) est la propriété que l'on s'attend à trouver. Une statistique est une formule qui à partir des observations fournit un nombre [1]. Le seuil de confiance est un indicateur de la marge d'erreur que l'on peut commettre.

La clef de voûte est que la loi de probabilité de la statistique est connue [2] si l'hypothèse nulle est vraie. Le seuil de confiance est la probabilité de refuser à tort l'hypothèse nulle [3].

2.1.2 Backtesting de la VaR

Fonction indicatrice

Si l'on considère la définition de la VaR, on se rend compte qu'elle se fonde sur un événement du type « les pertes dépassent tel seuil ». Les procédures de backtesting utilisent toutes cette idée en introduisant ce que l'on appelle la fonction indicatrice de dépassement de la VaR. Cette fonction vaut tout simplement 1 si les pertes dépassent la VaR et 0 sinon.

1. Rigoureusement, une statistique est une variable aléatoire mesurable qui s'écrit comme fonction des variables aléatoires observables.

2. En générale, il s'agit d'une connaissance asymptotique, c'est-à-dire lorsque le nombre d'observations tend vers l'infini.

3. On parle de risque de première espèce; le risque de second espèce est d'accepter à tort l'hypothèse non-nulle. En pratique ces deux risques jouent en sens opposé : le praticien doit faire un arbitrage entre ces deux risques.

Définition 2.1 (Fonction indicatrice)**.**

$$I_t(\gamma) := \mathbb{1}_{PnL_t<-VaR_{t|t-1}(\gamma)} \tag{2.1}$$

◇

La définition de la VaR se traduit directement comme une propriété de la fonction indicatrice. Elle doit valoir en moyenne le seuil de confiance en la VaR (γ) et tout dépassement doit être indépendant des précédents. Cette dernière caractéristique provient du fait que la VaR est calculée en connaissant toute l'information disponible à la date actuelle.

Proposition 2.1 (Propriété de la fonction indicatrice)**.**

$$\mathbb{E}(I_t(\gamma) - \gamma|\mathcal{F}_{t-1}) = 0 \tag{2.2}$$

◇

C'est cette propriété qui constitue l'hypothèse nulle fondamentale à tous les tests que l'on note (H_0).

Hypothèses nulles

Malheureusement, cette propriété est très délicate à tester. En effet, la notion financière « toute l'information disponible » traduit le fait que la prévision doit prendre en compte les évolutions du marchés, les notes de conjonctures,... Mais ces éléments constitutifs de l'information disponible sont très difficilement modélisables.

Par conséquent, les mathématiciens-financiers ont proposé des tests portants sur des *conséquences* de l'hypothèse nulle fondamentale. Les deux conséquences majeures testées sont d'une part, la moyenne de $I_t(\gamma)$ vaut γ et d'autre part, les violations sont indépendantes. La première conséquence constitue l'hypothèse de couverture conditionnelle tandis que la seconde l'hypothèse d'indépendance.

Certains auteurs introduisent d'autres hypothèses se fondant sur d'autres conséquences. Les démonstrations d'implications entre l'hypothèse nulle fondamentale et ces hypothèses sont présentes en annexes.

Principe des tests utilisés pour le backtesting

Lorsqu'un auteur propose une procédure de backtesting de la VaR, sa démarche est en générale la suivante :

1. Choisir une approche soit de couverture inconditionnelle, soit d'indépendance soit les deux simultanément [4]
2. Construire une conséquence de l'hypothèse nulle fondamentale compatible avec cette approche : cette conséquence détermine l'hypothèse nulle retenue ainsi que son « contraire » que l'on appelle hypothèse alternative
3. Proposer une statistique
4. Déterminer la loi de la statistique sous l'hypothèse nulle

Le lecteur peut en première lecture passer directement à la conclusion de ce chapitre (p37) s'il souhaite éviter les aspects techniques. Au demeurant, il trouvera dans les sections suivantes les détails de différents tests.

Pour chaque test, on présente de manière rigoureuse l'hypothèse nulle testée, l'hypothèse alternative testée, la statistique utilisée, ainsi que la démarche algorithmique qui permet de transformer des formules théoriques en méthodes de calcul.

2.2 Test de Kupiec

2.2.1 Repères

Auteur : P Kupiec en 1995 ([2])
Sources : [13], [21], [3]
Hypothèse acceptée : couverture inconditionnelle

4. Dans ce cas, on parle de couverture conditionnelle.

2.2.2 Démarche théorique

L'hypothèse initiale est :

$$(H_{0,\mathcal{B}}) : \text{ « } (I_t(\gamma))_t \overset{iid}{\sim} \mathcal{B}(\gamma) \text{ »} \tag{2.3}$$

Kupiec ne retient qu'une conséquence :

$$(H^{(1)}_{0,uc}) : \text{ « } (I_t)iid \ ; \ \mathbb{E}(I_t) = \gamma \ ; \ \mathbb{V}(I_t) = \gamma(1-\gamma) \text{ »} \tag{2.4}$$

La statistique utilisée est :

$$z = \frac{\sqrt{T}}{\sqrt{\gamma(1-\gamma)}}\Big(\frac{1}{T}\sum_{t=1}^{T} I_t(\gamma) - \gamma\Big) \tag{2.5}$$

Sous $H^{(1)}_{0,uc}$, le théorème de limite centrale permet de donner la loi asymptotique de z :

$$z \overset{H^{(1)}_{0,uc}}{\rightsquigarrow} \mathcal{N}(0,1) \tag{2.6}$$

Ainsi la zone de rejet à 1% est $\{|z| > 2.57\}$ tandis que cella à 5% $\{|z| > 1.96\}$.

2.2.3 Démarche algorithmique

1. déclarer z réel, $z \leftarrow 0$
2. déclarer *indice* entier
3. pour *indice* allant de 1 à T, faire
4. $\quad z \leftarrow z + I(indice)$
5. fin pour
6. $z \leftarrow z/T$
7. $z \leftarrow \frac{\sqrt{T}}{\sqrt{\gamma(1-\gamma)}}(z-\gamma)$

2.3 Test de Markov

2.3.1 Repères

Auteur : P Christoffersen en 1998 ([15])
Sources : [17], [13],[26]
Hypothèses acceptées : couverture inconditionnelle, indépendance, couverture conditionnelle

2.3.2 Démarche théorique

L'idée de Christoffersen est de tester comme alternative un processus markovien d'ordre 1. En effet, H_0 implique que[5]

$$(I_t)_t \bowtie \begin{pmatrix} 1-\gamma & \gamma \\ 1-\gamma & \gamma \end{pmatrix} \tag{2.7}$$

Pour l'aspect « couverture inconditionnelle », Christoffersen propose comme hypothèse nulle :

$$(H^{(2)}_{0,uc}) : \text{ « } (I_t) \bowtie \begin{pmatrix} 1-\gamma & \gamma \\ 1-\gamma & \gamma \end{pmatrix} \text{ »} \tag{2.8}$$

Contre l'hypothèse alternative

$$(H^{(2)}_{1,uc}) : \text{ « } (I_t) \bowtie \begin{pmatrix} 1-\pi_1 & \pi_1 \\ 1-\pi_1 & \pi_1 \end{pmatrix} \text{ et } \pi_1 = \frac{card\{I_t = 1\}}{T} \text{ »} \tag{2.9}$$

La statistique utilisée est

$$LR^{(2)}_{uc} = 2ln\Big(\frac{\pi_1^{card\{I_t=1\}} \times (1-\pi_1)^{T-card\{I_t=1\}}}{\gamma^{card\{I_t=1\}} \times (1-\gamma)^{T-card\{I_t=1\}}}\Big) \tag{2.10}$$

5. On introduit la notation suivante : « $(X_t)_t \bowtie M$ »signifie que le processus $(X_t)_t$ est une chaîne de Markov stationnaire dont la matrice de transition est M.

Sous $H_{0,uc}^{(2)}$, ce rapport de vraisemblance suit asymptotiquement une loi du χ_1^2 :

$$LR_{uc}^{(2)} \overset{H_{0,uc}^{(2)}}{\rightsquigarrow} \chi_1^2 \tag{2.11}$$

Ainsi la zone de rejet à 1% est $\{LR_{uc}^{(2)} > 6.63\}$ tandis que cella à 5% $\{LR_{uc}^{(2)} > 3.87\}$.

Pour l'aspect « indépendance », Christoffersen propose comme hypothèse nulle[6] :

$$(H_{0,ind}^{(2)}) : \ \text{« } (I_t) \bowtie \begin{pmatrix} \pi_{00} & \pi_{01} \\ \pi_{10} & \pi_{11} \end{pmatrix} \text{ et } \pi_{ij} = \frac{card\{I_{t-1} = i \curlywedge I_t = j\}}{T} \text{ »} \tag{2.12}$$

Contre l'hypothèse alternative

$$(H_{1,uc}^{(2)}) : \ \text{« } (I_t) \bowtie \begin{pmatrix} 1-\pi_1 & \pi_1 \\ 1-\pi_1 & \pi_1 \end{pmatrix} \text{ et } \pi_1 = \frac{card\{I_t = 1\}}{T} \text{ »} \tag{2.13}$$

La statistique utilisée est

$$LR_{ind}^{(2)} = 2ln\Big(\frac{\pi_{00}^{card\{I_{t-1}=0 \curlywedge I_t=0\}} \pi_{01}^{card\{I_{t-1}=0 \curlywedge I_t=1\}} \times \pi_{10}^{card\{I_{t-1}=1 \curlywedge I_t=0\}} \pi_{11}^{card\{I_{t-1}=1 \curlywedge I_t=1\}}}{\pi_1^{card\{I_t=1\}} \times (1-\pi_1)^{T-card\{I_t=1\}}}\Big) \tag{2.14}$$

Sous $H_{0,uc}^{(2)}$, ce rapport de vraisemblance suit asymptotiquement une loi du χ_1^2 :

$$LR_{ind}^{(2)} \overset{H_{0,uc}^{(2)}}{\rightsquigarrow} \chi_1^2 \tag{2.15}$$

Ainsi la zone de rejet à 1% est $\{LR_{ind}^{(2)} > 6.63\}$ tandis que celle à 5% $\{LR_{ind}^{(2)} > 3.87\}$.

Pour l'aspect « couverture conditionnelle », Christoffersen propose comme hypothèse nulle :

$$(H_{0,cc}^{(2)}) : \ \text{« } (I_t) \bowtie \begin{pmatrix} \pi_{00} & \pi_{01} \\ \pi_{10} & \pi_{11} \end{pmatrix} \text{ et } \pi_{ij} = \frac{card\{I_{t-1} = i \curlywedge I_t = j\}}{T} \text{ »} \tag{2.16}$$

Contre l'hypothèse alternative

$$(H_{1,uc}^{(2)}) : \ \text{« } (I_t) \bowtie \begin{pmatrix} 1-\gamma & \gamma \\ 1-\gamma & \gamma \end{pmatrix} \text{ »} \tag{2.17}$$

La statistique utilisée est

$$LR_{cc}^{(2)} = 2ln\Big(\frac{\pi_{00}^{card\{I_{t-1}=0 \curlywedge I_t=0\}} \pi_{01}^{card\{I_{t-1}=0 \curlywedge I_t=1\}} \times \pi_{10}^{card\{I_{t-1}=1 \curlywedge I_t=0\}} \pi_{11}^{card\{I_{t-1}=1 \curlywedge I_t=1\}}}{\alpha^{card\{I_t=1\}} \times (1-\alpha)^{T-card\{I_t=1\}}}\Big) \tag{2.18}$$

On pourra remarquer que

$$LR_{cc}^{(2)} = LR_{ind}^{(2)} \times LR_{uc}^{(2)} \tag{2.19}$$

Sous $H_{0,uc}^{(2)}$, ce rapport de vraisemblance suit asymptotiquement une loi du χ_2^2 :

$$LR_{ind}^{(2)} \overset{H_{0,cc}^{(2)}}{\rightsquigarrow} \chi_2^2 \tag{2.20}$$

Ainsi la zone de rejet à 1% est $\{LR_{cc}^{(2)} > 9.21\}$ tandis que celle à 5% $\{LR_{cc}^{(2)} > 5.99\}$.

2.3.3 Démarche algorithmique

1. déclarer n_0, n_1, n_{01}, n_{11} entiers
2. $n_0 \leftarrow 0$
3. $n_1 \leftarrow 0$
4. $n_{01} \leftarrow 0$

6. Dans les formules, le symbole $\curlywedge$ correspond au connecteur logique « et ».

5. $n_{11} \leftarrow 0$
6. si $I(1) = 1$, alors
7. $\quad n_1 \leftarrow n_1 + 1$
8. sinon,
9. $\quad n_0 \leftarrow n_0 + 1$
10. fin si
11. déclarer *indice* réel
12. pour *indice* allant de 2 à T, faire
13. $\quad$ si $I(indice) = 1$, alors
14. $\quad\quad n_1 \leftarrow n_1 + 1$
15. $\quad\quad$ si $I(indice - 1) = 0$, alors
16. $\quad\quad\quad n_{01} \leftarrow n_{01} + 1$
17. $\quad\quad$ sinon,
18. $\quad\quad\quad n_{11} \leftarrow n_{11} + 1$
19. $\quad\quad$ fin si
20. $\quad$ fin si
21. fin pour
22. déclarer p_0, p_1, p_{01}, p_{11} entiers
23. $p_0 \leftarrow \frac{n_0}{T}$
24. $p_1 \leftarrow \frac{n_1}{T}$
25. $p_{01} \leftarrow \frac{n_{01}}{T-1}$
26. $p_{11} \leftarrow \frac{n_{11}}{T-1}$
27. déclarer L_1, L_2, L_3 réels
28. $L_1 \leftarrow \gamma^{n_1}(1-\gamma)^{n_0}$
29. $L_2 \leftarrow p_1^{n_1} p_0^{n_0}$
30. $L_3 \leftarrow (1-p_{01})^{n_0-n_{01}} p_{01}^{n_{01}} (1-p_{11})^{n_1-n_{11}} p_{11}^{n_{11}}$
31. déclarer $LR_{uc}, LR_{ind}, LR_{cc}$ réels
32. $LR_{uc} \leftarrow |2ln(\frac{L_2}{L_1})|$
33. $LR_{ind} \leftarrow |2ln(\frac{L_3}{L_2})|$
34. $LR_{cc} \leftarrow |2ln(\frac{L_1}{L_3})|$

2.4 Test de Portmanteau

2.4.1 Repères

Auteur : J Berkowitz en 2005 pour une formalisation dédiée au backtesting de la VaR ([8])
Sources : [26], [8], [17]
Hypothèse acceptée : indépendance

2.4.2 Démarche théorique

Il s'agit d'un test de Ljung-Box. L'idée consiste à remarquer que[7]

$$H_0 : \text{ « } \mathbb{E}(I_t(\gamma) - \gamma | \mathcal{F}_t) = 0 \text{ » } \equiv H_0^{(3)} : \text{ « } (I_t(\gamma) - \gamma) \text{ est un bruit blanc fort » } \tag{2.21}$$

L'indépendance est alors abordée par la nullité des coefficients d'autocorrélation que l'on note λ_j[8]. Christoffersen propose comme hypothèse nulle :

$$(H_{0,ind}^{(3)}) : \text{ « } \forall m > 0, \forall j \in [\![1; m]\!], \lambda_j = 0 \text{ » } \tag{2.22}$$

7. On dit que $(X_t)_{t \in [\![1;T]\!]}$ est bruit blanc fort lorsque X_t *iid*, $\forall t \in [\![1;T]\!], \mathbb{E}(X_t) = 0$ et $\forall (t,t') \in [\![1;T]\!]^2, \mathbb{V}(X_t) = \mathbb{V}(X_{t'}) < \infty$.
8. Par définition, λ_j est le coefficient de corrélation entre la série $(I_t)_t$ et la série $(I_{t+j})_t$.

La statistique utilisée est

$$LB_{ind}^{(3)}(m) = T(T+2)\sum_{k=1}^{m}\frac{\lambda_k^2}{T-k} \tag{2.23}$$

$$\Leftrightarrow \begin{cases} LB_{ind}^{(3)}(m) = T(T+2)\sum_{k=1}^{m}\left(\frac{1}{T-k}\left(\frac{\frac{1}{T-k}\sum_{t=1}^{T-k}(I_t-\overline{I})(I_{t+j}-\overline{I})}{\frac{1}{T-1}\sum_{t=1}^{T}(I_t-\overline{I})^2}\right)^2\right) \\ \text{avec } \overline{I} = \sum_{t=1}^{T} I_t \end{cases} \tag{2.24}$$

Sous $H_{0,ind}^{(3)}$, cette statistique suit asymptotiquement une loi du χ_m^2 :

$$LB(m)_{ind}^{(3)} \overset{H_{0,ind}^{(3)}}{\rightsquigarrow} \chi_2^m \tag{2.25}$$

Ainsi, pour $m = 5$ et $m = 10$, les seuils de rejet à 1% et 5% sont respectivement de 15.09 et 11.07 d'une part et de 23.21 et 18.31 d'autres parts.

2.4.3 Démarche algorithmique

1. déclarer *moyenne*, *variance* réels
2. $moyenne \leftarrow 0$
3. $variance \leftarrow 0$
4. déclarer X_1, X_2, λ vecteurs de réels
5. pour *indice* allant de 1 à T, faire
6. $\quad moyenne \leftarrow moyenne + I(indice)$
7. fin pour
8. $moyenne \leftarrow \frac{moyenne}{T}$
9. pour *indice* allant de 1 à T, faire
10. $\quad variance \leftarrow variance + (I(indice) - moyenne)^2$
11. fin pour
12. $variance \leftarrow \frac{variance}{T-1}$
13. pour j allant de 1 à m, faire
14. $\quad \lambda(j) \leftarrow 0$
15. $\quad X_1 \leftarrow I(1...(T-j)) - moyenne$
16. $\quad X_2 \leftarrow I((1+j)...T) - moyenne$
17. $\quad \lambda(j) \leftarrow \frac{X_1 \times X_2'}{(T-j)variance}$
18. fin pour
19. déclarer LB réel
20. $LB \leftarrow 0$
21. pour *indice* allant de 1 à m, faire
22. $\quad LB \leftarrow LB + \frac{\lambda(indice)^2}{T-indice}$
23. fin pour
24. $LB \leftarrow (T+2)TLB$

2.5 Test des durées

2.5.1 Repères

Auteur : P Christoffersen et D Pelletier en 2004 ([17])
Sources : [26], [8], [17]
Hypothèses acceptées : couverture inconditionnelle, indépendance, couverture conditionnelle

2.5.2 Démarche théorique

L'idée de Christoffersen est de considérer les durées entre deux violations successives de la VaR. L'hypothèse nulle centrale devient donc $H_{0,\mathcal{D}}$ qui s'exprime en notant D_i la variable aléatoire de la $i^{\text{ème}}$ durée entre deux violations successives :

$$(H_{0,\mathcal{D}}) \ : \ \text{« } D_i \overset{iid}{\sim} \mathcal{G}(\gamma) \text{ »} \tag{2.26}$$

Bien que cette loi soit discrète, on va comparer les distributions empiriques à des distributions de loi de Weibull. Une loi de Weibull est une loi continue dépendant de deux paramètres (a et b) dont la fonction de densité est :

$$f_W(D,a,b) = a^b b D^{b-1} e^{-(aD)^b} \qquad a>0 \ ; \ b>0 \tag{2.27}$$

On retrouve la loi géométrique $\mathcal{G}(\gamma)$ dans le cas où

$$a=\gamma \ \text{ et } \ b=1 \tag{2.28}$$

Pour l'aspect « couverture inconditionnelle », Christoffersen propose comme hypothèse nulle :

$$(H^{(4)}_{0,uc}) : \ \text{« } f_D(D) = \gamma e^{-\gamma D} \text{ »} \tag{2.29}$$

Contre l'hypothèse alternative

$$(H^{(4)}_{1,uc}) : \ \text{« } f_D(D) = p e^{-paD} \text{ »} \tag{2.30}$$

Où p est déterminé par maximisation de la vraisemblance. On note $\ell_{nominale}$ et ℓ_{expo} les vraisemblances associées respectivement à $(H^{(4)}_{0,uc})$ et $(H^{(4)}_{1,uc})$. La statistique utilisée est

$$LR^{(4)}_{uc} = 2ln\Big(\frac{\ell_{expo}}{\ell_{nominale}}\Big) \tag{2.31}$$

Sous $H^{(4)}_{0,uc}$, ce rapport de vraisemblance suit asymptotiquement une loi du χ^2_1 :

$$LR^{(4)}_{uc} \overset{H^{(4)}_{0,uc}}{\rightsquigarrow} \chi^2_1 \tag{2.32}$$

Ainsi le seuil de rejet à 1% est $\{LR^{(4)}_{uc} > 6.63\}$ tandis que celui à 5% $\{LR^{(4)}_{uc} > 3.87\}$.

Pour l'aspect « indépendance », Christoffersen propose comme hypothèse nulle :

$$(H^{(4)}_{0,ind}) : \ \text{« } f_D(D) = p e^{-pD} \text{ »} \ \ (\equiv (H^{(4)}_{1,uc})) \tag{2.33}$$

Contre l'hypothèse alternative

$$(H^{(4)}_{1,ind}) : \ \text{« } f_D(D) = a^b b D^{b-1} e^{-(aD)^b} \text{ »} \tag{2.34}$$

Où p, a et b sont déterminés par maximisation de la vraisemblance. On note ℓ_{expo} et $\ell_{Weibull}$ les vraisemblances associées respectivement à $(H^{(4)}_{0,ind})$ et $(H^{(4)}_{1,ind})$. La statistique utilisée est

$$LR^{(4)}_{uc} = 2ln\Big(\frac{\ell_{expo}}{\ell_{Weibull}}\Big) \tag{2.35}$$

Sous $H^{(4)}_{0,ind}$, ce rapport de vraisemblance suit asymptotiquement une loi du χ^2_1 :

$$LR^{(4)}_{ind} \overset{H^{(4)}_{0,ind}}{\rightsquigarrow} \chi^2_1 \tag{2.36}$$

Ainsi le seuil de rejet à 1% est $\{LR^{(4)}_{uc} > 6.63\}$ tandis que celui à 5% $\{LR^{(4)}_{uc} > 3.87\}$.

Pour l'aspect « couverture conditionnelle », Christoffersen propose comme hypothèse nulle :

$$(H^{(4)}_{0,cc}) : \ \text{« } f_D(D) = \gamma e^{-\gamma D} \text{ »} \ \ (\equiv (H^{(4)}_{0,uc})) \tag{2.37}$$

Contre l'hypothèse alternative

$$(H_{1,cc}^{(4)}) : \text{ « } f_D(D) = a^b b D^{b-1} e^{-(aD)^b} \text{ » } \quad (\equiv (H_{1,ind}^{(4)})) \tag{2.38}$$

Où a et b sont déterminés par maximisation de la vraisemblance. On note $\ell_{nominal}$ et $\ell_{Weibull}$ les vraisemblances associées respectivement à $(H_{0,cc}^{(4)})$ et $(H_{1,cc}^{(4)})$. La statistique utilisée est

$$LR_{cc}^{(4)} = 2ln\Big(\frac{\ell_{nominal}}{\ell_{Weibull}}\Big) \tag{2.39}$$

Sous $H_{0,ind}^{(4)}$, ce rapport de vraisemblance suit asymptotiquement une loi du χ_1^2 :

$$LR_{ind}^{(4)} \overset{H_{0,cc}^{(4)}}{\rightsquigarrow} \chi_2^2 \tag{2.40}$$

Ainsi seuil de rejet à 1% est $\{LR_{cc}^{(4)} > 9.213\}$ tandis que celui à 5% $\{LR_{cc}^{(4)} > 5.99\}$.

2.5.3 Démarche algorithmique

1. déclarer *distrib* liste d'entiers vide
2. déclarer $indice_inf$, $indice_sup$ entiers
3. $indice_inf \leftarrow 1$
4. $indice_sup \leftarrow T$
5. tant que $I(indice_inf) = 0$, faire
6. $\quad indice_inf \leftarrow indice_inf + 1$
7. fin tant que
8. tant que $I(indice_sup) = 0$, faire
9. $\quad indice_sup \leftarrow indice_sup - 1$
10. fin tant que
11. déclarer *duree* entier
12. $duree \leftarrow 1$
13. pour *indice* allant de $indice_inf$ à $indice_sup$, faire
14. $\quad$ si $I(indice) = 0$, alors
15. $\quad\quad duree \leftarrow duree + 1$
16. $\quad$ sinon
17. $\quad\quad distrib(indice) \leftarrow duree$
18. $\quad\quad indice \leftarrow indice + 1$
19. $\quad\quad duree \leftarrow 1$
20. $\quad$ fin si
21. fin pour
22. déclarer $param_wb$, $param_exp$
23. $param_wb \leftarrow optimisation(weibull, distrib)$
24. $param_expo \leftarrow optimisation(exponentielle, distrib)$
25. déclarer ll_wb, ll_expo, $ll_nominal$ réels
26. $ll_wb \leftarrow logvraisemblance(weibull, param_wb, distrib)$
27. $ll_expo \leftarrow logvraisemblance(exponentielle, param_expo, distrib)$
28. $ll_nominal \leftarrow logvraisemblance(weibull, [\frac{1}{\gamma}; 1], distrib)$
29. $LR_uc \leftarrow 2(ll_expo - ll_nominal)$
30. $LR_ind \leftarrow 2(ll_expo - ll_weibull)$
31. $LR_cc \leftarrow 2(ll_weibull - ll_nominal)$

2.6 Autres tests

2.6.1 Test des quantiles dynamiques

Auteurs : Engle et Manganelli en 2004 ([20])
Sources : [8], [28], [16]
Hypothèses acceptées : couverture conditionnelle

L'idée d'Engle et Manganelli est de se concentrer sur la modélisation de l'information disponible (terme « $\mathcal{F}_{t-1}$ »). Alors que la plupart des auteurs réduisent l'information à l'indicatrice à la date immédiatement précédente, Engle et Manganelli proposent de mener une régression de l'indicatrice à la date t sur l'ensemble des indicatrices passées et sur une fonction d'un jeu de variables dites explicatives.

$$I_t - \gamma = \delta + \underbrace{\sum_{k=1}^{K} \beta_k I_{t-k}}_{\text{indicatrices}} + \underbrace{\sum_{k=1}^{K} \xi_k g(..., z_t - k, ...)}_{\text{variables explicatives}} + \underbrace{\epsilon_t}_{\text{aléa}} \tag{2.41}$$

L'hypothèse nulle est la nullité de tous les coefficients :

$$(H_0) : \ « \ \delta = \beta_1 = ... = \beta_K = \xi_1 = ... = \xi_k = 0 \ » \tag{2.42}$$

Les auteurs utilisent ainsi une statistique de Wald :

$$\begin{cases} \Psi := (\delta, \beta_1, ..., \beta_K, \xi_1, ..., \xi_K) \\ Z := \begin{pmatrix} z_1^{(1)} & ... & z_1^{(m)} \\ ... & z_i^{(j)} & ... \\ z_n^{(1)} & ... & z_n^{(m)} \end{pmatrix} \\ DQ = \frac{\hat{\Psi}' Z' Z \hat{\Psi}}{\gamma(1-\gamma)} \overset{H_0}{\rightsquigarrow} \chi^2_{2K+1} \end{cases} \tag{2.43}$$

Le point fort de cette approche est de pouvoir modéliser plus finement l'indépendance de toute information. Le revers est que ce modèle possède plus de paramètres à estimer. Par conséquent, à taille d'échantillon donnée, la précision sera moindre qu'un modèle plus élémentaire.

2.6.2 Test sur plusieurs niveaux conjoints

Auteurs : Crnkovic et Drachman en 1997 ([18]))
Sources : [13]
Hypothèses acceptées : couverture inconditionnelle, indépendance, couverture conditionnelle

Il existe plusieurs types de tests qui cherchent à ne pas dépendre du seuil de confiance γ de la VaR. Pour cela, Crnokovic et Drachman proposent de considérer la série des quantiles des réalisations effectives des pertes et profits :

$$z_t := F^{-1}(PnL_t | \mathcal{F}_t) \tag{2.44}$$

La qualité du modèle sous-jacent à l'expression analytique de la VaR est présente dans la modélisation de $F^{-1}(.|\mathcal{F}_t)$.
L'approche couverture inconditionnelle correspond à une distribution des z_t uniforme sur $[0;1]$ tandis que l'approche indépendance se traduit par l'indépendance de cette distribution.
En pratique, on va utiliser une statistique de Pearson. On se donne $(u_i, l_i)_i$ une subdivision de l'intervalle $[0;1]$ pour construire la statistique suivante :

$$Q = \sum_{i=1}^{K} \frac{\Big(card\{z_i \in [l_i, u_i[\} - card\{z_i\}(u_i - l_i)\Big)^2}{card\{z_i\}(u_i - l_i)} \overset{H_{0,cc}}{\rightsquigarrow} \chi^2_{K-1} \tag{2.45}$$

2.6.3 Test sur les moments

Auteurs : Bontemps et Maddahi en 2005 ([11])
Sources : [14]

Hypothèses acceptées : couverture inconditionnelle, indépendance, couverture conditionnelle

L'idée initiale est d'utiliser les durées entre deux violations successives de la VaR. Au lieu de tester directement ces grandeurs, on chercher à vérifier des relations polynomiales sur les moments d'ordres successifs de la distribution empirique[9].

On définit la suite $(d_i)_{i\in\mathbb{N}_n}$ comme la suite des durées entres deux violations successives de la VaR. On définit les moments de la distribution géométrique de paramètre β par :

$$\begin{cases} M_{-1}(d,\beta) = 0 \\ M_0(d,\beta) = 1 \\ M_{j+1}(d,\beta) = \dfrac{(1-\beta)(2j+1)+\beta(j-d+1)}{(j+1)\sqrt{1-\beta}} M_j(d,\beta) \\ \qquad\qquad - \dfrac{j}{j+1} M_{j-1}(d,\beta) \quad j \in \mathbb{N} \end{cases}$$

On note $M(d,\beta)$ le vecteur des p premiers polynômes $M_j(d,\beta)$.
On définit alors la statistique suivante :

$$J_{cc}(p) = \Big(\frac{1}{\sqrt{N}} \sum_{i=1}^{N} M(d_i,\gamma)\Big)^2$$

Sous $(H_{0,cc})$: ogℙ$(I_t(\gamma)|\mathcal{F}_{t-1}) = \gamma$fg, on a :

$$J_{cc}(p) \overset{H_{0,cc}}{\rightsquigarrow} \chi^2_p$$

On définit la statistique suivante :

$$J_{uc}(p) = \Big(\frac{1}{\sqrt{N}} \sum_{i=1}^{N} M_1(d_i,\gamma)\Big)^2$$

Sous $(H_{0,uc})$: ogℙ$(I_t(\gamma)) = \gamma$fg, on a :

$$J_{uc} \overset{H_{0,uc}}{\rightsquigarrow} \chi^2$$

On définit la statistique suivante :

$$J_{ind}(p) = \Big(\frac{1}{\sqrt{N}} \sum_{i=1}^{N} M(d_i,\beta)\Big)^2$$

Sous $(H_{0,ind})$: og$t \neq t' \Rightarrow I_t(\gamma) \perp I_{t'}(\gamma)$fg, on a :

$$J_{ind} \overset{H_{0,ind}}{\rightsquigarrow} \chi^2$$

2.7 Remarques sur l'implementation informatique

Les sections précédentes sont des présentations théoriques des outils utilisés. Les détails « algorithmiques » peuvent sembler superflus dans la mesure où ils ne semblent traduire que les formules mathématiques présentées précédemment. Pourtant, ils traduisent une formulation qui est en pratique fiable.
Le problème est que si l'on considère des plages de données réalistes (une année de cotation correspond à environ 250 dates), le nombre de dépassement de la VaR est faible (le seuil de confiance est de 1% en général). Les observations sont donc très peu nombreuses. L'imprécision engendrée joue différemment selon les formulations utilisées.
Par exemple, pour calculer la vraisemblance par rapport à un processus de Markov d'ordre 1, les deux formules sont *mathématiquement équivalentes* :

$$LR^{(2)}_{uc} = 2ln\Big(\frac{\pi_1^{card\{I_t=1\}} \times (1-\pi_1)^{T-card\{I_t=1\}}}{\gamma^{card\{I_t=1\}} \times (1-\gamma)^{T-card\{I_t=1\}}}\Big) \text{ avec } \pi_1 = \frac{card\{I_t=1\}}{T} \tag{2.46}$$

9. Le principe d'une approche par moments généralisés est que chaque distribution vérifie un système particulier de relations polynomiales entre ces moments. Par exemple, les moments m_i de la loi normale vérifient : $m_1 = \mu$, $m_2 = \sigma^2$, puis $m_{2k} = (2k-1)m_{2k-2}$.

$$\Leftrightarrow LR_{uc}^{(2)} = 2ln\Big(\frac{(1-\pi_0)^{T-card\{I_t=0\}} \times \pi_0 card\{I_t=^0\}}{\gamma^{card\{I_t=1\}} \times (1-\gamma)^{T-card\{I_t=1\}}}\Big) \text{ avec } \pi_0 = \frac{card\{I_t=0\}}{T} \tag{2.47}$$

Mais il n'est pas évident qu'elles soient *équivalentes en pratique*. Par exemple, si $X = 0,5 \pm 10\%$ et $Y = 0,5 \pm 1\%$ alors $X^Y = 0,71 \pm 5\%$ tandis que $Y^X = 0.71 \pm 3\%$.
On signalera aussi pour le test des durées le fait de devoir contourner les valeurs aux bords de l'échantillon.

Les algorithmes présentés ont été implémentes et testés. Pour plus de détails, le lecteur peut consulter les codes sources présents en annexes (c.f. p97).

2.8 Conclusion tests

Les différents tests abordent la vérification de la VaR selon trois axes : la couverture inconditionnelle, l'indépendance et la couverture conditionnelle. Chacun met en place selon un principe propre.
Il est important de noter qu'un test de « couverture conditionnelle »n'est pas nécessairement meilleur que deux tests l'un portant sur la couverture inconditionnelle et l'autre sur l'indépendance. Premièrement, en cas de rejet de la validité de la VaR dans le premier cas on ne sait pas quelle composante n'est pas vérifiée. Deuxièmement, de manière empirique, les tests de couverture conditionnelle sont moins performants que les autres ([13]).
Le tableau suivant présente de manière synthétique les caractéristiques des tests utilisés dans ce projet de recherche.

	Couverture Inconditionnelle	Indépendance	Couverture Conditionnelle	Principe
Kupiec	✓			Ecart à la moyenne
Markov	✓	✓	✓	Processus de Markov d'ordre 1
Durées	✓	✓	✓	Durées entre deux violations
Portmanteau		✓		Nullité des coefficients d'auto-corrélation

FIGURE 1 – Tableau Récapitulatif Tests-Hypothèses

Troisième partie

Effet d'estimation

Une des étapes de la procédure de mesure du risque est l'estimation du modèle. Cette estimation introduit une marge d'erreur qui n'est jamais prise en compte dans les autres étapes. Le but de cette partie est d'étudier les effets de cette erreur : une faible erreur sur l'estimation a-t-elle un impact fort sur les conclusions de la procédure de contrôle ?

L'estimation du modèle joue principalement sur le calcul de la VaR et sur les tests. Au niveau de la VaR, il n'est pas évident que substituer la volatilité par la volatilité estimée est satisfaisant : c'est simple mais est-ce probant ? Au niveau des tests, le critère de décision se fonde sur la probabilité que la statistique considérée suive une loi de probabilité précise. Mais cette règle n'est vraie que lorsque la taille de l'échantillon tend vers l'infini : les tailles considérées en pratique sont-elles, *au sens de ces tests*, proches de l'infini ?

Un premier chapitre présente des résultats empiriques qui mettent en exergue le fait qu'en pratique l'effet d'estimation n'est pas sans conséquence. Les deux chapitres suivants traitent de l'effet de l'estimation au niveau de la VaR et au niveau des tests : la correction de la VaR est bien plus délicate que la correction d'un test.

L'objectif final est d'avoir un état des lieux de l'effet d'estimation, de différents méthodes de correction ainsi que des compromis que l'on doit mener en pratique.

Chapitre 3

Éléments empiriques

La notion de risque recouvre une grande variété de phénomènes si on songe aux sources de risque. La théorie financière a donc cherché à qualifier les risques en fonction de leurs sources. Le risque de modèle renvoie à la notion d'une trop grande distance entre le modèle choisi et la réalité. Ce risque est à compléter par l'effet d'estimation -que certains auteurs appellent risque de paramètres- : même si le modèle choisi est correct, l'effet d'estimation peut fausser les prévisions.
Une première section présente des résultats empiriques qui montrent que l'effet d'estimation joue de manière notable sur les résultats des procédures de backtesting. Les sections suivantes analysent ces résultats selon plusieurs axes afin d'en déduire le maximum d'éléments.

3.1 Résultats

Les résultats de cette section sont issus de l'expérience n°2 (p81).
Le principe est de générer 50 000 trajectoires d'un processus GARCH(1,1) :

$$\begin{cases} \varepsilon_t = \sigma_t \eta_t \\ \sigma_t^2 = 0.0001 + 0.08\varepsilon_{t-1}^2 + 0.8\sigma_{t-1}^2 \\ \eta_t \overset{iid}{\sim} \mathcal{N}(0,1) \\ \varepsilon_0 = \sigma_0 = 0 \end{cases} \tag{3.1}$$

Pour chaque trajectoire, on considère que les log-rendements de la fonction de pertes et profits de la banque suivent cette trajectoire. Une partie de l'échantillon est utilisée pour estimer les valeurs des paramètres (fenêtre d'estimation) tandis que la partie restante est utilisée pour mener les procèdures de backtesting (fenêtre de backtesting). Pour chaque test, on calcule deux statistiques : une à partir des paramètres estimés ($\hat{\theta}$) et une à partir de paramètres de la génération ($\theta_0 = (0.0001; 0.08; 0.8)$). Ces deux statistiques sont comparées au seuil de la loi asymptotique sous l'hypothèse nulle : le seuil de confiance nominal est fixé à 1%.

3.2 Présentation des résultats sur les niveaux

Les tableaux suivants – fig.1 p44, fig.2 p44 et fig.3 p45 – présentent les niveaux empiriques de rejets [1] pour une fenêtre d'estimation respectivement de 250, 500 et 1000 dates.

Chaque tableau se compose de deux parties : sur le gauche sont présents les résultats pour une fenêtre de backtesting de 250 dates tandis que sur la droite sont présents les résultats pour une fenêtre de backtesting de 500 dates.

Pour chaque test (Kupiec, Markov, Durée, Portmanteau d'horizon 5, Portmanteau d'horizon 10), les trois lignes associées correspondent au niveau de rejet constaté avec le modèle estimé (ligne $\hat{\theta}$), au niveau de rejet constaté avec le modèle théorique (ligne θ_0) et au niveau nominal fixé à 1%.

Les colonnes correspondent aux types d'hypothèses testées : « UC »pour la couverture inconditionnelle, « IND »pour l'indépendance et « CC »pour la couverture conditionnelle.

1. Un niveau empirique de rejet est la proportion de statistiques dépassant la valeur seuil fixé par le choix du niveau ; par exemple, pour une statistique sensée suivre une loi normale centrée réduite, le niveau empirique de rejet à 5% est $card\{|z^{(i)}| > 1.96\}/card\{z^{(i)}\}$

Estimation d'un GARCH(1,1)–$\mathcal{N}$ de paramètres $(10^{-4}; 0.08; 0.8)$ **sur 250 dates**
Seuil nominal des tests fixé à 1 %

		Backtesting sur 250 dates			Backtesting sur 500 dates		
Hypothèses		UC	IND	CC	UC	IND	CC
Kupiec	$\hat{\theta}$	6.01%			8.71%		
	θ_0	1.45%			1.28%		
Markov	$\hat{\theta}$	3.14%	1.26%	2.44%	6.64%	1.58%	4.87%
	θ_0	0.44%	1.08%	1.16%	1.15%	0.96%	1.71%
Durée	$\hat{\theta}$	10.22%	6.84%	23.54%	8.51%	4.22%	16.78%
	θ_0	3.33%	4.12%	8.28%	1.40%	2.19%	3.84%
PM-5	$\hat{\theta}$		6.91%			9.85%	
	θ_0		5.21%			7.74%	
PM-10	$\hat{\theta}$		10.90%			12.27%	
	θ_0		4.19%			4.05%	

FIGURE 1 – Tableau des niveaux empiriques de rejets - Estimation sur 250 dates

Idem Fig. 1 p44 sauf **estimation sur 500 dates**

Hypothèses		UC	IND	CC	UC	IND	CC
Kupiec	$\hat{\theta}$	3.26%			4.47%		
	θ_0	1.40%			1.37%		
Markov	$\hat{\theta}$	1.47%	1.11%	1.56%	3.36%	1.18%	2.81%
	θ_0	0.45%	1.01%	1.06%	1.16%	0.95%	1.69%
Durée	$\hat{\theta}$	6.46%	7.50%	19.49%	4.39%	4.49%	12.33%
	θ_0	3.31%	3.81%	8.56%	1.46%	2.12%	3.69%
PM-5	$\hat{\theta}$		5.68%			9.01%	
	θ_0		5.26%			7.98%	
PM-10	$\hat{\theta}$		9.87%			11.15%	
	θ_0		4.20%			4.14%	

FIGURE 2 – Tableau des niveaux empiriques de rejets - Estimation sur 500 dates

3.3 Analyse générale des niveaux

Si on compare les lignes $\hat{\theta}$ et θ_0, on constate que l'effet de paramètre est en général sensible (rapport de l'ordre de l'unité) mais que sauf exception il joue toujours dans le sens de l'excès : l'effet de paramètre pousse à une prudence exagérée. Le test de Kupiec semble être le plus sensible à l'effet de paramètres.
La qualité des tests est très variable (comparaison de la ligne $\hat{\theta}$ et du nominal) : le test de Markov est proche du nominal tandis que les tests fondés sur des grandeurs intermédiaires (durées et Portmanteau) en sont éloignés.

3.4 Analyse par taille de la fenêtre d'estimation des niveaux

Il s'agit d'analyser tableau par tableau l'évolution des niveaux de rejets lorsque la taille de la fenêtre de backtesting passe de 250 dates à 500 dates. On se concentre sur les lignes $\hat{\theta}$.
Lorsque la fenêtre d'estimation est de 250 dates (fig.1 p44) ou de 500 dates (fig.2 p44) l'augmentation de la taille de la fenêtre de backtesting dégrade en général les niveaux excepté pour le test des durées.
Lorsque l'estimation est effectuée sur 1000 dates, les niveaux obtenus avec un backtesting sur 250 et sur 500 dates sont très proches les uns des autres.

Idem Fig. 1 p44 sauf **estimation sur 1000 dates**

Hypothèses		UC	IND	CC	UC	IND	CC
Kupiec	$\hat{\theta}$	2.23%			2.65%		
	θ_0	1.39%			1.39%		
		1.00%			1.00%		
Markov	$\hat{\theta}$	0.78%	1.00%	1.18%	1.98%	0.97%	2.10%
	θ_0	0.39%	0.95%	1.06%	1.23%	0.92%	1.74%
		1.00%	1.00%	1.00%	1.00%	1.00%	1.00%
Durée	$\hat{\theta}$	4.60%	7.54%	18.37%	2.59%	4.89%	10.64%
	θ_0	3.21 %	3.87%	7.99%	2.48%	2.08%	3.71%
		1.00%	1.00%	1.00%	1.00%	1.00%	1.00%
PM-5	$\hat{\theta}$		5.71%			8.46%	
	θ_0		5.32%			7.95%	
			1.00%			1.00%	
PM-10	$\hat{\theta}$		9.54%			10.57%	
	θ_0		4.26%			4.15%	
			1.00%			1.00%	

FIGURE 3 – Tableau des niveaux empiriques de rejets - Estimation sur 1000 dates

3.5 Analyse par hypothèse des niveaux

De manière générale, l'augmentation de la taille de la fenêtre d'estimation diminue l'effet de paramètre : le niveau empirique (lignes $\hat{\theta}$) se rapproche du niveau nominal (1.00%).

L'hypothèse de couverture inconditionnelle (« UC ») voit ses niveaux empiriques baisser d'environ 50% lorsque l'on passe d'une estimation sur 250 dates à une estimation sur 500 dates et d'une estimation sur 500 dates à une estimation sur 1000 dates.

Les hypothèses d'indépendance (« IND ») et de couverture conditionnelle (« CC ») ont un comportement moins monotone. Les variations sont de l'ordre de 10 à 20 % mais même si la tendance générale est à l'amélioration, certains tests voient leurs niveaux légèrement augmenter.

3.6 Conclusion

Les principaux résultats de cette étude sont :

- L'effet de paramètre est sensible pour des tailles d'échantillons utilisés en pratique.
- Les tests, hypothèses et statistiques, n'ont pas tous le même comportement vis-à-vis de cet effet.
- L'effet de paramètre diminue en général avec l'augmentation de la taille des échantillons.

Chapitre 4

Correction de la VaR

4.1 Comment estimer la VaR ?

La VaR se définit comme le quantile des pertes conditionnelles à l'information disponible :

$$\mathbb{P}\Big(P_t - P_{t-1} < -VaR_{t|t-1}(\gamma)|\mathcal{F}_{t-1}\Big) = \gamma$$

Pour l'estimer, il est nécessaire de choisir un degré de formalisation.

Une approche avec un minimum de formalisation est dite du type non-paramètrique : on ne suppose aucune forme sur les lois suivies par les variables aléatoires. A l'opposé, on peut spécifier avec précision les lois suivies ainsi que leurs paramètres ; on parle alors d'approche paramétrique. La méthode non-paramètrique la plus usitée est la simulation dite historique qui consiste à approcher la VaR par le γ-quantile de la distribution empirique des pertes passées.

Si l'approche paramétrique permet d'obtenir des résultats intéressants, elle expose le flanc à l'estimation des paramètres.

Chaque méthode donnera une estimation différente de la VaR. Il y a un compromis entre la contrainte des hypothèses et la qualité de la VaR prédite.

4.2 Méthode semi-paramétrique

Dans cette méthode, on relâche les hypothèses du modèle utilisé pour l'évolution des log-rendements. On ne spécifie plus la loi suivie par les innovations :

$$\begin{cases} \varepsilon_t = \sigma_t \eta_t \\ \sigma_t^2 = \omega + \alpha \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2 \\ \eta_t \overset{iid}{\sim} \mathcal{L} \\ \varepsilon_0 = \sigma_0 = 0 \end{cases} \tag{4.1}$$

On commence par estimer son modèle sur n dates : on obtient les valeurs $\hat{\theta}_n = (\hat{\omega}_n, \hat{\alpha}_n, \hat{\beta}_n)$. Ces valeurs permettent de construire une suite $(\tilde{\sigma}_t)_t$ à partir des observations ε_t.

On construit alors la repartition empirique des innovations approchées en posant :

$$\tilde{\eta}_t = \frac{\varepsilon_t}{\tilde{\sigma}_t} \tag{4.2}$$

Cette distribution empirique permet de déterminer le quantile d'ordre γ des innovations $(\eta)_t$: $\hat{\Phi}^{-1}(\gamma)$ (en toute rigueur, on devrait noter $\widehat{\Phi^{-1}(\gamma)}$). On obtient alors une formule approchée de la VaR :

$$\widehat{VaR}^{(2)}_{t|t-1}(\gamma) = P_{t-1}(1 - e^{\tilde{\sigma}_t \hat{\Phi}^{-1}(\gamma)}) \tag{4.3}$$

4.3 Méthode paramétrique : cas général

Un méthode sera dite paramétrique lorsqu'elle fait intervenir une modélisation jouant sur des paramètres (facteurs multiplicatifs, paramètres d'une loi,...). Dans cette partie, on considère un modèle

général : la composante « paramètre » n'est présente que de manière formelle sous le symbole θ.

Cela permet d'obtenir une expression de la VaR corrigée de l'effet d'estimation dans le cas général.

4.3.1 Modèle

On considère un modèle général où les rendements (ε_t) sont fonction des rendements passés (synthétisés par ε_{t-1}), de paramètre (θ_0) et d'un facteur aléatoire (η_t). Formellement, cela correspond à ceci :

$$\begin{cases} \varepsilon_t = g(\varepsilon_{t-1}, \theta_0; \eta_t) \\ \eta_t \overset{iid}{\sim} \mathcal{L} \\ \text{à } \eta_{t-1} \text{ et } \theta \text{ fixé, } g \text{ définit une relation croissante bijective de } \eta_t \text{ vers } \varepsilon_t \end{cases} \tag{4.4}$$

On introduit les notations suivantes qui découlent naturellement de la formalisation choisie :

$$\begin{cases} \Phi \text{ la fonction de repartition de la loi } \mathcal{L} \\ a(\varepsilon_{t-1}, \theta_0; .) \text{ la fonction réciproque de } g \text{ sur la troisième composante} \end{cases} \tag{4.5}$$

4.3.2 VaR théorique

On considère la VaR d'horizon 1 jour de niveau de confiance γ. Dans notre cas général, on peut démontrer :

Proposition 4.1 (VaR théorique dans le cas d'un modèle paramètrique)**.**

$$\mathbb{P}\Big(\varepsilon_t < VaR\Big) = \gamma \Leftrightarrow VaR = g(\varepsilon_{t-1}, \theta_0, \Phi^{-1}(\gamma)) \tag{4.6}$$

◇ *Démonstration p93*

4.3.3 VaR estimée

Si l'on se place du point de vue d'un praticien, les aléas sont inobservables. On peut estimer le modèle sur des données pour reconstruire la série des aléas, que l'on note $(\tilde{\eta}_t)_t$. L'effet d'estimation joue au sens où ces aléas construits ne sont pas les aléas réels : la différence introduite va engendrer une altération de la distribution anticipée des pertes. Il faut donc prendre en compte dans le calcul de la VaR cette altération. On peut donc considérer une nouvelle grandeur, notée $\widetilde{VaR}$, qu'on appellera VaR corrigée du biais d'estimation.
Cette VaR vérifie la propriété suivante :

$$\mathbb{P}\Big(\tilde{\eta}_t < a(\varepsilon_{t-1}, \hat{\theta}_n; \widetilde{VaR})\Big) = \gamma \tag{4.7}$$

Le problème est donc de déterminer le $\gamma-$quantile de la variable aléatoire $\tilde{\eta}_t$. Le principe de la méthode utilisée est que la convergence de $\hat{theta}_n$ vers θ permet de voir $\tilde{\eta}_t$ comme la somme de η_t - l'aléa réel- et un terme d'approximation qui tend vers 0. Formellement, cela s'écrit :

$$\left.\begin{array}{l} \eta_t = a(\varepsilon_{t-1}, \theta_0; \varepsilon_t) \\ \tilde{\eta}_t = a(\varepsilon_{t-1}, \hat{\theta}_n; \varepsilon_t) \end{array}\right\} \Rightarrow \tilde{\eta}_t \simeq \eta_t + \overbrace{\frac{\partial a}{\partial \theta}(\varepsilon_{t-1}, \theta_0, \varepsilon_t) \times \underbrace{(\hat{\theta}_n - \theta_0)}_{=\frac{1}{\sqrt{n}}\Omega^{\frac{1}{2}} \underbrace{V_n}_{\rightsquigarrow \mathcal{N}(0, I_p)}}}^{\text{terme d'approximation}} \tag{4.8}$$

La VaR apparait alors comme le quantile de la somme de deux variables aléatoires dont l'une tend en probabilité vers 0. Cette formulation permet d'appliquer le lemme de Rau-Bredow ([4]) qui prolonge le lemme de Gouriéroux-Laurent-Scaillet ([23]) :

Proposition 4.2 (Lemme de Rau-Bredow [4]). *Soient X et Y deux variables aléatoires réelles indépendantes.*

$$\begin{cases} \mathbb{P}(X+hY<Q(\epsilon,\alpha))=\alpha \Rightarrow Q(h,\alpha)=Q(0,\alpha)+h\times\dfrac{\partial Q}{\partial h}(0,\alpha)+\dfrac{h^2}{2}\dfrac{\partial^2 Q}{\partial h^2}(0,\alpha)+o(h^2) \\ avec\ \dfrac{\partial Q}{\partial h}(h,\alpha)=\mathbb{E}(Y|X+hY=Q(h,\alpha)) \\ et\quad \dfrac{\partial^2 Q}{\partial h^2}(\epsilon,\alpha)=-\dfrac{\partial \mathbb{V}(Y|X+hY=s)}{\partial s}\Big]_{s=Q(h,\alpha)}-\Big(\mathbb{V}(Y|X+hY=s)\dfrac{\partial ln f_{X+hY}(s)}{\partial s}\Big)\Big]_{s=Q(h,\alpha)} \end{cases} \tag{4.9}$$

◇ *c.f. A.11 p96*

La première partie de ce lemme est un développement limité d'un quantile. La partie intéressante est l'expression de ces termes qui ont l'avantage de se simplifier en $h=0$.
Ce lemme permet d'obtenir l'expression de la VaR corrigée du biais d'estimation :

Proposition 4.3 (VaR corrigée du biais d'estimation).

$$\widetilde{VaR}=g\Big(\eta_{t-1},\hat{\theta}_n;\ (\Phi^{-1}(\gamma)+\frac{1}{2n}\frac{\partial^2 Q}{\partial h^2}(0,\gamma))\ \Big) \tag{4.10}$$

Avec l'expression suivante pour $\frac{\partial^2 Q}{\partial h^2}(0,\gamma))$:

$$\begin{cases} \dfrac{\partial^2 Q}{\partial h^2}(0,\gamma)=-\Big[\dfrac{\partial}{\partial s}\Big(\dfrac{\partial a}{\partial \theta'}(\varepsilon_{t-1},\theta_0,g(\varepsilon_{t-1},\theta_0,s)\times\Omega\times\dfrac{\partial a}{\partial\theta}(\varepsilon_{t-1},\theta_0,g(\varepsilon_{t-1},\theta_0,s)\Big)\Big]_{s=\Phi^{-1}(\gamma)} \\ \qquad -\dfrac{\partial a}{\partial\theta'}(\varepsilon_{t-1},\theta_0,VaR)\times\Omega\times\dfrac{\partial a}{\partial\theta}(\varepsilon_{t-1},\theta_0,VaR)\times\Big[\dfrac{\partial ln f_{\eta_t}(s)}{\partial s}\Big]_{s=\Phi^{-1}(\gamma)} \end{cases} \tag{4.11}$$

◇ *Démonstration p94*

4.4 Modèle paramètrique : cas d'un GARCH(1,1)$-\mathcal{N}(0,1)$

La partie précédente aboutit à une formulation générale de la VaR corrigée de l'effet d'estimation. Cette partie complète l'approche théorique par l'application au cas d'un GARCH(1,1)$-\mathcal{N}(0,1)$.

4.4.1 Expression générale

Pour l'instant, on va réduire la portée générale du modèle en considérant un modèle de type GARCH : l'aléa n'intervient que de manière multiplicative (il y a en quelque sorte séparabilité d'une variable). Formellement on va donc retenir Dans le cas d'un modèle de type GARCH :

$$\begin{cases} \varepsilon_t=\sigma_t(\theta_0)\eta_t \\ \eta_t \overset{iid}{\sim} \mathcal{L} \text{ avec } \mathcal{L} \text{ d'espérance nulle} \end{cases} \tag{4.12}$$

Ces hypothèses permet d'exprimer les différents termes intervenant dans l'expression de la VaR corrigée de l'effet d'estimation :

Proposition 4.4 (VaR corrigée de l'effet d'estimation dans le cas d'un modèle de type GARCH).

$$\widetilde{VaR}=\sigma_t(\hat{\theta}_n)\Phi^{-1}(\gamma)\Big(1-\frac{1}{n}\frac{1}{2\sigma_t(\theta)^2}\frac{\partial\sigma_t(\theta)}{\partial\theta'}(\theta_0)\Omega\frac{\partial\sigma_t(\theta)}{\partial\theta}(\theta_0)\times(2-\Phi^{-1}(\gamma)^2)\Big) \tag{4.13}$$

◇ *Démonstration p95*

4.4.2 Estimateur de la VaR corrigée

La formule précédente n'est pas applicable en pratique pour deux raisons. Premièrement, elle fait intervenir des grandeurs en θ_0 qui est inconnu. Deuxièmement, la matrice Ω qui traduit la variabilité de la méthode d'estimation de θ_0 par $\hat{\theta}_n$ est aussi inconnue.

Pour le premier problème, la solution la plus simple à mettre en place pour un praticien consiste à considérer les valeurs en $\hat{\theta}_n$ c'est-à-dire avec la meilleure approximation des paramètres dont il dispose. Il s'agit effectivement d'estimer l'erreur que l'on fait en estimant : évidemment, si on estime mal le terme correctif, il se peut que corriger le biais d'estimation soit moins performant que de négliger ce biais. Ce point sera abordé dans la partie suivante.

Concernant le second problème, il se résout en deux étapes. Tout d'abord la forme asymptotique de Ω est déterminée par la méthode d'estimation. Dans le cas usuel, ie par maximisation de la vraisemblance, cette forme asymptotique peut être estimée. Ici aussi, on estime la correction du biais d'estimation.

Le choix d'une estimation par maximisation de la vraisemblance et d'une loi pour les résidus -ici une gaussienne centrée réduite- permet d'obtenir une estimation de la VaR corrigée du biais d'estimation, que l'on note $\widehat{\widetilde{VaR}}$:

Proposition 4.5 (Estimateur de la VaR corrigée du biais d'estimation)**.**

$$\widehat{\widetilde{VaR}} = \tilde{\sigma}_t(\hat{\theta}_n)\Phi^{-1}(\gamma)\Big(1 - \frac{1}{n}\frac{1}{2\tilde{\sigma}_t(\hat{\theta})^2}\frac{\partial\tilde{\sigma}_t(\hat{\theta})}{\partial\theta'}(\hat{\theta}_n)\hat{\Omega}_n\frac{\partial\tilde{\sigma}_t(\hat{\theta})}{\partial\theta}(\hat{\theta}_n) \times (2 - \Phi^{-1}(\gamma)^2)\Big) \tag{4.14}$$

Avec comme estimateur de Ω la statistique suivante :

$$\begin{cases} \hat{\Omega}_n = (\hat{\kappa} - 1)(\hat{J}_n)^{-1} \\ \hat{\kappa}_n = \dfrac{1}{n}\displaystyle\sum_{t=1}^{n}\left(\frac{\varepsilon_t}{\tilde{\sigma}_t}\right)^4 \\ \hat{J}_n = \dfrac{1}{n}\displaystyle\sum_{t=1}^{n}\left(\frac{1}{\tilde{\sigma}_t^4}\frac{\partial\tilde{\sigma}_t^2(\hat{\theta})}{\partial\theta'}(\hat{\theta}_n)\frac{\partial\tilde{\sigma}_t^2(\hat{\theta})}{\partial\theta}(\hat{\theta}_n)\right) \end{cases} \tag{4.15}$$

◇

4.5 Comparaison des différentes VaR

Cette partie a pour objectif de comparer les différentes manières de corriger le biais d'estimation. Pour cela, on a calculé les différentes VaR sur un modèle parfaitement connu.

4.5.1 Modèle sous-jacent

Le modèle qui nous a servi à générer les trajectoires est le modèle qui sert de fil directeur à cette étude. Ainsi, les log-rendements suivent le processus GARCH(1,1) suivant :

$$\begin{cases} \varepsilon_t = ln(\frac{P_t}{P_{t-1}}) \\ \varepsilon_t = \sigma_t\eta_t \\ \sigma_t^2 = 0.0001 + 0.08\varepsilon_{t-1}^2 + 0.8\sigma_{t-1}^2 \\ \eta_t \overset{iid}{\sim} \mathcal{N}(0,1) \\ \eta_0 = \sigma_0 = 0 \end{cases} \tag{4.16}$$

4.5.2 Procédure

Une trajectoire correspond à la réalisation des variables aléatoires pour $t \in [\![1;T]\!]$. Cette plage est découpée en deux : sur $[\![1;n]\!]$ on mène une estimation des paramètres du modèles GARCH tandis que la VaR est prédite sur la plage $[\![n+1;T]\!]$ (plage de backtesting).

On calcule les 4 VaR suivantes :

- VaR : Il s'agit de la VaR théorique :

$$VaR(\gamma) = P_{t-1}(1 - e^{\sigma_t\Phi^{-1}(\gamma)}) \tag{4.17}$$

- $\widehat{VaR}^{(1)}$: Il s'agit d'une estimation de la VaR où on remplace σ_t par sa valeur estimé $\tilde{\sigma}_t$.

$$\widehat{VaR}^{(1)}(\gamma) = P_{t-1}(1 - e^{\tilde{\sigma}_t \Phi^{-1}(\gamma)}) \tag{4.18}$$

- $\widehat{VaR}^{(2)}$: Il s'agit de la VaR déterminée par l'approche semi-paramétrique c'est-à-dire où aucune hypothèse n'est faite sur la loi des résidus empiriques.

$$\widehat{VaR}^{(2)} = P_{t-1}(1 - e^{\tilde{\sigma}_t \hat{\Phi}^{-1}(\gamma)}) \tag{4.19}$$

- $\widehat{\widehat{VaR}}$: Il s'agit de la VaR où l'on estime la correction à apporter du fait de l'estimation par un développement du quantiles.

$$\widehat{\widehat{VaR}}^{(1)} = P_{t-1}(1 - e^{\Phi^{-1}(\gamma)(1+\frac{1}{2n}\frac{\partial^2 Q}{\partial h^2}(0,\gamma))}) \tag{4.20}$$

4.5.3 Résultats

On a effectué 10 000 simulations. Pour chaque simulations, on a calculé les différente VaR. On a ensuite fait la moyenne temporelle des écarts relatifs à la VaR théoriques. Le tableau suivant (fig.1 p51) présente les grandeurs suivantes :

- Moyenne des écarts relatifs moyens avec la VaR usuelle :

$$\Delta^{(1)} = \frac{1}{M}\sum_{i=1}^{M} \underbrace{\frac{1}{T-n}\sum_{t=n+1}^{T} \frac{\left|\widehat{VaR}_i^{(1)}(t) - VaR_i(t)\right|}{VaR_i(t)}}_{\text{moyenne temporelle}} \tag{4.21}$$

- Moyenne des écarts relatifs moyens avec la VaR semi-paramétrique :

$$\Delta^{(2)} = \frac{1}{M}\sum_{i=1}^{M} \underbrace{\frac{1}{T-n}\sum_{t=n+1}^{T} \frac{\left|\widehat{VaR}_i^{(2)}(t) - VaR_i(t)\right|}{VaR_i(t)}}_{\text{moyenne temporelle}} \tag{4.22}$$

- Moyenne des écarts relatifs moyens avec la VaR corrigée de l'effet d'estimation :

$$\Delta^{(3)} = \frac{1}{M}\sum_{i=1}^{M} \underbrace{\frac{1}{T-n}\sum_{t=n+1}^{T} \frac{\left|\widehat{\widehat{VaR}}_i^{(1)}(t) - VaR_i(t)\right|}{VaR_i(t)}}_{\text{moyenne temporelle}} \tag{4.23}$$

Estimation d'un GARCH(1,1)–$\mathcal{N}$

de paramètres $(10^{-4}; 0.08; 0.8)$

fenêtre d'estimation n	fenêtre de backtesting $T-n$	Ecarts moyens $\Delta^{(1)}$	$\Delta^{(2)}$	$\Delta^{(3)}$
250	250	0.066 789 390	0.187 351 315	0.066 785 930
250	500	0.065 068 220	0.251 964 952	0.065 067 759
500	250	0.051 389 779	0.107 725 268	0.051 538 528
500	500	0.048 667 383	0.115 717 731	0.048 667 896

FIGURE 1 – Ecarts relatifs des VaR - 10000 dates

Premièrement, la méthode semi-paramétrique est peu efficace. L'écart par rapport à la VaR théorique est de l'ordre de 10%. Néanmoins, on peut supposer que par sa construction elle est plus robuste en cas de mauvaise spécification du modèle.
Deuxièmement, l'écart entre les écarts relatifs de la VaR usuelle et de la VaR corrigée de l'effet d'estimation est très faible : sur un écart à la VaR théorique de l'ordre de 5%, seule la 6ème décimale varie. Le sens de variation est celui d'une amélioration car la VaR corrigée a tendance à être plus proche de la VaR théorique mais l'effet est insensible.

4.6 Conclusion sur la correction de la VaR

La correction de l'effet d'estimation de la VaR correspond à la compréhension des fondements théoriques de la mesure de risque. Cela offre une application originale des travaux sur la granularité des portefeuilles dont est issu le principal résultat (Lemme de Rau-Bredow [4]). Néanmoins, d'un point de vue pratique les résultats ne sont guère intéressants : la correction est tellement minime qu'il est difficile de savoir si les écarts ne sont pas dus aux aléas des expériences.
Le seul avantage est qu'on peut malgré tout estimer que cette méthode fournit une meilleure estimation de la VaR. Comme elle a tendance à être plus petite que la VaR usuelle, les provisions obligatoires qui en découlent sont plus faibles. L'écart peut sembler négligeable mais multiplié par le bilan d'une institution financière, la somme n'est pas si minime.

Chapitre 5

Variance asymptotique

5.1 Éléments théoriques

Cette partie (Éléments théoriques) se fonde sur l'article *Estimation Risk Effects on Backtesting For Parametric Value-at-Risk Models* ([21]) de Juan-Carlos Escanciano et Jose Olmo, paru en 2007 au Center for Applied Economics and Policiy Research de l'Université de l'Indiana (États-Unis d'Amérique).

Cet article se concentre sur une statistique apparentée à celle de Kupiec. Les auteurs analysent le comportement de cette statistique à distance finie et mènent une analyse de l'effet d'estimation. On peut distinguer trois résultats dans la partie théorique de l'article : la décomposition risque de paramètre-risque de modèle, la correction de la variance pour un test de l'hypothèse de couverture inconditionnelle et un résultat sur un test de l'hypothèse d'indépendance.

5.1.1 Décomposition de la statistique

Prise en compte des paramètres

Dans cette partie, on ne considère que la VaR à 1 jour. Une fois un modèle paramétrique arrêté, son calcul dépend de la valeur des paramètres ; ainsi on note :

- $VaR_{t,\gamma}(\theta_0)$: la VaR de niveau γ calculée pour la date t avec les valeurs idéales des paramètres (θ_0).
- $VaR_{t,\gamma}(\theta_n)$: la VaR de niveau γ calculée pour la date t avec les valeurs estimées des paramètres sur la fenêtre $[\![1;n]\!]$ (θ_n).

On définit alors les indicatrices de dépassement associées :

$$\begin{cases} I_{t,\gamma}(\theta_0) := \mathbb{1}_{PnL_t < -VaR_{t,\gamma}(\theta_0)} \\ I_{t,\gamma}(\theta_n) := \mathbb{1}_{PnL_t < -VaR_{t,\gamma}(\theta_n)} \end{cases} \tag{5.1}$$

Statistique

On pose la statistique S_n, fortement apparentée à la statistique de z de Kupiec :

$$S_n = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} (I_{t,\gamma}(\theta_n) - \gamma) \tag{5.2}$$

On remarquera que dans cette approche il n'y a pas de distinction entre une fenêtre d'estimation des paramètres et une fenêtre de backtesting.

On complète les notations en introduisant la fonction de répartition des pertes conditionnelles :

$$F_{t-1}(y) = \mathbb{P}(PnL_t \leq y | \mathcal{F}_{t-1}) \tag{5.3}$$

Décomposition

Sous certaines hypothèses, vérifiées pour les modèles usuels, les auteurs établissent la formule suivante :

$$\begin{cases} S_n = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \Big(I_{t,\gamma}(\theta_0) - F_{t-1}(VaR_{t,\gamma}(\theta_0)) \Big) + EP + RM + o_{\mathbb{P}}(1) \\ EP = \underbrace{\sqrt{n}(\hat{\theta}_n - \theta_0)' \mathbb{E}\Big(\frac{\partial VaR_{t,\gamma}(\theta_0)}{\partial \theta} F'_{t-1}(VaR_{t,\gamma}(\theta_0)) \Big)}_{\text{effet de paramétres}} \\ RM = \underbrace{\frac{1}{\sqrt{n}} \sum_{t=1}^{n} \Big(F_{t-1}(VaR_{t,\gamma}(\theta_0)) - \gamma \Big)}_{\text{risque de modèle}} \end{cases} \tag{5.4}$$

Cette décomposition fait clairement apparaître la grandeur cible (premier terme), l'effet de paramètre (terme EP) qui tend vers 0 lorsque n tend vers l'infini puisque θ_0 est la limite de $\hat{\theta}_n$ et le risque de modèle qui traduit un mauvaise spécification.
Les auteurs remarquent que le risque de modèle est plus important qu'il n'y paraît parce que il s'agit d'une somme de variable aléatoire et non de leurs espérances.

5.1.2 Couverture inconditionnelle

L'hypothèse de couverture inconditionnelle s'écrit :

$$H_{0,uc} : \text{« } \mathbb{E}(I_{t,\gamma}(\theta_0)) = \gamma \text{ »} \tag{5.5}$$

J.-C. Escanciano et J. Olmo démontrent la convergence suivante :

$$\begin{cases} S_n = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \Big(I_{t,\gamma}(\hat{\theta}_n) - \gamma \Big) \overset{H_{0,uc}}{\rightsquigarrow} \mathcal{N}(0, \varsigma_c^2) \\ \varsigma_c^2 = \underbrace{\gamma(1-\gamma)}_{\text{variance usuelle}} + \underbrace{2A\rho + AVA'}_{\text{terme complémentaire}} \end{cases} \tag{5.6}$$

Les termes A, V et ρ nécessitent d'être explicités. Pour cela, on doit exprimer le fait que les estimations $(\theta_n)_n$ suivent asymptotiquement une expansion de Bahadur (c'est une des hypothèse de travail des auteurs) :

$$\sqrt{n}(\hat{\theta}_n - \theta_0) = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \ell(PnL_t, \mathcal{F}_{t-1}, \theta_0) + o_{\mathbb{P}}(1) \tag{5.7}$$

On peut alors exprimer A, V et ρ ainsi :

$$\begin{cases} \rho = \mathbb{E}\Big((I_{t,\gamma}(\theta_0) - \gamma) \ell(PnL_t, \mathcal{F}_{t-1}, \theta_0) \Big) \\ A = \mathbb{E}\Big(\frac{\partial VaR_{t,\gamma}(\theta_0)}{\partial \theta} F'_{t-1}(VaR_{t,\gamma}(\theta_0)) \Big) \\ V = \mathbb{E}\Big(\frac{\partial \ell(PnL_t, \mathcal{F}_{t-1}, \theta_0)}{\partial \theta} \frac{\partial \ell(PnL_t, \mathcal{F}_{t-1}, \theta_0)}{\partial \theta'} \Big) \end{cases} \tag{5.8}$$

L'article de J.-C. Escanciano et J. Olmo propose donc une procédure de backtesting exempt du risque de paramètre.

5.1.3 Indépendance

Le dernière résultat théorique majeur de l'article porte sur un test de l'indépendance. Les auteurs proposent dans une approche de type Portmanteau de tester la nullité des coefficients d'auto-corrélation. Les auteurs proposent d'utiliser comment "coefficients d'auto-corrélation" les grandeurs suivantes :

$$\gamma_{n,j} = \frac{1}{n-j} \sum_{t=j+1}^{n} (I_{t,\gamma}(\hat{\theta}_n) - \gamma)(I_{t-j,\gamma}(\hat{\theta}_n) - \gamma) \tag{5.9}$$

On remarquera que par rapport au test de Portemanteau présenté précédemment, ces coefficients présupposent que l'espérance de $I_{t,\gamma}(\hat{\theta}_n)$ est γ.
Le résultat principale est que le test d'indépendance fondé sur la statistique suivante

$$C_{n,m} = \sum_{j=1}^{m} \left(\frac{\gamma_{n,j}}{\gamma(1-\gamma)} \right)^2 \tag{5.10}$$

n'est pas soumis à un effet de paramètre.
Mais, il suffit de calculer $C_{n,m}$ avec des coefficients d'auto-corrélation obtenus par

$$\gamma^{(2)}_{n,j} = \frac{1}{n-j} \sum_{t=j+1}^{n} I_{t,\gamma}(\hat{\theta}_n) I_{t-j,\gamma}(\hat{\theta}_n) - \gamma^2 \tag{5.11}$$

pour que ce résultat ne soit plus valable

5.2 Correction de la variance

Cette section renvoie à l'expérience n°3 (p84).

5.2.1 Mise-en-oeuvre de la correction de la variance

Démarche théorique

Le principal point est d'avoir des estimateurs des grandeurs A, V et ρ nécessaires au calcul de la variance corrigée.
Dans le cas d'un modèle GARCH estimé par maximisation de la vraisemblance, on peut établir la formule suivante :

$$\begin{cases} \sqrt{n}(\theta_n - \theta_0) \xrightarrow{d} \mathcal{N}(0, (\kappa - 1)J^{-1}) \\ J = \mathbb{E}(\frac{\partial \ell_t}{\partial \theta} \frac{\partial \ell_t}{\partial \theta'}) \\ \kappa = \mathbb{E}(\eta_t^4) \\ \ell_t = -\frac{1}{2} ln(2\pi) - \frac{1}{2}(ln\sigma_t^2 + \frac{\varepsilon_t^2}{\sigma_t^2}) \end{cases} \tag{5.12}$$

Puisque $V = (\kappa - 1)J^{-1}$, on peut donc construire un estimateur de V en combinant des estimateurs de κ et de J.

Les calculs aboutissent à :

$$\begin{cases} \hat{\kappa}_n = \frac{1}{n}\sum_{t=1}^{n}\left(\frac{\varepsilon_t}{\tilde{\sigma}_t}\right)^4 \\ \hat{J}_n = \frac{1}{n}\sum_{t=1}^{n}\left(\frac{1}{\tilde{\sigma}_t^4}\frac{\partial\tilde{\sigma}_t^2}{\partial\theta'}\frac{\partial\tilde{\sigma}_t^2}{\partial\theta}\right) \\ \hat{V}_n = (\hat{\kappa}-1)(\hat{J}_n)^{-1} \\ \hat{\rho}_n = \frac{1}{n}\sum_{t=1}^{n}\left((\mathbb{1}_{PnL_t<-VaR(\hat{\theta})(\gamma)}(t)-\gamma)\times(\hat{J}_n)^{-1}\frac{1}{\tilde{\sigma}_t^2}(1-\frac{\varepsilon_t^2}{\tilde{\sigma}_t^2})\frac{\partial\tilde{\sigma}_t^2}{\partial\theta}\right) \\ \hat{A}_n = \frac{1}{n}\sum_{t=1}^{n}\left(\frac{\Phi^{-1}(\gamma)e^{-\Phi^{-1}(\gamma)^2}}{2\sqrt{2\pi}\tilde{\sigma}_t^2}\frac{\partial\tilde{\sigma}_t^2}{\partial\theta}(\hat{\theta}_n)\right) \\ \text{avec } \frac{\partial\tilde{\sigma}_t^2}{\partial\theta} = \begin{pmatrix}1\\ \varepsilon_t^2\\ \tilde{\sigma}_t^2\end{pmatrix}\tilde{\beta}\times\frac{\partial\tilde{\sigma}_{t-1}^2}{\partial\theta} \\ \quad \frac{\partial\tilde{\sigma}_0^2}{\partial\beta} = 0 \\ \quad \tilde{\sigma}_0^2 = 0 \quad \tilde{\sigma}_t^2 = \hat{\omega}_n + \hat{\alpha}_n\varepsilon_{t-1}^2 + \hat{\beta}_n\tilde{\sigma}_{t-1}^2 \end{cases} \tag{5.13}$$

On a alors un estimateur de la variance corrigée de la statistique de Kupiec :

$$\widehat{\varsigma_c^2} = \gamma(1-\gamma) + 2\hat{A}_n\hat{\rho}_n + \hat{A}_n'\hat{V}_n\hat{A}_n \tag{5.14}$$

Démarche algorithmique

1. Choix de ω, α et β
2. Choix de n et de T $(n < T)$
3. Génération de σ_t^2 et ε_t pour $t \in [\![1;T[\![$
4. Estimation des paramètres $\hat{\omega}_n$, $\hat{\alpha}_n$ et $\hat{\beta}_n$ pour sur la plage $[\![1;n]\!]$
5. Déclarer vecteur $\tilde{\sigma}^2$
6. $\tilde{\sigma}^2(1) \leftarrow 0$
7. Déclarer vecter VaR
8. $VaR(1) \leftarrow 0$
9. pour i allant de 2 à T, faire
10. $\quad \tilde{\sigma}^2(i) \leftarrow \hat{\omega}_n + \hat{\alpha}_n\varepsilon(i-1)^2 + \hat{\beta}_n\tilde{\sigma}^2(i-1)$
11. $\quad VaR(i) \leftarrow cours(i-1)(1-exp(\sqrt{\tilde{\sigma}^2(i)}\Phi^{-1}(\gamma)))$
12. fin pour
13. Calcul de l'indicatrice $I(t)$ pour $t \in [\![n+1;T]\!]$
14. Déclarer κ réel
15. $kappa \leftarrow 0$
16. pour i allant de 2 à n, faire
17. $\quad \kappa \leftarrow \kappa + \varepsilon(i)^4/(\tilde{\sigma}_t^2)^2$
18. fin pour
19. $\kappa \leftarrow \kappa/(n-1)$
20. Déclarer der_omega, der_alpha et der_beta vecteurs réels initialisés à 0
21. pour i allant de 2 à n, faire
22. $\quad der_omega(i) = 1 + \hat{\beta}_n \times der_omega(i-1)$
23. $\quad der_alpha(i) = \varepsilon(i-1)^2 + \hat{\beta}_n \times der_alpha(i-1)$
24. $\quad der_beta(i) = \tilde{\sigma}^2(i-1) + \hat{\beta}_n \times der_beta(i-1)$
25. fin pour
26. $der_omega(1) \leftarrow moyenne(der_omega)$

27. $der_alpha(1) \leftarrow moyenne(der_alpha)$
28. $der_beta(1) \leftarrow moyenne(der_beta)$
29. pour i allant de 2 à n, faire
30. $\quad der_omega(i) = 1 + \hat{\beta}_n \times der_omega(i-1)$
31. $\quad der_alpha(i) = \varepsilon(i-1)^2 + \hat{\beta}_n \times der_alpha(i-1)$
32. $\quad der_beta(i) = \tilde{\sigma}^2(i-1) + \hat{\beta}_n \times der_beta(i-1)$
33. fin pour
34. Déclarer l matrice réelle 3 lignes sur n colonnes initialisée à 0
35. pour i allant de 2 à n, faire
36. $\quad l(.,i) \leftarrow (1/\tilde{\sigma}^2(i)) \times [der_omega(i) der_alpha(i) der_beta(i)]'$
37. fin pour
38. Déclarer J matrice réelle 3×3 initialisée à 0
39. pour i allant de 2 à n, faire
40. $\quad J \leftarrow J + l(.,i) \times l(.,i)'$
41. fin pour
42. Déclarer J_inv
43. $J_inv \leftarrow inverse(J) \times (n-1)$
44. Déclarer ρ vecteur réel de taille 3 initialisé à 0
45. pour i allant de 2 à n, faire
46. $\quad \rho \leftarrow \rho + (I(i) - \gamma) \times J_inv \times l(.,i)$
47. fin pour
48. $\rho \leftarrow \rho/(n-1)$
49. Déclarer A vecteur réel de taille 3 initialisé à 0
50. Déclarer $facteur_commun$
51. $facteur_commun \leftarrow \Phi^{-1}(\gamma) \times exp(-\Phi^{-1}(\gamma)^2)/(2\sqrt{2\pi})$
52. pour i allant de 2 à n, faire
53. $\quad A \leftarrow A + facteur_commun \times l(.,i)$
54. fin pour
55. $A \leftarrow A/(n-1)$
56. Déclarer $variance_corrigee$ réel
57. $variance_corrigee \leftarrow \gamma(1-\gamma) + 2A'\rho + A'(\kappa - 1)VA$

5.2.2 Résultats sur la variance en fonction des paramètres

Les éléments théoriques montrent que ς_c^2 ne dépend que des valeurs des paramètres du modèles. Malheureusement, il n'existe pas d'expression calculable des termes de la correction. Au demeurant, on peut utiliser les estimateurs avec les valeurs correctes des paramètres (sans estimation) sur une très grande plage.
Les éléments fournissent pour quelques valeurs de paramètres une estimation de la variance corrigée du test de Kupiec que l'on note $\widetilde{\varsigma_c^2}$. Chaque estimation a été menée sur une plage de 1 000 000 dates

5.2.3 Résultats sur la variance en fonction de la taille de l'échantillon

Le tableau suivant (fig.2 p58) présente les diverses variances en fonction de la taille de la fenêtre d'estimation et de la taille de la fenêtre de backtesting (première et deuxième colonne). La troisième colonne indique la variance nominale ($\gamma(1-\gamma)$), ie celle qui est spontanément utilisée. La quatrième colonne renseigne la variance théorique estimée ($\widetilde{\varsigma_c^2}$) calculée selon la méthode de la sous- section précédente. Sur la cinquième colonne, on peut lire la variance constatée de la distribution des statistiques de Kupiec sur l'ensemble des simulations ($\mathbb{V}(z^{(i)})$). La sixième colonne présente la variance corrigée moyenne ($\widehat{\varsigma_c^2}$)
Le résultat est très net : en dessous de 750 dates pour la fenêtre d'estimation, la variance estimée théorique est largement éloignée de la variance constatée. Les résultats sur les niveaux ne sont pas présentés dans la mesure où ils sont très similaires entre eux et très mauvais malgré la correction.
Au demeurant, l'estimation de la variance corrigée ($\widehat{\varsigma_c^2}$) est tout-à-fait correcte comme le montre le faible écart entre la colonne $\widehat{\varsigma_c^2}$ et la colonne $\widetilde{\varsigma_c^2}$.

ω	α	β	$\alpha+\beta$	$\widetilde{\varsigma}_c^2$
0.0001	0.08	0.	0.08	0.009 908 734
0.0001	0.08	0.8	0.88	0.009 908 644
0.0001	0.1	0.8	0.90	0.009 908 242
0.0001	0.05	0.9	0.95	0.009 908 612
0.0000011	0.0754	0.9181	0.9935	0.009 908 971
0.0001	0.1	0.9	1.00	0.009 910 026

Figure 1 – Variances du test de Kupiec en fonction de la taille de l'échantillon

Correction du test de Kupiec pour GARCH(1,1)–$\mathcal{N}$ de paramètres $(10^{-4}; 0.08; 0.8)$

10 000 simulations

fenêtre d'estimation n	fenêtre de backtesting T-n	Variance nominale $\gamma(1-\gamma)$	Variance théorique estimée $\widetilde{\varsigma}_c^2$	Variance constatée $\mathbb{V}(z^{(i)})$	Variance corrigée estimée $\widehat{\varsigma}_c^2$
250	250	0.009 900	0.009 908	0.027 669	0.009 947
250	500	0.009 900	0.009 908	0.048 269	0.009 949
500	250	0.009 900	0.009 908	0.031 315	0.009 945
500	500	0.009 900	0.009 908	0.021 343	0.009 946
750	250	0.009 900	0.009 908	0.011 883	0.009 951

Figure 2 – Variances du test de Kupiec en fonction de la taille de l'échantillon

5.3 Elimination du caractère « à distance finie »

Les résultats précédents aboutissent à un double constat : pour chaque trajectoire, l'estimation de la variance corrigée est convenable mais au global la variance constatée est éloignée de la variance corrigée asymptotique. Cette ambivalence s'explique par le fait que la variance corrigée asymptotique ne repose pas sur la construction des statistiques de Kupiec et est donc insensible à la faible taille de la fenêtre de backtesting.
L'effet d'estimation concerne uniquement le fait que la fenêtre d'estimation est réduite mais le caractère asymptotique de la convergence des statistiques de Kupiec persiste. Afin de vérifier que c'est la faible taille de la fenêtre de backtesting qui dilue nettement la correction apportée on a mené l'expérience en autorisant une très grande plage de backtesting. Cela ne rend pas compte d'une situation pratique mais vise à identifier l'origine de la piètre qualité de la procédure.

5.4 Conclusion : risque de paramètre

La correction du test de Kupiec proposée par Escanciano et Olmo ([21]) n'est pas pertinente pour un modèle de type GARCH(1,1) avec des tailles d'échantillon envisageables en pratique.

Premièrement, on notera que l'illustration donnée dans l'article proposé n'est pas un modèle GARCH(1,1). Ensuite, la correction apportée ne concerne que l'effet d'estimation, ie la faible taille de la fenêtre d'estimation, mais ne corrige en rien le fait que la fenêtre de backtesting est elle-même très réduite. Cela altère peut-être suffisamment la distribution des statistiques de Kupiec pour diluer la correction apportée.

Chapitre 6

Conclusion III

Cette partie a traité deux conséquences problématiques disjointes de l'effet d'estimation : d'une part la prédiction de la VaR et d'autre par la variance asymptotique du test de Kupiec.

Concernant la prédiction de la VaR, l'analyse mise en oeuvre est progressive de manière à ce que chaque conséquence du choix d'un modèle GARCH(1,1) à aléas gaussiens soit clairement identifiée. D'un point de vue théorique, la correction de la VaR est un développement limité dont les facteurs ont l'avantage d'avoir une expression simple dans le cas de la VaR. Néanmoins, la mise en oeuvre sur machine n'a pas été concluante. On aurait espéré que la proportion de dépassement se rapproche sensiblement du seuil de confiance. Aucun bénéfice au niveau du backtesting n'est à attendre : seul la volonté d'un calcul « au plus juste » peut justifier l'implementation de cette formule.

La correction de la variance asymptotique du test de Kupiec est une application des éléments développés par Escanciano et Olmo ([21]) avec un modèle sous-jacent différent. Les résultats sont très différents de ces auteurs puisqu'aucune amélioration n'est perceptible. Néanmoins, il apparaît que le caractère à distance finie est peut-être diluée dans l'approche de ces auteurs. Dans leurs article, ils proposent d'utiliser la même plage de données pour l'estimation et pour le backtesting. Ce choix ne correspond à aucun cas pratique des procédures de backtesting dans les institutions financières. Dans ce mémoire de recherche, on a préféré découpler la période d'estimation de la période de backtesting, cela augmente peut-être l'effet de perturbation dû à la faible taille de la fenêtre de backtesting.

Quatrième partie

Conclusion

La problématique de ce mémoire de recherche était de déterminer l'influence de l'effet d'estimation. Cet objectif a permis d'analyser les différentes procédures de backtesting et de dégager quelques points importants.

Premièrement, même si cela ne répond pas directement à la problématique, ce mémoire détaille avec précision les tests usuels de backtesting. Si chacun cherche à tester la validité de la VaR, il est intéressant de constater que les hypothèses nulles et/ou alternatives sont en réalités des cas spécifiques et très disparates même pour des tests portant sur le même objet (indépendance,...).
Deuxièmement, la VaR estimée a une tendance conservatrice dans la mesure où elle est plus prudente que la VaR réelle. Ce décalage persiste au niveau des procédures de backtesting. De manière générale, les procédures ont tendance à être bien trop strictes. L'effet d'estimation joue donc grandement.

Troisièmement, l'influence de l'effet d'estimation étant établie, on a proposé deux pistes de correction. Ces deux pistes sont indépendantes et concernent d'une part une meilleure estimation de la VaR et d'autre part une meilleure qualité au niveau du backtesting. Chaque piste possède une cadre théorique propre mais repose au fond sur une décomposition additive : la valeur calculée n'est que la somme de la « vraie » valeur et d'un terme parasite dont on connaît le comportement (il tend vers 0 lorsque la taille de l'échantillon d'estimation tend vers l'infini). Il faut noter que la connaissance de ce comportement est du à des résultats sur la convergence de méthodes d'estimation par maximisation de la vraisemblance. Malgré des qualités théoriques, ces corrections ont en pratique été inefficaces : pour des situations similaires à celles observées en pratique dans les banques, le gain est minime, voire nul.

Pour prolonger cette étude, on pourrait essayer de neutraliser l'effet « à distance fini » en augmentant artificiellement la fenêtre de backtesting. Certes on s'écarterait d'un cas similaire aux cas pratiques mais cela permettrait de déterminer l'origine de la piètre performance de ces améliorations.

Cinquième partie

Références

Bibliographie

[1] Monfort A (1982). *Cours de statistique mathématique*. Economica.

[2] Kupiec P (1995). Techniques for Verifying the Accuracy of Risk Measurement Models. *Journal of Derivatives*, pages 73–84.

[3] Jorion P (2000). *Value-at-Risk : the New benchmark of Managing Financial Risk* . Springer Boston, McGraw-Hill Compagnies.

[4] Rau-Bredow H (2005). Unsystematic Credit Risk and Coherent Risk Measure.

[5] RiskGlossary (2009). Basel Committee on Banking Supervision. http ://www.riskglossary.com/.

[6] Sollow R (2009). L'économie entre empirisme et mathématisation. *Le Monde*.

[7] Abramowitz M, Stegun A (1964). *Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables*. http ://www.math.ucla.edu/ cbm/aands/.

[8] Berkowitz J, Chrisoffersen P, Pelletier D (2008). Evaluating Value-at-Risk Models with Desk-Level Data.

[9] Berkowitz J, O'Brien J (2005). *Risks of Financial Institutions*, chapter Bank Trading Revenues, VaR and Market Risk. R. Stulz, M. Carey, Cambridge, MA.

[10] Black F, Scholes M (1973). The Princing of Options and Corporate Liabilities. *Journal of Political Economy*, 81(3) :637–654.

[11] Bontemps C, Meddahi N (2005). Testing normality : A GMM approach. *Journal of Econometrics*, 124 :149–186.

[12] Boorstin D (1992). *Les Découvreurs*. Robert Lafont.

[13] Campbell S (2005). A Review of Backtesting and Backtesting Procedures. Finance and Economics Discussion Series, Federal Reserve Board.

[14] Candelon B, Colletaz G, Hurlin C, Tokpavi S (2008). Backetesting Value-at-risk : a GMM duration-based test.

[15] Christoffersen P (1998). Evaluating Interval Forecasts. *International Economic Review*, 39 :841–862.

[16] Christoffersen P (2008). *Encyclopedia of Quantitative finance*, chapter Backtesting. Rama Cont, John Wiley&Sons, Ltd.

[17] Christoffersen P Pelletier D (2004). Backtesting Value-at-Risk : A Duration-Based Approach. *Journal of Financial Econometrics*, 2(1) :84–118.

[18] Crnkovic D (1997). *VaR : understandig and applying Value-at-Risk*, chapter Quality Control. Risk Publication, Londres.

[19] Dostaler G (2009). Thomas d'Aquin et le péché de l'usure. *Alternatives Economiques*, 277.

[20] Engle RF Manganelli S (2004). CAViAR : Conditional Autoregressive Value-at-Risk by Regression Quantiles. *Journal of Business and Economic Statistics*, 22 :367–381.

[21] Escanciano JC, Olmo J (2007). Estimation Risk Effect on Backtesting For Parametric Value-at-Risk Models. *Center for Appliec Economics and Policy Research, Indiana University, Working Paper*.

[22] Francq C, Zakoïan JM (2006). *Modèles GARCH et à volatilité stochastique*. Université Lille III.

[23] Gourieroux C, Laurent JP, Scaillet O (2000). Sensitivity Analysis of Values-at-Risk.

[24] Gourieroux C, Monfort A (1996). *Statistiques et modèles économétriques, Tome 2, Tests, Régions de confiance, Choix de modèles, Théorie Asymptotique*. Economica.

[25] Gourieroux C, Zakoïan JM (2009). *Mesure de Risque*. ENSAE.

[26] Hurlin C, Tokpavi S (2006). Backtesting VaR Accuracy : A new simple test. Preliminary Version.

[27] Hurlin C, Tokpavi S (2007). Un test de validité de la VaR. *Revue Economique*, 3-58(ISBN 2-7249-3073-2) :599–608. consulté en Avril 2009.

[28] Hurlin C, Tokpavi S (2007). Une évaluation des procédures de backtesting. *HALSHS*, 00159846.

Sixième partie

Table des figures

Table des figures

1 Exemple de publications - Rapport annuel du Crédit Suisse 2008 7
2 Exemple de publications - Rapport annuel du Crédit Suisse 2008 8
3 Illustration de la notion de mesure de risque 9
4 Illustration de la notion de modélisation 9
5 Illustration des notions de modélisation et de mesure de risque 9
6 Illustration des différences entres les mesures 10
7 Illustration de l'effet de paramètre 10
8 Illustration de l'effet de paramètre sur les mesures de risques 10
9 Illustration de l'effet de paramètre sur le backtesting 11

1 Graphique de l'indice S&P500 entre mai 1999 et mai 2009 18
2 Graphique des accroissements journaliers de l'indice S&P500 entre mai 1999 et mai 2009 . 18
3 Graphique des taux de croissance journaliers de l'indice S&P500 entre mai 1999 et mai 2009 19
4 Graphique des log-rendements journaliers de l'indice S&P500 entre mai 1999 et mai 2009 19
5 Estimation des paramètres à partir du S&P500 21
6 Tableau de convergence des paramètres estimés en moyenne 21
7 Graphique de convergence du rapport moyen paramètres estimés-paramètres théoriques . 22
8 Boxplot de la répartition des erreurs entre ω et ω_0 selon la taille de la fenêtre d'estimation 23
9 Boxplot de la répartition des erreurs entre α et α_0 selon la taille de la fenêtre d'estimation 23
10 Boxplot de la répartition des erreurs entre β et β_0 selon la taille de la fenêtre d'estimation 24

1 Tableau Récapitulatif Tests-Hypothèses 37

1 Tableau des niveaux empiriques de rejets - Estimation sur 250 dates 44
2 Tableau des niveaux empiriques de rejets - Estimation sur 500 dates 44
3 Tableau des niveaux empiriques de rejets - Estimation sur 1000 dates 45

1 Ecarts relatifs des VaR - 10000 dates 51

1 Variances du test de Kupiec en fonction de la taille de l'échantillon 58
2 Variances du test de Kupiec en fonction de la taille de l'échantillon 58

1 Rapport paramètres moyens estimés sur paramètres théoriques en fonction de la taille de la fenêtre d'estimation 77
2 Boxplot de l'erreur d'estimation de ω en fonction de la taille de la fenêtre d'estimation (de 50 à 1000 dates) 78
3 Boxplot de l'erreur d'estimation de α en fonction de la taille de la fenêtre d'estimation (de 50 à 1000 dates) 78
4 Boxplot de l'erreur d'estimation de β en fonction de la taille de la fenêtre d'estimation (de 50 à 1000 dates) 79
5 Boxplot de l'erreur d'estimation de ω en fonction de la taille de la fenêtre d'estimation (de 1000 à 5000 dates) 79
6 Boxplot de l'erreur d'estimation de α en fonction de la taille de la fenêtre d'estimation (de 1000 à 5000 dates) 80
7 Boxplot de l'erreur d'estimation de β en fonction de la taille de la fenêtre d'estimation (de 1000 à 5000 dates) 80
8 Moyenne empirique 83
9 Tableau croisé des violations 83

10 Moyenne des écarts relatifs entre VaR 83
11 Synthèse d'un calibrage d'un modèle GARCH sur les accroissements du SP500 88
12 Synthèse d'un calibrage d'un modèle GARCH sur les taux d'accroissements du SP500 . . 89
13 Synthèse d'un calibrage d'un modèle GARCH sur les log-rendements du SP500 90
14 Organisation des différentes hypothèses de backtesting 91

Septième partie

Annexes

Annexe A

Éléments techniques

A.1 Expérience 1 : Convergence de la méthode d'estimation

Objectif : estimer la vitesse de convergence d'estimation des paramètres d'un modèle d'actif

A.1.1 Démarche

On souhaite ici analyse la vitesse de convergence d'estimateurs. Même si cela ne rend pas compte immédiatement de la qualité des modèles, cela permet d'avoir une idée de la précision empirique pour des jeux de données utilisées dans le milieu financier.

A.1.2 Formalisation

On considère un modèle théorique parfait suivant : les log-rendements des positions sur les marchés financiers d'une banque suivent un processus GARCH(1,1).
Si on note P_t la valeur à l'instant t de la position, on peut alors écrire :

$$\varepsilon_t = ln(\frac{P_t}{P_{t-1}}) \tag{A.1}$$

La modélisation s'écrit alors :

$$\begin{cases} \varepsilon_t = \sigma_t u_t \\ u_t \overset{iid}{\sim} \mathcal{N}(0,1) \\ \sigma_t^2 = \omega + \alpha\sigma_t^2 + \beta\varepsilon_t^2 \\ \omega \in \mathbb{R}^{*+} \quad \alpha \in \mathbb{R}^{+} \quad \beta \in \mathbb{R}^{+} \quad \alpha + \beta < 1 \end{cases} \tag{A.2}$$

A.1.3 Mise en oeuvre

1. Choix (arbitraire) de ω, α, β, σ_0 et ε_0.
2. Pour i allant de i_0 à i_{max}, faire :
   (a) Pour j allant de 1 à *nb_simulation* (Boucle de Monte Carlo), faire :
      i. Génération de la série ε_t pour $t \in [1; i]$
      ii. Estimation des paramètres
      iii. Ajout de la contribution de cette trajectoire
   (b) sauvegarder les valeurs
3. Affichage de l'évolution de $\tilde{\omega}$, $\tilde{\alpha}$ et $\tilde{\beta}$ en fonction de i. En moyenne et sous forme de boxplot.

A.1.4 Analyse

- L'analyse en moyenne montre clairement une convergence assez rapide. La vitesse de convergence est plus faible pour le paramètre ω.
- L'affichage sous forme de Box plot [1] permet de remarquer des comportements différents.
- Le paramètre ω converge de manière correcte car elle concerne la médiane, les quartiles centraux et l'étendue.
- Le paramètre α converge très bien : la réduction des quartiles centraux et de l'étendue de la distribution est très claire. Il y effectivement un effet de paramètre.
- Le paramètre β a un comportement délicat. Si la médiane (et la moyenne) tend à se rapprocher de la valeur théorique, en deçà de 200 dates pour l'estimation, la variance (estimée par les quartiles centraux) ne se réduit pas beaucoup. Pour l'étendue, il faut attendre 300 dates pour l'estimation pour voir une réduction. L'effet de paramètre est donc assez fort et à peu près persistant pour des fenêtres d'estimation modestes (moins de 300 dates).

1. les extrémités du rectangles représentent les frontières des quartiles inférieur et supérieur ; la ligne en pointillé l'étendue de la répartition, les croix les points extrêmes ; la ligne verticale la médiane

A.1.5 Tableaux et graphiques

n	ω (10^{-4})	α	β	Simulations
100	3.26	0.113	0.437	10 000
200	2.94	0.100	0.519	10 000
300	2.33	0.094	0.614	10 000
400	1.98	0.091	0.661	10 000
500	1.76	0.088	0.693	10 000
600	1.68	0.087	0.704	10 000
700	1.62	0.086	0.714	10 000
800	1.54	0.086	0.724	10 000
∞	1.00	0.080	0.80	

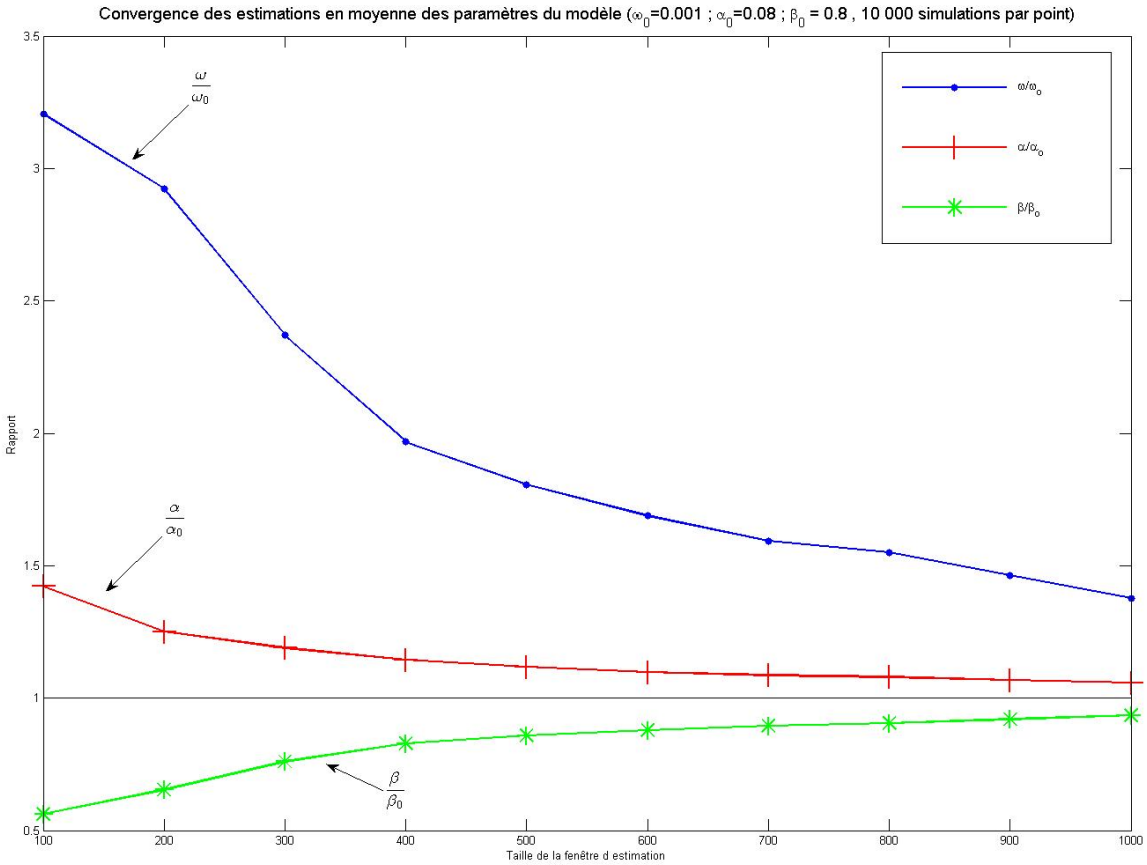


FIGURE 1 – Rapport paramètres moyens estimés sur paramètres théoriques en fonction de la taille de la fenêtre d'estimation

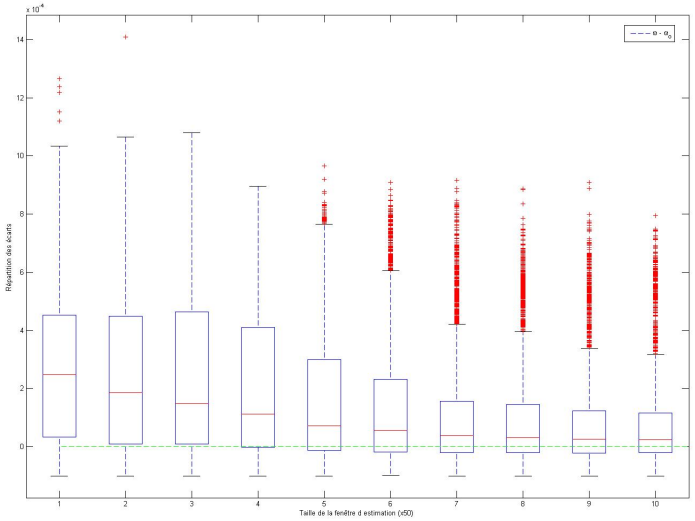


FIGURE 2 – Boxplot de l'erreur d'estimation de ω en fonction de la taille de la fenêtre d'estimation (de 50 à 1000 dates)

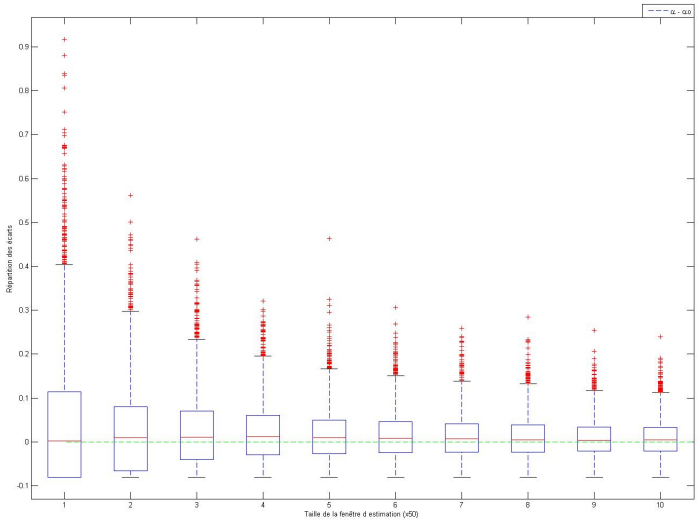


FIGURE 3 – Boxplot de l'erreur d'estimation de α en fonction de la taille de la fenêtre d'estimation (de 50 à 1000 dates)

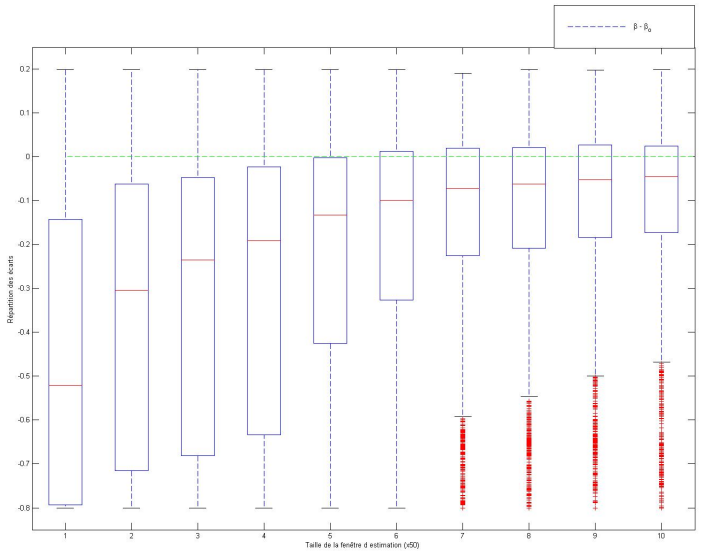


FIGURE 4 – Boxplot de l'erreur d'estimation de β en fonction de la taille de la fenêtre d'estimation (de 50 à 1000 dates)

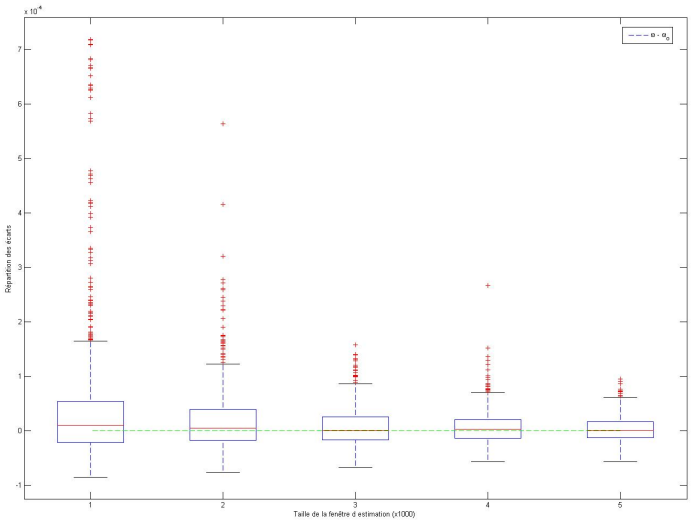


FIGURE 5 – Boxplot de l'erreur d'estimation de ω en fonction de la taille de la fenêtre d'estimation (de 1000 à 5000 dates)

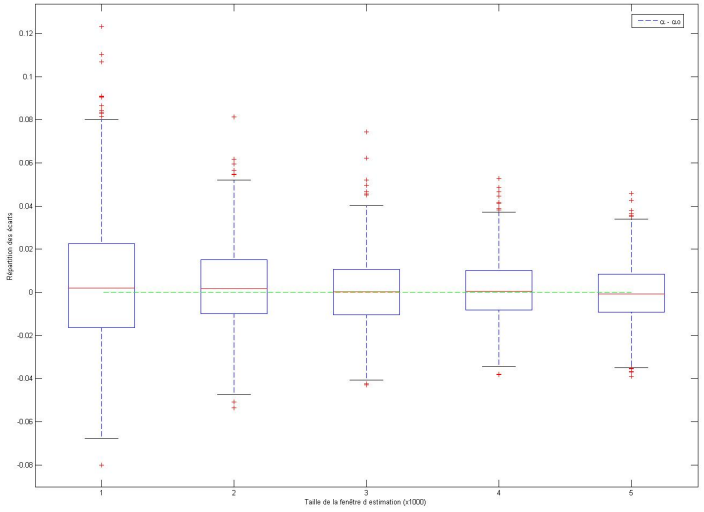


FIGURE 6 – Boxplot de l'erreur d'estimation de α en fonction de la taille de la fenêtre d'estimation (de 1000 à 5000 dates)

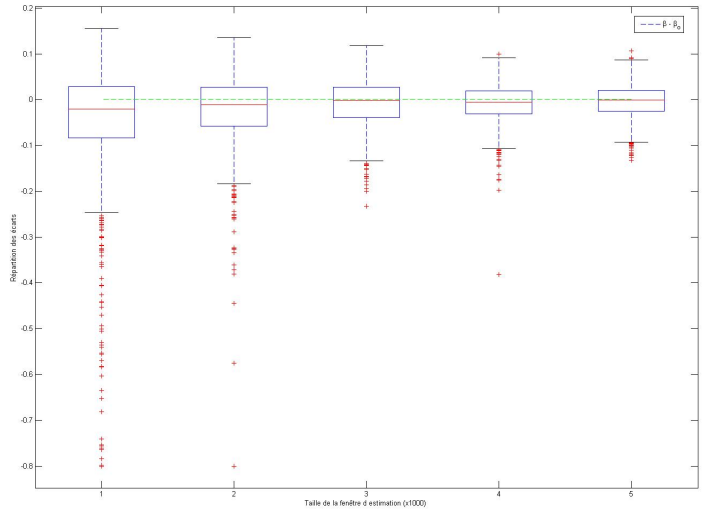


FIGURE 7 – Boxplot de l'erreur d'estimation de β en fonction de la taille de la fenêtre d'estimation (de 1000 à 5000 dates)

A.2 Expérience 2 : Tests usuels

Objectif : estimer l'effet de paramètre dans les procédures de backtesting de la VaR

A.2.1 Démarche

La VaR, comme toute mesure de risque fondée sur un modèle paramétrique, donne une information qui regroupe trois termes : le risque propre, le risque de modèle (importance de l'écart entre la réalité et le modèle) et le risque de paramètre (spécification approximative du modèle).
Asymptotiquement, le risque de paramètre est nul et donc le risque mesuré devient uniquement le risque propre et le risque de modèle. La précision de la mesure ne dépend plus que de la justesse du modèle.
Pour essayer d'estimer le risque de paramètre, on construit une expérience où le risque de modèle est nul, ie on se donne la définition du processus de perte. Il est alors possible d'estimer le risque de paramètre en faisant la différence entre le risque estimé (risque propre, risque de modèle, risque de paramètre) et le risque de propre.
Ce premier calcul permet de quantifier l'influence de la composante "paramètres" dans la précision de la mesure de risque.
Dans un second temps, on cherche à estimer cet effet "en aval" ie dans les procédures de backtesting. Pour cela, on compare les résultats des tests usuels (couverture inconditionnelle, indépendance d'horizon 1 et durée) sur les indicatrices de dépassement.

A.2.2 Formalisation

La VaR considérée sera toujours la value-at-risk au seuil de 1%.

On considère $(P_t)_t \geq 0$ le cours de l'actif principal 'une banque à la date t. On note $(\varepsilon_t)_t = (ln\frac{P_t}{P_{t-1}})_t$ le log-rendement du cours. On suppose que ce processus suit un processus $GARCH(1,1) - \mathcal{N}$:

$$\begin{cases} \varepsilon_t = \sigma_t \eta_t \\ \sigma_t^2 = \omega + \alpha\sigma_{t-1}^2 + \beta\varepsilon_{t-1} \\ \omega > 0 \ ; \ \alpha \geq 0 \ ; \ \beta \geq 0 \\ \eta_t \overset{iid}{\sim} \mathcal{N}(0,1) \end{cases} \tag{A.3}$$

On suppose que σ_0 et ε_0 sont données. La série est considéré pour $t \in [1;T]$

On introduit les notations suivantes :

- $\tilde{\alpha}, \tilde{\beta}, \tilde{\omega}$ les estimateurs des paramètres du modèle utilisée par la banque.
- $\tilde{\sigma}_t$ l'estimateur de la volatilité
- $VaR^a_{t|t-1}$ la VaR empirique (ou estimée) publiée par la banque pour la date t sachant connaissant toutes les pertes jusqu'à la date t-1.
- $VaR^{th}_{t|t-1}$ la VaR théorique (ou réelle) pour la date connaissant tous les processus jusqu'à la date t-1.
- $I^{th}_t = \mathbb{1}_{X_t > VaR^{th}_{t|t-1}}$ l'indicatrice de violation de la VaR théorique.
- $I^a_t = \mathbb{1}_{X_t > VaR^a_{t|t-1}}$ l'indicatrice de violation de la VaR empirique.
- n la taille de la fenêtre à partir de laquelle sont estimés les paramètres du modèle ($n < T$).

A.2.3 Méthode de calcul de la VaR par la banque

La banque dispose des pertes journalières sur n dates successives. Pour publier $VaR^a_{t|t-1}$, elle procède ainsi :

1. Estimation de α, β et ω par maximisation de la vraisemblance.
2. Calcul de $\tilde{\sigma}_{n+1}$ par
$$\tilde{\sigma}_{n+1} = \hat{\omega} + \hat{\alpha}\tilde{\sigma}_n^2 + \hat{\beta}\varepsilon_n$$
3. Publication de $VaR^a_{n+1|n} = VaR = P_n(1 - e^{\tilde{\sigma}_{n+1}\phi^{-1}(\gamma)})$ puisque :
$$\mathbb{P}(P_{n+1} - P_n < -VaR) = \gamma \tag{A.4}$$
$$\mathbb{P}(P_n(\frac{P_{n+1}}{P_n} - 1) < -VaR) = \gamma \tag{A.5}$$

$$\mathbb{P}(\frac{P_{n+1}}{P_n} < 1 - \frac{VaR}{P_n}) = \gamma \tag{A.6}$$

$$\mathbb{P}(\frac{1}{\tilde{\sigma}_{n+1}} ln \frac{P_{n+1}}{P_n} < ln(1 - \frac{VaR}{P_n})) = \gamma \tag{A.7}$$

$$or \quad \frac{1}{\tilde{\sigma}_{n+1}} ln \frac{P_{n+1}}{P_n} \sim \mathcal{N}(0,1) \tag{A.8}$$

$$\frac{1}{\tilde{\sigma}_{n+1}} ln \frac{P_{n+1}}{P_n} = \phi^{-1}(\gamma) \tag{A.9}$$

$$VaR = P_n(1 - e^{\tilde{\sigma}_{n+1}\phi^{-1}(\gamma)}) \tag{A.10}$$

A.2.4 Mise en oeuvre

1. Choix (arbitraire) de ω, α, β, σ_0 et ε_0.
2. Choix de T
3. Choix de n (fenêtre d'estimation : $n < T$)
4. Pour i allant de i_0 à *nb_simulation* (boucle de Monte Carlo), faire :
   (a) Génération de la série ε_t pour $t \in [1; T]$
   (b) Estimation des paramètres sur les n premières valeurs : $\hat{\omega}$, $\hat{\alpha}$ et $\hat{\beta}$.
   (c) Calcul des VaR (à 1 jour) théorique et approchée pour les dates entre n+1 et T.
   (d) Calcul des indicatrices de violations
   (e) Procédure de backtesting simple :
   - Moyenne de l'indicatrice des violations approchée : $\overline{I_t^a}$
   - Moyenne de l'indicatrice des violations théoriques : $\overline{I_t^{th}}$
   - Tableau croisé des proportions : $I_t^a \times I_t^{th}$
   - Moyenne de l'écart relatif entre les VaR : $(1/T)\int (VaR^a - VaR^{th})/VaR^{th}$

   (f) Test de Kupiec (proportion de violations)
   - calcul de la statistique de Kupiec pour l'indicatrice de violations approchée
   - calcul de la statistique de Kupiec pour l'indicatrice de violations théorique

   (g) Autres tests
   - calcul de la p-value pour l'indicatrice de violation approchée : tests de Markov, tests sur les durées, tests de portemanteau.
   - calcul de la p-value pour l'indicatrice de violation théorique : tests de Markov, tests sur les durées, tests de portemanteau.
   - sauvegarde du choix de rejet ou non de l'hypothèse nulle en fonction de la p-value (seuil de rejet fixé à 1%).

A.2.5 Backtesting simple

On précise les formules suivantes :

$$\begin{cases} \overline{I_t^a} := \sum\limits_{i \in \text{simulation}} \Big(\dfrac{1}{T-n-1} \sum\limits_{t=n+1}^{T} I_t^{a(i)} \Big) \\ \overline{I_t^{th}} := \sum\limits_{i \in \text{simulation}} \Big(\dfrac{1}{T-n-1} \sum\limits_{t=n+1}^{T} I_t^{th(i)} \Big) \\ (I_t^a = 0 \times I_t^{th} = 1) = \dfrac{1}{\text{nombre de simulations}} \sum\limits_{i \in \text{simulation}} \Big(\sum\limits_{t=n+1}^{T} \mathbb{1}_{I_t^{a(i)}=0} \cdot \mathbb{1}_{I_t^{th(i)}=0} \Big) \end{cases} \tag{A.11}$$

fenêtre	250	1000	∞
$\overline{I_t^a}$	0.0169	0.0151	0.010
$\overline{I_t^{th}}$	0.0146	0.0144	0.010
Simulations	10000	5000	

FIGURE 8 – Moyenne empirique

fenêtre	250		1000	
	$I_t^{th}=0$	$I_t^{th}=1$	$I_t^{th}=0$	$I_t^{th}=1$
$I_t^a=0$	98.4%	0.1%	98.5%	0.1%
$I_t^a=1$	0.2%	1.3%	0.1%	1.3%

FIGURE 9 – Tableau croisé des violations

fenêtre	250	1000	∞
Ecart	-1.47%	-0.49%	0

FIGURE 10 – Moyenne des écarts relatifs entre VaR

A.3 Expérience 3 : Correction de la variance asymptotique du test de Kupiec

Objectif : Apporter la correction au test de Kupiec présentée dans Escanciano, Olmo (2008)

A.3.1 Position du problème

On considère une banque ne possédant qu'un seul actif dont le cours à la date t est P_t. On note ε_t le log-rendement de cet actif ($\varepsilon_t := ln\frac{P_t}{P_{t-1}}$).
On suppose que les log-rendements suivent un processus GARCH(1,1) :

$$\begin{cases} \varepsilon_t = \sigma_t \eta_t \\ \sigma_t^2 = \omega_0 + \alpha_0 \varepsilon_{t-1}^2 + \beta_0 \sigma_{t-1}^2 \\ \omega_0 > 0 \ ; \ \alpha_0 \geq 0 \ ; \ \beta_0 \geq 0 \\ \eta_t \overset{iid}{\sim} \mathcal{N}(0,1) \end{cases} \tag{A.12}$$

A.3.2 Estimation des paramètres

On dispose d'une plage de constations des valeurs du cours de l'actif : $\{P_t / t \in [1;n]\}$. Avec ces données, on cherche à estimer les coefficients du modèle GARCH(1,1) par maximisation de la log-vraisemblance.
La log-vraisemblance s'écrit :

$$ln\Big(\prod_{t=1}^{n} \mathcal{L}_t\Big) = \sum_{t=1}^{n} \underbrace{ln\mathcal{L}_t}_{\ell_t} \tag{A.13}$$

Avec par définition,

$$\ell_t = ln\mathcal{L}_t = ln(\frac{1}{\sqrt{2\pi}\sigma_t} e^{-\frac{\varepsilon_t^2}{2\sigma_t^2}}) = -\underbrace{\frac{1}{2} ln(2\pi)}_{constante} - \frac{1}{2}\Big(ln\sigma_t^2 + \frac{\varepsilon_t^2}{\sigma_t^2}\Big) \tag{A.14}$$

La maximisation s'écrit donc :

$$max\Big\{\sum_{t=1}^{n} \ell_t\Big\} \equiv min\Big\{\sum_{t=1}^{n} (ln\sigma_t^2 + \frac{\varepsilon_t^2}{\sigma_t^2})\Big\} \tag{A.15}$$

On note les valeurs des paramètres réalisant cette maximisation sur un ensemble compact noté Θ (par exemple $[0;1]^3$) ainsi :

$$\theta_n = \begin{pmatrix} \omega_n \\ \alpha_n \\ \beta_n \end{pmatrix} = \overset{(\omega,\alpha,\beta)\in\Theta}{argmin} \Big\{\sum_{t=1}^{n} (ln\sigma_t + \frac{\varepsilon_t^2}{\sigma_t^2})\Big\} \tag{A.16}$$

A.3.3 Convergence

On peut démontrer la propriété suivante :

$$\begin{cases} \sqrt{n}(\theta_n - \theta_0) \xrightarrow{d} \mathcal{N}(0, (\kappa - 1)J^{-1}) \\ J = \mathbb{E}(\frac{\partial \ell_t}{\partial \theta} \frac{\partial \ell_t}{\partial \theta'}) \\ \kappa = \mathbb{E}(\eta_t^4) \end{cases} \tag{A.17}$$

Estimation de κ :
Le vecteur θ_n permet de calculer pour tout t l'écart-type instantané du log-rendement : $\hat{\sigma}_t$. Comme on connaît les réalisations de ε_t, on peut construire une estimation des réalisations de la loi gaussienne :

$$\hat{\eta}_t = \frac{\varepsilon_t}{\hat{\sigma}_t} \tag{A.18}$$

On estime alors κ par :

$$\boxed{\kappa_n = \frac{1}{n}\sum_{t=1}^{n} \hat{\eta}_t^4 = \frac{1}{n}\sum_{t=1}^{n} (\frac{\varepsilon_t}{\hat{\sigma}_t})^4} \tag{A.19}$$

Estimation de J :
On construire l'estimateur de J suivant :

$$J_n = \frac{1}{n}\sum_{t=1}^{n} \frac{\partial \ell_t}{\partial \theta}(\theta_n)\frac{\partial \ell_t}{\partial \theta'}(\theta_n)) \tag{A.20}$$

On a :

$$J_n = \frac{1}{n}\sum_{t=1}^{n} \frac{1}{\sigma_t^4}\frac{\partial \sigma_t^2}{\partial \theta'}\frac{\partial \sigma_t^2}{\partial \theta} \tag{A.21}$$

La dérivée de σ_t^2 s'écrit :

$$\frac{\partial \sigma_t^2}{\partial \theta} = \begin{pmatrix} 1 \\ \varepsilon_{t-1}^2 \\ \sigma_{t-1}^2 \end{pmatrix} + \beta \frac{\partial \sigma_{t-1}^2}{\partial \theta} \tag{A.22}$$

D'où

$$\boxed{J_n = \frac{1}{n}\sum_{t=1}^{n} \frac{1}{\sigma_t^4}\frac{\partial \sigma_t^2}{\partial \theta'}\frac{\partial \sigma_t^2}{\partial \theta}} \tag{A.23}$$

Expansion de Bahadur :
La méthode ci-dessous cherche à maximiser la log-vraisemblance. Pour obtenir l'expansion de Bahadur, on effectue un développement de Taylor avec reste de Laplace à l'ordre 1 de la différrentielle de la log-vraisemblance :

$$\frac{\partial}{\partial \theta}\sum_{t=1}^{n} \ell_t(\theta_n) = \frac{\partial}{\partial \theta}\sum_{t=1}^{n} \ell_t(\theta_0) + (\theta_n - \theta_0)\frac{\partial^2}{\partial \theta \partial \theta'}\sum_{t=1}^{n} \ell_t(\theta^*) \qquad \theta^* \in [\theta_0, \theta_n] \tag{A.24}$$

$$\underbrace{\sum_{t=1}^{n} \frac{\partial \ell_t}{\partial \theta}(\theta_n)}_{0} = \sum_{t=1}^{n} \frac{\partial \ell_t(\theta_0)}{\partial \theta} + (\theta_n - \theta_0)\sum_{t=1}^{n} \frac{\partial^2 \ell_t(\theta^*)}{\partial \theta \partial \theta'} \qquad \theta^* \in [\theta_0, \theta_n] \tag{A.25}$$

Or en θ_n la log-vraisemblance est maximale, donc la différentielle est nulle. Par conséquent, on peut écrire :

$$(\theta_n - \theta_0) = -\Big(\sum_{t=1}^{n} \frac{\partial^2 \ell_t(\theta^*)}{\partial \theta \partial \theta'}\Big)^{-1}\sum_{t=1}^{n} \frac{\partial \ell_t}{\partial \theta}(\theta_0) \tag{A.26}$$

On multiplie par $\sqrt{n}$ à gauche et à droite :

$$\sqrt{n}(\theta_n - \theta_0) = -\sqrt{n}\Big(\sum_{t=1}^{n} \frac{\partial^2 \ell_t(\theta^*)}{\partial \theta \partial \theta'}\Big)^{-1}\sum_{t=1}^{n} \frac{\partial \ell_t}{\partial \theta}(\theta_0) \tag{A.27}$$

$$\sqrt{n}(\theta_n - \theta_0) = -\underbrace{\frac{1}{\frac{1}{n}}\frac{1}{\sqrt{n}}}_{\sqrt{n}}\Big(\sum_{t=1}^{n} \frac{\partial^2 \ell_t(\theta^*)}{\partial \theta \partial \theta'}\Big)^{-1}\sum_{t=1}^{n} \frac{\partial \ell_t}{\partial \theta}(\theta_0) \tag{A.28}$$

$$\sqrt{n}(\theta_n - \theta_0) = -\Big(\frac{1}{n}\sum_{t=1}^{n} \frac{\partial^2 \ell_t(\theta^*)}{\partial \theta \partial \theta'}\Big)^{-1} \times \frac{1}{\sqrt{n}}\sum_{t=1}^{n} \frac{\partial \ell_t}{\partial \theta}(\theta_0) \tag{A.29}$$

De plus, comme θ_n converge vers θ_0, θ^* tend vers θ_0. On peut donc écrire :

$$\sqrt{n}(\theta_n - \theta_0) = \underbrace{-\Big(\frac{1}{n}\sum_{t=1}^{n} \frac{\partial^2 \ell_t(\theta_0)}{\partial \theta \partial \theta'}\Big)^{-1}}_{\to J^{-1}} \times \frac{1}{\sqrt{n}}\sum_{t=1}^{n} \frac{\partial \ell_t}{\partial \theta}(\theta_0) + o_{\mathbb{P}}(1) \tag{A.30}$$

$$\boxed{\sqrt{n}(\theta_n - \theta_0) = \frac{1}{\sqrt{n}}\sum_{t=1}^{n} J^{-1}\frac{\partial \ell_t}{\partial \theta}(\theta_0) + o_{\mathbb{P}}(1)} \tag{A.31}$$

A.3.4 Correction de la variance dans le test de Kupiec

Escanciano et Olmo donnent comme variance de la moyenne empirique des violations :

$$\hat{\sigma}_c^2 = \gamma(1-\gamma) + 2A\rho + AVA' \tag{A.32}$$

Avec les notations suivantes :

- γ : le seuil de confiance de la VaR ($\gamma = 1\%$ ou 5%)
- V : matrice de variance-covariance de θ ($V = (\kappa - 1)J^{-1}$)
- $\rho = \mathbb{E}((I_t - \alpha)l(Y_t, \mathcal{F}_{t-1}, \theta_0))$
- $A = \mathbb{E}(\partial_\theta(VaR(\gamma)_{t|t-1}) \times f_{\mathcal{F}_{t-1}}(VaR(\gamma)_{t|t-1}))$
- I_t : indicatrice de violation ($I_t = \mathbb{1}_{\{Y_t < -VaR(\gamma)_{t|t-1}\}}$)
- $l(Y_t, \mathcal{F}_{t-1}, \theta_0)$: terme sous le signe somme dans l'écriture de l'expansion de Bahadur ($l(Y_t, \mathcal{F}_{t-1}, \theta_0 = J^{-1}\frac{\partial \ell_t}{\partial \theta}(\theta_0))$.
- $f_{\mathcal{F}_{t-1}}$: fonction de densité des pertes conditionnelles à l'information disponible en t-1.

Estimation de ρ :

$$\boxed{\rho_n = \frac{1}{n}\sum_{t=1}^{n}(I_t - \gamma)J_n^{-1}\frac{\partial \ell_t}{\partial \theta}(\theta_n)} \tag{A.33}$$

Estimation de A :

Calcul de la VaR :

$$\mathbb{P}(P_t - P_{t-1} < -VaR) = \gamma \tag{A.34}$$

$$\mathbb{P}(P_{t-1}(\frac{P_t}{P_{t-1}} - 1) < -VaR) = \gamma \tag{A.35}$$

$$\mathbb{P}(\frac{P_t}{P_{t-1}} < 1 - \frac{VaR}{P_{t-1}}) = \gamma \tag{A.36}$$

$$\mathbb{P}(\frac{1}{\tilde{\sigma}_t} ln\frac{P_t}{P_{t-1}} < \frac{1}{\tilde{\sigma}_t} ln(1 - \frac{VaR}{P_{t-1}})) = \gamma \tag{A.37}$$

$$or \quad \frac{1}{\tilde{\sigma}_t} ln\frac{P_t}{P_{t-1}} \sim \mathcal{N}(0,1) \tag{A.38}$$

$$\frac{1}{\tilde{\sigma}_t} ln(1 - \frac{VaR}{P_{t-1}}) = \phi^{-1}(\gamma) \tag{A.39}$$

$$\boxed{VaR_{t|t-1}(\gamma) = P_{t-1}(1 - e^{\tilde{\sigma}_t \phi^{-1}(\gamma)})} \tag{A.40}$$

A :

$$F(y) = \mathbb{P}(PnL_t \leq y) = \mathbb{P}(\frac{1}{\sigma_t} ln\frac{P_t}{P_{t-1}} \leq \frac{1}{P_{t-1}} ln(1 + \frac{y}{P_{t-1}})) \tag{A.41}$$

$$F(y) = \int_{-\infty}^{\frac{1}{\sigma_t} ln(1+\frac{y}{P_{t-1}})} \frac{1}{\sqrt{2\pi}} e^{-x^2} dx \tag{A.42}$$

D'où

$$f(y) = \frac{1}{\sqrt{2\pi}\sigma_t(P_{t-1} + y)} exp\Big(-\frac{1}{\sigma_t^2}(ln(1 + \frac{y}{P_{t-1}}))^2\Big) \tag{A.43}$$

On applique en $y = -VaR = -P_{t-1}(1 - e^{\sigma_t \Phi^{-1}(\gamma)})$:

$$f(-VaR) = \frac{e^{-\sigma_t \Phi^{-1}(\gamma)}}{\sqrt{2\pi}\sigma_t P_{t-1}} e^{-\Phi^{-1}(\gamma)^2} \tag{A.44}$$

On trouve alors :

$$A = \mathbb{E}\Big(f(-VaR) \times \frac{\partial VaR}{\partial \theta}\Big) \tag{A.45}$$

Donc :

$$A = \mathbb{E}\Big(\frac{e^{-\sigma_t \Phi^{-1}(\gamma)}}{\sqrt{2\pi}\sigma_t P_{t-1}} e^{-\Phi^{-1}(\gamma)^2} \times -P_{t-1} e^{\sigma_t \Phi^{-1}(\gamma)} \sigma_t \times \frac{1}{2\sigma_t}\frac{\partial \sigma_t^2}{\partial \theta}\Big) \tag{A.46}$$

$$A = \mathbb{E}\Big(\frac{\Phi^{-1}(\gamma)e^{-\Phi^{-1}(\gamma)^2}}{2\sqrt{2\pi}\sigma_t^2}\frac{\partial\sigma_t^2}{\partial\theta}\Big) \tag{A.47}$$

D'où :

$$\hat{A}_n = \frac{1}{n}\sum_{t=1}^{n}\Big(\frac{\Phi^{-1}(\gamma)e^{-\Phi^{-1}(\gamma)^2}}{2\sqrt{2\pi}\tilde{\sigma}_t^2}\frac{\partial\tilde{\sigma}_t^2}{\partial\theta}(\hat{\theta}_n)\Big) \tag{A.48}$$

A.4 S&P 500 - GARCH sur les accroissements

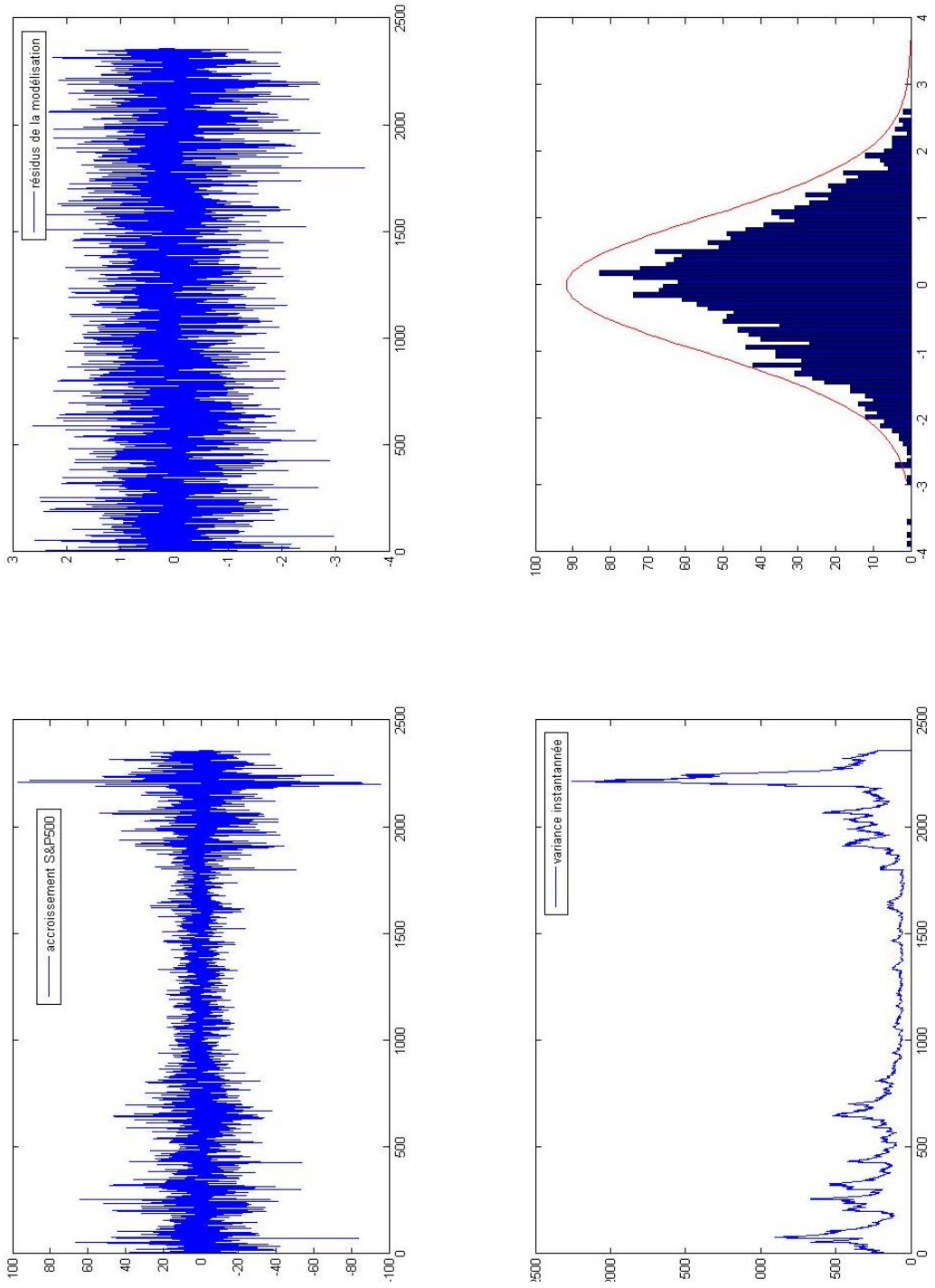


Figure 11 – Synthèse d'un calibrage d'un modèle GARCH sur les accroissements du SP500

A.5 S&P 500 - GARCH sur les taux d'accroissement

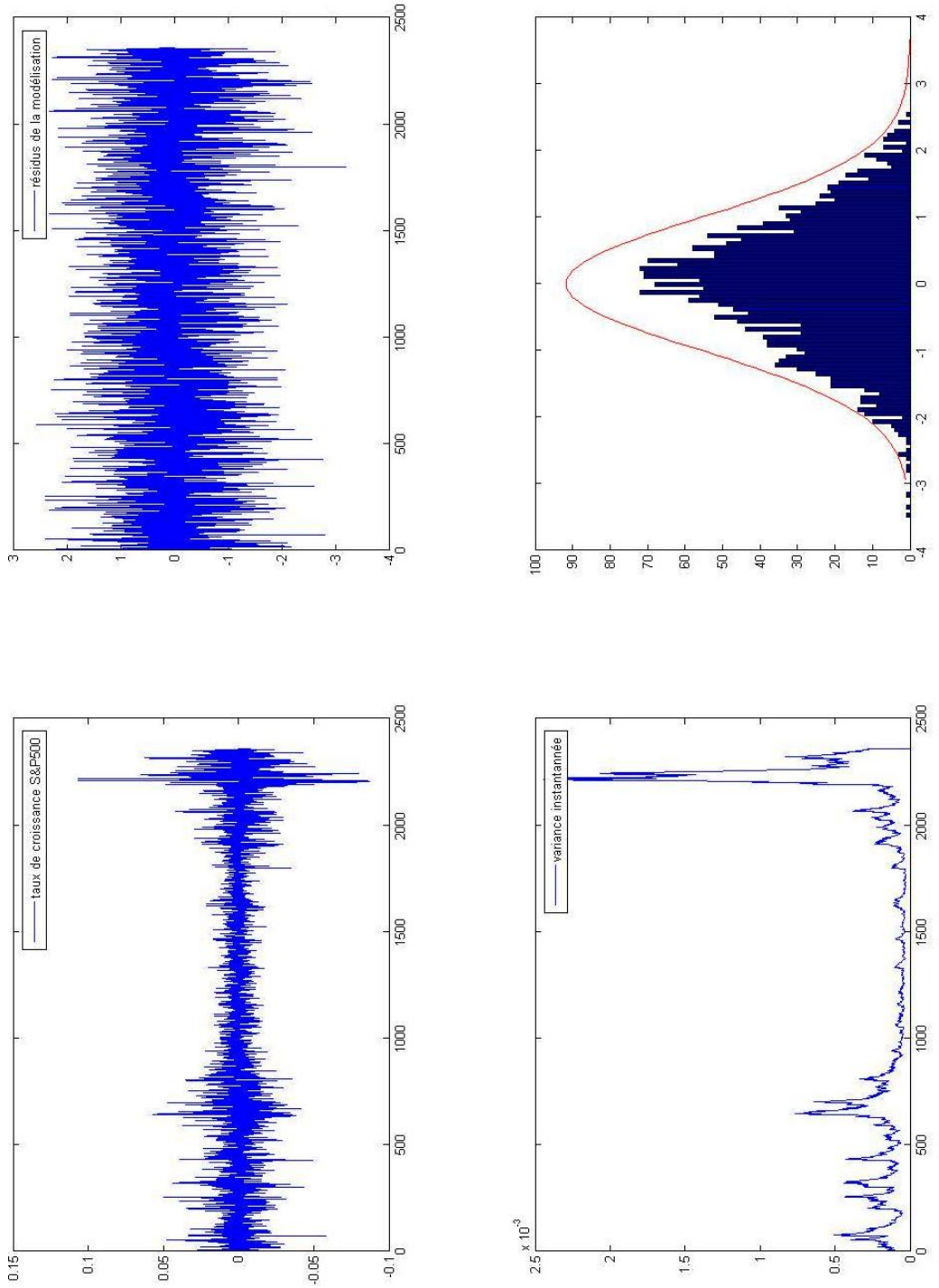


FIGURE 12 – Synthèse d'un calibrage d'un modèle GARCH sur les taux d'accroissements du SP500

A.6 S&P 500 - GARCH sur les log-rendements

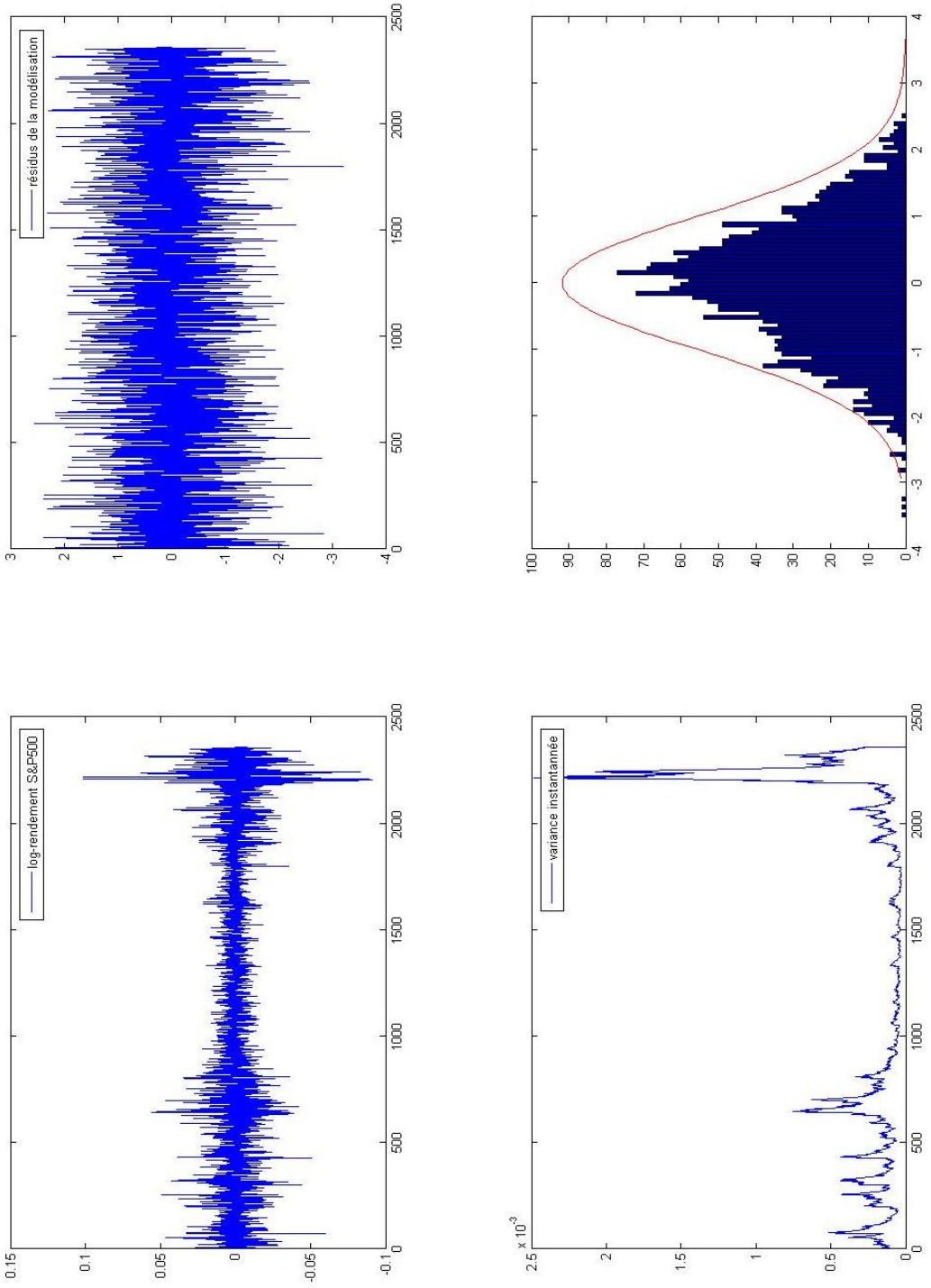


Figure 13 – Synthèse d'un calibrage d'un modèle GARCH sur les log-rendements du SP500

A.7 Démonstration des relations entre les hypothèses des tests

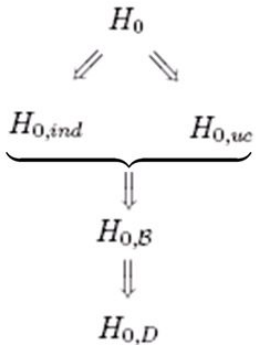


FIGURE 14 – Organisation des différentes hypothèses de backtesting

A.7.1 $H_0 \Rightarrow H_{0,uc}$

On rappelle :

$$\begin{cases} (H_0) : \text{ « } \mathbb{E}(I_t(\alpha) - \alpha | \mathcal{F}_{t-1}) = 0 \ mbox \ ˇ \\ (H_{0,uc}) : \text{ « } \mathbb{E}(I_t(\alpha) - \alpha) = 0 \text{ »} \end{cases} \tag{A.49}$$

Supposons (H_0). On peut écrire :

$$(H_0) \Leftrightarrow \mathbb{E}(I_t(\alpha)|\mathcal{F}_{t-1}) = \alpha \tag{A.50}$$

On prend l'espérance de part et d'autre :

$$\Rightarrow \underbrace{\mathbb{E}\Big(\mathbb{E}(I_t(\alpha)|\mathcal{F}_{t-1})\Big)}_{\mathbb{E}(I_t(\alpha))} = \underbrace{\mathbb{E}(\alpha)}_{\alpha} \tag{A.51}$$

$$\Rightarrow \mathbb{E}(I_t(\alpha)) = \alpha \Rightarrow H_{0,uc} \tag{A.52}$$

L'implication réciproque n'est pas vraie. Il suffit de considérer :

$$\begin{cases} \tilde{I}_1 \sim \mathcal{B}(\alpha) \\ \tilde{I}_t \sim \mathcal{B}(\tilde{I}_1) \quad t > 1 \end{cases} \tag{A.53}$$

On a alors :

$$\mathbb{E}(\tilde{I}_t) = \alpha \quad \text{et} \quad \mathbb{E}(\tilde{I}_t|\mathcal{F}_{t-1}) = \tilde{I}_{t-1} \ t > 1 \tag{A.54}$$

A.7.2 $H_0 \Rightarrow H_{0,ind}$

On rappelle :

$$\begin{cases} (H_0) : \text{ « } \mathbb{E}(I_t(\alpha) - \alpha | \mathcal{F}_{t-1}) = 0 \text{ »} \\ (H_{0,ind}) : \text{ « } \forall (t,t'), t \neq t' \Rightarrow Cov(I_t(\alpha), I_{t'}(\alpha)) = 0 \text{ »} \end{cases} \tag{A.55}$$

Supposons (H_0). On peut écrire :

$$(H_0) \Leftrightarrow \mathbb{P}(I_t(\alpha) = 1 | \mathcal{F}_{t-1}) = \alpha \tag{A.56}$$

$$\Rightarrow \begin{cases} \mathbb{P}(I_t(\alpha) = 1 | I_{t-1}(\alpha) = 0) = \alpha \\ \mathbb{P}(I_t(\alpha) = 1 | I_{t-1}(\alpha) = 1) = \alpha \end{cases} \tag{A.57}$$

$$\Rightarrow \begin{cases} \mathbb{P}(I_t(\alpha) = 0 | I_{t-1}(\alpha) = 0) = 1 - \alpha \\ \mathbb{P}(I_t(\alpha) = 0 | I_{t-1}(\alpha) = 1) = 1 - \alpha \end{cases} \tag{A.58}$$

$$\Rightarrow \forall t, (\delta, \delta') \in \{0,1\}^2, \quad \{I_t(\alpha) = \delta\} \perp \{I_{t-1}(\alpha) = \delta'\} \tag{A.59}$$

On prolonge la propriété par récurrence,

$$\Rightarrow \forall t, t', t \neq t', \forall (\delta, \delta') \in \{0,1\}^2, \quad \{I_t(\alpha) = \delta\} \perp \{I_{t'}(\alpha) = \delta'\} \tag{A.60}$$

$$\Rightarrow \forall (t, t'), t \neq t' \Rightarrow Cov(I_t(\alpha), I_{t'}(\alpha)) = 0 \Leftrightarrow (H_{0,ind}) \tag{A.61}$$

En considérant le processus suivant $I_t \overset{iid}{\sim} \mathcal{B}(\alpha'), \alpha' \neq \alpha$, on montre que la réciproque est fausse.

A.7.3 $H_0 \Rightarrow H_{0,\mathcal{B}}$

On rappelle :

$$\begin{cases} (H_0) : \text{ « } \mathbb{E}(I_t(\alpha) - \alpha | \mathcal{F}_{t-1}) = 0 \text{ »} \\ (H_{0,\mathcal{B}}) : \text{ « } (I_t(\alpha))_t \overset{iid}{\sim} \mathcal{B}(\alpha) \text{ »} \end{cases} \tag{A.62}$$

Supposons (H_0). Le processus $(I_t)_t$ est à valeur dans $\{0,1\}$. Or son espérance est α, donc $\forall t, I_t \sim \mathcal{B}(\alpha)$. L'indépendance a été démontrée dans la partie ci-dessus. Par conséquent, on a bien :

$$H_0 \Rightarrow H_{0,\mathcal{B}} \tag{A.63}$$

A.7.4 $H_0 \Rightarrow H_{0,D}$

On rappelle :

$$\begin{cases} (H_0) : \text{ « } \mathbb{E}(I_t(\alpha) - \alpha | \mathcal{F}_{t-1}) = 0 \text{ »} \\ (H_{0,D}) : \text{ « les durées entre deux violations successives suivent une loi géométrique de paramètre } \alpha \text{ »} \end{cases} \tag{A.64}$$

$H_{0,D}$ découlent de $H_{0,\mathcal{B}}$. Si $(I_t)_t \overset{iid}{\sim} \mathcal{B}(\alpha)$, on peut écrire en notant D la variable indicatrice de durée entre deux violations successives,

$$\begin{cases} \mathbb{P}(D=0) = \mathbb{P}(I_{t+1} = 1 | I_t = 1) = \alpha \\ \mathbb{P}(D=1) = \mathbb{P}(I_{t+2} = 1 \curlywedge I_{t+1} = 0 | I_t = 1) = \underbrace{\mathbb{P}(I_{t+2} = 1) \times \mathbb{P}(I_{t+1} = 0)}_{I_{t+2} \perp I_{t+1}} = \alpha \times \alpha = \alpha^2 \\ \ldots \\ \mathbb{P}(D=d) = \mathbb{P}(I_{t+d} = 1 \curlywedge I_{t+d-1} = 0 \curlywedge \ldots \curlywedge I_{t+1} = 0 | I_t = 1) \\ \qquad = \underbrace{\mathbb{P}(I_{t+d} = 1) \times \mathbb{P}(I_{t+d-1} = 0) \times \ldots \times \mathbb{P}(I_{t1} = 0}_{I_{t+d} \perp I_{t+d-1} \perp \ldots \perp I_{t+1}} = \alpha \times \ldots \times \alpha = \alpha^d \end{cases} \tag{A.65}$$

Ce résultat est valable pour tout $t > 0$. Donc

$$H_0 \Rightarrow H_{0,D} \tag{A.66}$$

A.8 Annexes techniques du chapitre

A.8.1 Complément : intérêt des log-rendements

Par exemple, pour définir des rendements d'horizon supérieur à 1, il suffit de sommer les rendements journaliers :

$$ln \frac{X_t}{X_{t-2}} = ln \frac{X_t}{X_{t-1}} + ln \frac{X_{t-1}}{X_{t-2}} \tag{A.67}$$

A.8.2 Démonstration : le théorème 1.2, p 25

Démonstration.

$$\mathbb{P}\Big(PnL_t < -\underbrace{VaR(t, t-1, \gamma)}_{VaR(\gamma)} | \mathcal{F}_{t-1}\Big) = \gamma \tag{A.68}$$

Utilisons le fait que $\varepsilon_t \sim \mathcal{N}(0,1)$ pour exprimer la VaR :

$$\mathbb{P}\Big(P_t - P_{t-1} < -VaR(\gamma)|\mathcal{F}_{t-1}\Big) = \gamma \tag{A.69}$$

$$\Leftrightarrow \mathbb{P}\Big(P_{t-1}(\frac{P_t}{P_{t-1}} - 1) < -VaR(\gamma)|\mathcal{F}_{t-1}\Big) = \gamma \tag{A.70}$$

$$\Leftrightarrow \mathbb{P}\Big(\frac{P_t}{P_{t-1}} < 1 - \frac{VaR(\gamma)}{P_{t-1}}|\mathcal{F}_{t-1}\Big) = \gamma \tag{A.71}$$

$$\Leftrightarrow \mathbb{P}\Big(\underbrace{\frac{1}{\sigma_t} ln \frac{P_t}{P_{t-1}}}_{\sim\mathcal{N}(0,1)} < \frac{1}{\sigma_t} ln(1 - \frac{VaR(\gamma)}{P_{t-1}})|\mathcal{F}_{t-1}\Big) = \gamma \tag{A.72}$$

$$\Leftrightarrow \frac{1}{\sigma_t} ln(1 - \frac{VaR(\gamma)}{P_{t-1}}) = \Phi^{-1}(\gamma) \tag{A.73}$$

$$\Leftrightarrow P_{t-1} - VaR(\gamma) = P_{t-1} e^{\sigma_t \Phi^{-1}(\gamma)} \tag{A.74}$$

□

A.8.3 Démonstration : Viabilité de la VaR par un modèle GARCH

Calculons la probabilité de dépassement :

$$\mathbb{P}\Big(P_t - P_{t-1} < -VaR(\gamma)\Big) = \mathbb{P}\Big(P_t - P_{t-1} < -P_{t-1}(1 - e^{\sigma_t \Phi^{-1}(\gamma)})\Big) \tag{A.75}$$

$$\mathbb{P}\Big(P_t - P_{t-1} < -VaR(\gamma)\Big) = \mathbb{P}\Big(P_t < P_{t-1} e^{\sigma_t \Phi^{-1}(\gamma)}\Big) \tag{A.76}$$

$$\mathbb{P}\Big(P_t - P_{t-1} < -VaR(\gamma)\Big) = \mathbb{P}\Big(ln(\frac{P_t}{P_{t-1}} < \sigma_t \Phi^{-1}(\gamma)\Big) \tag{A.77}$$

$$\mathbb{P}\Big(P_t - P_{t-1} < -VaR(\gamma)\Big) = \mathbb{P}\Big(\frac{1}{\sigma_t} ln(\frac{P_t}{P_{t-1}} < \Phi^{-1}(\gamma)\Big) \tag{A.78}$$

$$\mathbb{P}\Big(P_t - P_{t-1} < -VaR(\gamma)\Big) = \mathbb{P}\Big(\varepsilon_t < \Phi^{-1}(\gamma)\Big) \tag{A.79}$$

Les ε_t sont indépendants donc les violations successives sont indépendantes malgré la dépendance temporelle du modèle choisi. La proportion de dépassements est clairement égale à γ.
Cette formule exprime donc une VaR viable.

A.9 Démonstration pour le chapitre Correction de la VaR

A.9.1 Démonstration : le théorème 4.1 p48

Démonstration.

$$\mathbb{P}\Big(\varepsilon_t < VaR\Big) = \gamma \Leftrightarrow \mathbb{P}\Big(g(\varepsilon_{t-1}, \theta_0; \eta_t) < VaR\Big) = \gamma \tag{A.80}$$

$$\Leftrightarrow \mathbb{P}\Big(\eta_t < a(\varepsilon_{t-1}, \theta_0; VaR)\Big) = \gamma \tag{A.81}$$

$$\Leftrightarrow VaR = g(\varepsilon_{t-1}, \theta_0, \Phi^{-1}(\gamma)) \tag{A.82}$$

□

A.9.2 Démonstration : le théorème 4.3 p49

Démonstration. Dans le cas qui nous intéresse, les expressions du lemme de Rau-Bredow (A.11 p96) se réduisent :

$$\begin{cases} \dfrac{\partial Q}{\partial h}(0,\alpha) = \mathbb{E}(Y|X = Q(0,\alpha)) \\ \dfrac{\partial^2 Q}{\partial h^2}(0,\alpha) = -\dfrac{\partial \mathbb{V}(Y|X=s)}{\partial s}(s = Q(0,\alpha)) + \mathbb{V}(Y|X = Q(0,\alpha)) \times \dfrac{\partial ln f_X(s)}{\partial s}(s = Q(0,\alpha)) \end{cases} \tag{A.83}$$

S'il transcrit les notations du lemme au cas que l'on étudie, on obtient :

$$\mathbb{P}\Big(\underbrace{\eta_t}_{X} + \underbrace{\frac{\partial a}{\partial \theta}(\varepsilon_{t-1},\theta_0,\varepsilon_t)\Omega^{\frac{1}{2}} V_n}_{Y} \times \underbrace{\frac{1}{\sqrt{n}}}_{\epsilon} > \underbrace{a(\varepsilon_{t-1},\hat{\theta}_n;\widetilde{VaR})}_{Q(h,\alpha)} \Big) = 1-\alpha \tag{A.84}$$

La première étape est de calculer $Q(0,\alpha)$:

$$Q(0,\alpha) = a(\varepsilon_{t-1},\theta_0;VaR) = \Phi^{-1}(\alpha) \tag{A.85}$$

On peut alors calculer la dérivée première :

$$\frac{\partial Q}{\partial h}(0,\alpha) = \mathbb{E}\Big(\Omega^{\frac{1}{2}} V_n \frac{\partial a}{\partial \theta}(\varepsilon_{t-1},\theta_0,\varepsilon_t)\Big|\eta_t = \Phi^{-1}(\alpha)\Big) \tag{A.86}$$

$$\Leftrightarrow \frac{\partial Q}{\partial h}(0,\alpha) = \frac{\partial a}{\partial \theta}(\varepsilon_{t-1},\theta_0,VaR)\Omega^{\frac{1}{2}}\mathbb{E}\Big(\underbrace{V_n}_{\rightsquigarrow \mathcal{N}(0,1)} \Big) \tag{A.87}$$

Par conséquent, la dérivée première du quantile est nulle :

$$\frac{\partial Q}{\partial h}(0,\alpha) = 0 \tag{A.88}$$

Pour le calcul de la dérivée seconde, il est nécessaire de déterminer au préalable l'expression de la variance conditionnelle :

$$\mathbb{V}(Y|X=s) = \mathbb{V}\Big(\Omega^{\frac{1}{2}} V_n \frac{\partial a}{\partial \theta}(\varepsilon_{t-1},\theta_0,\varepsilon_t)\Big|\eta_t = s\Big) \tag{A.89}$$

$$\mathbb{V}(Y|X=s) = \mathbb{V}\Big(\Omega^{\frac{1}{2}} V_n \frac{\partial a}{\partial \theta}(\varepsilon_{t-1},\theta_0,g(\varepsilon_{t-1},\theta_0,s))\Big) \tag{A.90}$$

Cette variance est calculée à la date t donc ici aussi la dérivée partielle de a se comporte comme une constante, et $V_n \rightsquigarrow \mathcal{N}(0,1)$:

$$\mathbb{V}(Y|X=s) = \frac{\partial a}{\partial \theta'}(\varepsilon_{t-1},\theta_0,g(\varepsilon_{t-1},\theta_0,s) \times \Omega \times \frac{\partial a}{\partial \theta}(\varepsilon_{t-1},\theta_0,g(\varepsilon_{t-1},\theta_0,s) \tag{A.91}$$

Le second terme du développement limité du quantile s'écrit :

$$\begin{cases} \dfrac{\partial^2 Q}{\partial h^2}(0,\gamma) = -\Big[\dfrac{\partial}{\partial s}\Big(\dfrac{\partial a}{\partial \theta'}(\varepsilon_{t-1},\theta_0,g(\varepsilon_{t-1},\theta_0,s) \times \Omega \times \dfrac{\partial a}{\partial \theta}(\varepsilon_{t-1},\theta_0,g(\varepsilon_{t-1},\theta_0,s)\Big)\Big]_{s=\Phi^{-1}(\gamma)} \\ \qquad -\dfrac{\partial a}{\partial \theta'}(\varepsilon_{t-1},\theta_0,\underbrace{g(\varepsilon_{t-1},\theta_0,\Phi^{-1}(\gamma))}_{VaR}) \times \Omega \times \dfrac{\partial a}{\partial \theta}(\varepsilon_{t-1},\theta_0,\underbrace{g(\varepsilon_{t-1},\theta_0,\Phi^{-1}(\gamma))}_{VaR}) \times \Big[\dfrac{\partial ln f_{\eta_t}(s)}{\partial s}\Big]_{s=\Phi^{-1}(\gamma)} \end{cases} \tag{A.92}$$

$$\Leftrightarrow \begin{cases} \dfrac{\partial^2 Q}{\partial h^2}(0,\gamma) = -\Big[\dfrac{\partial}{\partial s}\Big(\dfrac{\partial a}{\partial \theta'}(\varepsilon_{t-1},\theta_0,g(\varepsilon_{t-1},\theta_0,s) \times \Omega \times \dfrac{\partial a}{\partial \theta}(\varepsilon_{t-1},\theta_0,g(\varepsilon_{t-1},\theta_0,s)\Big)\Big]_{s=\Phi^{-1}(\gamma)} \\ \qquad -\dfrac{\partial a}{\partial \theta'}(\varepsilon_{t-1},\theta_0,VaR) \times \Omega \times \dfrac{\partial a}{\partial \theta}(\varepsilon_{t-1},\theta_0,VaR) \times \Big[\dfrac{\partial ln f_{\eta_t}(s)}{\partial s}\Big]_{s=\Phi^{-1}(\gamma)} \end{cases} \tag{A.93}$$

Par conséquent, la VaR approchée corrigée de l'effet d'estimation est :

$$a(\eta_{t-1},\hat{\theta}_n,\widetilde{VaR}) = \Phi^{-1}(\gamma) + \frac{1}{2n}\frac{\partial^2 Q}{\partial h^2}(0,\gamma) \tag{A.94}$$

$$\Leftrightarrow \widetilde{VaR} = g\Big(\eta_{t-1},\hat{\theta}_n;\ (\Phi^{-1}(\gamma) + \frac{1}{2n}\frac{\partial^2 Q}{\partial h^2}(0,\gamma))\ \Big) \tag{A.95}$$

□

A.9.3 Démonstration : le théorème 4.4 p49

Démonstration. On identifie alors les fonctions « a »et « g » :

$$\begin{cases} g(\varepsilon_{t-1},\theta,x) = \tilde{g}(\varepsilon_{t-1},\theta)x = \sigma_t(\theta)x \\ a(\varepsilon_{t-1},\theta,x) = \frac{x}{\tilde{g}(\varepsilon_{t-1},\theta)} = \frac{x}{\sigma_t(\theta)} \end{cases} \tag{A.96}$$

On peut alors calculer le terme de la dérivée seconde du quantile :

$$\frac{\partial a}{\partial \theta'}(\varepsilon_{t-1},\theta_0,g(\varepsilon_{t-1},\theta_0,s)) = \frac{\partial a}{\partial \theta'}(\varepsilon_{t-1},\theta_0,\sigma_t(\theta)s)) = \frac{\partial}{\partial \theta'}\frac{(\sigma_t(\theta_0)s)}{\sigma_t(\theta_0)} = -\frac{(\theta_0 s)}{\sigma_t(\theta_0)^2}\frac{\partial \sigma_t(\theta)}{\partial \theta'}(\theta_0) = -\frac{s}{\sigma_t(\theta_0)}\frac{\partial \sigma_t(\theta)}{\partial \theta'}(\theta_0) \tag{A.97}$$

Donc le premier terme s'écrit :

$$\frac{\partial}{\partial s}\Big(\frac{s^2}{\sigma_t(\theta_0)^2}\frac{\partial \sigma_t(\theta)}{\partial \theta'}(\theta_0)\Omega\frac{\partial \sigma_t(\theta)}{\partial \theta}(\theta_0)\Big) = \frac{2s}{\sigma_t(\theta_0)^2}\frac{\partial \sigma_t(\theta)}{\partial \theta'}(\theta_0)\Omega\frac{\partial \sigma_t(\theta)}{\partial \theta}(\theta_0) \tag{A.98}$$

Pour le second, on peut poser cet intermédiaire :

$$\frac{\partial ln f_{\eta_t}(s)}{\partial s} = \frac{\partial}{\partial s}\Big(-\frac{1}{2}ln(2\pi) - \frac{s^2}{2}\Big) = -s \tag{A.99}$$

D'où :

$$\frac{\partial^2 Q}{\partial h^2}(0,\gamma) = -\frac{2\Phi^{-1}(\gamma)}{\sigma_t(\theta_0)^2}\frac{\partial \sigma_t(\theta)}{\partial \theta'}(\theta_0)\Omega\frac{\partial \sigma_t(\theta)}{\partial \theta}(\theta_0) + \frac{\Phi^{-1}(\gamma)\Phi^{-1}(\gamma)^2}{\sigma_t(\theta_0)^2}\frac{\partial \sigma_t}{\partial \theta'}(\theta_0)\Omega\frac{\partial \sigma_t}{\partial \theta}(\theta_0) \tag{A.100}$$

Par conséquent,

$$\widetilde{VaR} = \sigma_t(\hat{\theta}_n)\Phi^{-1}(\gamma)\Big(1 - \frac{1}{2n}\frac{2}{\sigma_t(\theta_0)^2}\frac{\partial \sigma_t(\theta)}{\partial \theta'}(\theta_0)\Omega\frac{\partial \sigma_t(\theta)}{\partial \theta}(\theta_0) + \frac{1}{2n}\frac{\Phi^{-1}(\gamma)^2}{\sigma_t(\theta_0)^2}\frac{\partial \sigma_t}{\partial \theta'}(\theta_0)\Omega\frac{\partial \sigma_t}{\partial \theta}(\theta_0)\Big) \tag{A.101}$$

□

A.10 Estimation par maximisation de la vraisemblance

Principe et formalisation

Le statisticien dispose d'un ensemble de réalisations d'une seule variable aléatoire. Il suppose que ces réalisations sont indépendantes. Notons $(X_i)_{i\in[\![1;m]\!]}$ ces réalisations. Le statisticien cherche à déterminer la vraisemblance de l'échantillon vis-à-vis d'une hypothèse de travail : « la variable aléatoire X suit une loi dont la fonction de densité est f_X ». On écrit alors la vraisemblance ℓ ainsi :

$$\ell(X_{i\in[\![1;m]\!]}, f_X) = \prod_{j=1}^{m} f_X(X_i) \tag{A.102}$$

La maximisation de la vraisemblance est alors la recherche de la loi f_X (ou de ses paramètres) qui rend ℓ maximale.
On appelle maximisation de la quasi-vraisemblance (parfois pseudo-vraisemblance) lorsque dans le cas d'un modèle à innovations, on pré-suppose que les innovations sont gaussiennes (afin de simplifier les calculs).

Convergence

Si on suppose que l'estimateur $\hat{\theta}_n$ est obtenu par maximisation de la quasi-vraisemblance, on peut obtenir les équations suivantes :

$$\begin{cases} \sqrt{n}(\theta_n - \theta_0) \xrightarrow{d} \mathcal{N}(0,(\kappa-1)J^{-1}) \\ J = \mathbb{E}(\frac{\partial \ell_t}{\partial \theta}\frac{\partial \ell_t}{\partial \theta'}) \\ \kappa = \mathbb{E}(\eta_t^4) \\ \ell_t = -\frac{1}{2}ln(2\pi) - \frac{1}{2}(ln\sigma_t^2 + \frac{\varepsilon_t^2}{\sigma_t^2}) \end{cases} \tag{A.103}$$

Résultats empiriques

Afin d'obtenir une estimation de la variance de l'estimateur dans le cas d'un GARCH(1,1)$-\mathcal{N}$ de paramètres $(0,0001; 0,08; 0.8)$, on calcule ces grandeurs sur une plage de 3 000 000 dates.

$$\begin{cases} J = 10^{13} \times \begin{pmatrix} 4.107 & 2.727 \times 10^{-3} & 3.245 \times 10^{-3} \\ 2.727 \times 10^{-3} & 3.390 \times 10^{-6} & 2.306 \times 10^{-6} \\ 3.245 \times 10^{-3} & 2.306 \times 10^{-6} & 2.636 \times 10^{-6} \end{pmatrix} \\ V = \begin{pmatrix} 2.511 \times 10^{-6} & 7.480 \times 10^{-4} & -3.745 \times 10^{-3} \\ 7.480 \times 10^{-4} & 7.574 \times 10^{-1} & -1.583 \\ -3.745 \times 10^{-3} & -1.583 & 6.071 \end{pmatrix} \\ A = \begin{pmatrix} -1.302 \times 10^{1} \\ -9.608 \times 10^{-3} \\ -1.054 \times 10^{-2} \end{pmatrix} \\ \rho = \begin{pmatrix} -7.743 \times 10^{-8} \\ -2.223 \times 10^{-5} \\ 1.216 \times 10^{-4} \end{pmatrix} \\ \varsigma_c^2 = 0.00990844 \end{cases} \tag{A.104}$$

A.11 Propositions

Proposition A.1 (Lemme de Gouriéroux-Laurent-Scaillet [23]). *Soient X et Y deux variables aléatoires réelles indépendantes.*

$$\begin{cases} \mathbb{P}(X + hY < Q(\epsilon, \alpha)) = \alpha \Rightarrow Q(h, \alpha) = Q(0, \alpha) + h \times \dfrac{\partial Q}{\partial h}(0, \alpha) + o(h) \\ \\ avec \ \dfrac{\partial Q}{\partial h}(h, \alpha) = \mathbb{E}(Y | X + hY = Q(h, \alpha)) \end{cases} \tag{A.105}$$

c.f. Annexe ◇

Proposition A.2 (Lemme de Rau-Bredow reférence [4]). *Soit X et Y deux variables aléatoires. Pour un niveau de confiance γ donné, on note la Value-at-Risk (VaR) et l'Expected-Shortfall (ES) ainsi :*

$$\begin{cases} VaR = VaR_{1-\alpha}(X + hY) \\ ES = ES_{1-\alpha}(X + hY) = \int_0^\gamma VaR_{1-u}(X + hY)du \end{cases} \tag{A.106}$$

Alors :

$$\begin{cases} \dfrac{\partial VaR}{\partial h} = \mathbb{E}(Y | X + hY = VaR) \\ \dfrac{\partial^2 VaR}{\partial h^2} = -\Big(\dfrac{\partial \mathbb{V}(Y | X + hY = s)}{\partial s} + \mathbb{V}(Y | X + hY = s) \dfrac{\partial ln f_{X+hY}(s)}{\partial s} \Big)_{s=VaR} \\ \dfrac{\partial ES}{\partial h} = \mathbb{E}(Y | X + hY > VaR) \\ \dfrac{\partial^2 ES}{\partial h^2} = \dfrac{1}{\alpha} \mathbb{V}(Y | X + hY = VaR) f_{X+hY}(VaR) \end{cases} \tag{A.107}$$

◇

Annexe B

Code Matlab

B.1 Code - ToolBox de Kevin Sheppard

L'ensemble de ces fichiers sont le travail exclusif de M Kevin Sheppard, Professeur au département d'économie de l'Université d'Oxford, Royaume-Uni. Ils sont disponibles gratuitement à l'adresse internet suivante :
http://www.kevinsheppard.com/wiki/UCSD_GARCH

B.1.1 Fichier garchcore.m

```
function h=garchcore(data,parameters,covEst,p,q,m,T)
% PURPOSE:
%     Forward recursion to construct h's
%
% USAGE:
%     h=garchcore(data,parameters,covEst,p,q,m,T);
%
% INPUTS:
%     See garchlikelihood
%
% OUTPUTS:
%     See garchlikelihood
%
% COMMENTS:
%     Helper function part of UCSD_GARCH toolbox. Used if you do not use the MEX file.
%     You should use the MEX file.
%
%
% Author: Kevin Sheppard
% kevin.sheppard@economics.ox.ac.uk
% Revision: 2    Date: 12/31/2001
h=zeros(size(data)); h(1:m)=covEst; for t = (m + 1):T
   h(t) = parameters' * [1 ; data(t-(1:p)).^2;  h(t-(1:q)) ];
end
```

B.1.2 Fichier garchpq.m

```
function [parameters, likelihood, ht, stderrors, robustSE, scores,
grad] = garchpq(data , p , q , startingvals, options)
% PURPOSE:
%     GARCH(P,Q) parameter estimation with normal innovations using analytic derivatives
%
% USAGE:
%     [parameters, likelihood, ht, stderrors, robustSE, scores, grad] = garchpq(data , p , q , start
```

```
%
% INPUTS:
%     data: A single column of zero mean random data, normal or not for quasi likelihood
%
%     P: Non-negative, scalar integer representing a model order of the ARCH
%       process
%
%     Q: Positive, scalar integer representing a model order of the GARCH
%       process: Q is the number of lags of the lagged conditional variances included
%       Can be empty([]) for ARCH process
%
%     startingvals: A (1+p+q) vector of starting vals.  If you do not provide, a naieve guess of 1/(1
%       used for the arch and garch parameters, and omega is set to make the real unconditional varia
%       to the garch expectation of the expectation.
%
%     options: default options are below.  You can provide an options vector.  See HELP OPTIMSET
%
% OUTPUTS:
%     parameters : a [1+p+q X 1] column of parameters with omega, alpha1, alpha2, ..., alpha(p)
%                  beta1, beta2, ... beta(q)
%
%     likelihood = the loglikelihood evaluated at he parameters
%
%     ht = the estimated time varying VARIANCES
%
%     stderrors = the inverse analytical hessian, not for quasi maximum liklihood
%
%     robustSE = robust standard errors of form A^-1*B*A^-1*T^-1
%                where A is the analytic hessian
%                and B is the covariance of the scores
%
%     scores = the list of T scores for use in M testing
%
%     grad = the average score at the parameters
%
% COMMENTS:
%
%   GARCH(P,Q) the following(wrong) constratins are used(they are right for the (1,1) case or any Ar
%     (1) Omega > 0
%     (2) Alpha(i) >= 0 for i = 1,2,...P
%     (3) Beta(i)  >= 0 for i = 1,2,...Q
%     (4) sum(Alpha(i) + Beta(j)) < 1 for i = 1,2,...P and j = 1,2,...Q
%
%   The time-conditional variance, H(t), of a GARCH(P,Q) process is modeled
%   as follows:
%
%     H(t) = Omega + Alpha(1)*r_{t-1}^2 + Alpha(2)*r_{t-2}^2 +...+ Alpha(P)*r_{t-p}^2+...
%                    Beta(1)*H(t-1)+ Beta(2)*H(t-2)+...+ Beta(Q)*H(t-q)
%
%   Default Options
%
%   options  =  optimset('fmincon');
%   options  =  optimset(options , 'TolFun'      , 1e-003);
%   options  =  optimset(options , 'Display'     , 'iter');
%   options  =  optimset(options , 'Diagnostics' , 'on');
%   options  =  optimset(options , 'LargeScale'  , 'off');
%   options  =  optimset(options , 'MaxFunEvals' , '400*numberOfVariables');
%   options  =  optimset(options , 'GradObj'     , 'on');
```

```
%
%
%  uses GARCH_LIKELIHOOD and GARCHCORE.  You should MEX, mex 'path\garchcore.c', the MEX source
%  The included MEX is for R12, 12.1 and 11 Windows and was compiled with Intel Compiler 5.01.
%  It gives a 10-15 times speed increase
%
% Author: Kevin Sheppard
% kevin.sheppard@economics.ox.ac.uk
% Revision: 2    Date: 12/31/2001

if size(data,2) > 1
    error('Data series must be a column vector.')
elseif isempty(data)
    error('Data Series is Empty.')
end


if (length(q) > 1) | any(q < 0)
    error('Q must ba a single positive scalar or an empty vector for ARCH.')
end

if (length(p) > 1) | any(p <  0)
    error('P must be a single positive number.')
elseif isempty(p)
    error('P is empty.')
end

if isempty(q)
    q=0;
    m=p;
else
    m  =  max(p,q);
end


if nargin<=3 | isempty(startingvals)
    guess  =  1/(2*m+1);
    alpha  =  .15*ones(p,1)/p;
    beta   =  .75*ones(q,1)/q;
    omega  =  (1-(sum(alpha)+sum(beta)))*cov(data);  %set the uncond = to its expection
else
    omega=startingvals(1);
    alpha=startingvals(2:p+1);
    beta=startingvals(p+2:p+q+1);
end


LB         =  []; UB         =  []; sumA =  [-eye(1+p+q); ...
        0  ones(1,p)  ones(1,q)];
sumB =  [zeros(1+p+q,1);...
        1];


if (nargin <= 4) | isempty(options)
    options  =  optimset('fmincon');
    options  =  optimset(options , 'TolFun'      , 1e-003);
    %options  =  optimset(options , 'Display'     , 'iter');
    options  =  optimset(options , 'Display'     , 'off');
```

```
    %options  =  optimset(options , 'Diagnostics' , 'on');
    options  =  optimset(options , 'Diagnostics' , 'off');
    options  =  optimset(options , 'LargeScale'  , 'off');
    options  =  optimset(options , 'MaxFunEvals' , 400*(1+p+q));
    options  =  optimset(options , 'GradObj'     , 'on');
    %options  =  optimset(options , 'Algorithm'     , 'active-set');
end sumB = sumB - [zeros(1+p+q,1); 1]*2*optimget(options, 'TolCon',
1e-6);


stdEstimate =  std(data,1); data        =  [stdEstimate(ones(m,1)) ;
data];
% Estimate the parameters.
[parameters, LLF, EXITFLAG, OUTPUT, LAMBDA, GRAD] =
fmincon('garchlikelihood', [omega ; alpha ; beta] ,sumA  , sumB ,[]
, [] , LB , UB,[],options,data, p , q, m, stdEstimate);


if EXITFLAG<=0
    EXITFLAG
    fprintf(1,'Not Sucessful! \n')
end

parameters(find(parameters    <  0)) = 0;
parameters(find(parameters(1) <= 0)) = realmin; if nargout>1
    [likelihood, grad, hessian, ht, scores, robustSE] = garchlikelihood(parameters , data , p , q, m
    stderrors=hessian^(-1);
    likelihood=-likelihood;
end
```

B.1.3 Fichier garchlikelihood.m

```
function [LLF, grad, hessian, h, scores, robustse] =
garchlikelihood(parameters , data , p , q,  m, stdEstimate)
% PURPOSE:
%     Likelihood and analytic derivatives for garchpq
%
% USAGE:
%     [LLF, grad, hessian, h, scores, robustse] = garchlikelihood(parameters , data , p , q, m, stdE
%
%
% INPUTS:
%     parameters:   A vector of GARCH process aprams of the form [constant, arch, garch]
%     data:         A set of zero mean residuals
%     p:            The lag order length for ARCH
%     q:            The lag order length for GARCH
%     m:            The max of p and q
%     stdEstimate:  The sample standard deviation of the data
%
%
% OUTPUTS:
%     LLF:          Minus 1 times the log likelihood
%     grad:         The analytic gradient at the parameters
%     hessian:      The analytical hessian at the parameters
%     h:            The time series of conditional variances implied by the parameters and the data
%     scores:       A matrix, T x #params of the individual scores
%     robustse:     Quasi-ML Robust Standard Errors(Bollweslev Wooldridge)
%
```

```
%
% COMMENTS:
%     This is a helper function for garchpq
%
%
% Author: Kevin Sheppard
% kevin.sheppard@economics.ox.ac.uk
% Revision: 2    Date: 12/31/2001

parameters(find(parameters <= 0)) = realmin;

if isempty(q)
    m=p;
else
    m  =  max(p,q);
end T           =  size(data,1);
h=garchcore(data,parameters,stdEstimate^2,p,q,m,T);

t    = (m + 1):T; h=h(t); LLF  =  0.5 * (sum(log(h)) +
sum((data(t).^2)./h)  +  (T - m)*log(2*pi));

if nargout >= 2
%    keyboard
    h  =  [ones(m,1)*data(1)^2 ; h];
    garchp=parameters(2+p:1+p+q);
    [d,e,f]=garchgrad(garchp,p,q,data,h,m,T,stdEstimate);
    Base=(((data.^2)./h)-1)./(2*h);
    Base=repmat(Base,1,(1+p+q));
    gradsum=Base.*[d';e';f']';
    dht=[d';e';f']';
    grad=-sum(gradsum(t,:))';
end

if nargout >= 3
    scores=gradsum(t,:);
    hessian=zeros(1+p+q);
    for i=t
        hessian=hessian+(dht(i,:)'*dht(i,:))*(data(i)^2/(2*h(i)^3));
    end
    Tau=T-m;
    A=hessian/Tau;
    Bscores=scores-repmat(mean(scores),Tau,1);
    B=(Bscores'*Bscores)/Tau;
    robustse=(A^-1)*B*(A^-1)*(T^-1);
    h=h(t);
end
```

B.1.4 Fichier garchgrad.m

```
function [d,e,f]=garchgrad(garchp,p,q,data,h,m,T,stdEstimate)
% PURPOSE:
%     Analytic gradiedn m file for garchpq(use MEX file)
%
% USAGE:
%     [d,e,f]=garchgrad(garchp,p,q,data,h,m,T,stdEstimate)
%
% INPUTS:
```

```
%     See garchlikelihood
%
% OUTPUTS:
%     See garchlikelihood
%
% COMMENTS:
%     Helper function part of UCSD_GARCH toolbox. Used if you do not use the MEX file.
%     You should use the MEX file.
%
%
% Author: Kevin Sheppard
% kevin.sheppard@economics.ox.ac.uk
% Revision: 2    Date: 12/31/2001

t=m+1:T; d=ones(size(h)); e=ones(size(h,1),p)*stdEstimate^2;
f=ones(size(h,1),q)*stdEstimate^2; for i=t
    d(i)=1+d(i-1:-1:i-q)'*garchp;
    for j=1:p
        e(i,j)=data(i-j)^2+garchp'*e(i-1:-1:i-q,j);
    end
    for j=1:q
        f(i,j)=h(i-j)+garchp'*f(i-1:-1:i-q,j);
    end
end
```

B.2 Code - Expérience n°1

B.2.1 Arborescence

Dans un même répertoire,

- exp1.m
- garchpq.m
- garchore.m
- garchlikelihood.m

B.2.2 Fichier exp1.m

```
%=========================================================================
%=================  Expérience n°1 ======================================
%=========================================================================

%nécessite les fichiers suivant garchpq.m, garchcore.m, garchgrad.m,
%gardelikelihood.m disponibles gratuitement à
%http://www.kevinsheppard.com/wiki/UCSD_GARCH

%But de l'expérience :
%Evaluer la vitesse de convergence des paramètres d'un modèle en fonction
%de la taille de l'échantillon

% Date : 14-05-2009
%=========================================================================

clear all close all format long

%Déclaration ============================================================

% Nombre de simulations (Monte Carlo)
nb_simu = 3000 ;
```

```
%Choix des paramètres du modèle de la fonction de ln(PnL)
omega = 0.0001; alpha = 0.08; beta = 0.8; rendement_0 =
omega*(1-alpha-beta); sigma2_0 = rendement_0 ;

% Définition des suites des estimations des paramètres
alpha_chap = zeros(10,nb_simu) ; beta_chap = zeros(10,nb_simu) ;
omega_chap = zeros(10,nb_simu) ;

%Partie Principale ==========================================================

fenetre = [50 100 150 200 250 300 350 400 450 500];
for n=50:50:1000  %enchainement des fenêtres
    ligne = n/50;
    for i=1:nb_simu  %Boucle de Monte Carlo
        %Affichage de l'avancement
        n
        i

        %Calcul de la trajectoire
        rendement = [1:n+1];
        sigma2 = [1:n+1] ;
        rendement(1) = rendement_0 ;
        sigma2(1) = sigma2_0 ;
        for j=2:(n+1)
            sigma2(j) = omega + alpha*rendement(j-1)^2 + beta * sigma2(j-1) ;
            rendement(j) = sqrt(sigma2(j))* normrnd(0,1) ;
        end

        %Estimation du modèle
        echantillon = rendement(2:n+1) ;
        [parameters, likelihood, ht, stderrors, robustSE, scores, grad] = garchpq(echantillon',1,1)
        omega_chap(ligne,i) = parameters(1)   ;   %sauvegarde de la valeur
        alpha_chap(ligne,i) = parameters(2)  ;   %sauvegarde de la valeur
        beta_chap(ligne,i) = parameters(3)     ;   %sauvegarde de la valeur
    end
end




%Erreur ======================================================================

omega_err = zeros(10,nb_simu) ; alpha_err = zeros(10,nb_simu) ;
beta_err  = zeros(10,nb_simu) ;

for ligne = 1:10
    for colonne =1:nb_simu
        omega_err(ligne,colonne) = omega_chap(ligne,colonne) - omega ;
        alpha_err(ligne,colonne) = alpha_chap(ligne,colonne) - alpha ;
        beta_err(ligne,colonne)  = beta_chap(ligne,colonne) - beta ;
    end
end



save('sauvegarde','omega_err','alpha_err','beta_err')
```

```
%Affichage
figure(1) boxplot(omega_err') hold on plot(zeros(1,13),'g--')
legend('\omega - \omega_{o}') xlabel('Taille de la fenêtre d
estimation (x50)'); ylabel('Répartition des écarts');

figure(2) boxplot(alpha_err') hold on plot(zeros(1,13),'g--')
legend('\alpha - \alpha{o}') xlabel('Taille de la fenêtre d
estimation (x50)'); ylabel('Répartition des écarts');

figure(3) boxplot(beta_err') hold on plot(zeros(1,13),'g--')
legend('\beta - \beta_{o}') xlabel('Taille de la fenêtre d
estimation (x50)'); ylabel('Répartition des écarts');
```

B.3 Code - Expérience n°2

B.3.1 Arborescence

Dans un même répertoire,

- exp2MC.m (fichier principal)
- exp2_fonction.m
- test_christoffersen.m
- test_portmanteau.m
- weibull.m
- garchgrad.m
- garchchore.m
- garchlikelihood.m
- garchpq.m

B.3.2 Fichier exp2MC.m

```
%=========================================================================
%=================  Expérience n°2 - fichier principal ==================
%=========================================================================
%fichier : MCexp2.m
%
%
%But de l'expérience :
%Evaluer les p-values des différents tests de backtesting pour la VaR
%approchée et pour la VaR théorique par une méthode de Monte Carlo
%
%nécesseite les fichiers suivants :
% + exp2_fonction.m => fonction de calcul des p-values
% + test_christofersen.m => fonction de test de Markov
% + test_portemanteau. => fonction de test de portemanteau
% + weibull.m => fonction de test sur les durée
%On utilise de plus : garchcore.m, garchgrad.m, garchlikelihood.m,
%garchpq.m disponibles gratuitement  à
%http://www.kevinsheppard.com/wiki/UCSD_GARCH

% Date : 11-05-2009
%=========================================================================
clear all close all format long

%Paramètres de l'expérience ==============================================
nb_simu = 10000 ;
```

```
%Déclaration des variables
% _a pour la valeur estimée
% _th pour la valeur théorique

%Grandeurs générales
moy_a = 0 ; moy_th = 0 ; indicatrices =  [[0 0];[0 0]] ;

%Pour le test de Kupiec
kupiec_a = 0 * [1:nb_simu] ; kupiec_th = 0 * [1:nb_simu] ; kup_a=0;
kup_th=0;

%Pour le test de Markov
% Nombre de rejet et nombre de test, en fonction de l'hypothèse
m_uc_a = 0 ; m_ind_a = 0 ; m_cc_a = 0 ; m_uc_th = 0 ; m_ind_th = 0 ;
m_cc_th = 0 ;

m_uc_a_t = 0 ; m_ind_a_t = 0 ; m_cc_a_t = 0 ; m_uc_th_t = 0 ;
m_ind_th_t = 0 ; m_cc_th_t = 0 ;


%Pour le test des durées (w pour weibull)
% Nombre de rejet et nombre de test, en fonction de l'hypothèse
w_uc_a = 0 ; w_ind_a = 0 ; w_cc_a = 0 ; w_uc_th = 0 ; w_ind_th = 0;
w_cc_th = 0 ;

w_uc_a_t = 0 ; w_ind_a_t = 0 ; w_cc_a_t = 0 ; w_uc_th_t = 0 ;
w_ind_th_t = 0 ; w_cc_th_t = 0 ;

%Pour le test de portmanteau (lb pour ljung-bexin)
% Nombre de rejet et nombre de test, en fonction de l'hypothèse
% test sur les 5 ou 10 premiers coefficients

lb5_ind_a = 0; lb10_ind_a = 0; lb5_ind_th = 0; lb10_ind_th = 0 ;

lb5_ind_a_t = 0; lb10_ind_a_t = 0 ; lb5_ind_th_t = 0 ; lb10_ind_th_t
= 0 ;




%Partie principale ======================================================
for i=1:nb_simu
    i
[ m_a, m_th, indicatrices_croisees, z_th, z_a, LR_m_uc_a,
LR_m_ind_a, LR_m_cc_a, LR_m_uc_th, LR_m_ind_th, LR_m_cc_th,
LR_w_uc_a, LR_w_ind_a, LR_w_cc_a, LR_w_uc_th, LR_w_ind_th,
LR_w_cc_th, LR_p5_uc_a, LR_p5_uc_th, LR_p10_uc_a, LR_p10_uc_th] =
exp2_fonction() ;

    % Moyenne temporelle de dépassement
    moy_a = moy_a + m_a /nb_simu ;
    moy_th = moy_th + m_th / nb_simu ;

    % Tableau croisées des indicatrice
    indicatrices(1,1) = indicatrices(1,1) + indicatrices_croisees(1,1)/nb_simu ;
    indicatrices(1,2) = indicatrices(1,2) + indicatrices_croisees(1,2)/nb_simu ;
    indicatrices(2,1) = indicatrices(2,1) + indicatrices_croisees(2,1)/nb_simu ;
```

```
indicatrices(2,2) = indicatrices(2,2) + indicatrices_croisees(2,2)/nb_simu ;

% Sauvegarde des statistiques deKupiec
kupiec_a(i) = z_a ;
kupiec_th(i) = z_th  ;

% Calcul du résultat de la règle de décision pour Kupiec
if( abs(z_a) > 2.5758 * sqrt(0.01*(1-0.01)))
    kup_a = kup_a+1 ;
end
if( abs(z_th) > 2.5758* sqrt(0.01*(1-0.01)))
    kup_th = kup_th+1 ;
end

% Calcul du résultat de la règle de décision pour Markov
if(isnan(LR_m_uc_a)<1)
    m_uc_a_t = m_uc_a_t +1 ;
    if( LR_m_uc_a > 6.63)
       m_uc_a = m_uc_a +1 ;
    end
end
if(isnan(LR_m_ind_a)<1)
   m_ind_a_t = m_ind_a_t +1 ;
   if( LR_m_ind_a > 6.63)
       m_ind_a = m_ind_a +1 ;
   end
end
if(isnan(LR_m_cc_a)<1)
    m_cc_a_t = m_cc_a_t +1 ;
    if( LR_m_cc_a > 9.21)
        m_cc_a = m_cc_a +1;
    end
end
if(isnan(LR_m_uc_th)<1)
    m_uc_th_t = m_uc_th_t + 1 ;
     if( LR_m_uc_th > 6.63)
        m_uc_th = m_uc_th +1 ;
     end
end
if(isnan(LR_m_ind_th)<1)
    m_ind_th_t = m_ind_th_t + 1;
   if( LR_m_ind_th > 6.63)
        m_ind_th = m_ind_th +1 ;
   end
end
if(isnan(LR_m_cc_th)<1)
    m_cc_th_t = m_cc_th_t + 1 ;
   if( LR_m_cc_th > 9.21)
       m_cc_th = m_cc_th +1;
   end
end

% Calcul du résultat de la règle de décision pour les durées
if(isnan(LR_w_uc_a)<1)
    w_uc_a_t = w_uc_a_t + 1;
    if( LR_w_uc_a > 6.63)
        w_uc_a = w_uc_a +1 ;
    end
```

```
   end
   if(isnan(LR_w_ind_a)<1)
       w_ind_a_t = w_ind_a_t + 1;
       if( LR_w_ind_a > 6.63)
           w_ind_a = w_ind_a +1 ;
       end
   end
   if(isnan(LR_w_cc_a)<1)
       w_cc_a_t = w_cc_a_t + 1;
       if( LR_w_cc_a > 6.63)
           w_cc_a = w_cc_a +1 ;
       end
   end
   if(isnan(LR_w_uc_th)<1)
       w_uc_th_t = w_uc_th_t + 1;
       if( LR_w_uc_th > 6.63)
           w_uc_th = w_uc_th +1 ;
       end
   end
   if(isnan(LR_w_ind_th)<1)
       w_ind_th_t = w_ind_th_t + 1;
       if( LR_w_ind_th > 9.21)
           w_ind_th = w_ind_th +1 ;
       end
   end
   if(isnan(LR_w_cc_th)<1)
       w_cc_th_t = w_cc_th_t + 1 ;
       if( LR_w_cc_th > 9.21)
           w_cc_th = w_cc_th +1 ;
       end
   end

   % Calcul du résultat de la règle de décision pour Portmanteau
   if( isnan(LR_p5_uc_a)<1)
       lb5_ind_a_t = lb5_ind_a_t + 1 ;
       if(LR_p5_uc_a > 15.09)
           lb5_ind_a = lb5_ind_a +1 ;
       end
   end
   if( isnan(LR_p10_uc_a)<1)
       lb10_ind_a_t = lb10_ind_a_t + 1 ;
       if(LR_p10_uc_a > 23.21)
           lb10_ind_a = lb10_ind_a +1 ;
       end
   end
    if( isnan(LR_p5_uc_th)<1)
        lb5_ind_th_t = lb5_ind_th_t + 1 ;
        if(LR_p5_uc_th > 15.09)
            lb5_ind_th = lb5_ind_th +1 ;
        end
    end
    if( isnan(LR_p10_uc_th)<1)
        lb10_ind_th_t = lb10_ind_th_t + 1 ;
        if(LR_p5_uc_th > 23.21)
            lb10_ind_th = lb10_ind_th +1 ;
        end
    end
end
```

```
%Affichage pour le test de Kupiec ======================================
figure(1) subplot(1,2,1) hist(kupiec_a) subplot(1,2,2)
hist(kupiec_th)


%Affichage résultats du backtesting ====================================

%Test de Kupiec
kup_a kup_th [a,b]=normfit(kupiec_a) [a,b]=normfit(kupiec_th)

%Test de Markov
m_uc_a m_uc_a_t m_ind_a m_ind_a_t m_cc_a m_cc_a_t m_uc_th m_uc_th_t
m_ind_th m_ind_th_t m_cc_th m_cc_th_t

%Test des durées
w_uc_a w_uc_a_t w_ind_a w_ind_a_t w_cc_a w_cc_a_t w_uc_t w_uc_th_t
w_ind_th w_ind_th_t w_cc_th w_cc_th_t

%Test de Portmanteau
lb5_ind_a lb5_ind_a_t lb10_ind_a lb10_ind_a_t lb5_ind_th
lb5_ind_th_t lb10_ind_th lb10_ind_th_t
```

B.3.3 Fichier exp2_fonction.m

```
%=========================================================================
%======================== Expérience n°2 =================================
%=========================================================================
% fichier :   exp_2.m
%
%Fonction :
%cette fonction prend aucun argument. Elle consiste à générer une
%trajectoire selon un modèle GARCH(1,1) dont les paramètres sont définis
%dans le code (omega = 10-4, alpha = 0.08 et beta=0.8). Les différents
%retours sont indicés par "_a" lorsque la grandeur porte sur la VaR
%approchée et par "_th" lorsque la grandeur porte sur la VaR théorique.
%Chaque p-value comporte une lettre de référence pour le test
%(par ex "_k"=> Kupiec) et un groupement pour l'hypothèse nulle (par ex
% "_ind" => indépendance).
%
%
%nécesseite les fichiers suivants :
% + test_christofersen.m => fonction de test de Markov
% + test_portemanteau. => fonction de test de portemanteau
% + VaR.m => fonction de calcul de la VaR (pour un modèle GARCH(1,1))
% + weibull.m => fonction de test sur les durée
%On utilise de plus : garchcore.m, garchgrad.m, garchlikelihood.m,
%garchpq.m disponibles gratuitement  à
%http://www.kevinsheppard.com/wiki/UCSD_GARCH

% Date : 13-05-2009
%=========================================================================
```

```
function [ m_a, m_th, indicatrices_croisees, z_th, z_a, LR_m_uc_a,
LR_m_ind_a, LR_m_cc_a, LR_m_uc_th, LR_m_ind_th, LR_m_cc_th,
LR_w_uc_a, LR_w_ind_a, LR_w_cc_a, LR_w_uc_th, LR_w_ind_th,
LR_w_cc_th, LR_p5_uc_a, LR_p5_uc_th, LR_p10_uc_a, LR_p10_uc_th] =
exp2_fonction() ;


format long

%Paramètres de l'expérience :
T = 750 ;            % Taille maximale de l'échantillon
fenetre = 500 ; % Taille de la fenêtre d'estimation des paramètres
seuil = 0.01 ;  % seuil de confiance en la VaR



%============   Génération du modèle théorique ============================
%============================================================================

%Choix des paramètres du modèle
omega = 0.0001; alpha = 0.08; beta = 0.8; log_rendement_0 = 0;
sigma2_0 = 0 ;


%Génération du modèle
log_rendement = [1:T];                % Vecteur des rendements log
log_rendement(1) = log_rendement_0 ;

cours = [1:T];                        % Vecteur du cours de l'actif
cours(1) = 1 ;

var_th=[1;T] ;                        % Vecteur de la VaR théorique
var_th(1) = 0 ;

PnL = [1:T];                          % Vecteur des PnL
PnL(1) =0 ;

sigma2 = [1:T] ; sigma2(1) = sigma2_0 ;


for i=2:T
    alea = randn(1) ;
    sigma2(i) = omega + alpha * log_rendement(i-1)^2 + beta * sigma2(i-1) ;
    log_rendement(i) = sqrt(sigma2(i)) * alea ;
end

for i=2:T
    cours(i) = cours(i-1) * exp(log_rendement(i)) ;
end

for i=2:T
    PnL(i) = cours(i) - cours(i-1) ;
end

for i=2:T
    var_th(i) = cours(i-1) * (1- exp(sqrt(sigma2(i))*norminv(seuil))) ;
```

```
end


%================== Estimation et Prédiction ============================
%========================================================================

% Définition des suites des estimations des paramètres
echantillon = log_rendement(1:fenetre) ; [parameters, likelihood,
ht, stderrors, robustSE, scores, grad] = garchpq(echantillon',1,1) ;
omega_chap = parameters(1) ; alpha_chap =parameters(2) ; beta_chap =
parameters(3) ;

% Prédiction :
var_a = [1:T] ; var_a(1) = 0 ; sigma2_chap = [1:T]; sigma2_chap(1) =
0 ;

for i=2:(T+1)
    sigma2_chap(i) = omega_chap + alpha_chap * log_rendement(i-1)^2 + beta_chap * sigma2_chap(i-1) ;
    var_a(i) = cours(i-1) * (1 - exp(sqrt(sigma2_chap(i))*norminv(seuil))) ;
end


%===================== Backtesting simple ===============================
%========================================================================

% Indicatrices de violations
I_a = 0*[fenetre+1:T] ; I_th = 0*[fenetre+1:T] ;

for i=(fenetre+1):1:T
    if( PnL(i) < - var_a(i))
        I_a(i) = 1 ;
    else
        I_a(i) = 0 ;
    end
    if( PnL(i) < - var_th(i))
        I_th(i) = 1 ;
    else
        I_th(i) = 0 ;
    end
end


% Tableau croisé des indicatrices de violations
indicatrices_croisees = zeros(2,2) ; for i =fenetre+1:1:T
    if( I_a(i)>0 )
        indicatrices_croisees(2,1) = indicatrices_croisees(2,1) + (1-I_th(i))/(T-fenetre-1) ;
        indicatrices_croisees(2,2) = indicatrices_croisees(2,2) + I_th(i)/(T-fenetre-1) ;
    else
        indicatrices_croisees(1,1) = indicatrices_croisees(1,1) + (1-I_th(i))/(T-fenetre-1) ;
        indicatrices_croisees(1,2) = indicatrices_croisees(1,2) + I_th(i)/(T-fenetre-1) ;
    end
end indicatrices_croisees ;


% Seuil empirique
m_a=0; m_th=0; for i=(fenetre+1):1:T
    m_a = m_a + I_a(i) ;
    m_th = m_th + I_th(i) ;
```

```
end

m_a=m_a/(T-fenetre-1) ; m_th=m_th/(T-fenetre-1) ;



%=================== Backtesting : Tests Statistiques ====================
%=========================================================================


%Test de Kupiec ==========================================================

%Test pour la VaR approchée
z_a = sqrt(T-fenetre-1) * (m_a - seuil) ; %/ sqrt( seuil*(1-seuil) ) ;
%Test pour la VaR théorique
z_th = sqrt(T-fenetre-1) * (m_th - seuil); % / sqrt( seuil*(1-seuil) ) ;




%Test : processus markovien d'ordre 1 ====================================

[LR_m_uc_a, LR_m_ind_a, LR_m_cc_a ] =
test_christoffersen(seuil,I_a((fenetre+1):T)) ; [LR_m_uc_th,
LR_m_ind_th, LR_m_cc_th ] =
test_christoffersen(seuil,I_th((fenetre+1):T))  ;



%Test : durée Weibull ====================================================

[LR_w_uc_a, LR_w_ind_a, LR_w_cc_a ] =
weibull(seuil,I_a((fenetre+1):T)); [LR_w_uc_th, LR_w_ind_th,
LR_w_cc_th ] = weibull(seuil,I_th((fenetre+1):T));


%Test : portmanteau ======================================================
[LR_p5_uc_a] = test_portmanteau(I_a((fenetre+1):T), seuil, 5) ;
[LR_p5_uc_th] = test_portmanteau(I_th((fenetre+1):T), seuil, 5) ;
[LR_p10_uc_a] = test_portmanteau(I_a((fenetre+1):T), seuil, 10) ;
[LR_p10_uc_th] = test_portmanteau(I_th((fenetre+1):T), seuil, 10) ;



end
```

B.3.4 fichier test_portmanteau.m

```
%=========================================================================
%======================== Expérience n°2 =================================
%=========================================================================
% fichier :   test_portmanteau.m
%
%Fonction :
%cette fonction prend en argument le seuil théorique de la VaR (alpha), la
%suite des indicatrices (I), la taille de la fenêtre d'estimation (fenetre)
%et la taille totale de l'échantillon (T) et le degré de nullite (m).
%Cette fonction renvoie les p-value selon un test de portemanteau sur les m
```

```
%premiers coefficients d'autocorrélation du processus I.
%
% Source : Backtesting Value-at-Risk : A Duration based Approcah,
% Christoffersen, Pelletier, 2004
%
% Date : 11-05-2009
%=========================================================================


function [ LB ] = test_portmanteau(I, alpha, m )

[ligne,T] = size(I) ;

gamma = 1:m ;

moyenne = 0 ;
for t=1:T
    moyenne = moyenne + I(t) ;
end
moyenne = moyenne / T ;

if(moyenne >0)
    variance = 0 ;
    for t=1:T
    variance = variance + (I(t)-moyenne)^2 ;
    end
    variance = variance / (T-1)  ;

    for j=1:m
            gamma(j) = 0 ;
            X1 = I(1:T-j) - moyenne ;
            X2 = I(1+j:T) - moyenne ;
            gamma(j) = (X1 * X2')/((T-j)*variance) ;
    end

    LB = 0 ;
    for j=1:m
        LB = LB + gamma(j)^2/(T-j) ;
    end

    LB = (T+2)*T*LB ;
else
    LB = 0/0 ;
end
```

B.3.5 Fichier test_christoffersen.m

```
%=========================================================================
%======================== Expérience n°2 =================================
%=========================================================================
% fichier :   test_christoffersen
%
%Fonction :
%cette fonction prend en argument le seuil théorique de la VaR (alpha), la
%suite des indicatrices (I), la taille de la fenêtre d'estimation (fenetre)
%et la taille totale de l'échantillon (T).
%Cette fonction renvoie les p-value selon un test correspondant à une
```

```
%hypothèse nulle où le processus suit un procecessus de Markov d'ordre 1
%avec les cas possible : couverture inconditionnelle ("uc"), indépendance
%("ind") et couverture conditionnelle ("cc").
%
%
% Source : Backtesting Value-at-Risk : A Duration based Approcah,
% Christoffersen, Pelletier, 2004
%
% Date : 18-05-2009
%=========================================================================
function [LR_uc, LR_ind, LR_cc ] = test_christoffersen(alpha, I )

[ligne,T] = size(I) ;
% f et T sont pairs !
%
% |==========|=|(==)(==)(==)(==)(==)(==)|=|
% 1           f |               |          | T
%               f+1             |          T-1
%                               |
%                     couple pour test

%Déclaration des nombres de 0, de 1, de...
n0 = 0 ;
n1 = 0 ;
nb_markov=zeros(2,2) ;

p0 = 0 ;
p1 = 0 ;
markov=zeros(2,2) ;

if( I(1) > 0 )
    n1 = 1 ;
else
    n0 = 1 ;
end

for t=2:T
    if( I(t) == 1 )
        n1 = n1 + 1 ;
        if( I(t-1) == 1 )
            nb_markov(2,2) = nb_markov(2,2) + 1 ;
        else
            nb_markov(1,2) = nb_markov(1,2) + 1 ;
        end
    else
        n0 = n0 + 1 ;
        if( I(t-1) == 1 )
            nb_markov(2,1) = nb_markov(2,1) + 1 ;
        else
            nb_markov(1,1) = nb_markov(1,1) + 1 ;
        end
    end
end

markov = nb_markov / (T-1) ;
p0 = n0 / T ;
p1 = n1 / T ;
```

```
%Coucerture unconditionnelle
L1 = alpha^n1 * (1-alpha)^n0 ;
L2 = p1^n1 * p0^n0 ;
L3 =  (1-markov(1,2))^(n0 - nb_markov(1,2)) * markov(1,2)^nb_markov(1,2) * (1 - markov(2,2))^(n1 - nb

LR_uc = abs(2 * log( L2/L1 )) ;
LR_ind = abs( 2 * log(L2/L3)) ;
LR_cc = abs(2 * log(L1/L3)) ;
```

B.3.6 Fichier weibull.m

```
%========================================================================
%======================== Expérience n°2 ================================
%========================================================================
% fichier :   weibull.m
%
%Fonction :
%cette fonction prend en argument le seuil théorique de la VaR (alpha), la
%suite des indicatrices (I), la taille de la fenêtre d'estimation (fenetre)
%et la taille totale de l'échantillon (T).
%Cette fonction renvoie les p-value pour l' indépendance ("ind") et la
%couverture conditionnelle ("cc"). Le test repose sur le fait que les
%durées entre deux violations successives de la VaR suivent une loi
%géométrique.
%
%
% Source : Backtesting Value-at-Risk : A Duration based Approcah,
% Christoffersen, Pelletier, 2004
%
% Date : 11-05-2009
%========================================================================


function [ LR_uc, LR_ind, LR_cc ] = weidbull ( alpha, I)


%Détermination de la distribution des durées
%=========================================
[ligne,T] = size(I) ;

ok = 0 ;
for t=1:T
    if( I(t)==1)
        ok = ok +1 ;
    end
end

if( ok>2 )
    distrib=[1:1]*0 ;
    indice_inf = 1 ;
    while( I(indice_inf) == 0 )
        indice_inf = indice_inf + 1 ;
    end

    indice_sup = T ;
    while(I(indice_sup)==0)
        indice_sup = indice_sup - 1 ;
```

```
    end

    duree = 1 ;
    indice = 1 ;
    en_cours = 1 ; %1 on est en train de compter
    for t=(indice_inf+1):indice_sup
        if( I(t) == 0)
            duree = duree + 1;
        else
            distrib(indice) = duree ;
            indice = indice + 1 ;
            duree = 1 ;
        end
    end

    L_distrib = wbllike(wblfit(distrib'),distrib') ; %empi
    L_expo = explike(expfit(distrib'),distrib') ; %indépendance
    L_ideal = wbllike([1/alpha,1],distrib') ;  % parfait

    %UC
    LR_uc = 2 * abs( L_expo - L_ideal) ;

    %IND
    LR_ind = 2 * abs( L_distrib - L_expo) ;

    %CC
    LR_cc = 2 *abs( L_distrib - L_ideal) ; %==> totalement out
else
   LR_uc = 0/0 ;
   LR_ind = 0/0 ;
   LR_cc = 0/0 ;
end
end
```

B.4 Code - Expérience n°3

B.4.1 Arborescence

Dans un même répertoire,

- exp3bis_fonction.m
- MCexp3.m
- garchpq.m
- garchore.m
- garchlikelihood.m
- garchpq.m
- test_portmanteau.m
- test_christoffersen.m
- weibull.m

B.4.2 Fichier exp3bis_fonction.m

```
%=========================================================================
%======================== Expérience n°3 ================================
%=========================================================================
% fichier :   exp3_fonction.m
%
```

```
%
%
%nécesseite les fichiers suivants :
% + test_christofersen.m => fonction de test de Markov
% + test_portemanteau. => fonction de test de portemanteau
% + VaR.m => fonction de calcul de la VaR (pour un modèle GARCH(1,1))
% + weibull.m => fonction de test sur les durée
%On utilise de plus : garchcore.m, garchgrad.m, garchlikelihood.m,
%garchpq.m disponibles gratuitement  à
%http://www.kevinsheppard.com/wiki/UCSD_GARCH

% Date : 11-05-2009
%========================================================================




function [z, vc] = exp3_fonction()


format long

%Paramètres de l'expérience :
T = 5500 ;            % Taille maximale de l'échantillon
seuil = 0.01 ; % seuil de confiance en la VaR

fenetre = 500;


%============   Génération du modèle théorique ==========================
%========================================================================

%Choix des paramètres du modèle
omega = 0.0001;
alpha = 0.08;
beta = 0.8;
log_rendement_0 = 0;
sigma2_0 = 0 ;


%Génération du modèle
log_rendement = [1:T];                % Vecteur des rendements log
log_rendement(1) = log_rendement_0 ;

cours = [1:T];                        % Vecteur du cours de l'actif
cours(1) = 1 ;

PnL = [1:T];                          % Vecteur des PnL
PnL(1) =0 ;

sigma2 = [1:T] ;
sigma2(1) = sigma2_0 ;


for i=2:T
    alea = randn(1); % trnd(30) ;
    sigma2(i) = omega + alpha * log_rendement(i-1)^2 + beta * sigma2(i-1) ;
    log_rendement(i) = sqrt(sigma2(i)) * alea ;
```

```
end
for i=2:T
    cours(i) = cours(i-1) * exp(log_rendement(i)) ;
end
for i=2:T
    PnL(i) = cours(i) - cours(i-1) ;
end


%=================== Estimation et Prédiction ============================
%=========================================================================

% Définition des suites des estimations des paramètres
echantillon = log_rendement(1:fenetre) ;
[parameters, likelihood, ht, stderrors, robustSE, scores, grad] = garchpq(echantillon',1,1) ;
omega_chap = parameters(1) ;
alpha_chap =parameters(2) ;
beta_chap = parameters(3) ;


% Prédiction :
var_a = [1:T] ;
var_a(1) = 0 ;
sigma2_chap = [1:T];
sigma2_chap(1) = 0 ;

for i=2:T
    sigma2_chap(i) = omega_chap + alpha_chap * log_rendement(i-1)^2 + beta_chap * sigma2_chap(i-1) ;
    var_a(i) = cours(i-1) * (1 - exp(sqrt(sigma2_chap(i))*norminv(seuil))) ;
end


% Indicatrices de violations
I_a = 0*[1:T] ;

for i=(fenetre+1):T
    if( PnL(i) < - var_a(i))
        I_a(i) = 1 ;
    else
        I_a(i) = 0 ;
    end
end



%====================== Convergence ======================================
%=========================================================================

%Estimation de kappa

kappa = 0 ;
for i=2:fenetre
    kappa = kappa + log_rendement(i)^4 / sigma2_chap(i)^2 ;
end
kappa = kappa / (fenetre-1) ;


%estimation de derivée de sigma2 par rapport à beta
```

```
der_omega = [1:fenetre] ;
der_alpha = [1:fenetre] ;
der_beta = [1:fenetre] ;

der_beta(1) = 0;
der_omega(1) = 0;
der_alpha(1) = 0;

for i=2:fenetre
    der_omega(i) = 1 + beta_chap * der_omega(i-1)  ;
    der_alpha(i) = log_rendement(i-1)^2 + beta_chap * der_alpha(i-1)  ;
    der_beta(i) = sigma2(i-1) + beta_chap * der_beta(i-1) ;
end

der_beta(1) = mean(der_beta);
der_omega(1) = mean(der_omega);
der_alpha(1) =mean(der_alpha);
for i=2:fenetre
    der_omega(i) = 1 + beta_chap * der_omega(i-1)  ;
    der_alpha(i) = log_rendement(i-1)^2 + beta_chap * der_alpha(i-1)  ;
    der_beta(i) = sigma2_chap(i-1) + beta_chap * der_beta(i-1) ;
end


% estimation derivée de l ronde
ell = zeros(3,fenetre) ;
for i=2:fenetre
    l = [der_omega(i)  der_alpha(i)  der_beta(i) ]' ;
    ell(:,i) = (1/sigma2_chap(i)) * l ;
end


% estimation de J
J = zeros(3,3) ;
for i=2:fenetre
    J= J + (ell(:,i)*ell(:,i)') ;
end
J_inv = inv(J)* (fenetre-1) ;


% estimation de rho
rho = zeros(3,1) ;
for i=2:fenetre
    rho = rho + (I_a(i)-seuil) * J_inv * ell(:,i) ;
end
rho = rho / (fenetre-1) ;

% estimation de A
A = zeros(3,1) ;
facteur_commun = norminv(seuil,0,1) * exp(-norminv(seuil,0,1)^2) /(2*sqrt(2*pi)) ;
for i = 2:fenetre
    A = A + facteur_commun  * ell(:,i) ;
end

A = A / (fenetre-1) ;
```

```
vnc = seuil*(1-seuil) ;
vc = seuil*(1-seuil) + 2 * A' * rho + A' * (kappa-1) * J_inv * A ;


% Calcul de la statistique z

m_a=0;
for i=(fenetre+1):T
    m_a = m_a + I_a(i) ;
end
m_a=m_a/(T-fenetre) ;


z = sqrt(T-fenetre) * (m_a - seuil) ;









end



%
% %====================== Convergence théorique ============================
% %=========================================================================
%
% %Estimation de kappa
%
% kappa_th = 0 ;
% for i=1:fenetre
%     kappa_th = kappa_th + log_rendement(i)^4 / sigma2(i)^2 ;
% end
% kappa_th = kappa_th / fenetre ;
%
%
% %estimation de derivée de sigma2 par rapport à beta
%
% der_omega_th = [1:T] ;
% der_alpha_th = [1:fenetre] ;
% der_beta_th = [1:fenetre] ;
%
% der_beta_th(1) = 0;
% der_omega_th(1) = 0;
% der_alpha_th(1) = 0;
%
% for i=2:fenetre
%     der_omega_th(i) = 1 + beta * der_omega_th(i-1)  ;
%     der_alpha_th(i) = log_rendement(i-1)^2 + beta * der_alpha_th(i-1)  ;
%     der_beta_th(i) = sigma2(i-1) + beta * der_beta_th(i-1) ;
% end
%
```

```
% der_beta_th(1) = der_beta_th(T/2);
% der_omega_th(1) = der_omega_th(T/2);
% der_alpha_th(1) =der_alpha_th(T/2);
% for i=2:fenetre
%     der_omega_th(i) = 1 + beta * der_omega_th(i-1)  ;
%     der_alpha_th(i) = log_rendement(i-1)^2 + beta * der_alpha_th(i-1)  ;
%     der_beta_th(i) = sigma2(i-1) + beta * der_beta_th(i-1) ;
% end
%
% % estimation derivée de l ronde
%
% ell_th = zeros(3,fenetre) ;
%
% for i=1:fenetre
%     l_th = [der_omega_th(i) der_alpha_th(i)  der_beta_th(i) ]' ;
%     ell_th(:,i) = (1/sigma2(i)) * l_th ;
% end
%
%
% % estimation de J
%
% J_th = zeros(3,3) ;
%
% for i=1:fenetre
%     J_th= J_th + (ell_th(:,i)*ell_th(:,i)') ;
% end
%
% J_th = J_th  ;
% J_th_inv = inv(J_th) *(T-1) ;
%
%
% % estimation de rho
%
% rho_th = zeros(3,1) ;
% for i=1:fenetre
%     rho_th = rho_th + (I_th(i)-seuil) * J_th_inv * ell_th(:,i) ;
% end
% rho_th = rho_th / (T-1)  ;
%
% % estimation de A
%
% A_th = zeros(3,1) ;
% facteur_commun = norminv(seuil,0,1) * exp(-norminv(seuil,0,1)^2) /(2*sqrt(2*pi)) ;
% for i = 1:fenetre
%     A_th = A_th + facteur_commun  * ell_th(:,i) ;
% end
%
% A_th = A_th / fenetre ;
%
% a_rho = 2 * A_th' * rho_th;
% a_v_a = A_th' * (kappa_th-1) * J_th_inv * A_th;
%
% vc_th = seuil*(1-seuil) + 2 * A_th' * rho_th + A_th' * (kappa_th-1) * J_th_inv * A_th ;
%
```

B.4.3 Fichier MCexp3.m

```
%=========================================================================
%=================  Expérience n°3 - fichier principal ==================
%=========================================================================
%fichier : MCexp3.m
%
%
%But de l'expérience :
%Evaluer les p-values des différents tests de backtesting pour la VaR
%approchée et pour la VaR théorique par une méthode de Monte Carlo
%
%nécesseite les fichiers suivants :
% + exp2_fonction.m => fonction de calcul des p-values
% + test_christofersen.m => fonction de test de Markov
% + test_portemanteau. => fonction de test de portemanteau
% + VaR.m => fonction de calcul de la VaR (pour un modèle GARCH(1,1))
% + weibull.m => fonction de test sur les durée
%On utilise de plus : garchcore.m, garchgrad.m, garchlikelihood.m,
%garchpq.m disponibles gratuitement  à
%http://www.kevinsheppard.com/wiki/UCSD_GARCH

% Date : 11-05-2009
%=========================================================================
clear all
close all

nb_simu = 5000 ;

table = zeros(2,nb_simu) ;
nb_rejet_1pc_nc = 0 ;
nb_rejet_1pc_c = 0 ;
nb_rejet_5pc_nc = 0 ;
nb_rejet_5pc_c = 0 ;
nb_rejet_10pc_nc = 0 ;
nb_rejet_10pc_c = 0 ;


vnc = 0.01*(1-0.01) ;
vc_th = 0.009908643640275 ;

for i=1:nb_simu
    i
    [z, vc ] = exp3bis_fonction() ;

    table(1,i) = z ;
    table(2,i)  = vc ;

    if( abs(z) > 2.5758 * sqrt(vnc))
        nb_rejet_1pc_nc = nb_rejet_1pc_nc  + 1;
    end
    if( abs(z) > 2.5758 * sqrt(vc))
        nb_rejet_1pc_c = nb_rejet_1pc_c  + 1;
    end

    if( abs(z) > 1.96 * sqrt(vnc))
```

```
        nb_rejet_5pc_nc = nb_rejet_5pc_nc  + 1;
    end
    if( abs(z) > 1.96 * sqrt(vc))
        nb_rejet_5pc_c = nb_rejet_5pc_c  + 1;
    end

    if( abs(z) > 1.6449 * sqrt(vnc))
        nb_rejet_10pc_nc = nb_rejet_10pc_nc  + 1;
    end
    if( abs(z) > 1.6449 * sqrt(vc))
        nb_rejet_10pc_c = nb_rejet_10pc_c  + 1;
    end

end


nb_rejet_1pc_nc
nb_rejet_1pc_c
nb_rejet_5pc_nc
nb_rejet_5pc_c
nb_rejet_10pc_nc
nb_rejet_10pc_c



droite = [1:nb_simu];
for i=1:nb_simu
    droite(i) = vnc ;
end

droite2 = [1:nb_simu];
for i=1:nb_simu
    droite2(i) = vc_th ;
end

a=mean(table(2,:))


figure(1)
plot(droite,'r')
hold on
plot(droite2,'r--')
hold on
plot(table(2,:),'b')
legend('variance asymptotique','variance corrigée théorique asymptotique','variance corrigée estimée

figure(2)
hist((table(1,:)),50)


vnc
vc_th
[a,b] = normfit(table(1,:))
```

Table des matières

I Introduction : l'importance de la VaR comme mesure de risque 5

II Procédures de backtesting de la VaR 13

1 Modélisation 17
1.1 Modélisation financière . . . 17
1.1.1 Données empiriques . . . 17
1.1.2 GARCH . . . 17
1.1.3 Conclusion . . . 22
1.2 VaR : théorie et estimation . . . 24
1.2.1 Définition . . . 24
1.2.2 Développement - approximation . . . 25
1.2.3 Estimation . . . 26
1.2.4 Conclusion . . . 26
1.3 Conclusion sur le Modélisation . . . 26

2 Tests statistiques 27
2.1 Principes des tests . . . 27
2.1.1 Test statistique . . . 27
2.1.2 Backtesting de la VaR . . . 27
2.2 Test de Kupiec . . . 28
2.2.1 Repères . . . 28
2.2.2 Démarche théorique . . . 29
2.2.3 Démarche algorithmique . . . 29
2.3 Test de Markov . . . 29
2.3.1 Repères . . . 29
2.3.2 Démarche théorique . . . 29
2.3.3 Démarche algorithmique . . . 30
2.4 Test de Portmanteau . . . 31
2.4.1 Repères . . . 31
2.4.2 Démarche théorique . . . 31
2.4.3 Démarche algorithmique . . . 32
2.5 Test des durées . . . 32
2.5.1 Repères . . . 32
2.5.2 Démarche théorique . . . 33
2.5.3 Démarche algorithmique . . . 34
2.6 Autres tests . . . 35
2.6.1 Test des quantiles dynamiques . . . 35
2.6.2 Test sur plusieurs niveaux conjoints . . . 35
2.6.3 Test sur les moments . . . 35
2.7 Remarques sur l'implementation informatique . . . 36
2.8 Conclusion tests . . . 37

III Effet d'estimation **39**

3 Éléments empiriques **43**
3.1 Résultats . . . 43
3.2 Présentation des résultats sur les niveaux . . . 43
3.3 Analyse générale des niveaux . . . 44
3.4 Analyse par taille de la fenêtre d'estimation des niveaux . . . 44
3.5 Analyse par hypothèse des niveaux . . . 45
3.6 Conclusion . . . 45

4 Correction de la VaR **47**
4.1 Comment estimer la VaR ? . . . 47
4.2 Méthode semi-paramétrique . . . 47
4.3 Méthode paramétrique : cas général . . . 47
4.3.1 Modèle . . . 48
4.3.2 VaR théorique . . . 48
4.3.3 VaR estimée . . . 48
4.4 Modèle paramètrique : cas d'un GARCH(1,1)$-\mathcal{N}(0,1)$ . . . 49
4.4.1 Expression générale . . . 49
4.4.2 Estimateur de la VaR corrigée . . . 50
4.5 Comparaison des différentes VaR . . . 50
4.5.1 Modèle sous-jacent . . . 50
4.5.2 Procédure . . . 50
4.5.3 Résultats . . . 51
4.6 Conclusion sur la correction de la VaR . . . 52

5 Variance asymptotique **53**
5.1 Éléments théoriques . . . 53
5.1.1 Décomposition de la statistique . . . 53
5.1.2 Couverture inconditionnelle . . . 54
5.1.3 Indépendance . . . 55
5.2 Correction de la variance . . . 55
5.2.1 Mise-en-oeuvre de la correction de la variance . . . 55
5.2.2 Résultats sur la variance en fonction des paramètres . . . 57
5.2.3 Résultats sur la variance en fonction de la taille de l'échantillon . . . 57
5.3 Elimination du caractère « à distance finie » . . . 58
5.4 Conclusion : risque de paramètre . . . 58

6 Conclusion III **59**

IV Conclusion **61**

V Références **65**

VI Table des figures **69**

VII Annexes **73**

A Éléments techniques **75**
A.1 Expérience 1 : Convergence de la méthode d'estimation . . . 76
A.1.1 Démarche . . . 76
A.1.2 Formalisation . . . 76
A.1.3 Mise en oeuvre . . . 76
A.1.4 Analyse . . . 76
A.1.5 Tableaux et graphiques . . . 77
A.2 Expérience 2 : Tests usuels . . . 81

A.2.1 Démarche . . . 81
A.2.2 Formalisation . . . 81
A.2.3 Méthode de calcul de la VaR par la banque . . . 81
A.2.4 Mise en oeuvre . . . 82
A.2.5 Backtesting simple . . . 82
A.3 Expérience 3 : Correction de la variance asymptotique du test de Kupiec . . . 84
A.3.1 Position du problème . . . 84
A.3.2 Estimation des paramètres . . . 84
A.3.3 Convergence . . . 84
A.3.4 Correction de la variance dans le test de Kupiec . . . 86
A.4 S&P 500 - GARCH sur les accroissements . . . 88
A.5 S&P 500 - GARCH sur les taux d'accroissement . . . 89
A.6 S&P 500 - GARCH sur les log-rendements . . . 90
A.7 Démonstration des relations entre les hypothèses des tests . . . 91
A.7.1 $H_0 \Rightarrow H_{0,uc}$ . . . 91
A.7.2 $H_0 \Rightarrow H_{0,ind}$ . . . 91
A.7.3 $H_0 \Rightarrow H_{0,\mathcal{B}}$ . . . 92
A.7.4 $H_0 \Rightarrow H_{0,D}$ . . . 92
A.8 Annexes techniques du chapitre . . . 92
A.8.1 Complément : intérêt des log-rendements . . . 92
A.8.2 Démonstration : le théorème 1.2, p 25 . . . 92
A.8.3 Démonstration : Viabilité de la VaR par un modèle GARCH . . . 93
A.9 Démonstration pour le chapitre Correction de la VaR . . . 93
A.9.1 Démonstration : le théorème 4.1 p48 . . . 93
A.9.2 Démonstration : le théorème 4.3 p49 . . . 94
A.9.3 Démonstration : le théorème 4.4 p49 . . . 95
A.10 Estimation par maximisation de la vraisemblance . . . 95
A.11 Propositions . . . 96

B Code Matlab **97**
B.1 Code - ToolBox de Kevin Sheppard . . . 97
B.1.1 Fichier garchcore.m . . . 97
B.1.2 Fichier garchpq.m . . . 97
B.1.3 Fichier garchlikelihood.m . . . 100
B.1.4 Fichier garchgrad.m . . . 101
B.2 Code - Expérience n°1 . . . 102
B.2.1 Arborescence . . . 102
B.2.2 Fichier exp1.m . . . 102
B.3 Code - Expérience n°2 . . . 104
B.3.1 Arborescence . . . 104
B.3.2 Fichier exp2MC.m . . . 104
B.3.3 Fichier exp2_fonction.m . . . 108
B.3.4 fichier test_portmanteau.m . . . 111
B.3.5 Fichier test_christoffersen.m . . . 112
B.3.6 Fichier weibull.m . . . 114
B.4 Code - Expérience n°3 . . . 115
B.4.1 Arborescence . . . 115
B.4.2 Fichier exp3bis_fonction.m . . . 115
B.4.3 Fichier MCexp3.m . . . 121