
Mesure du risque de marché d'un portefeuille de type Actions (Value-At-Risk, Value-At-Risk Conditionnelle)

Mémoire de fin d'études présenté en vue de l'obtention du titre d'Actuaire
par

Raoul FOKOU

Promotion 2006

Remerciements

J'adresse tous mes remerciements à l'EURO-Institut d'Actuariat JEAN DIEU-DONNÉ et à Monsieur Hervé LE BORGNE pour nous avoir offert l'opportunité d'obtenir le titre d'Actuaire EURIA, d'adhérer à l'Institut des Actuaire et d'achever un cycle d'études commencé depuis quelques années.

Résumé

Cette étude s'intéresse au risque de perte d'un portefeuille d'actifs financiers en calculant sa Value-at-Risk (VaR) et sa Value-at-Risk Conditionnelle (CVaR). Ces calculs nécessitent d'estimer la distribution des rendements d'un portefeuille d'actifs. La difficulté provient alors de la manière de prendre en compte les dépendances entre les rendements des actifs et de modéliser les queues de distributions des rendements.

Nous proposons ici de résoudre ces deux obstacles en utilisant les deux techniques suivantes :

- l'analyse en composantes indépendantes, technique issue du traitement du signal, qui permet de transformer, sous certaines hypothèses, un vecteur multivarié en un autre vecteur multivarié dont les composantes sont statistiquement indépendantes,
- la modélisation par mélanges de lois gaussiennes afin de modéliser les queues de distributions empiriques plus épaisses que ne l'autorise la loi normale.

Nous prenons l'exemple d'un portefeuille d'actifs composé des trois indices boursiers Bel20, Cac40 et Dax30. À partir de ces deux techniques, il est possible d'établir des formules analytiques pour le calcul de la VaR et de la CvaR, puis d'obtenir des résultats numériques de VaR et CvaR à un, cinq et dix jours. Enfin, nous procédons à une comparaison sommaire de différentes méthodes d'estimation de VaR et CvaR en réalisant des simulations de rendements à partir des données historiques.

Les premiers résultats obtenus montrent qu'en combinant ces deux techniques, les valeurs estimées pour la VaR et la CvaR ont des taux d'échecs inférieurs aux valeurs calculées selon les approches historique et analytique. Il serait alors intéressant de pouvoir comparer cette méthodologie avec la méthode des copules qui s'attache à modéliser les dépendances entre les rendements.

Mots-clés : value-at-risk, value-at-risk conditionnelle, analyse en composantes indépendantes, mélanges gaussiens

Abstract

This survey aims at computing Value-at-Risk (VaR) and Conditional value-at-Risk (CvaR) of a portfolio of stocks. This computation requires asset returns model but difficulty comes with modeling dependencies and tail distribution of returns.

We propose to apply two methods :

- the independent component analysis used in signal processing which allows us to deal with dependencies by building asset returns time series which are statistically independent,
- the finite gaussian mixture to model asset returns densities functions in order to capture tail distributions observed on our data.

Applying this two techniques, we can obtain analytical expressions and numerical values for the VaR and the CvaR over one, five and ten days.

Using historical data, we perform some back testing in order to roughly assess the quality of our estimations. Then results are compared with estimations using classical methods such historical or variance-covariance techniques. First simulations show that our estimates have a probability of failure lower than previous two methods. Then we could enhance this survey by comparing estimates obtained using other dependency modeling such copula theory.

Keywords : value-at-risk, conditional value-at-risk, independent components analysis, gaussian mixtures

Table des matières

Résumé	ii
Abstract	iii
Liste des tableaux	vii
Liste des figures	ix
Introduction	1
1 Paradigme Rendement-Risque	3
1.1 Le rendement	3
1.1.1 Pour une action	3
1.1.2 Pour un portefeuille d'actions	5
1.2 Le risque de marché	5
1.2.1 La volatilité historique	6
1.2.2 La semi-variance	7
1.2.3 Les mesures du risque de perte	9
2 Value-at-Risk VaR et Value-at-risk Conditionnelle CvaR	10
2.1 Introduction	10
2.2 Définition	11
2.3 Les méthodes de calcul	12
2.3.1 Analyse historique (ou simulation historique)	12
2.3.2 La simulation de Monte-Carlo	12
2.3.3 La méthode des variances-covariances	13
2.4 Expression en terme de rendement	14
2.5 Value-at-risk conditionnelle	15

3	Données financières	17
3.1	Les indices actions BEL20, CAC40 et DAX30	17
3.1.1	L'indice CAC40	18
3.1.2	L'indice DAX30	18
3.1.3	L'indice BEL20	18
3.1.4	Évolution des indices entre le 04/01/99 et le 15/05/06	19
3.2	Tests statistiques	20
3.2.1	Introduction	20
3.2.2	Tests de stationnarité Philipps-Perron et KPSS	20
3.2.3	Tests d'indépendance et de distribution identique	21
3.2.4	Tests de normalité	25
3.3	Corrélations	28
4	L'Analyse en Composantes Indépendantes	31
4.1	Description	31
4.2	Lien avec l'Analyse en Composantes Principales	34
4.3	Mesure de l'indépendance statistique	35
4.3.1	Définition	35
4.3.2	Les Cumulants	35
4.4	Applications	36
4.4.1	Série des rendements à 5 jours	37
4.4.2	Série des rendements à 1 et 10 jours	39
4.5	Conclusion	40
5	Modélisation par mélange fini de lois gaussiennes	41
5.1	Modélisation	41
5.1.1	Formalisation mathématique	42
5.1.2	Modélisation mathématique	42
5.1.3	Algorithme EM	43
5.2	Applications	45
5.2.1	Composantes sources à 5 jours ${}_5S$	46
5.2.2	Composantes sources à 1 et 10 jours ${}_1S$ et ${}_{10}S$	47
5.3	Conclusion	48
6	VaR, CVaR et Test historique	52
6.1	Densité de rendement du portefeuille	52
6.1.1	Exploitation de l'ACI	52
6.1.2	Exploitation des mélanges gaussiens	53

<i>Table des matières</i>	vi
6.2 Applications numériques	56
6.2.1 Méthode historique	56
6.2.2 Méthode variance-covariance	57
6.2.3 Méthode ACI et mélange gaussien	58
6.3 Tests historiques	59
Conclusion	60
Bibliographie	63
Annexe 1. Histogrammes des rendements à 1 et 10 jours	65
Annexe 2. Séries des rendements à 1 et 10 jours	66
Annexe 3. Tests de stationnarité à 1 et 10 jours	67
Annexe 4. Signaux sources à 1 et 10 jours	72
Annexe 5. Histogrammes des sources à 1 et 10 jours	73

Liste des tableaux

1.1	Exemple d'évolution des prix	4
1.2	Exemple de calcul de rendement	4
3.1	Seuils critiques - Statistiques PP et KPSS	21
3.2	Tests de stationnarité sur les rendements des indices boursiers	23
3.3	Tests de la médiane et des "Up and Down" observés sur les rendements des indices boursiers	25
3.4	Statistiques observées sur les rendements des indices boursiers	26
3.5	Test Normalité Rendement Indice BEL20	27
3.6	Test Normalité Rendement Indice CAC40	28
3.7	Test Normalité Rendement Indice DAX30	28
3.8	Matrice des corrélations - Rendements à 1 jour	29
3.9	Matrice des corrélations - Rendements à 5 jours	29
3.10	Matrice des corrélations - Rendements à 10 jours	29
4.1	Matrice des corrélations - Rendements Sources à 5 jours	38
4.2	Matrice des corrélations - Rendements Sources à 1 jour	39
4.3	Matrice des corrélations - Rendements Sources à 10 jours	39
5.1	Paramètres du mélange gaussien de la composante IC1_5J	46
5.2	Paramètres du mélange gaussien de la composante IC2_5J	47
5.3	Paramètres du mélange gaussien de la composante IC3_5J	47
5.4	Paramètres du mélange gaussien de la composante IC1_1J	48
5.5	Paramètres du mélange gaussien de la composante IC2_1J	48
5.6	Paramètres du mélange gaussien de la composante IC3_1J	48
5.7	Paramètres du mélange gaussien de la composante IC1_10J	51
5.8	Paramètres du mélange gaussien de la composante IC2_10J	51
5.9	Paramètres du mélange gaussien de la composante IC3_10J	51
6.1	Rappel des paramètres du mélange gaussien de la composante IC1_5J	55

6.2	Rappel des paramètres du mélange gaussien de la composante IC2.5J	55
6.3	Rappel des paramètres du mélange gaussien de la composante IC3.5J	55
6.4	VaR - méthode historique	57
6.5	CvaR - méthode historique	57
6.6	VaR - méthode multinormale	58
6.7	CvaR - méthode multinormale	58
6.8	VaR - méthode ACI-Mélange gaussien	58
6.9	CvaR - méthode ACI-Mélange gaussien	58
6.10	Taux d'échec - Rendements simulés à 1, 5 et 10 jours	59

Liste des figures

2.1	Représentation graphique de la VaR	11
3.1	Évolution des indices boursiers du 04 janvier 1999 au 15 mai 2006	19
3.2	Séries des rendements à 5 jours des indices Bel20, Cac40 et Dax30	22
3.3	Test PP & KPSS des rendements à 5 jours de l'indice BEL20	23
3.4	Histogramme des rendements à 5 jours de l'indice BEL20	27
3.5	Histogramme des rendements à 5 jours de l'indice CAC40	27
3.6	Histogramme des rendements à 5 jours de l'indice DAX30	28
4.1	Problème de la séparation de sources vu par <i>Gotlib</i> dans les <i>Dingodossiers</i> ([1])	32
4.2	Schéma de l'ACI. Des signaux sources s sont mélangés aux travers d'une matrice A pour former des signaux observés x . La matrice de séparation W transforme les signaux observés x en composantes indépendantes formant le signal y	32
4.3	Principe de l'ACI dans le cas du cocktail party problem	33
4.4	Séries des signaux sources à 5 jours	38
5.1	Histogrammes des signaux sources à 5 jours	45
5.2	Modélisation des densités du rendement source IC1_5J	46
5.3	Modélisation des densités du rendement source IC2_5J	47
5.4	Modélisation des densités du rendement source IC3_5J	48
5.5	Modélisation des densités des rendements sources à 1 jour	49
5.6	Modélisation des densités des rendements sources à 10 jours	50
6.1	Rendements historiques observés à 1,5 et 10 jours	56
2	Histogramme des rendements à 1 et 10 jours	65
3	Séries des rendements à 1 et 10 jours	66

4	Test PP des rendements à 1 jour de l'indice BEL20	67
5	Test KPSS des rendements à 1 jour de l'indice BEL20	67
6	Test PP des rendements à 10 jours de l'indice BEL20	67
7	Test KPSS des rendements à 10 jours de l'indice BEL20	68
8	Test PP des rendements à 1 jour de l'indice CAC40	68
9	Test KPSS des rendements à 1 jour de l'indice CAC40	68
10	Test PP des rendements à 5 jours de l'indice CAC40	68
11	Test KPSS des rendements à 5 jours de l'indice CAC40	69
12	Test PP des rendements à 10 jours de l'indice CAC40	69
13	Test KPSS des rendements à 10 jours de l'indice CAC40	69
14	Test PP des rendements à 1 jour de l'indice DAX30	70
15	Test KPSS des rendements à 1 jour de l'indice DAX30	70
16	Test PP des rendements à 5 jours de l'indice DAX30	70
17	Test KPSS des rendements à 5 jours de l'indice DAX30	70
18	Test PP des rendements à 10 jours de l'indice DAX30	71
19	Test KPSS des rendements à 10 jours de l'indice DAX30	71
20	Séries des signaux sources à 1 et 10 jours	72
21	Histogrammes des signaux sources à 1 et 10 jours	73

Introduction

L'une des problématiques majeures en gestion financière est la détermination du risque financier d'un portefeuille d'actifs. Depuis quelques années, la Value at Risk (VaR) s'est imposée comme l'une des mesures du risque de marché. Elle vise à résumer de manière concise par un montant (2 500 000 euros par exemple) la perte probable d'un portefeuille d'actifs sur un horizon de temps déterminé, par exemple 1 jour, 10 jours ou 1 mois, étant donné un indice de confiance choisi (par exemple 95 % ou 99 %). Si la période de base est un jour de bourse, la VaR à 95 % est le niveau de perte qui sera dépassé cinq jours sur 100.

Il existe plusieurs méthodes de calcul de la Value at Risk classique et conditionnelle (simulation historique, simulation de Monte-Carlo, méthode matrice variance-covariance ...) qui présentent toutes des avantages et inconvénients. Notre approche tente d'appliquer certaines techniques issues de la science du traitement du signal aux séries financières utilisées pour le calcul de VaR. Pour calculer la VaR, il faut modéliser la distribution des pertes et profits d'un portefeuille d'actifs. Toutefois, comme toute somme de variables aléatoires, la prise en compte des corrélations et dépendances entre actifs rend l'estimation de cette distribution relativement délicate.

Nous prenons l'exemple d'un portefeuille d'actifs de type actions. Pour modéliser la distribution des rendements de ce portefeuille, nous proposons l'utilisation de l'Analyse en Composantes Indépendantes (ACI), technique provenant de la science dite traitement du signal. En considérant les séries de rendements financiers comme des signaux physiques, cette technique vise, par transformation linéaire, à trouver un autre ensemble de signaux indépendants au sens statistique. À partir de ces signaux sources indépendants, nous modélisons leurs distributions par un mélange de lois normales afin de prendre en compte les queues de distributions observées qui sont plus épaisses que ne l'autorise l'hypothèse de normalité. Nous choisissons classiquement des mélanges gaussiens car les expressions analytiques sont plus faciles à manipuler. L'utilisation du produit de convolution, dans le cas de variables aléatoires ayant des fonctions de densités indépendantes, nous permet ensuite d'obtenir la fonction de densité du rendement de notre portefeuille composé de plusieurs actifs. Ceci nous permet

enfin d'effectuer un calcul de Value at Risk et de Value at Risk Conditionnelle.

Ce mémoire est divisé en six chapitres regroupés en trois parties : les chapitres un à trois sont des chapitres généraux. Ils rappellent les concepts de risque, de rendement, de Value at Risk (classique et conditionnelle) et présentent les données financières utilisées. Les chapitres quatre et cinq présentent et appliquent les techniques d'analyse en composantes indépendantes et de mélanges gaussiens à nos données financières. Enfin le chapitre six présente les premiers résultats numériques obtenus.

Chapitre 1

Paradigme Rendement-Risque

Ce premier chapitre est un rappel du cadre théorique qui oriente la recherche et la réflexion en finance, le couple rendement et risque.

1.1 Le rendement

La gestion des portefeuilles d'actifs s'appuie sur la notion de rendement ou de taux de rentabilité. Cette grandeur mesure, pendant un intervalle de temps donné, l'appréciation ou la dépréciation relative de la valeur d'un actif financier ou d'un portefeuille d'actifs. Nous limiterons notre étude à des actifs financiers de type action.

1.1.1 Pour une action

Pour une période donnée, on définit usuellement deux types de rendement¹ d'un actif financier A_i de type action :

1. le rendement arithmétique comme la variation relative de son prix $P_i(t)$ entre les instants t et $t - 1$ par :

$$R_i^{arith}(t) = \frac{P_i(t) - P_i(t-1)}{P_i(t-1)} = \frac{\Delta P_i(t)}{P_i(t-1)} \quad (1.1)$$

2. le rendement logarithmique, également appelé rendement géométrique, entre les instants t et $t - 1$ par :

$$R_i^{ln}(t) = \ln \frac{P_i(t)}{P_i(t-1)} \quad (1.2)$$

L'inconvénient du rendement arithmétique est qu'il n'est pas additif. En effet, son calcul privilégie les deux instants t et $t - 1$, et néglige ainsi l'évolution de la valeur

¹nous ne prenons pas en compte les éventuels dividendes ou autres flux financiers

t	1	2	3
$P(t)$	100	102	100

TABLEAU 1.1 – Exemple d'évolution des prix

$\Delta(t)$	$R^{arith}(t)$	$r^{\ln}(t)$
$(2 - 1)$	2,00 %	1,98 %
$(3 - 2)$	-1,96 %	-1,98 %
$(2 - 1) + (3 - 2)$	0,04 %	0,00 %
$(3 - 1)$	0,00%	0,00 %

TABLEAU 1.2 – Exemple de calcul de rendement

pendant la période qui les sépare l'un de l'autre (voir [3]).

En effet, supposons qu'un actif financier prenne successivement les valeurs 100, 102 et 100 aux instants 1, 2 et 3 (tableau 1.1). Nous calculons les rendements arithmétiques et logarithmiques entre les instants 1, 2 et 3, puis 1 et 3 (tableau 1.2).

La rentabilité sur la période totale est nulle comme l'indique le rendement logarithmique. On montre facilement l'additivité des rendements logarithmiques sur une période $[0, T]$ par :

$$\begin{aligned}
 R_{[0,T]}^{\ln} &= \ln \left(\frac{P(T)}{P(0)} \right) = \ln \left(\frac{P(T)}{P(0)} \frac{P(T-1)}{P(T-1)} \frac{P(T-2)}{P(T-2)} \cdots \frac{P(1)}{P(1)} \right) \\
 &= \ln \left(\frac{P(T)}{P(T-1)} \right) + \dots + \ln \left(\frac{P(2)}{P(1)} \right) + \ln \left(\frac{P(1)}{P(0)} \right) \\
 &= R_{[T-1,T]}^{\ln} + \dots + R_{[1,2]}^{\ln} + R_{[0,1]}^{\ln}
 \end{aligned}$$

En supposant un intervalle de temps suffisamment petit, on peut définir un rendement instantané comme étant la valeur moyenne des valeurs prises par l'actif entre les instants t et $t + \delta(t)$, soit :

$$R_t^{\ln} = \int_t^{t+\delta_t} \frac{1}{P} dP = \ln \left(\frac{P(t + \delta_t)}{P(t)} \right) = \ln \left(\frac{P(t) + \delta P(t)}{P(t)} \right)$$

d'où, à partir du développement limité à l'ordre 1 de la fonction \ln :

$$R_t^{\ln} = \ln \left(1 + \frac{\delta P(t)}{P(t)} \right) \approx \frac{\delta P(t)}{P(t)} \tag{1.3}$$

L'équation (1.3) montre ainsi que la rentabilité arithmétique et la rentabilité logarithmique sont égales en première approximation lorsque le 'pas de temps' est faible. Dans cette étude, nous utilisons des rendements logarithmiques à 1 jour, 5 jours et 10 jours. Les rendements à 5 et 10 jours sont obtenus en effectuant la somme des rendements logarithmiques quotidiens sur la période considérée. On considérera que la sous-période de 1 jour est suffisamment faible pour appliquer l'approximation de l'équation (1.3).

1.1.2 Pour un portefeuille d'actions

Nous considérons un investisseur financier disposant d'un portefeuille P composé de trois actifs de type action $A_i, i = 1, 2, 3$. La rentabilité du portefeuille s'exprime alors comme une combinaison linéaire des rendements R_i de chaque actif pondéré par le poids de chaque actif dans le portefeuille P .

En notant ω_i ce poids, le rendement du portefeuille s'écrit :

$$R_P = \sum_{i=1}^3 \omega_i R_i \quad (1.4)$$

1.2 Le risque de marché

Le risque de marché d'un portefeuille provient des variations des prix des actifs financiers et de leur effet négatif sur la valeur financière totale du portefeuille. Traduit en terme de rendement, les rentabilités des actifs deviennent alors des variables aléatoires dont l'évolution future n'est pas, à priori, connue et quantifiable. Dans le cas d'une obligation², les fluctuations des taux d'intérêts modifient leur valeur de marché. Ainsi le rendement R_i de l'actif i est considéré comme une variable aléatoire pouvant prendre, à l'instant t , des valeurs aléatoires $r_i(t)$ dans un intervalle plus ou moins déterminé.

Une action³, comme tout autre produit financier, comporte des risques dont les plus connus sont les risques liés à l'entreprise (liquidation judiciaire, baisse du chiffre d'affaire, risques juridiques . . .). On distingue habituellement deux grandes catégories de risques, le risque systémique et le risque spécifique.

²titre négociable représentant tout ou partie d'un prêt octroyée à une société lors de l'émission d'un emprunt

³titre de propriété représentant une part d'une société

Risque systémique

Le risque systémique, également appelé systématique, correspond au risque qu'a chaque investisseur quand il décide d'acheter une action et ce quelque soit la société à laquelle se rattache cette action. Il s'agit en fait du risque de marché dans son ensemble que nous avons l'habitude de 'mesurer' par des indices boursiers (Cac40, Dow-jones, Nikkei . . .) censés reflétés l'activité d'un marché. Ce risque systémique est considéré comme un risque irréductible, intrinsèque au marché des actions.

Risque spécifique

Contrairement au risque systémique, le risque spécifique est lié à une action en particulier. Supposez que vous achetiez une valeur dite "valeur Internet", outre le risque systématique lié à ce type d'actions, la société peut très bien ne pas atteindre les objectifs de croissance qu'elle s'était fixée. Ceci aura un impact, le plus souvent négatif, sur le cours de bourse de cette société.

Ce risque spécifique peut être réduit au minimum en détenant un portefeuille identique à la composition d'un indice de marché comme le Cac40. Ainsi seul subsistera le risque systématique. C'est typiquement le cas de la gestion dite par benchmark⁴ qui gère des portefeuilles répliquant des indices boursiers.

La gestion du risque est une composante très importante de la gestion de portefeuille. Elle intervient de plus en plus dans les processus de sélection de portefeuille des gérants de portefeuille et les différentes méthodes de prévision du risque vont de l'utilisation des simples volatilités et corrélations historiques entre actions aux modèles à plusieurs facteurs.

1.2.1 La volatilité historique

La volatilité a été introduite par MARKOWITZ ([33]) dans son modèle fondateur d'allocation d'actifs dit modèle "Moyenne-Variance". Dans ce cadre, le risque d'un investissement provient du fait que les espérances de rendement ne sont pas toujours réalisées. La dispersion des rentabilités autour de leur moyenne estimée traduit alors l'incertitude du placement. L'écart-type σ , appelé également volatilité⁵, constitue une mesure de cette incertitude.

⁴indice de référence

⁵lorsque les données financières sont annualisées

Si le rendement R d'un actif A possède une fonction de densité f_r , nous avons :

$$\sigma_i^2 = \int_{-\infty}^{+\infty} (r - \bar{r})^2 f_r(r) dr \quad (1.5)$$

Dans le cas discret, nous pouvons également utiliser l'expression :

$$\sigma_i^2 = \frac{1}{T} \sum_{t=1}^T (r - \bar{r}(t))^2 \quad (1.6)$$

MARKOWITZ a utilisé la variance en raison de ses propriétés statistiques qui permettent d'explicitier la variance σ_P d'un portefeuille à partir des variances individuelles σ_i des N actifs le composant :

$$\sigma_P^2 = \sum_{i=1}^N \omega_i^2 \sigma_i^2 + 2 \sum_{j=1, j \neq i}^N \omega_i \omega_j \sigma_i \sigma_j \rho_{i,j} \quad (1.7)$$

où $\rho_{i,j}$ ⁶ est le coefficient de corrélation entre les rendements des actifs A_i et A_j .

Ce modèle a fait l'objet de deux principales critiques, du fait de deux hypothèses fortes :

- les rendements des actifs boursiers sont des variables aléatoires indépendantes et identiquement distribuées suivant une loi normale (hypothèse utilisée pour exprimer la variance d'un portefeuille) ;
- la prise en compte de manière identique des variations 'positives' autant que 'négatives' des rendements autour de leur moyenne.

1.2.2 La semi-variance

MARKOWITZ avait déjà suggéré, en 1959, l'usage de la semi-variance, qui de son point de vue, est une mesure du risque plus proche de la réalité financière :

“Since an investor worries about underperformance rather than overperformance, semi-deviation is a more appropriate measure of investor's risk than variance”, MARKOWITZ 1992.

⁶ $\rho_{i,j} = \frac{1}{T} \sum_{t=1}^T \frac{1}{\sigma_i \sigma_j} (r_i - \bar{r}_i) (r_j - \bar{r}_j)$

La semi-variance, ou risque de perte quadratique, s'écrit dans le cas discret :

$$LPM(2, \bar{r}_i) = \frac{1}{T} \sum_{t=1}^T (\min(r_i(t) - \bar{r}_i; 0))^2 \quad (1.8)$$

Elle traduit l'idée que l'éventualité d'obtenir un rendement inférieur à celui espéré doit être considérée comme risquée. Par contre, la possibilité d'obtenir un rendement supérieur au rendement attendu est un événement favorable et non un risque.

La semi-variance est un cas particulier d'une classe de mesure asymétrique du risque appelée les moments partiels inférieurs⁷, concept introduit par Bawa et Fishburn([9], [23]).

Le moment partiel inférieur d'ordre α et de rendement cible r_{cible} d'un actif A , se définit dans le cas continu par :

$$LPM(\alpha, r_{cible}) = \int_{-\infty}^{r_{cible}} (r - r_{cible})^\alpha f_r(r) dr \quad (1.9)$$

et dans le cas discret par

$$LPM(\alpha, r_{cible}) = \frac{1}{T} \sum_{t=1}^T (\min(r(t) - r_{cible}; 0))^\alpha \quad (1.10)$$

À l'opposé de la variance, aucune hypothèse n'est faite sur la distribution statistique des rendements des actifs. Ainsi la semi-variance représente en quelque sorte l'écart-type des rendements inférieurs à un rendement cible. Elle ne pénalise pas les rendements supérieurs à un rendement cible et différencie le risque, obtenir des rendements inférieurs à un rendement minimal, de l'incertitude, variabilité des rendements.

Le paramètre α détermine en quelque sorte l'aversion au risque de l'investisseur, plus celui-ci est élevé, plus l'investisseur est averse au risque. Cette mesure du risque semble également appropriée pour la gestion des actifs des compagnies d'assurance car :

- certains produits d'épargne sont assortis d'une garantie de taux de rendement,
- ou dans le cas de produit de retraite, il s'agit d'obtenir un taux de rendement supérieur au taux d'actualisation des engagements des régimes de retraite, et
- de façon générale, pour tout investissement dont le risque se détermine en fonction d'un rendement minimal cible.

⁷Lower Partial Moments - LPM

1.2.3 Les mesures du risque de perte

Nous venons de présenter deux types de risques :

- le risque défini par la variabilité ou variance des revenus futurs, et
- le risque défini par la probabilité de ne pas atteindre un objectif minimal de rendement.

On peut définir le risque financier d'une troisième façon, le risque de perte. On tente alors de quantifier la perte maximale à laquelle on peut s'attendre. Dans ce cadre, la *Value at Risk*, que l'on peut traduire par valeur à risque ou perte potentielle probable, est l'un des outils le plus utilisé pour mesurer et contrôler ce risque. Le chapitre suivant est consacré à sa présentation ainsi qu'au concept de *Conditional Value at Risk* (valeur à risque conditionnelle) qui lui a succédé.

Chapitre 2

Value-at-Risk VaR et Value-at-risk Conditionnelle CvaR

Historiquement, la mesure du risque la plus utilisée a toujours été la variance ou une de ses dérivées directes (écart-type, semi-variance...). Puis, les progrès et les recherches effectuées dans le domaine ont permis d'utiliser des indicateurs plus aboutis comme la Value-at-Risk (VaR) ou la Value-at-Risk conditionnelle (CvaR). Nous présentons le concept de VaR qui correspond à la notion statistique de quantile.

2.1 Introduction

Le concept de Valeur à Risque (Value at Risk, VaR) est devenu la norme en matière de mesure des risques de marché au sein des établissements financiers. En effet, la réglementation internationale impose désormais aux établissements bancaires de mesurer les risques de marché, et d'avoir un niveau de fonds propres satisfaisant. En offrant un indicateur simple et global de l'exposition aux risques de marché, elle permet un suivi et un contrôle centralisés des risques conformes aux recommandations des organismes régulateurs.

Elle est ainsi devenue la mesure du risque privilégiée dans le cadre des accords de Bâle visant à établir un cadre réglementaire pour la gestion des risques de marchés. De plus, en tant qu'outil de reporting, elle fournit des informations sur les concentrations de risques par type de marché, par trader, par produit financier, etc. Elle permet également d'allouer un capital disponible et de fixer des limites de négociation.

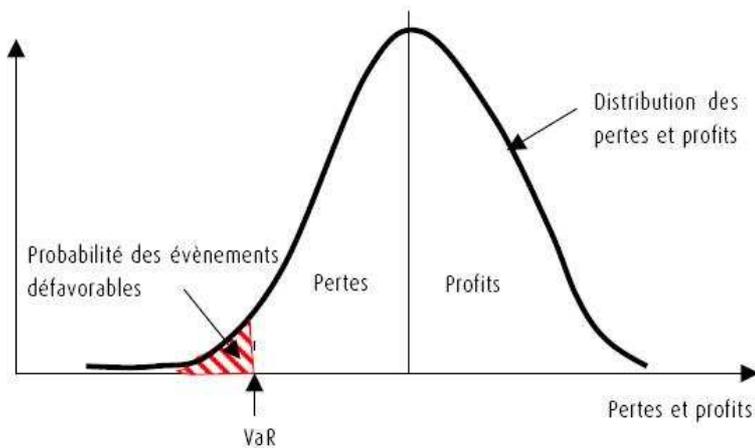


FIGURE 2.1 – Représentation graphique de la VaR

2.2 Définition

La VaR est une mesure des risques correspondant, pour un horizon de gestion donné, à un montant de perte probable d'un portefeuille ou d'un ensemble de portefeuilles d'instruments financiers. La VaR exprime la perte liée à des variations défavorables attendues sur les différents marchés financiers. Sa définition admise par l'ensemble des investisseurs financiers est la suivante :

“La VaR est la perte potentielle maximale qu'un portefeuille peut subir, pour un horizon de temps donné et un niveau de probabilité donné, en supposant que ce portefeuille reste inchangé durant l'horizon spécifié.”

Par exemple, dire que la VaR d'un portefeuille est de 25 millions d'euros sur 1 jour avec une probabilité de 5 % signifie qu'on estime qu'il y a seulement 5 % de chance pour que la valeur de marché du portefeuille diminue de plus de 25 millions d'euros entre aujourd'hui et demain. La traduction française “Valeur à Risque” est souvent trompeuse. En réalité, il s'agit d'un niveau de perte atteint avec un certain niveau de confiance de 95 % par exemple. Il ne s'agit donc pas d'une perte maximale, mais d'une perte maximale dans 95 % des cas.

Ce concept a initialement été mis en place par les grandes banques d'investissement américaines (JP Morgan, Chase Manhattan ...) et son utilisation s'est particulièrement développée au cours des années 90.

2.3 Les méthodes de calcul

2.3.1 Analyse historique (ou simulation historique)

Cette méthode est la plus intuitive et la plus simple à utiliser. On la nomme également méthode non paramétrique. Elle suppose une stationnarité des rendements, ce qui signifie que la distribution des rendements observée à partir des données historiques se reproduira à l'avenir.

Cette approche utilise des données de marché plus ou moins récentes pour déterminer la distribution des variations des facteurs de risque qui peuvent simplement être les prix des actifs dans le cas des actions. Ces variations historiques sont ensuite appliquées au portefeuille actuel dont on calcule la valeur pour chaque scénario historique. La distribution des valeurs possibles du portefeuille est ainsi estimée puis on extrait le quantile correspondant à un niveau de probabilité choisi.

L'élément clé de cette approche reste donc la sélection de l'historique utilisé pour les calculs. Les résultats seront différents selon la fenêtre historique retenue et si l'on souhaite par exemple donner plus d'importance aux données de marché récentes ou à des périodes "agitées" (scénarii extrêmes). La simulation historique reste la méthode la plus couramment utilisée car elle est la plus facile à mettre en place malgré la nécessité de maintenir des historiques nombreux et longs selon le nombre d'actifs pris en compte.

2.3.2 La simulation de Monte-Carlo

La méthode de simulation de Monte-Carlo reprend la démarche introduite par la simulation historique. En revanche, l'estimation de la distribution des facteurs de risque diffère. Au lieu de l'approcher par des scénarii passés, la simulation de Monte-Carlo utilise la modélisation de la distribution suivant les études et les hypothèses retenues par l'investisseur.

À partir de la distribution choisie, la génération de nombres aléatoires permet d'obtenir la valeur des facteurs de risque ; on en déduit alors la valeur du portefeuille à la fin de l'horizon de temps fixé. Cette simulation est répétée des milliers de fois pour approcher la distribution des valeurs possibles du portefeuille, de laquelle on extrait le quantile voulu, c'est-à-dire la VaR.

Le principal inconvénient de cette méthode est la complexité de sa mise en œuvre. Elle nécessite en effet un important effort de recherche et de développement pour la modélisation des facteurs de risque, la validation de ce modèle ainsi que sa mise en place informatique. Cependant, une fois ce temps de développement effectué, le temps

de calcul de la VaR est sensiblement équivalent au temps de calcul nécessaire pour la simulation historique.

La possibilité de paramétrer le modèle théorique en fonction des anticipations sur les évolutions futures des facteurs de risque est à la fois un avantage et un inconvénient. En effet, mal paramétrées, les hypothèses utilisées par la méthode de Monte-Carlo peuvent être contraires à la réalité du marché.

2.3.3 La méthode des variances-covariances

Cette méthode, mise en place par la banque JP Morgan, s'est répandue à travers le monde sous l'impulsion de sa filiale RiskmetricsTM. La méthode des variances-covariances, aussi connue sous le nom de méthode analytique, s'est vite imposée comme le standard de calcul de la VaR du fait des facilités de développement disponibles dans tous les progiciels financiers. Mais, petit à petit, les hypothèses de calcul retenues par cette méthode ont été vivement critiquées et sa notoriété a diminué.

En effet, la méthode analytique suppose que les variations des facteurs de risque suivent une loi normale et utilise les propriétés statistiques de ce type de loi pour estimer la VaR. Or, cette hypothèse ne se vérifie pas pour tous les instruments financiers et notamment les produits dérivés complexes, comme les options.

L'objectif de la méthode analytique est d'estimer la variance du portefeuille à partir de variances et de corrélations standardisées pour certaines échéances et certains produits. Cette étape supplémentaire de décomposition du portefeuille en facteurs de risques standards s'appelle le *mapping*. Par exemple, une position en obligation de maturité 2 ans sera décomposée en une position en zéro-coupon 1 an, $V1$, et une position en zéro-coupon 2 ans, $V2$. Les positions standards en produits de taux d'intérêts sont les positions en zéro-coupon.

Pour déterminer la VaR, il faut alors prendre en compte les corrélations entre les facteurs de risque standards, c'est la matrice des variances-covariances. Les hypothèses de cette méthode sont :

1. la stationnarité des rendements des actifs,
2. l'hypothèse de normalité des rendements, et
3. une décomposition linéaire des prix des actifs en facteurs de risques (*mapping*¹).

Appliquée à un portefeuille d'actifs, la matrice des variances-covariances peut devenir incalculable car de dimension trop élevée. Pour traiter le cas des actions, on évalue plutôt les actions par rapport à des indices de références. Ainsi, la matrice des

¹Dans le cas d'un portefeuille d'actions, cela correspond simplement à l'écriture de la valeur du portefeuille en la moyenne pondérée des prix de chaque actif

variances-covariances s'exprime par rapport aux variations des indices et non plus par rapport aux variations du prix des actions.

2.4 Expression en terme de rendement

La méthode proposée tout au long de ce travail est basée sur la modélisation des distributions statistiques des rendements. On exprimera alors la VaR de notre portefeuille comme étant le rendement négatif potentiel maximal (en valeur absolue)² qu'un portefeuille peut subir, pour un horizon de temps donné et un niveau de probabilité donné, en supposant que ce portefeuille reste inchangé durant l'horizon spécifié. Nous nous proposons de calculer une VaR à 1, 5 et 10 jours pour un portefeuille composé de trois actions.

Par définition, pour un seuil de confiance α donné (95 % par exemple), la Value at Risk correspond au montant de la perte potentielle sur une période de temps fixée qui ne sera dépassé que dans $(1 - \alpha)$ % des cas (5 %).

Considérons un portefeuille P_{ptf} dont la valeur de marché est P et R la rentabilité (logarithmique) de ce portefeuille sur un horizon h . À l'issue de cette période de temps h , la valeur du portefeuille P_{ptf} est $P(1 + R)$. Si le rendement obtenu R est négatif, nous constatons une perte égale à RP . Pour un portefeuille fixe sur la période h , le calcul de la perte maximale probable dépend alors essentiellement de la distribution de la variable aléatoire R . Il faut donc être capable d'obtenir le quantile qui correspond au niveau de confiance voulu pour la VaR.

En notant α ce niveau de confiance choisi, ce montant vérifie la probabilité (2.1) :

$$\mathbb{P}(R < R^{VaR}) = 1 - \alpha \quad (2.1)$$

Ainsi, la VaR peut être définie comme étant le rendement négatif maximal (en valeur absolue) sur un horizon de h jours avec une probabilité $p^{VaR} = 1 - \alpha$. En supposant que nous ayons modélisé la fonction de densité du rendement d'un portefeuille, la VaR se détermine alors en intégrant cette fonction de densité sur l'intervalle $[-\infty, r^{VaR}]$ selon l'équation (2.2) :

$$p^{VaR} = \int_{-\infty}^{r^{VaR}} f_r(r) dr = 1 - \alpha \quad (2.2)$$

Pour un portefeuille de valeur P à l'instant t , on obtient alors le montant de la VaR

²il s'agit en fait du rendement négatif minimal

par :

$$VaR = r^{VaR} * P \quad (2.3)$$

2.5 Value-at-risk conditionnelle

La principale critique faite à l'encontre de la VaR est qu'elle n'est pas une mesure de risque additive. ARTZNER ([5]) énonce un ensemble d'axiomes qu'une mesure de risque dite cohérente doit vérifier. Il montre ainsi que la VaR ne respecte pas l'axiome d'additivité. Prenons l'exemple d'un investisseur qui gère deux portefeuilles distincts. Sous certaines conditions, on montre que la VaR de la somme des deux portefeuilles est supérieur à la somme des VaR de chaque portefeuille. Cette caractéristique est contraire au principe de diversification en finance.

Une seconde critique de la VaR est due à sa définition même. C'est la mesure d'une perte maximale potentielle avec un certain niveau de confiance. Une VaR à 95 % ne donne aucune idée de l'ampleur de la perte si cette perte potentielle dépasse le quantile fixé à 5 %. On a donc défini une mesure complémentaire du risque de perte, la VaR conditionnelle CvaR³. Pour une variable aléatoire X , elle est définie par :

$$CvaR_\alpha = E[-X | X < -VaR_\alpha(X)] \quad (2.4)$$

En d'autres termes, la VaR conditionnelle peut se définir comme l'espérance de la perte lorsque cette perte dépasse la VaR. Puisque la VaR mesure la valeur qui sépare les $(1 - \alpha)$ % de la distribution, on cherche à se focaliser sur la queue de distribution de la perte, les α % restant, dont on ne connaît ni la distribution, ni l'espérance.

Contrairement à la VaR, la CvaR est une mesure cohérente du risque qui respecte les axiomes définis par Artzner ([5], [39]).

Traduit en terme de rendement par Chin ([16]), il s'agit de calculer l'espérance moyenne des rendements sachant que ces rendements sont inférieurs au rendement

³Conditional Value at Risk appelée également Expected Shortfall ou Conditional Tail Expectation

r^{VaR} , soit

$$r^{CvaR} = \mathbb{E} [r | r < r^{VaR}] \quad (2.5)$$

$$= \int_{-\infty}^{r^{VaR}} r f_{r^{VaR}}(r) dr \quad (2.6)$$

$$= \int_{-\infty}^{r^{VaR}} r f_r(r) \frac{1}{\int_{-\infty}^{r^{VaR}} f_r(r) dr} dr \quad (2.7)$$

$$= \frac{1}{\int_{-\infty}^{r^{VaR}} f_r(r) dr} \int_{-\infty}^{r^{VaR}} r f_r(r) dr \quad (2.8)$$

En posant

$$c = \frac{1}{\int_{-\infty}^{r^{VaR}} f_r(r) dr} = \frac{1}{p^{VaR}} \quad (2.9)$$

on obtient alors :

$$r^{CvaR} = \frac{1}{p^{VaR}} \int_{-\infty}^{r^{VaR}} r f_r(r) dr \quad (2.10)$$

Chapitre 3

Données financières

Ce chapitre est consacré à la présentation et à l'étude statistique des données financières que nous allons utiliser tout au long de cette étude. Dans un premier temps, nous présentons les trois indices de marché retenus. Ensuite, nous effectuons une analyse statistique de leurs séries de rendements calculés pour des périodes de détention de 1, 5 et 10 jours.

Les tests statistiques utilisés visent à déterminer si ces séries ont les propriétés suivantes :

- stationnarité,
- les observations forment un échantillon indépendant et identiquement distribués, hypothèse dite *iid*,
- de distributions gaussiennes.

En effet, nous rappelons que nous travaillons sur des séries temporelles. L'estimation des grandeurs statistiques à l'aide des grandeurs temporelles nécessitent la vérification d'une part de la stationnarité et d'autre part que l'échantillon est *iid*. Enfin, le dernier test de stationnarité nous conduit à proposer des modélisations non gaussiennes des rendements des actifs.

3.1 Les indices actions BEL20, CAC40 et DAX30

L'objectif d'un indice de marché est de représenter de manière synthétique et fidèle l'évolution d'un marché ou d'une place boursière. Un indice de marché est à la fois un instrument de mesure du comportement du marché et un instrument de comparaison très utile pour les gestionnaires d'actifs. Il sert ainsi de référence permettant la comparaison des performances d'un gérant par rapport au marché. Certains indices donnent également lieu à la création de produits dérivés. Dans le

cadre de ce travail, nous supposerons que les supports d'investissement se limitent à trois actions représentées par les indices boursiers Bel20, Cac40 et Dax30. Ainsi, notre portefeuille d'actifs P est composé de trois actifs de type action $A_i, i = 1, 2, 3$ dont les prix sont reflétés par les valeurs de ces indices.

3.1.1 L'indice CAC40

L'indice Cac40¹ est le principal indicateur boursier des actions françaises sur la place parisienne. Créé par la Compagnie des Agents de Change, il est déterminé à partir des cours de quarante actions cotées en continu² sur le Premier Marché parmi les 100 sociétés les plus capitalisées. Il a été lancé officiellement en 1988 avec pour base une valeur de 1 000 points le 31 décembre 1987. Il a atteint son maximum en l'an 2000 (plus de 6 900 points) avant de s'effondrer après l'éclatement de la bulle spéculative relative internet. Suite à cet effondrement, il a atteint son plus bas niveau, historique sur la décennie 1996 - 2006, en mars 2003 à 2 400 points. Courant janvier 2006, il a franchi la barre des 4 800 points avant d'amorcer un léger repli, puis de dépasser enfin la barre des 5 000 points, revenant ainsi à un son niveau historique du début de l'année 2000.

3.1.2 L'indice DAX30

L'indice allemand Dax30³, indicateur de la bourse de Francfort, inclut trente valeurs représentant environ 70 % de la capitalisation totale et 80 % des transactions sur les actions. Les critères déterminant l'entrée d'une action dans le DAX30 sont le poids boursier de la valeur, l'importance du volume des transactions et le chiffre d'affaire de la société. Il a été lancé le 1er janvier 1989 et a également pour base 1 000 points au 31 décembre 1987. À la différence de l'indice CAC40, il est ajusté pour tenir compte des dividendes versés par les sociétés qui le composent.

3.1.3 L'indice BEL20

L'indice Bel20 est l'indice des actions des vingt entreprises belges les plus importantes admises à la cote de la bourse des valeurs mobilières de Bruxelles. Les vingt sociétés sont sélectionnées sur la base de leur volume de transaction et de leur capitalisation boursière totale (capital flottant). Il porte également le nom d'indice Euronext

¹Compagnie des Agents de Change ou également Cotation Assistée par Ordinateur des quarante valeurs

²l'indice est mis à jour toutes les 30 secondes pendant la journée de 9h00 à 17h30

³Deutsche Aktienindex



FIGURE 3.1 – Évolution des indices boursiers du 04 janvier 1999 au 15 mai 2006

Bruxelles. Cet indice est de moins en moins utilisé depuis la fusion des bourses de Paris, Bruxelles et Amsterdam regroupées au sein d'Euronext. Cet indice a pour base 1 000 au 30 décembre 1990.

3.1.4 Évolution des indices entre le 04/01/99 et le 15/05/06

Le graphique (3.1) trace l'évolution des indices sur la période considérée. On peut constater la chute des indices Cac40 et Dax30 à partir du second trimestre 2000 suite à l'éclatement de la bulle Internet. Nos trois indices ont également atteint leur plus bas niveau, historique entre 1996 et 2006, à des dates proches de mars 2003. À l'opposé, on note la relative 'robustesse' de l'indice Bel20 qui ne perd que 2 200 points sur cette période de baisse continue que l'on peut expliquer par sa faible exposition aux secteurs de la communication et de l'internet avant l'éclatement de la bulle spéculative.

Afin d'établir nos séries temporelles de rendement des indices, la fenêtre d'observation choisie correspond à un historique d'environ 2,5 années de cotations quotidiennes et s'étend du 2 janvier 2004 au 12 mai 2006. Ceci correspond à un échantillon de 618 observations ⁴ quotidiennes pour chaque indice de marché⁵.

Par la suite, nous considérons un portefeuille fictif P composé de trois actifs financiers de même poids dans le portefeuille. Ces trois actifs de type

⁴le choix de cette période est entièrement subjectif et résulte d'un compromis entre statistique significative et stationnarité des observations

⁵ceci afin de nous permettre d'obtenir 608 rendements d'horizon 1 jour, 5 jours et 10 jours

action seront représentés par nos indices financiers dont les prix seront égaux à leur indice.

3.2 Tests statistiques

3.2.1 Introduction

À partir des indices précédents, sur la période allant du 2 janvier 2004 au 12 mai 2006, nous avons calculé les rendements logarithmiques à 1 jour, 5 jours et 10 jours en utilisant la formule (3.1).

$${}_hR_i(t) = \ln \left(\frac{P_i(t)}{P_i(t-h)} \right), \quad h^6 = 1,5,10; i^7 = 1,2,3 \quad (3.1)$$

Les tests statistiques ont été réalisés à l'aide du logiciel d'économétrie EViews. Nous prendrons en général un niveau de confiance de 95 %. Dans la suite de ce chapitre, nous présentons les test statistiques en utilisant la notation générique $(R_t)_{t=1..T}$.

3.2.2 Tests de stationnarité Philipps-Perron et KPSS

Pour pouvoir travailler avec des données temporelles, celles-ci doivent conserver une distribution constante dans le temps⁸. De façon générale, on dit qu'une série temporelle est stationnaire lorsque les valeurs prises par cette série sont issues d'un processus dont les propriétés en terme de moments (ou de loi) sont invariantes dans le temps. Ceci signifie que la loi du couple $(R_t, R_{t+\theta})$ dépend de θ mais pas de t . Un exemple de série stationnaire est la série R^9 définie par $R_t = \beta + \varepsilon_t$ qui possède une valeur moyenne β autour de laquelle elle fluctue. À l'opposé la série définie par $R_t = R_{t-1} + \beta + \varepsilon_t$, où ε_t est une suite de variables aléatoires centrées de même variance et indépendantes¹⁰, n'est pas stationnaire.

En économétrie, on montre que si une série de données possède une racine unitaire (i.e. est non-stationnaire), cela implique qu'un choc sur cette série a un effet persistant dans le temps. À l'opposé, si la série est stationnaire, les chocs ne peuvent qu'avoir un effet temporaire.

La présentation des tests sera succincte car leur explication complète nécessite l'in-

⁶1 jour, 5 jours et 10 jours

⁷1 pour l'indice Bel20, 2 pour l'indice Cac40 et 3 pour l'indice Dax30

⁸afin de pouvoir utiliser les valeurs empiriques comme estimateurs

⁹on l'appelle également marché aléatoire ou processus intégré d'ordre 1, noté $I(1)$

¹⁰on l'appelle également bruit blanc

α	PP	KPSS
1 %	-3,973	0,216
5 %	-3,417	0,146
10 %	-3,131	0,119

TABLEAU 3.1 – Seuils critiques - Statistiques PP et KPSS

roduction de concepts économétriques¹¹ sortant du cadre de cette étude.

Nous étudions les séries temporelles des rendements à 1 jour, 5 jours et 10 jours (figures 3.2 et 3-annexe 2).

Les tests PP et KPSS

Pour valider la stationnarité des données, nous avons utilisé deux tests classiques couramment utilisés en économétrie, le test PP¹² et le test KPSS¹³ :

- Le test PP considère l’hypothèse nulle H_0 “la série présente des racines unitaires” (c’est-à-dire est non-stationnaire) contre l’hypothèse alternative. C’est un test unilatéral à gauche qui rejettera l’hypothèse nulle pour de faibles valeurs de la statistique de test,
- Le test KPSS teste l’hypothèse nulle H'_0 “la série est stationnaire” contre l’hypothèse alternative. C’est un test unilatéral à droite qui rejettera l’hypothèse nulle pour de grandes valeurs de la statistique de test.

Pour chacun de ces tests, les valeurs critiques pour les différents seuils significatifs α sont indiquées dans le tableau (3.1). En utilisant ces valeurs critiques ainsi que les règles de rejet précédentes, nous obtenons les résultats indiqués dans le tableau (3.2). La figure (3.3) montre les calculs réalisés sous EViews pour les séries de rendements à 5 jours de l’indice Bel20. On trouvera en annexe 3 les tests de stationnarité des séries à 1 et 10 jours. Au seuil $\alpha = 5\%$, nous concluons donc que les séries de rendements à 1, 5 et 10 jours sont toutes stationnaires.

3.2.3 Tests d’indépendance et de distribution identique

Nous avons utilisé les tests classiques nommés test de la Médiane et test des “Up and Down”. Nous exposons ces deux tests pour une série de rendements $R(t)_{t=1\dots T}$. L’hypothèse nulle H_0 commune au deux tests est : les $(R_t)_{t=1\dots T}$ forment un échantillon de rendements indépendants et identiquement distribués. À partir de cette suite, on

¹¹analyse de séries temporelles avec volatilité stochastique de type ARCH

¹²PHILIPPS-PERRON, 1988

¹³KWIOTOWSKI, PHILLIPS, SCHMIDT, SHIN, 1992

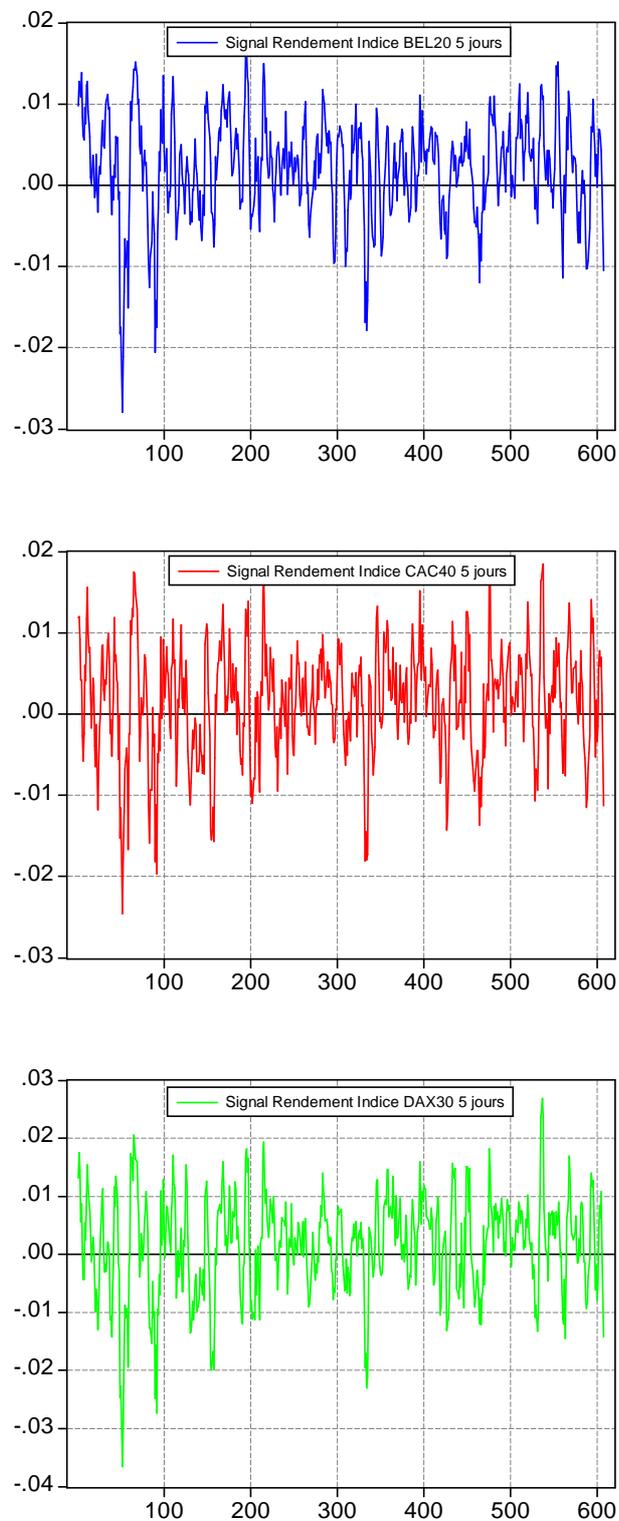


FIGURE 3.2 – Séries des rendements à 5 jours des indices Bel20, Cac40 et Dax30

Indice	h	Statistique PP	Statistique KPSS	Décision PP	Décision KPSS
BEL20	1j	-24,711	0,044	Rejet H_0	Accep. H'_0
BEL20	5j	-5,876	0,050	Rejet H_0	Accep. H'_0
BEL20	10j	-6,551	0,060	Rejet H_0	Accep. H'_0
CAC40	1j	-27,538	0,027	Rejet H_0	Accep. H'_0
CAC40	5j	-6,975	0,025	Rejet H_0	Accep. H'_0
CAC40	10j	-7,343	-0,031	Rejet H_0	Accep. H'_0
DAX30	1j	-25,094	0,0279	Rejet H_0	Accep. H'_0
DAX30	5j	-5,959	0,0229	Rejet H_0	Accep. H'_0
DAX30	10j	-6,836	0,0277	Rejet H_0	Accep. H'_0

TABLEAU 3.2 – Tests de stationnarité sur les rendements des indices boursiers

Null Hypothesis: RDT_BEL20_05J has a unit root
 Exogenous: Constant, Linear Trend
 Bandwidth: 37 (Newey-West using Bartlett kernel)

	Adj. t-Stat	Prob.*
Phillips-Perron test statistic	-5.875941	0.0000
Test critical values:		
1% level	-3.973245	
5% level	-3.417239	
10% level	-3.131010	

*MacKinnon (1996) one-sided p-values.

Null Hypothesis: RDT_BEL20_05J is stationary
 Exogenous: Constant, Linear Trend
 Bandwidth: 13 (Newey-West using Bartlett kernel)

	LM-Stat.
Kwiatkowski-Phillips-Schmidt-Shin test statistic	0.050375
Asymptotic critical values*:	
1% level	0.216000
5% level	0.146000
10% level	0.119000

*Kwiatkowski-Phillips-Schmidt-Shin (1992, Table 1)

FIGURE 3.3 – Test PP & KPSS des rendements à 5 jours de l'indice BEL20

définit la suite des différences positives $Z(t)_{t=1\dots T-1}$ par :

$$Z(t) = \mathbf{1}_{R_{t+1}-R_t > 0}$$

où $\mathbf{1}$ est la fonction indicatrice.

Test de la médiane

On calcule ensuite la quantité R_Z qui est le nombre total de différences positives dans la suite $Z(t)_{t=1\dots T-1}$. Sous l'hypothèse nulle H_0 , la statistique de test $Z = \frac{R_Z - E(R_Z)}{\sqrt{V(R_Z)}}$ suit une loi normale centrée réduite $\aleph(0; 1)$ avec $E(R_Z) = \frac{T+1}{2}$ (espérance de R_Z sous H_0) et $V(R_Z) = \frac{T-1}{4}$ (variance de R_Z sous H_0); T étant le nombre d'observations.

Nous refusons donc H_0 si la statistique est dans la région de rejet, soit $\frac{R_Z - E(R_Z)}{\sqrt{V(R_Z)}} > q_{1-\frac{\alpha}{2}}$ où $q_{1-\frac{\alpha}{2}}$ est le quantile d'ordre $(1 - \alpha)$ d'une loi normale $\aleph(0; 1)$ ¹⁴.

Test des Up and Down

Dans ce test, on calcule la quantité R_W qui est le nombre total de séquences de 1 et de 0 dans la suite $Z(t)_{t=1\dots T-1}$. Sous l'hypothèse H_0 , la statistique de test $W = \frac{R_W - E(R_W)}{\sqrt{V(R_W)}}$ suit une loi normale centrée réduite $\aleph(0; 1)$ avec $E(R_W) = \frac{2T-1}{3}$ (espérance de R_W sous H_0) et $V(R_W) = \frac{16T-29}{90}$ (variance de R_W sous H_0); T étant le nombre d'observations.

Nous refusons donc H_0 si la statistique est dans la région de rejet, soit $\frac{R_W - E(R_W)}{\sqrt{V(R_W)}} > q_{1-\frac{\alpha}{2}}$ où $q_{1-\frac{\alpha}{2}}$ est le quantile d'ordre $(1 - \alpha)$ d'une loi normale $\aleph(0; 1)$.

Résultats

Le tableau (3.3) indique les valeurs prises par les statistiques de test Z et W pour nos 9 séries de rendements. Bien que le test des Up and Down conduise à rejeter l'hypothèse nulle pour les rendements à 10 jours des indices Cac40 et Dax30, nous décidons d'accepter l'hypothèse nulle d'indépendance et de distribution identique pour nos rendements à 1, 5 et 10 jours, ceci avec un niveau de confiance de 95 %.

¹⁴pour $\alpha = 5\%$, nous comparons Z à 1,96

Indice	h	Z	W	Test Médiane Z	Test Up-Down W
BEL20	1j	0,041	0,482	Acceptation	Acceptation
BEL20	5j	0,446	1,021	Acceptation	Acceptation
BEL20	10j	0,746	1,925	Acceptation	Acceptation
CAC40	1j	1,339	0,674	Acceptation	Acceptation
CAC40	5j	1,015	0,809	Acceptation	Acceptation
CAC40	10j	1,203	2,028	Acceptation	Rejet
DAX30	1j	1,496	0,693	Acceptation	Acceptation
DAX30	5j	1,664	0,925	Acceptation	Acceptation
DAX30	10j	1,784	2,325	Acceptation	Rejet

TABLEAU 3.3 – Tests de la médiane et des "Up and Down" observés sur les rendements des indices boursiers

3.2.4 Tests de normalité

Le test le plus classique, dit test Jarque-Bera est un test de normalité fondé sur les propriétés d'asymétrie *skewness*¹⁵ et d'aplatissement *kurtosis*¹⁶. Soit $(R_t)_{t=1..T}$ une série temporelle et $\mu_k = \frac{1}{T} \sum_{t=1}^T (R_t - \bar{R})^k$ le moment centré d'ordre k .

Le coefficient d'asymétrie est $\beta_1^{1/2} = \frac{\mu_3}{\mu_2^{3/2}}$ et le coefficient d'aplatissement est $\beta_2 = \frac{\mu_4}{\mu_2^2}$.

Si la distribution est gaussienne et le nombre d'observations T grand, nous avons les convergences en loi suivantes :

$$\beta_1^{1/2} \rightarrow \mathfrak{N} \left(0; \sqrt{\frac{6}{T}} \right)$$

et

$$\beta_2 \rightarrow \mathfrak{N} \left(3; \sqrt{\frac{24}{T}} \right)$$

JARQUE ET BERA montrent que sous l'hypothèse nulle H_0 (hypothèse de normalité de la distribution), la statistique de test T_{JB} converge en loi vers la loi du χ^2 (Chi-deux) à deux degrés de liberté :

$$T_{JB} = \frac{T}{6} \left(\beta_1^{1/2} \right)^2 + \frac{T}{24} (\beta_2 - 3)^2 \xrightarrow[T \rightarrow \infty]{L} \chi_2^2$$

Plus la valeur du χ^2 est importante, moins l'hypothèse nulle de normalité des données est vraisemblable. À l'aide de ce test et pour un degré de signification α , l'hypothèse H_0 sera rejetée si la valeur de la statistique T_{JB} calculée sur les échantillons $R_i(t)$ est

¹⁵moment d'ordre 3

¹⁶moment d'ordre 4

Indice	h	Moyenne	Ecart-type	Skewness ²⁰	Kurtosis ²¹	T_{JB}
BEL20	1j	0,0392 %	0,2753 %	-0,6553	5,0841	153,55
BEL20	5j	0,2045 %	0,6148 %	-0,6781	4,4752	101,73
BEL20	10j	0,4149 %	0,8801 %	-0,7415	4,6036	120,86
CAC40	1j	0,0264 %	0,3396 %	-0,3560	3,8459	30,97
CAC40	5j	0,1429 %	0,6711 %	-0,4041	3,4832	22,46
CAC40	10j	0,2869 %	0,8930 %	-0,7250	3,5850	61,93
DAX30	1j	0,0286 %	0,3825 %	-0,4168	3,8736	36,94
DAX30	5j	0,1563 %	0,8455 %	-0,5023	4,019	51,85
DAX30	10j	0,3156 %	0,1151 %	-0,7874	4,1639	97,15
$\mathcal{N}(\mu, \sigma)$	-	μ	σ	0	3	

TABLEAU 3.4 – Statistiques observées sur les rendements des indices boursiers

supérieure au quantile d'ordre α d'une distribution suivant un χ^2 à deux degrés de liberté. Pour un seuil de confiance de 95 % ($\alpha = 5\%$), la valeur de la statistique de test T_{JB} ne doit pas dépasser 5,991¹⁷. Pour un seuil de 99 % ($\alpha = 1\%$), elle ne doit pas dépasser 9,21¹⁸. Les histogrammes nous indiquent les différents moments (moyenne, écart-type, ...) des séries étudiées ainsi que la probabilité que les statistiques T_{JB} soient égales aux valeurs observées¹⁹. Nous rappelons que pour une distribution normale, le coefficient d'asymétrie (*skewness*) est égal à 0 et le coefficient d'aplatissement (*kurtosis*) est égal à 3.

Les distributions empiriques sont indiquées par les figures (3.4, 3.5, 3.6 et 2-annexe 1), et résumées dans le tableau (3.4).

Rendements Indice BEL20

Dans le cas de l'indice BEL20, l'histogramme des rendements nous indique la présence d'une queue de distribution importante pour des rendements négatifs. Les tests de normalité de Jarque-Bera du tableau 3.5²² indiquent des valeurs largement supérieures à la valeur critique du χ^2 . Ceci nous conforte dans l'idée de rejeter l'hypothèse de normalité des rendements à 1, 5 et 10 jours de l'indice Bel20.

Rendements Indice CAC40

Dans le cas de l'indice CAC40, le *skewness* et le *kurtosis* sont plus faibles et voisins d'une loi gaussienne. Toutefois, ils restent statistiquement élevés et on note la

¹⁷table de la loi du χ^2 , Gilbert SAPORTA, Probabilités, analyse des données et statistique

¹⁸Table du χ^2 , [36], p.450

¹⁹c'est la p-value ou probabilité d'observer la valeur de la statistique de test. On rejette l'hypothèse nulle H_0 lorsque cette p-value est inférieure au degré de signification α

²²voir figure 2-annexe 1

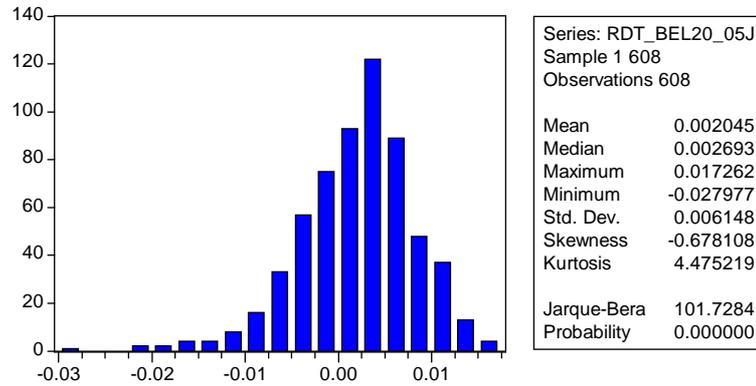


FIGURE 3.4 – Histogramme des rendements à 5 jours de l’indice BEL20

h (en jour)	$h t_{i=1}^{JB}$	Résultat
1	153,55	Rejet
5	101,73	Rejet
10	120,86	Rejet

TABLEAU 3.5 – Test Normalité Rendement Indice BEL20

présence d’une queue de distribution épaisse pour des valeurs négatives et positives. La statistique de test indiquée dans le tableau (3.6) nous conduit également à rejeter l’hypothèse d’une distribution gaussienne des rendements à 1, 5 et 10 jours de l’indice Cac40.

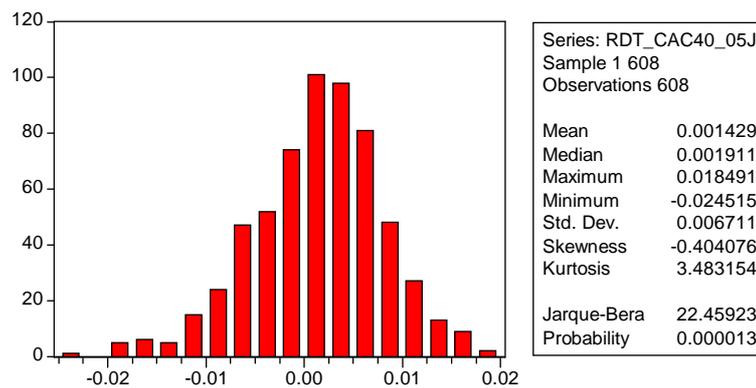


FIGURE 3.5 – Histogramme des rendements à 5 jours de l’indice CAC40

h (en jour)	$h t_{i=2}^{JB}$	Résultat
1	30,97	Rejet
5	22,46	Rejet
10	61,93	Rejet

TABLEAU 3.6 – Test Normalité Rendement Indice CAC40

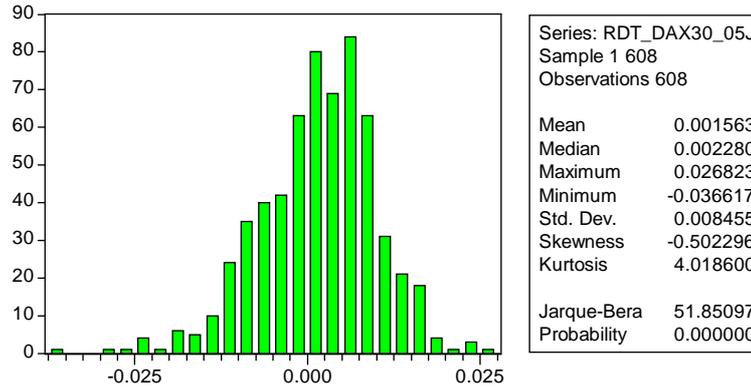


FIGURE 3.6 – Histogramme des rendements à 5 jours de l'indice DAX30

Rendements Indice DAX30

Dans le cas de l'indice Dax30, la forme de l'histogramme s'éloigne nettement d'une densité gaussienne. Ceci est essentiellement dû à la présence d'une queue de distribution significative. Nous aboutissons donc à la même conclusion que pour les précédents indices et rejetons l'hypothèse de normalité des rendements à 1, 5 et 10 jours de l'indice Dax30.

3.3 Corrélations

Nous terminons cette analyse statistique en établissant les matrices des corrélations empiriques (figures 3.8, 3.9 et 3.10) qui nous indiquent la présence de liens forts de dépendances de nos rendements boursiers, ceci quelque soit la période de 1, 5 ou 10 jours. Les tests statistiques précédents nous conduisent à conclure que :

h (en jour)	$h t_{i=3}^{JB}$	Résultat
1	36,94	Rejet
5	51,85	Rejet
10	97,15	Rejet

TABLEAU 3.7 – Test Normalité Rendement Indice DAX30

	${}_1R_{i=1}$ - BEL20	${}_1R_{i=2}$ - CAC40	${}_1R_{i=3}$ - DAX30
${}_1R_{i=1}$ - BEL20	100 %		
${}_1R_{i=2}$ - CAC40	80,39 %	100 %	
${}_1R_{i=3}$ - DAX30	54,26 %	61,59 %	100 %

TABLEAU 3.8 – Matrice des corrélations - Rendements à 1 jour

	${}_5R_{i=1}$ - BEL20	${}_5R_{i=2}$ - CAC40	${}_5R_{i=3}$ - DAX30
${}_5R_{i=1}$ - BEL20	100 %		
${}_5R_{i=2}$ - CAC40	80,73 %	100 %	
${}_5R_{i=3}$ - DAX30	77,66 %	92,01 %	100 %

TABLEAU 3.9 – Matrice des corrélations - Rendements à 5 jours

	${}_{10}R_{i=1}$ - BEL20	${}_{10}R_{i=2}$ - CAC40	${}_{10}R_{i=3}$ - DAX30
${}_{10}R_{i=1}$ - BEL20	100 %		
${}_{10}R_{i=2}$ - CAC40	82,15 %	100 %	
${}_{10}R_{i=3}$ - DAX30	78,42 %	92,03 %	100 %

TABLEAU 3.10 – Matrice des corrélations - Rendements à 10 jours

- les séries temporelles des rendements sont stationnaires,
- les observations sont indépendantes et identiquement distribués,
- les rendements ne suivent pas des lois normales.

Nous rappelons que nous souhaitons modéliser le rendement de notre portefeuille P composé des trois actifs Bel20, Cac40 et Dax30. Pour atteindre cet objectif, nous pourrions proposer la méthodologie suivante :

1. Modéliser chacune des distributions marginales observées :
 - soit en recherchant une loi de densité connue ajustant au mieux les densités empiriques,
 - ou bien en modélisant numériquement ces densités par des estimations non paramétriques de type estimateurs par noyaux de densité²³.
2. Puis modéliser la ‘dépendance’ entre ces lois marginales afin d’obtenir la distribution conjointe du rendement du portefeuille, moyennant un ajustement en fonction du poids de chaque actif dans le portefeuille. On pourra ensuite s’appuyer sur la théorie des copules.

Le copule, contrairement à la corrélation, modélise de manière plus complète la dépendance statistique entre deux variables aléatoires. En particulier, il permet de déterminer une distribution jointe à partir des distributions marginales. Cependant, les copules restent des outils difficiles à manipuler.

²³kernel density estimation

Nous avons choisi d'explorer une autre voie en utilisant une technique utilisée en traitement du signal. En considérant nos séries de rendements comme des signaux physiques (voir les figures 3.2) dont les ingénieurs en traitement du signal sont familiers, nous allons faire usage de la technique dite d'analyse en composantes indépendantes. Le chapitre suivant est consacré à l'exposition de cette technique et son application à nos séries de rendements.

Chapitre 4

L'Analyse en Composantes Indépendantes

L'Analyse en Composantes Indépendantes¹ ACI est une technique issue du traitement des signaux. Elle consiste en l'obtention, par opération linéaire, de variables indépendantes au sens statistique, ceci à partir d'un ensemble de signaux physiques. Il s'agit donc de rechercher une transformation linéaire qui minimise la dépendance statistique entre les composantes d'un vecteur aléatoire.

Dans ce chapitre, nous allons utiliser l'ACI afin de trouver une combinaison linéaire des trois séries de rendements qui les rendent indépendants au sens statistiques. Par abus de langage, nous appellerons ces séries indépendantes les rendements sources.

Pour une présentation complète de l'analyse en composantes indépendantes, nous invitons le lecteur à consulter [18], [25], [26], [28] ou [29].

Cette méthode est couramment utilisée en traitement du signal et a été appliquée pour la première fois sur des données financières par Back et Weigend dans [7] et [8].

4.1 Description

L'analyse en composantes indépendantes, également appelée séparation aveugle de sources², est une méthode traitant des observations vectorielles (multivariées) afin d'en extraire, linéairement, des composantes qui soient aussi "indépendantes" possibles.

Classiquement, on l'illustre par le problème de séparation de sources dit "*cocktail party problem*" problème rencontré en traitement du signal de parole et illustré par la figure (4.1). Dans une pièce où sont présentes trois personnes qui parlent ou chantent

¹Independent Components Analysis, ICA

²Blind Source Separation

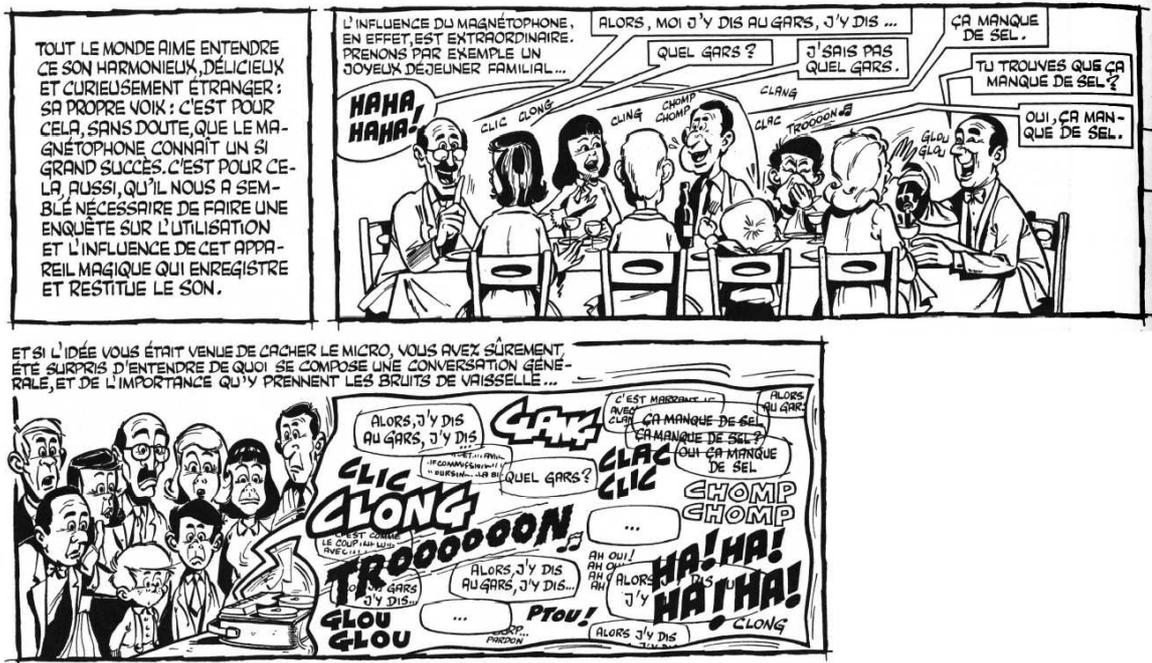


FIGURE 4.1 – Problème de la séparation de sources vu par Gotlib dans les *Dingodossiers* ([1])

indépendamment les unes des autres, nous plaçons aussi trois microphones. Sur les enregistrements sonores, nous allons pouvoir entendre les voix des trois personnes superposées les unes sur les autres. L'objectif de la séparation aveugle est de retrouver (ou séparer) les voix à partir des enregistrements sonores. Il s'agit alors de reproduire la capacité de l'ouïe humaine à discriminer des sources sonores dans des environnements plus ou moins bruités.

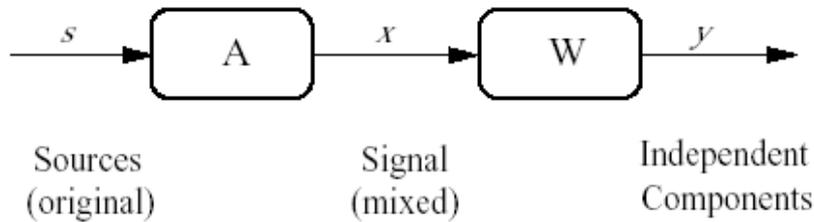


FIGURE 4.2 – Schéma de l'ACI. Des signaux sources s sont mélangés aux travers d'une matrice A pour former des signaux observés x . La matrice de séparation W transforme les signaux observés x en composantes indépendantes formant le signal y

De manière plus formelle, on considère qu'un ensemble de signaux observés³ (ou vec-

³on supposera dans notre cas que les signaux ne sont pas bruités

teur observation) notés $x(t)$ est issu d'un mélange linéaire de sources originales notées $s(t)$ indépendantes statistiquement. Le mélange linéaire est représenté par une matrice de mélange notée A et illustré par la figure (4.2).

La figure (4.3) présente ce problème dans le cas de la séparation de sources sonores :

- les signaux perçus par chaque microphone $x_i(t)$ sont supposés être, à tout instant, la superposition linéaire⁴ des trois signaux sonores $s_i(t)$. Les voix captées sont des signaux mélangés pour lesquels on souhaite obtenir les voix originelles dans le but par exemple de ne suivre qu'une conversation entre deux personnes. On essaye alors de réaliser une "séparation" du son.
- après détermination de la matrice de mélange A et obtention de la matrice inverse $W = A^{-1}$, on peut reconstituer les signaux sources $s_i(t)$.

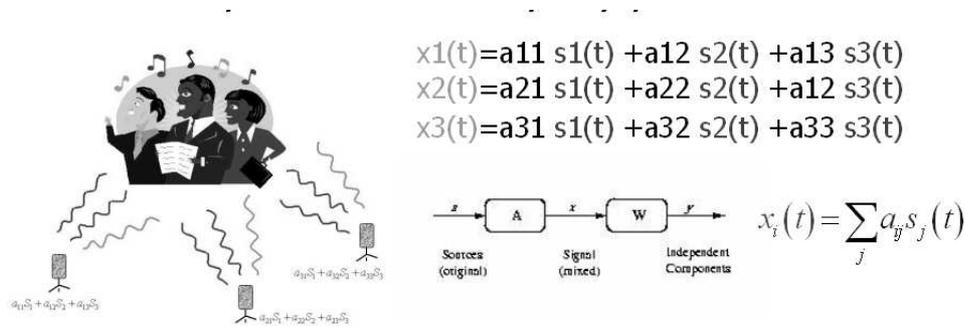


FIGURE 4.3 – Principe de l'ACI dans le cas du cocktail party problem

Nous utilisons alors l'ACI en supposant que la série temporelle multivariée x est le résultat d'un mélange linéaire de séries temporelles s statistiquement indépendantes. Soit, quelque soit t :

$$x(t) = As(t) \Leftrightarrow x_i(t) = \sum_{j=1}^n a_{ij} s_j(t) \tag{4.1}$$

On cherche alors à déterminer une matrice de séparation W qui nous permet de d'obtenir une série temporelle multivariée y dont les composantes sont statistiquement indépendantes.

En effet, Darmon a montré dans [20] qu'à partir de variables aléatoires dépendantes, il est toujours possible de construire une transformation non triviale de telle sorte que les variables aléatoires obtenues à la sortie du mélange soient indépendantes. Il s'agit de déterminer un espace de représentation S dont les axes sont indépendants selon un critère de dépendance que nous allons présenter par la suite. On peut formaliser

⁴on suppose dans ce cas qu'il n'y a pas de retard dans la propagation des ondes sonores

ce problème par :

$$\min_{A \in GL(n)} I(A^{-1}X)$$

où $I(Y)$ est une mesure de la dépendance des coordonnées d'un vecteur aléatoire Y , A est une matrice carrée inversible de taille n et n est le nombre de sources. Les méthodes de résolution aboutissent à trouver une matrice W dite matrice de séparation qui est l'inverse de la matrice A .

Avant d'aborder les principes mathématiques de l'ACI, et pour rendre le principe plus claire, nous pouvons faire une comparaison avec l'analyse en composantes principales ACP.

4.2 Lien avec l'Analyse en Composantes Principales

L'Analyse en Composantes Principales⁵ ACP est un outil d'analyse de données introduit par Karl Pearson en 1901. Cette méthode d'analyse statistique peut se résumer à des opérations matricielles, par définition linéaires, sur les observations. Lorsqu'il y existe des corrélations entre les variables étudiées, la dimension de l'espace des données est supérieure au nombre de variables caractéristiques nécessaires pour décrire ces données.

La transformation linéaire est destinée à réduire la dimension de la représentation des données tout en préservant au mieux l'information qu'elles véhiculent. L'ACP vise ainsi à trouver le meilleur espace, de dimension réduite, représentant de manière fiable les variables observées. L'ACI peut se rapprocher de l'ACP à la différence que l'ACP n'achève pas obligatoirement l'indépendance statistique des composantes obtenues car on se limite à trouver des axes (ou composantes) non corrélés.

L'analyse en composante indépendante poursuit une démarche similaire si ce n'est qu'elle recherche non seulement des axes non corrélés mais également indépendants. Nous verrons plus loin que l'ACI utilise le concept de cumulants pour quantifier la dépendance statistique. Parfois, on utilise le terme de moments centrés d'ordres supérieurs pour nommer les cumulants ([28]).

Ainsi en considérant des conditions nécessaires et suffisantes d'indépendance statistique, l'ACP ne tient compte que des conditions nécessaires de corrélation nulle entre les variables mais pas des conditions suffisantes. Nous verrons plus loin qu'une condition suffisante d'indépendance est que les cumulants croisés d'ordre supérieur à 2

⁵Principal Components Analysis PCA

doivent être nuls. D'un point de vue technique, l'ACP se limite alors à la minimisation des co-moments de second ordre, appelés covariances. Par contre, l'ACI minimise tous les co-moments d'ordre supérieur pour obtenir des axes indépendants même si en pratique on se limite à calculer les co-moments d'ordre 4.

4.3 Mesure de l'indépendance statistique

4.3.1 Définition

Dans le cadre de l'analyse en composantes indépendantes, il est nécessaire de savoir si des variables aléatoires sont indépendantes. Nous rappelons tout d'abord la définition statistique de l'indépendance : soit X un vecteur multivarié à d dimensions. $X = (X_1, \dots, X_d)'$ possède une loi multivariée de fonction de répartition F si pour tout $x = (x_1, \dots, x_d)' \in \mathfrak{R}^d$, nous avons la relation :

$$F(x) = P(X \leq x) = P(X \leq x_1, \dots, X \leq x_d) \quad (4.2)$$

La loi marginale de chaque composante j est telle que $F_j(x_j) = P(X_j \leq x_j)$. On dit alors que les X_1, \dots, X_d sont statistiquement indépendants si et seulement si

$$F(x_1, \dots, x_d) = F_1(x_1) \cdot F_2(x_2) \cdot \dots \cdot F_d(x_d) \quad (4.3)$$

Ainsi l'indépendance se traduit par le fait que la distribution conjointe est égale au produit des distributions marginales.

4.3.2 Les Cumulants

La première fonction caractéristique d'une variable aléatoire continue x est la transformée de sa densité de probabilité p_x et est notée $\phi(u)$:

$$\phi_x(u) = E(e^{jux}) = \int_{-\infty}^{+\infty} e^{jux} p_x dx \quad (4.4)$$

Un des principaux intérêts de la première fonction caractéristique est qu'elle donne, par son développement en série de McLaurin, tous les moments de la variable aléatoire x .

On définit également la seconde fonction caractéristique, notée ψ par :

$$\psi(u) = \ln \{\phi(u)\} \quad (4.5)$$

De même le développement en série de McLaurin de $\phi(u)$ introduit des coefficients c_k qui sont dénommés les cumulants de la distribution. On montre alors que les cumulants d'une distribution normale sont nuls, sauf éventuellement les deux premiers qui sont sa moyenne et sa variance. On montre également que les cumulants s'expriment comme des combinaisons de moments et quantifient l'écart entre la distribution d'une variable aléatoire et une distribution gaussienne (après normalisation par l'écart-type). En particulier, les cumulants d'ordres un et deux sont la moyenne et la variance des rendements. Les cumulants d'ordres trois et quatre sont le coefficient d'asymétrie skewness et le coefficient d'aplatissement kurtosis.

On généralise alors ces deux fonctions caractéristiques au cas de vecteurs aléatoires mais nous ne développerons pas cette approche car elle fait appel à des notions dépassant le cadre de ce travail⁶. Nous allons uniquement utiliser la propriété suivante. Des variables aléatoires sont indépendantes si et seulement si leurs cumulants croisés d'ordre supérieur à deux sont nuls. En effet, Mansour, dans [32], démontre que pour qu'un couple de variables aléatoires soit indépendant, il faut et il suffit que tous les cumulants d'ordre supérieur à deux soient nuls. Ce résultat se généralise à un vecteur aléatoire de dimension n . Les composantes du vecteur sont indépendantes au sens statistique si et seulement si tous les cumulants croisés d'ordre n sont nuls.

Pour appliquer l'ACI, nous avons utilisé l'algorithme jadeR développé par Cordoso ([14]) qui démontre qu'il suffit de trouver une matrice annulant les cumulants croisés d'ordre 4 pour satisfaire l'indépendance à la condition qu'au plus une source soit gaussienne et que les sources soient stationnaires. Ainsi l'indépendance statistique consiste à trouver une matrice de séparation annulant les cumulants à l'ordre 4, même si théoriquement il faudrait annuler l'infinité des cumulants de tout ordre (supérieur à 2).

4.4 Applications

Afin de faciliter les calculs formels que nous verrons au chapitre six, nous raisonnons sur des données centrées et réduites. Nous avons donc appliqué l'algorithme

⁶utilisation d'un formalisme tensoriel où un tenseur est une généralisation de la notion de matrice (tableau à plusieurs dimensions)

jadeR sur les séries des rendements, après centrage et réduction, notées

$${}_hR_i^{CR}$$

avec $i = 1$ pour l'indice Bel20, $i = 2$ pour l'indice Cac40 et $i = 3$ pour l'indice Dax30.

4.4.1 Série des rendements à 5 jours

Matrice de séparation

Pour les observations des rendements à 5 jours, les matrices estimées de séparation ${}_hW$ et de mélange ${}_hA$, $h=5$ obtenues sont les suivantes :

$${}_5W = \begin{bmatrix} 0,87388 & -0,8003 & 0,70924 \\ 1,4073 & -1,2911 & -0,23566 \\ 0,4411 & 2,0174 & -1,7227 \end{bmatrix} \quad (4.6)$$

et

$${}_5A = \begin{bmatrix} 0,9246 & 0,0179 & 0,3782 \\ 0,7948 & -0,6228 & 0,4124 \\ 1,1675 & -0,7247 & -0,0007 \end{bmatrix} \quad (4.7)$$

Sources obtenues

Nous supposons alors qu'il existe trois séries temporelles sources ${}_5S_1$, ${}_5S_2$ et ${}_5S_3$ qui, après l'opération linéaire ${}_5A_5S(t)$ nous donnent les rendements observés ${}_5R_1$, ${}_5R_2$ et ${}_5R_3$.

Ces séries sources peuvent alors se reconstruire par l'opération inverse :

$${}_5S(t) = {}_5A^{-1} \cdot {}_5R^{CR}(t) = {}_5W \cdot {}_5R^{CR}(t)$$

Ces séries que nous appellerons signaux sources sont représentées par les figures (4.4 et 20-annexe 5).

Mesure de l'indépendance

Afin de vérifier l'indépendance des sources, nous avons calculé les cumulants croisés à l'ordre 4 et avons obtenus :

- pour les rendements sources -4 000,7 (1 jour), -3 161 (5 jours) et -2 106 (10 jours)

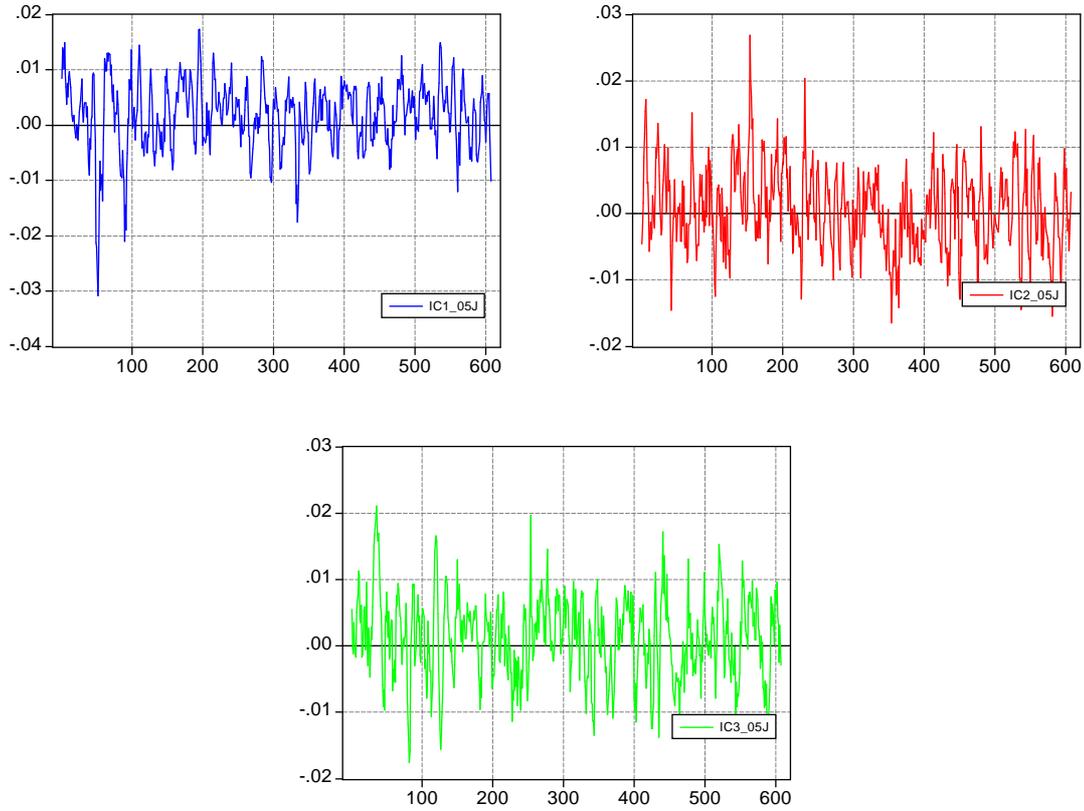


FIGURE 4.4 – Séries des signaux sources à 5 jours

	${}_5S_{i=1} - IC1_5J$	${}_5S_{i=2} - IC2_5J$	${}_5S_{i=3} - IC3_5J$
${}_5S_{i=1} - IC1_5J$	100 %		
${}_5S_{i=2} - IC2_5J$	0.0007 %	100 %	
${}_5S_{i=3} - IC3_5J$	0.0003 %	-0,0005 %	100 %

TABLEAU 4.1 – Matrice des corrélations - Rendements Sources à 5 jours

– pour les composantes indépendantes $-1,53e-08$ (1 jour), $-1,85e-7$ (5 jours) et $-5,84e-7$ (10 jours)

Nous constatons l'annulation des cumulants croisés et supposons l'indépendance des signaux sources.

Les matrices des corrélations sont indiquées au tableau (4.1) et confirment l'absence de corrélation entre les signaux sources.

	${}_1S_{i=1} - IC1_1J$	${}_1S_{i=2} - IC2_1J$	${}_1S_{i=3} - IC3_1J$
${}_1S_{i=1} - IC1_1J$	100 %		
${}_1S_{i=2} - IC2_1J$	-0.0000 %	100 %	
${}_1S_{i=3} - IC3_1J$	-0.0001 %	-0,0002 %	100 %

TABLEAU 4.2 – Matrice des corrélations - Rendements Sources à 1 jour

	${}_{10}S_{i=1} - IC1_10J$	${}_{10}S_{i=2} - IC2_10J$	${}_{10}S_{i=3} - IC3_10J$
${}_{10}S_{i=1} - IC1_10J$	100 %		
${}_{10}S_{i=2} - IC2_10J$	-0.0001 %	100 %	
${}_{10}S_{i=3} - IC3_10J$	-0.0004 %	0,0003 %	100 %

TABLEAU 4.3 – Matrice des corrélations - Rendements Sources à 10 jours

4.4.2 Série des rendements à 1 et 10 jours

Pour les rendements à 1 jour, les matrices de mélange et de séparation obtenues sont :

$${}_1W = \begin{bmatrix} 0,84397 & -0,10881 & 0,28475 \\ 0,0053327 & 0,91134 & -0,83163 \\ 1,466 & -1,1384 & -0,26845 \end{bmatrix} \quad (4.8)$$

et

$${}_1A = \begin{bmatrix} 0,94919 & 0,28153 & 0,13466 \\ 0,97022 & 0,5131 & -0,56041 \\ 1,0693 & -0,63837 & -0,61326 \end{bmatrix} \quad (4.9)$$

Pour les rendements à 10 jours, les matrices de mélange et de séparation sont :

$${}_{10}W = \begin{bmatrix} 0,86492 & -0,82042 & 0,7299 \\ 0,85788 & -2,5499 & 1,0234 \\ 1,2831 & 0,6905 & -1,5199 \end{bmatrix} \quad (4.10)$$

et

$${}_{10}A = \begin{bmatrix} 0,92816 & -0,21761 & 0,2992 \\ 0,76648 & -0,65933 & -0,075853 \\ 1,1317 & -0,48323 & -0,43981 \end{bmatrix} \quad (4.11)$$

Les matrices des corrélations sont indiquées aux tableaux (4.2) et (4.3). La figure (20-annexe 4) montre les signaux sources obtenus à partir des matrices (4.8) et (4.10) et des rendements centrés réduits ${}_1R^{CR}$ et ${}_{10}R^{CR}$.

4.5 Conclusion

Le présent chapitre avait pour objectif de présenter l'analyse en composantes indépendantes et son application pour contourner le problème de la modélisation des dépendances entre nos rendements. En effet, la simple utilisation de la corrélation implique l'hypothèse d'une distribution multinormale des rendements.

Nous avons également présenté la notion de cumulants qui est une mesure de la dépendance entre variables aléatoires. Selon cette approche par les cumulants, nous avons déterminé par opération linéaire, une série de rendements sources que nous considérons comme statistiquement indépendants au vue des valeurs des cumulants croisés.

L'objectif du chapitre suivant est de modéliser leur fonction de densité.

Chapitre 5

Modélisation par mélange fini de lois gaussiennes

Après avoir obtenu des séries temporelles ${}_hS$ jugées indépendantes, se pose la question de la modélisation de leurs distributions statistiques. Nous proposons ici d'utiliser la modélisation par mélange gaussien pour les raisons suivantes :

1. d'une part, tenter de modéliser les queues de distribution observées (voir les histogrammes empiriques (5.1)),
2. d'autre part, conserver une écriture mathématique pratique des densités des rendements.

Dans un premier temps, nous présentons le modèle dit des mélanges gaussiens finis, puis nous exposons l'algorithme EM¹ utilisé pour résoudre ce type de problème statistique. Cet exposé sera succinct car les théories sous-jacentes dépassent le cadre de ce mémoire. Enfin, nous présentons les modélisations obtenues à partir des données initiales.

5.1 Modélisation

Les modèles de mélanges permettent d'étudier des distributions complexes tout en s'appuyant sur des bases statistiques relativement simples comme la densité gaussienne. Ainsi la superposition de densités élémentaires nous permet d'obtenir la densité mélange.

¹Expectation-Maximization

5.1.1 Formalisation mathématique

Le modèle de mélange consiste à supposer que les observations $S(t)$ sont des réalisations de variables aléatoires mutuellement indépendantes et distribuées selon la même famille de loi. Ainsi, la densité d'un mélange correspond à la superposition de densités d'un même type de loi. En notant f la densité de S , nous avons :

$$f(s) = \sum_{k=1}^K \gamma_k f_k(s) \quad (5.1)$$

où K est le nombre de densités et γ_k le poids de chaque densité dans la densité finale. Nous souhaitons modéliser la densité s d'une composante indépendante S par un mélange de lois gaussiennes. Ceci revient à écrire l'équation (5.1) sous la forme :

$$s = \sum_{k=1}^K \gamma_k \mathcal{N}(\mu_k, \sigma_k) \quad (5.2)$$

Bien entendu, s'agissant d'une densité de probabilité, nous posons la contrainte (5.3) :

$$\sum_{k=1}^K \gamma_k = 1 \quad (5.3)$$

Il est utile de remarquer qu'il ne s'agit pas de modéliser S par une somme de variables aléatoires gaussiennes mais plutôt de modéliser sa densité s par une somme de densités gaussiennes. Ainsi, il s'agit de déterminer le nombre de densités K et les K triplets $(\gamma_k, \mu_k, \sigma_k)$

5.1.2 Modélisation mathématique

Comme le démontre [12], ce problème d'estimation appartient à la classe plus large des problèmes d'estimation en présence de données cachées. Supposons que le vecteur des observations S soit généré de la manière suivante :

1. on tire une variable aléatoire discrète u à valeur dans $1 \dots K$ où K désigne le nombre de classes du mélange. On note $\gamma_k = P\{u = k\}$, $k = 1 \dots K$ les probabilités respectives de tirer chacune des classes ;
2. conditionnellement à l'événement $\{u = k\}$, S est distribué selon la loi de densité de probabilité f_k . La densité conditionnelle du vecteur S peut s'écrire sous la forme concise suivante :

$$g(S|u) = \sum_{k=1}^K f_k(s) I_{\{U=k\}} \quad (5.4)$$

où I désigne la fonction indicatrice d'un événement. À partir de la règle de Bayes², la densité conjointe du vecteur S et de la fonction indicatrice I est :

$$h(s, u) = g(s|u) p(u) = \left(\sum_{k=1}^K f_k(s) I_{\{U=k\}} \right) \left(\sum_{k=1}^K \gamma_k I_{\{U=k\}} \right) = \sum_{k=1}^K \gamma_k f_k(s) I \quad (5.5)$$

La loi marginale du vecteur observé S s'obtient en sommant l'équation 5.5 sur le domaine de U et on obtient :

$$f(s) = \sum_{k=1}^K \gamma_k f_k(s) \quad (5.6)$$

Cette équation (5.6) correspond à l'équation (5.1) qui définit le modèle de mélange. Il est donc équivalent à un modèle dans lequel on suppose que les données sont réparties aléatoirement et indépendamment les unes des autres en K classes. Chaque classe est caractérisée par une distribution différente f_s . Ainsi la variable indicatrice U est une donnée constitutive du problème qui présente l'inconvénient majeur d'être manquante. On observe donc des réalisations du vecteur aléatoire S sans connaître quelle est la classe du mélange associée à chaque observation. Au sens de l'algorithme EM, la variable aléatoire U constitue une donnée manquante (ou cachée).

5.1.3 Algorithme EM

Fonction de vraisemblance

L'utilisation de la méthode statistique du maximum de vraisemblance conduit à écrire la fonction de vraisemblance sous la forme (voir [12]) :

$$F(S; \Theta) = \prod_{t=1}^T \sum_{k=1}^K \gamma_k f_k(s(t)) \quad (5.7)$$

où $S = (s_1, \dots, s_T)$ est l'ensemble des T valeurs dont on dispose pour estimer les paramètres du modèle et Θ regroupe les paramètres du modèle à estimer, c'est-à-dire les K triplets $(\gamma_k, \mu_k, \sigma_k)$. À ce stade, les difficultés d'estimation du modèle sont les

²le théorème de Bayes permet d'inverser les probabilités. C'est-à-dire que si l'on connaît les conséquences d'une cause, l'observation des effets permet de remonter aux causes

suivantes :

1. la détermination du nombre de densité K à prendre en compte
2. maximiser une fonction ayant l'expression (5.7) est délicat car elle n'est pas de forme convexe et n'admet donc pas, à priori, de maximum unique.

Principe

L'algorithme EM³ ou algorithme Espérance - Maximisation est l'une des méthodes de maximisation de la vraisemblance de la fonction $F(S; \Theta)$. Toutefois, il n'existe pas de méthode automatique permettant de déterminer le nombre de densité K à utiliser. Nous avons utilisé différents paramètres initiaux et avons conservé ceux pour lesquels l'algorithme EM converge.

Sans entrer dans les détails, l'algorithme EM est de type itératif où chaque itération comporte les deux étapes suivantes :

- Étape E-Expectation : lors de cette étape, on formule la quantité $Q_{\Theta_n}(\Theta)$ définie par $Q_{\Theta_n}(\Theta) = E[\log H(S, U, \Theta) | S; \Theta_n]$ qui est l'espérance de la log-vraisemblance,
- Étape M-Maximization : lors de cette étape, on recherche la valeur Θ_{n+1} qui maximise la quantité $Q_{\Theta_n}(\Theta)$

$$\Theta_{n+1} = \arg \max_{\Theta} Q_{\Theta_n}(\Theta)$$

où Θ_n désigne la valeur estimée des paramètres du modèle à la n^{e} itération. S désigne l'ensemble des données observées, U l'ensemble des données latentes associées et H la fonction de vraisemblance conjointe des données observées et latentes.

La propriété la plus importante de l'algorithme EM est que la suite des valeurs estimées Θ_n est construite de manière à ce que la vraisemblance des données observées augmente à chaque itération.

Une seconde propriété de l'algorithme EM est qu'il permet de calculer le gradient de la fonction de vraisemblance aux points Θ_n . Cette propriété permet d'assurer la convergence de l'algorithme puisqu'elle implique que les points stables de l'algorithme soient des points stationnaires de la vraisemblance. Toutefois, comme tout processus de recherche d'extrema, le comportement de l'algorithme EM dépend fortement de la valeur initiale Θ_1 de l'algorithme.

³Expectation - Maximization algorithm

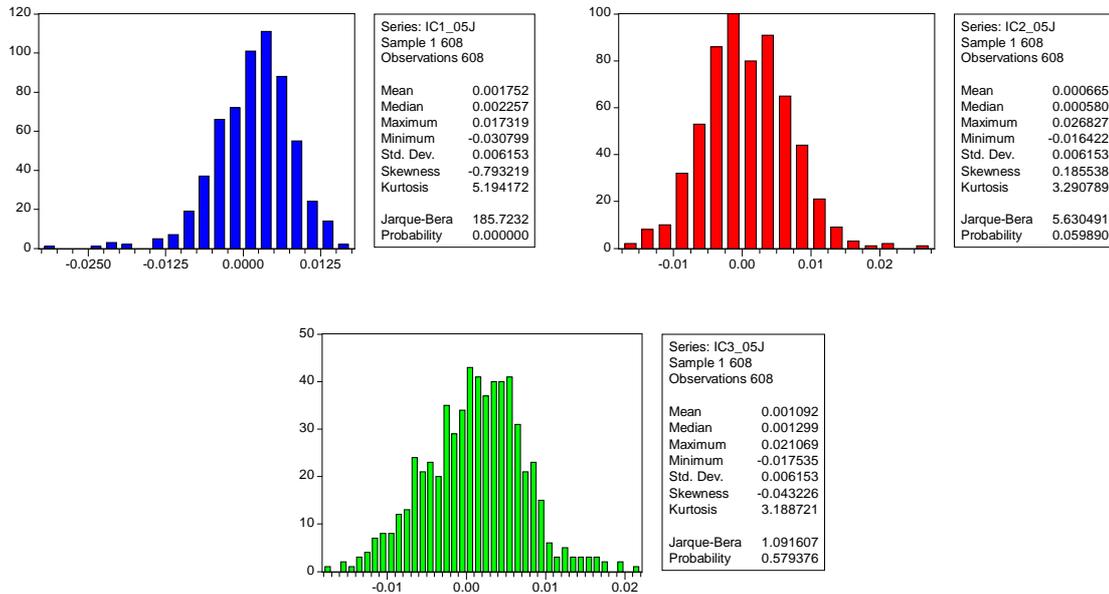


FIGURE 5.1 – Histogrammes des signaux sources à 5 jours

5.2 Applications

Les histogrammes empiriques des séries sources $S(t)$ pour une période de 5 jours sont indiquées par les figures (5.1) et (21-annexe 5).

Selon le test de normalité de Jarque-Bera, l'hypothèse de normalité de la deuxième (IC3_05J) et troisième composante indépendante (IC3_05J) peut être acceptée. Les p-values obtenues de 5,98 % et 57,94 % nous indiquent que cette hypothèse n'est pas rejetée avec un niveau de confiance confortable.

Les modélisations par mélange gaussien ont été réalisées à l'aide de la bibliothèque de fonction "mixdist"⁴ disponible sous le logiciel statistique libre R⁵. Comme tout problème d'optimisation, l'algorithme EM nécessite un bon paramétrage des conditions initiales. Pour chacune des modélisation par mélange gaussien, il existe peut-être d'autres représentations équivalentes ou meilleur. Le critère d'arrêt a été l'obtention d'une statistique du test d'adéquation à une loi (test du χ^2) satisfaisante.

En règle générale, les meilleurs ajustements ont été obtenus avec des mélanges à deux densités gaussiennes.

⁴<http://www.math.mcmaster.ca/peter/mix/mix.html>

⁵<http://www.r-project.org>

5.2.1 Composantes sources à 5 jours ${}_5S$

Composante n° 1 - IC1_05

Le meilleur ajustement a été obtenu avec un mélange de deux densités gaussiennes. En notant $\mathcal{N}(i, j, h)(\gamma_{i,j,h}, \mu_{i,j,h}, \sigma_{i,j,h})$ la densité gaussienne n° j , relative à la composante indépendante n° i pour un rendement à h jour(s), le tableau (5.1) indique les paramètres obtenus pour le mélange.

IC1_05J	$\mathcal{N}_{1,1,5}$	$\mathcal{N}_{1,2,5}$
γ	0,057017	0,94298
μ	-0,007984	0,002341
σ	0,0097468	0,0052915

TABLEAU 5.1 – Paramètres du mélange gaussien de la composante IC1_5J

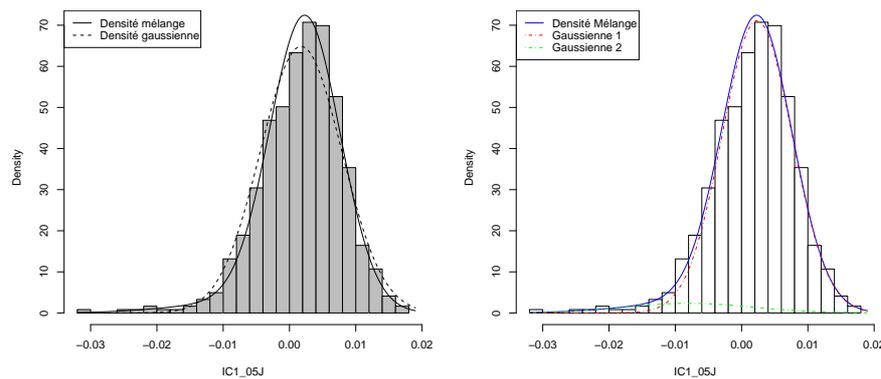


FIGURE 5.2 – Modélisation des densités du rendement source IC1_5J

La figure (5.2) représente graphiquement le mélange. Dans la première figure à droite, nous avons tracé l'histogramme de la série IC1.1J puis les deux fonctions de densités :

1. la densité gaussienne ayant pour valeur les moyennes et écart-types empiriques,
2. la densité mélange composée de deux densités gaussiennes.

Dans cette même figure, nous avons également représenté à droite les densités individuelles composant notre densité mélange.

Composante n° 2 - IC2_05

De même, l'ajustement optimal a été obtenu avec un mélange de deux densités gaussiennes. On observe que les densités mélange et gaussienne (basée sur les moments empiriques) sont proches. Ceci est conforme avec l'histogramme (5.1) pour

lequel la statistique Jarque-Bera ne conduit pas à rejeter l'hypothèse de normalité. La figure (5.3) représente graphiquement ce mélange.

IC2.05J	$\mathcal{N}_{2,1,5}$	$\mathcal{N}_{2,2,5}$
γ	0,77038	0,22962
μ	-0,000001	0,002893
σ	0,0056569	0,0071414

TABLEAU 5.2 – Paramètres du mélange gaussien de la composante IC2.5J

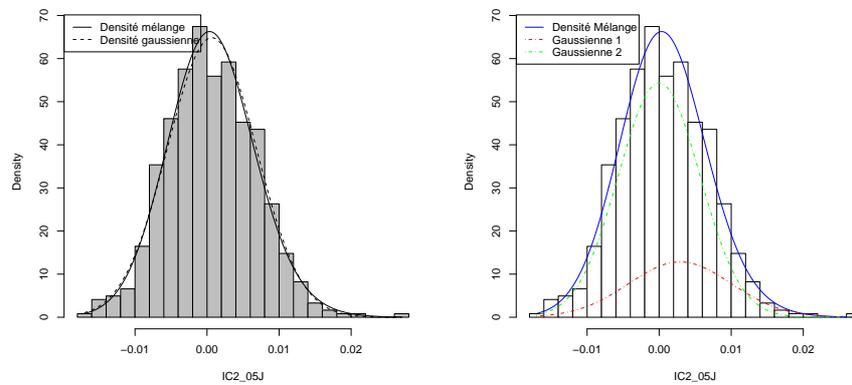


FIGURE 5.3 – Modélisation des densités du rendement source IC2.5J

Composante n ° 3 - IC3_05

Une modélisation basée sur deux densités gaussiennes semble être la meilleure et, comme pour la composante IC2.05, la modélisation par mélange gaussien ne semble pas, à première vue, apporter de valeur ajoutée par rapport à la densité gaussienne empirique.

IC3.05J	$\mathcal{N}_{3,1,5}$	$\mathcal{N}_{3,2,5}$
γ	0,63656	0,36344
μ	-0,002454	-0,001293
σ	0,0056569	0,006245

TABLEAU 5.3 – Paramètres du mélange gaussien de la composante IC3.5J

5.2.2 Composantes sources à 1 et 10 jours ${}_1S$ et ${}_{10}S$

Nous avons appliqué l'algorithme aux signaux sources à 1 et 10 jours et avons obtenu les résultats illustrés par les figures (5.5, 5.6), et résumés dans les tableaux (5.4) à (5.9).

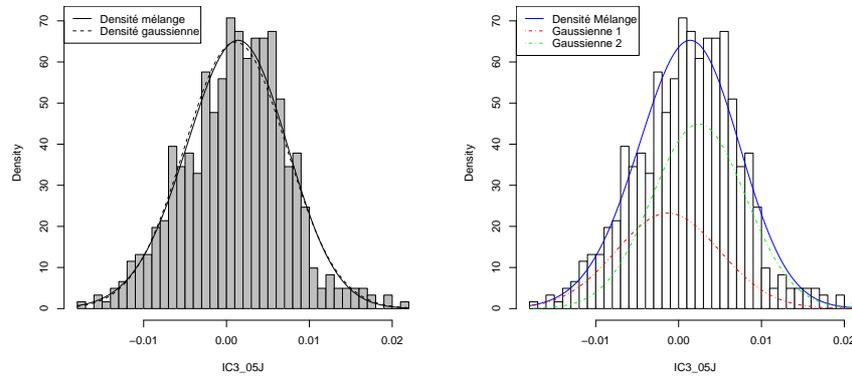


FIGURE 5.4 – Modélisation des densités du rendement source IC3_5J

IC1_01J	$\mathcal{N}_{1,1,1}$	$\mathcal{N}_{1,2,1}$
γ	0,1124	0,8876
μ	-0,002084	0,000696
σ	0,0045826	0,0022361

TABLEAU 5.4 – Paramètres du mélange gaussien de la composante IC1_1J

IC2_01J	$\mathcal{N}_{2,1,1}$	$\mathcal{N}_{2,2,1}$
γ	0,52821	0,47179
μ	-0,000049	0,001
σ	-0,00044	0,003873

TABLEAU 5.5 – Paramètres du mélange gaussien de la composante IC2_1J

IC3_01J	$\mathcal{N}_{3,1,1}$	$\mathcal{N}_{3,2,1}$
γ	0,50563	0,47179
μ	-0,00148	0,000552
σ	0,0022361	0,0031623

TABLEAU 5.6 – Paramètres du mélange gaussien de la composante IC3_1J

5.3 Conclusion

Nous avons modéliser les densités des rendements sources par des mélanges de densités gaussiennes. Toutefois, certains mélanges gaussiens n'apportent pas beaucoup plus d'information que la densité gaussienne empirique. C'est la cas des modélisations des composantes sources IC3_01J, IC3_05J.

Nous allons maintenant exploiter l'hypothèse d'indépendance afin d'obtenir des expressions analytiques de la VaR et de la CvaR. C'est l'objectif du chapitre suivant.

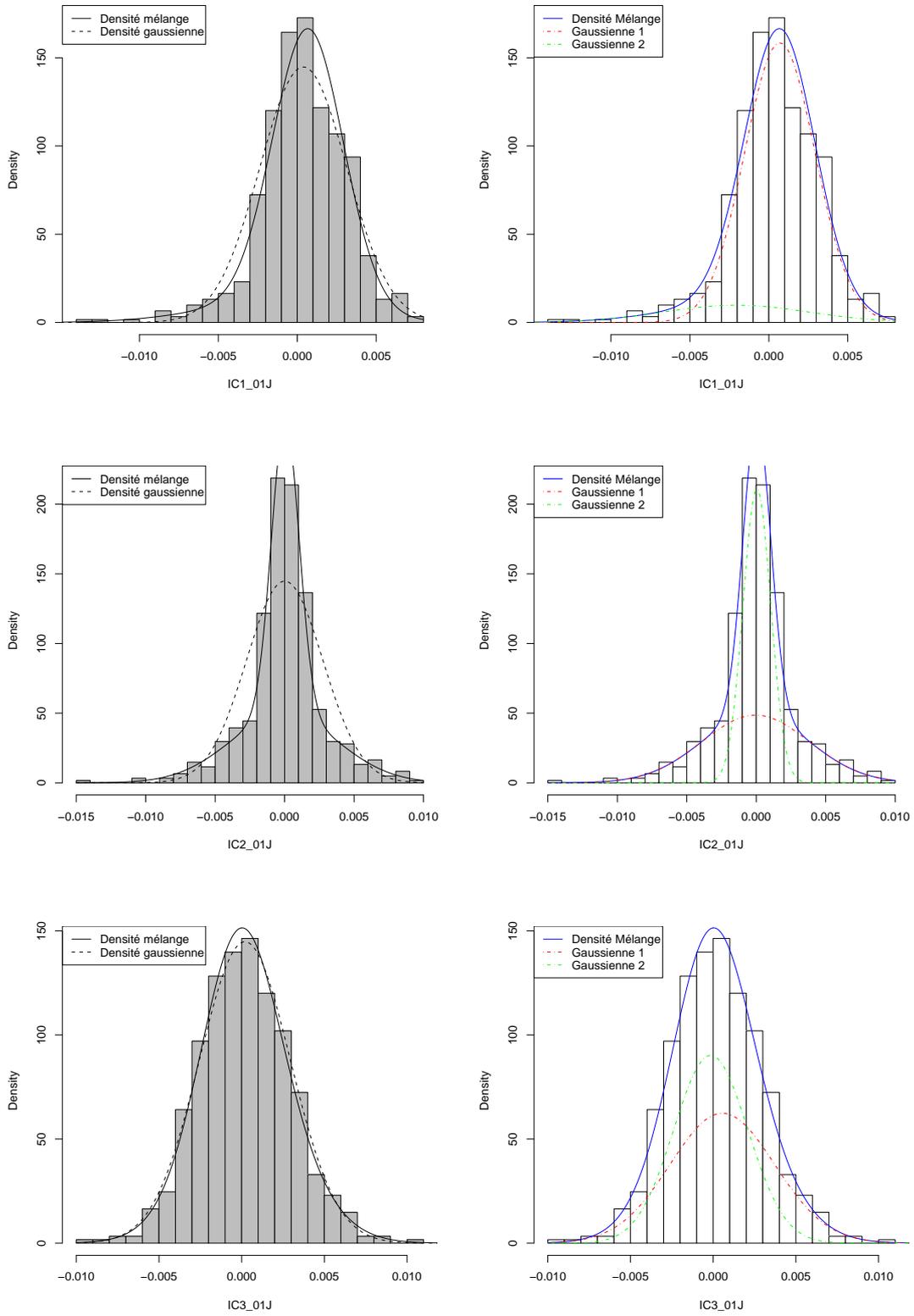


FIGURE 5.5 – Modélisation des densités des rendements sources à 1 jour

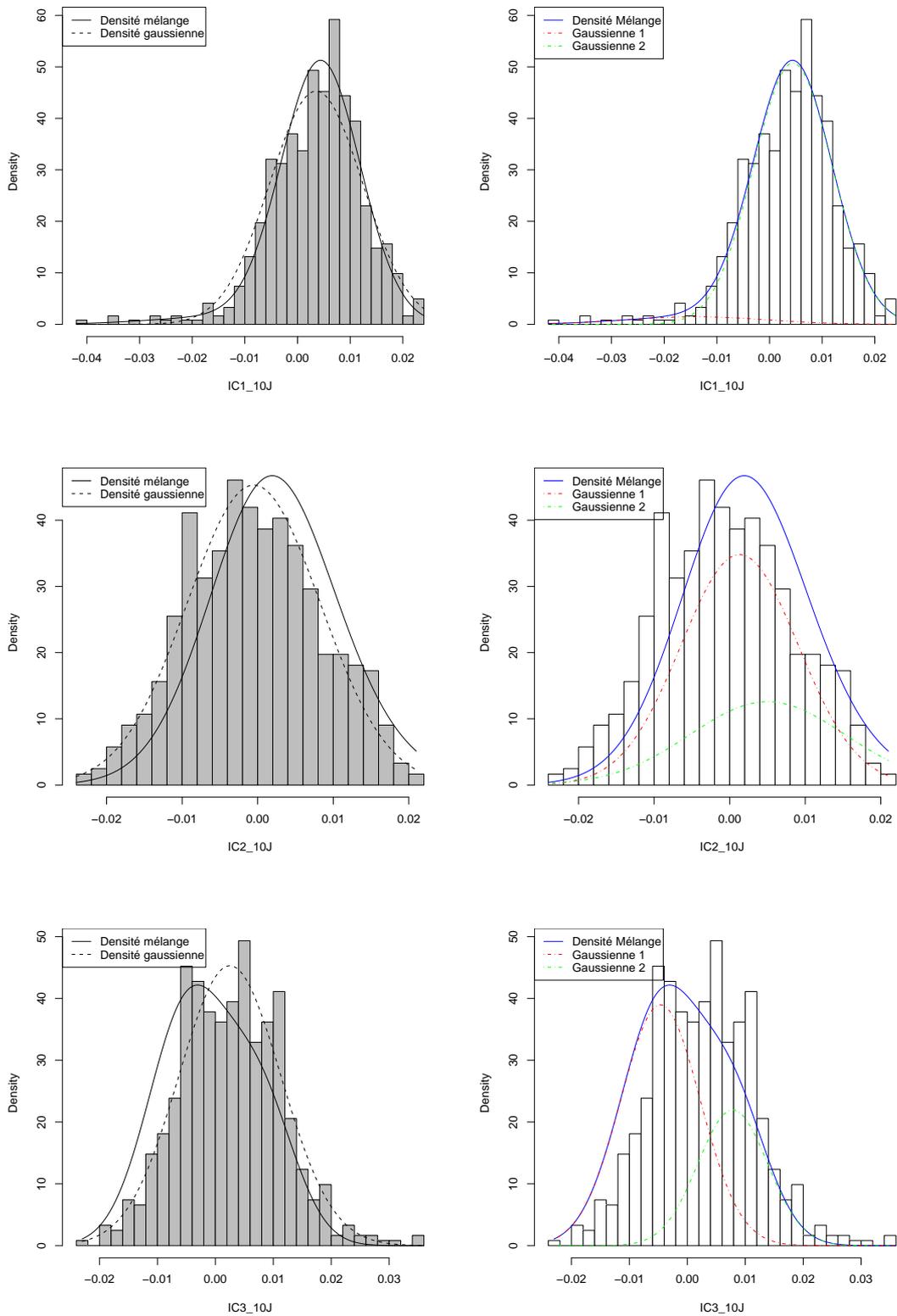


FIGURE 5.6 – Modélisation des densités des rendements sources à 10 jours

IC1_10J	$\aleph_{1,1,10}$	$\aleph_{1,2,10}$
γ	0,95125	0,04875
μ	0,004432	-0,13886
σ	0,0074833	0,013266

TABLEAU 5.7 – Paramètres du mélange gaussien de la composante IC1_10J

IC2_10J	$\aleph_{2,1,10}$	$\aleph_{2,2,10}$
γ	0,32316	0,67684
μ	0,005005	0,001315
σ	-0,010247	0,00776

TABLEAU 5.8 – Paramètres du mélange gaussien de la composante IC2_10J

IC3_10J	$\aleph_{3,1,10}$	$\aleph_{3,2,10}$
γ	0,33017	0,66983
μ	0,007833	-0,004646
σ	0,006	0,0068557

TABLEAU 5.9 – Paramètres du mélange gaussien de la composante IC3_10J

Chapitre 6

VaR, CVaR et Test historique

Nous avons utilisé l'analyse en composantes indépendantes sur les séries des rendements centrés réduits. Ensuite, nous avons modélisé les fonctions de densités de chaque composante indépendante obtenue. L'objectif de ce dernier chapitre est d'établir une formule analytique nous permettant d'exprimer la fonction de densité de notre portefeuille d'actifs.

Il s'agira donc de déterminer l'expression du rendement du portefeuille R_p et le quantile r^{VaR} vérifiant l'équation (2.2). La CvaR s'obtiendra en utilisant l'équation (2.10).

Pour cela, nous avons utilisé l'outil d'optimisation *Solver* inclus dans le tableur Microsoft ExcelTM.

6.1 Densité de rendement du portefeuille

Nous rappelons que notre portefeuille est composé des trois indices Bel20, Dax30 et Cac40. Chaque indice a un poids de 1/3.

6.1.1 Exploitation de l'ACI

Par application de la formule (1.4) vue au chapitre 1, nous avons :

$$R_P = \sum_{i=1}^3 \omega_i R_i \text{ avec } \omega_i = 1/3 \forall i \quad (6.1)$$

Puisque nous avons modélisé les rendements centrés réduits R_i^{CR} par une combinaison linéaire de 3 sources indépendantes S_1 , S_2 et S_3 , nous récrivons (6.1) sous la forme :

$$\begin{aligned} R_P &= \sum_{i=1}^3 \omega_i R_i = \sum_{i=1}^3 (\omega_i \bar{R}_i + \omega_i \sigma_i R_i^{CR}) \\ &= \sum_{i=1}^3 (\omega_i \bar{R}_i + \omega_i \sigma_i R_i^{CR}) \\ &= \sum_{i=1}^3 \omega_i \bar{R}_i + \sum_{i=1}^3 \omega_i \sigma_i R_i^{CR} \end{aligned}$$

En exploitant la formule (4.1) vue au chapitre 4, nous avons :

$$\sum_{i=1}^3 \omega_i \sigma_i R_i^{CR} = \sum_{i=1}^3 \omega_i \sigma_i \left(\sum_{j=1}^3 a_{ij} S_j \right)$$

En mettant en facteurs les termes relatifs aux sources indépendantes, l'équation précédente devient :

$$\sum_{i=1}^3 \omega_i \sigma_i R_i^{CR} = \sum_{j=1}^3 S_j \left(\sum_{i=1}^3 \omega_i \sigma_i a_{ij} \right)$$

d'où finalement

$$R_P = \sum_{i=1}^3 \omega_i \bar{R}_i + \sum_{j=1}^3 S_j \left(\sum_{i=1}^3 \omega_i \sigma_i a_{ij} \right) \quad (6.2)$$

6.1.2 Exploitation des mélanges gaussiens

À partir de cette étape, nous allons exploiter l'hypothèse d'indépendance statistique des composantes indépendantes et leur modélisation sous forme de mélange gaussien. En effet, la distribution du rendement de notre portefeuille devient alors égale à la distribution d'une somme pondérée de variables aléatoires dont nous connaissons les densités individuelles.

Rappel du produit de convolution

En exploitant le fait que la distribution d'une somme Z de deux variables aléatoires continues et indépendantes X et Y est égale au produit de convolution¹ des densités

¹Convolution, [36], p.49

x de X et y de Y , nous pouvons obtenir une formule analytique de la densité z :

$$f_z(z) = \int_{-\infty}^{\infty} f_x(x) f_y(z-x) dx = \int_{-\infty}^{\infty} f_y(y) f_x(z-y) dy \quad (6.3)$$

Nous développons (6.3) dans le cas de deux sources indépendantes :

$$\begin{aligned} f_z(z) &= \int_{-\infty}^{\infty} f_{s_1}(s_1) f_{s_2}(z-s_1) ds_1 \\ &= \int_{-\infty}^{\infty} \sum_{k_1=1}^{K_1} \gamma_{k_1} N(s_1; \mu_{k_1}, \sigma_{k_1}) \sum_{k_2=1}^{K_2} \gamma_{k_2} N(z-s_1; \mu_{k_2}, \sigma_{k_2}) ds_1 \\ &= \sum_{k_1=1}^{K_1} \sum_{k_2=1}^{K_2} \gamma_{k_1} \gamma_{k_2} \int_{-\infty}^{\infty} N(s_1; \mu_{k_1}, \sigma_{k_1}) N(z-s_1; \mu_{k_2}, \sigma_{k_2}) ds_1 \end{aligned}$$

où K_1 et K_2 sont les nombres de densités gaussiennes intervenants dans la modélisation des densités des sources S_1 et S_2 .

Or

$$\int_{-\infty}^{\infty} N(s_1; \mu_{k_1}, \sigma_{k_1}) N(z-s_1; \mu_{k_2}, \sigma_{k_2}) ds_1$$

est la distribution d'une somme de variables aléatoires gaussiennes indépendantes. Dans ce cas, nous savons que la densité de la somme est égale à une densité gaussienne dont :

- la moyenne est égale à la somme des moyennes,
- la variance est égale à la somme des variances.

On peut alors écrire la densité f_z sous la forme :

$$f_Z(z) = \sum_{k_1=1}^{K_1} \sum_{k_2=1}^{K_2} \gamma_{k_1} \gamma_{k_2} N\left(\mu_{k_1} + \mu_{k_2}, (\sigma_{k_1}^2 + \sigma_{k_2}^2)^{1/2}\right) \quad (6.4)$$

Cette dernière équation peut se généraliser à une somme pondérée de variables aléatoires gaussiennes :

- la moyenne est égale à la somme pondérée des moyennes $\mu_{\alpha x + \beta y} = \alpha \mu_x + \beta \mu_y$
- la variance est égale à la somme pondérée au carré des variances $\sigma_{\alpha x + \beta y}^2 = \alpha^2 \sigma_x^2 + \beta^2 \sigma_y^2$

Application au rendement du portefeuille

À partir du résultat précédent et en appliquant les équations précédentes, nous obtenons :

$$R_P = \sum_{i=1}^3 \omega_i \bar{R}_i + \sum_{j=1}^3 S_j \left(\sum_{i=1}^3 \omega_i \sigma_i a_{ij} \right)$$

$$= \sum_{i=1}^3 \omega_i \bar{R}_i + \sum_{j=1}^3 v_j s_j(t) \text{ avec } v_j = \sum_{i=1}^3 \omega_i a_{ij}$$

$$R_P = \sum_{i=1}^3 \omega_i \bar{R}_i + \sum_{k_1=1}^2 \sum_{k_2=1}^2 \sum_{k_3=1}^2 (\gamma_{k_1} \gamma_{k_2} \gamma_{k_3}) \aleph \left(v_1 \mu_{k_1} + v_2 \mu_{k_2} + v_3 \mu_{k_3}; \left(v_1^2 \sigma_{k_1}^2 + v_2^2 \sigma_{k_2}^2 + v_3^2 \sigma_{k_3}^2 \right) \right) \tag{6.5}$$

Nous rappelons que dans le cas des rendements à 5 jours, nous avons obtenu les mélanges gaussiens suivants :

IC1.05J	$\aleph_{1,1,5}$	$\aleph_{1,2,5}$
γ	0,057017	0,94298
μ	-0,007984	0,002341
σ	0,0097468	0,0052915

TABLEAU 6.1 – Rappel des paramètres du mélange gaussien de la composante IC1.5J

IC2.05J	$\aleph_{2,1,5}$	$\aleph_{2,2,5}$
γ	0,77038	0,22962
μ	-0,000001	0,002893
σ	0,0056569	0,0071414

TABLEAU 6.2 – Rappel des paramètres du mélange gaussien de la composante IC2.5J

IC3.05J	$\aleph_{3,1,5}$	$\aleph_{3,2,5}$
γ	0,63656	0,36344
μ	-0,002454	-0,001293
σ	0,0056569	0,006245

TABLEAU 6.3 – Rappel des paramètres du mélange gaussien de la composante IC3.5J

À partir de la modélisation des densités des sources indépendantes, nous avons pu obtenir une expression analytique de la distribution du rendement du portefeuille P que nous appliquons numériquement au cours des sections suivantes.

6.2 Applications numériques

Nous rappelons que notre portefeuille est composé des trois indices Bel20, Dax30 et Cac40. Chaque indice a un poids de 1/3.

6.2.1 Méthode historique

VaR

À partir des rendements de chaque indice, calculés sur la période allant du 4 janvier 2004 au 12 mai 2006, nous avons tracé l'histogramme des rendements observés pour des horizons de 1, 5 et 10 jours (figure 6.1).

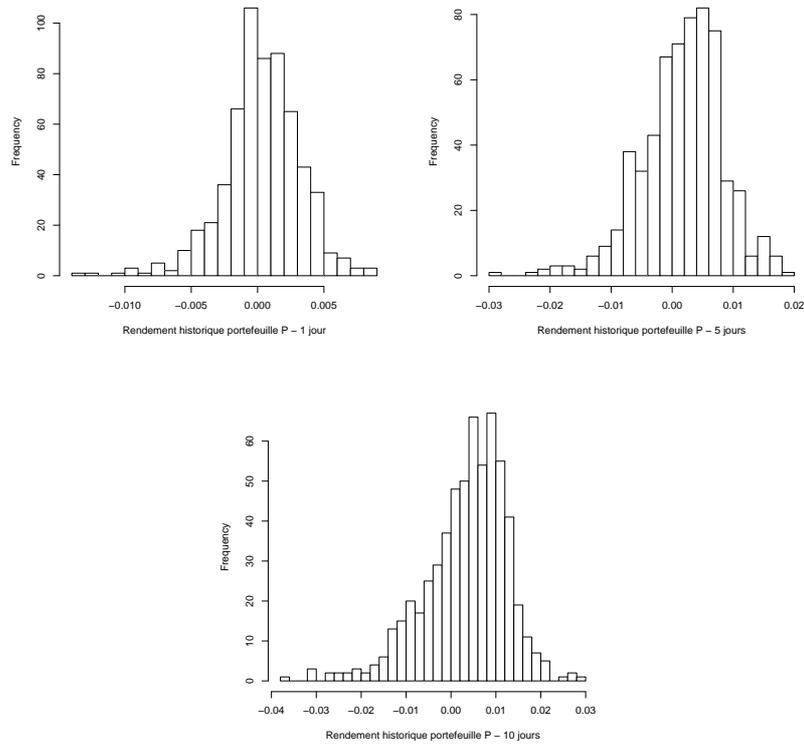


FIGURE 6.1 – Rendements historiques observés à 1,5 et 10 jours

Les quantiles empiriques (pour des niveaux de confiance de 95 %, 99 % et 99,9 %) associés à ces histogrammes sont résumés dans le tableau (6.4) et indiquent les rendements négatifs potentiels minimaux attendus selon la méthode historique.

Niveau de probabilité	α	1 jour	5 jours	10 jours
95 %	5 %	-0,451 %	-0,928 %	-1,273 %
99 %	1 %	-0,889 %	-1,813 %	-2,544 %
99,9 %	0,1 %	-1,294 %	-2,541 %	-3,364 %

TABLEAU 6.4 – VaR - méthode historique

Niveau de probabilité	α	1 jour	5 jours	10 jours
95 %	5 %	-0,662 %	-1,442 %	-2,030 %
99 %	1 %	-1,100 %	-2,327 %	-3,301 %
99,9 %	0,1 %	-1,505 %	-3,055 %	-4,121 %

TABLEAU 6.5 – CvaR - méthode historique

CvaR

À partir de ces VaR, nous pouvons également calculé les VaR conditionnelle qui sont les moyennes des rendements lorsque ces rendements sont inférieurs à la VaR. Ceci nous donne le tableau (6.5), pour un niveau de confiance donné.

6.2.2 Méthode variance-covariance

Rappel de la densité multinormale

On rappelle que cette méthode suppose implicitement que le rendement de chaque actif est distribué suivant une loi normale. Selon cette hypothèse, la fonction de densité du rendement du portefeuille suit une distribution multinormale qui est une généralisation à plusieurs dimensions d'une densité gaussienne. On note X le vecteur multivarié à d dimensions, $X = (X_1, \dots, X_d)' \rightsquigarrow \mathfrak{N}(\mu, \Sigma)$ et sa fonction de densité f s'écrit :

$$f(x) = \frac{1}{\sqrt{(2\pi)^d |\Sigma|}} \exp\left(-\frac{1}{2} (x - \mu)' \Sigma^{-1} (x - \mu)\right) \quad (6.6)$$

pour $x = (x_1, \dots, x_d)' \in \mathfrak{R}^d$, X de moyenne $\mu = (\mu_1, \dots, \mu_d)' \in \mathfrak{R}^d$ et de matrice de variance-covariances $\Sigma = \text{cov}(X)$.

Calcul de la VaR et CvaR

À partir de moyennes et variances empiriques, nous avons utilisé les matrices des corrélations vues au chapitre trois (figures 3.8, 3.9 et 3.10), nous avons calculé les valeurs de VaR et CvaR indiquées aux tableaux (6.6, 6.7).

Niveau de probabilité	α	1 jour	5 jours	10 jours
95 %	5 %	-0,647 %	-1,224 %	-1,648 %
99 %	1 %	-1,546 %	-2,423 %	-2,978 %
99,9 %	0,1 %	-1,853 %	-3,451 %	-4,016 %

TABLEAU 6.6 – VaR - méthode multinormale

Niveau de probabilité	α	1 jour	5 jours	10 jours
95 %	5 %	-0,941 %	-1,842 %	-2,214 %
99 %	1 %	-1,854 %	-2,751 %	-3,754 %
99,9 %	0,1 %	-2,121 %	-3,711 %	-4,213 %

TABLEAU 6.7 – CvaR - méthode multinormale

6.2.3 Méthode ACI et mélange gaussien

Nous avons repris l'expression (6.5) pour nos rendements à 1, 5 et 10 jours. Nous avons ensuite utilisé le solveur de Microsoft ExcelTM afin de déterminer les seuils de VaR et la CvaR selon les différents niveaux de probabilité. Nous avons obtenu les valeurs de VaR et CvaR des tableaux (6.8, 6.9).

Niveau de probabilité	α	1 jour	5 jours	10 jours
95 %	5 %	-1,225 %	-1,386 %	-3,767 %
99 %	1 %	-2,748 %	-2,673 %	-4,684 %
99,9 %	0,1 %	-3,251 %	-3,561 %	-4,756 %

TABLEAU 6.8 – VaR - méthode ACI-Mélange gaussien

Niveau de probabilité	α	1 jour	5 jours	10 jours
95 %	5 %	-1,355 %	-1,767 %	-4,045 %
99 %	1 %	-2,998 %	-2,861 %	-4,921 %
99,9 %	0,1 %	-3,846 %	-3,516 %	-5,210 %

TABLEAU 6.9 – CvaR - méthode ACI-Mélange gaussien

Pour des horizons de 1 et 10 jours, les valeurs de VaR et CvaR obtenues sont plus élevées que les valeurs issues de la méthode multinormale. Par contre, pour un horizon de 1 jour, les estimations sont relativement proches. Une première explication pourrait se trouver dans la modélisation des densités. Les figures (5.2, 5.3, 5.4) montrent que les mélanges sont très proches des densités normales empiriques. En revanche, les estimations obtenues à 1 et 10 jours nous semblent cohérentes avec notre volonté de prendre en compte les queues de distribution.

6.3 Tests historiques

Il est naturel de se demander quelle est la performance des méthodes utilisées. Nous nous sommes alors livrés à l'exercice suivant. Sur la période allant du 2 janvier 2004 au 15 mai 2006. Nous rappelons que les estimations des paramètres des modèles se basent également sur cette même période. Nous avons tiré au hasard des échantillons de 1, 5 et 10 rendements quotidiens. Ceci nous a permis de calculer des rendements simulés pour des horizons de 1, 5 et 10 jours. À partir de ces rendements aléatoires, nous avons calculé les rendements à 1, 5 et 10 jours. Puis nous avons dénombré les cas où les rendements simulés étaient inférieurs aux rendements VaR et CvaR indiqués dans les tableaux (6.4, 6.5, 6.6, 6.7, 6.8 et 6.9). Ces tests ont été réalisés en effectuant 10 000 tirages aléatoires et semblent montrer que la méthode basée sur l'ACI et les mélanges gaussiens soit un peu plus robuste que les méthodes historique et analytique.

Taux d'échec	1 jour	5 jours	10 jours
VaR - Méthode Historique	26,51 %	31,44 %	33,45 %
CvaR - Méthode Historique	34,45 %	37,21 %	28,21 %
VaR - Méthode multinormale	17,01 %	24,45 %	29,45 %
CvaR - Méthode Multinormale	16,21 %	19,21 %	26,21 %
VaR - Méthode ICA-Mélange	9,045 %	23,38 %	22,45 %
CvaR - Méthode ICA-Mélange	11,921 %	20,92 %	20,91 %

TABLEAU 6.10 – Taux d'échec - Rendements simulés à 1, 5 et 10 jours

Conclusion

Notre travail a consisté à nous intéresser aux mesures du risque de perte que sont la Value at Risk et la Value at Risk conditionnelle. Pour cela, nous avons pris l'exemple d'un portefeuille d'actifs composé des indices boursiers Bel20, Cac40 et Dax30. La problématique de ce type de mesure provient de la nature non-gaussienne des distributions de rendement et des dépendances entre ces rentabilités. Nous avons alors utilisé une technique issue de la science du traitement du signal qui permet, sous certaines hypothèses de trouver une transformation linéaire permettant d'obtenir des variables aléatoires statistiquement indépendantes, appelées signaux sources. Ceci nous a permis de contourner le problème délicat de la modélisation des dépendances.

Nous avons ensuite modélisé les distributions empiriques par des mélanges de densité afin de modéliser les queues de distributions observées.

Puis, à l'aide de l'hypothèse d'indépendance et des mélanges gaussiens, nous avons pu expliciter des formules analytiques d'expression de la VaR et de la CVaR d'un portefeuille.

Enfin, sur la base d'un historique de 608 observations, nous avons calculé les valeurs numériques de la VaR et de la CvaR pour des horizons à 1, 5 et 10 jours. Les premières analyses indiquent que selon la modélisation des densités mélanges, la méthode ICA et mélanges gaussiens soit plus robuste en présence de queues de distribution épaisses. Il serait alors intéressant de pouvoir la comparer avec une modélisation basée sur la théorie des copules qui s'attache à modéliser les liens de dépendances entre variables aléatoires.

Bibliographie

- [1] ACHARD S., Mesures de dépendance pour la séparation aveugle de sources, application aux mélanges post non-linéaires, Thèse de doctorat, Université Joseph Fourier Grenoble II, décembre 2003
- [2] ACERBI C., TASCHE D., Expected Shortfall : a natural coherent alternative to Value at Risk, mai 2001
- [3] AFTALION F., Statistiques & économétrie, Banque & Marchés, 25, novembre-décembre 1996
- [4] AFTALION F., Les rentabilités des actifs financiers, Banque & Marchés, 46, mai-juin 2000
- [5] ARTZNER P., DELBAEN F., EBER J.M., HEATH D., Coherent Measures of Risk, Mathematical Finance, 9, 1999
- [6] AIRES F., CHÉDIN A., NADAL J.-P., Analyse de séries temporelles dynamiques et théorie de l'information : l'analyse en composantes indépendantes, Revue Académie des Sciences, 9, 1999
- [7] BACK A., WEIGEND A.S., Discovering structure in finance using independent component analysis, Computational Finance, 1997
- [8] BACK A., WEIGEND A.S., A first application of independent component analysis to extract structure from stock returns, International Journal of Neural Systems, 8, 1997
- [9] BAWA V.S., Optimal rules for ordering uncertain prospects, Journal of Financial Economics, 2, 1975
- [10] BIOJOUT M., LEHMANN E., Application de la VaR pour le calcul de fonds propres en Assurance Vie, Mémoire d'actuariat, ENSAE, 2000
- [11] BOULIER J.-F., BRABANT A., DALAUD R., DIEU A.-L., Risque de marché : vue de profil, Direction de la Recherche et de l'Innovation, Crédit Commercial de France, 28, décembre 1997

- [12] CAPPÉ O., Modèles de mélange et modèles de Markov cachés pour le traitement automatique de la parole, 2000
- [13] CARDOSO J.-F., Source separation using higher-order moments, Proc. Internat. Conf. Acoust. Speech Signal Process., Glasgow, 1989
- [14] CARDOSO J.-F., On the performance of source separation algorithms, Proc. EUSIPCO9, Edinburgh, 1994
- [15] CHEN Y., HÄRDLE W., SPOKOINY V., Portfolio Value at Risk based on Independent Components Analysis, SFB Discussion Paper, 2005
- [16] CHIN E., WEIGEND A.S., ZIMMERMANN H., Computing portfolio risk using Gaussian mixtures and Independent Component Analysis, 1999
- [17] COMON P., Analyse en composantes indépendantes, Working paper
- [18] COMON P., Independent component analysis, a new concept ?, Signal Processing, 36, 1994
- [19] DACUNHA-CASTELLE D., GASSIAT E., The estimation of the order of a mixture model, Prépublication de l'équipe d'Analyse et de Probabilités, 1997
- [20] DARMOIS G., Analyse des liaisons de probabilités, In Proceedings Int. Stat. Conferences 1947, volume III A, Washington (D.C.), 1951
- [21] DURIEU C., KIEFFER M., Analyse en composantes indépendantes pour la séparation aveugle de sources, Working paper
- [22] ESCH L., KIEFFER R., LOPEZ T., Value at Risk - vers un risk management moderne, Édition de Boeck Université, 1997
- [23] FISHBURN P.C., Mean-risk analysis with risk associated with below-target returns, American Economic Review, 67, mars 1977
- [24] GORDON A., CVaR as a Measure of Risk : Implications for Portfolio Selection, février 2003
- [25] HYVARINEN, A., OJA E., Independent Component Analysis : a tutorial, Helsinki University of Technology, 1999
- [26] HYVARINEN, A., OJA E., Independent Component Analysis : algorithms and applications, Neural Networks, 13, 1999
- [27] HYVARINEN, A., KARHUNEN J., OJA E. Independent Component Analysis, Wiley Edition, 2001
- [28] LACOUME J.-L., AMBLARD P.-O., COMON P., Statistiques d'ordre supérieur pour le traitement du signal, Édition Masson, 1997

- [29] LAMBERT E., Étude des techniques d'optimisation et de différents réseaux de neurones appliquées à la séparation de signaux en traitement du signal, Rapport de DEA, Université de Mulhouse, 1999
- [30] LONGIN F., Value at Risk : une nouvelle approche fondée sur les valeurs extrêmes, Annales d'économie et de statistique , 52, 1998
- [31] MALAMANE F., Synthèse et identification aveugle de signaux pour la prospection électromagnétique, Mémoire, Université de Montréal, juillet 1996
- [32] MANSOUR A., Contribution à la séparation aveugle de sources, Thèse de doctorat, Institut National Polytechnique de Grenoble, janvier 1997
- [33] MARKOWITZ H., Portfolio selection, Journal of Finance, 7, 77-91, mars 1952
- [34] PERNOUD F., Asset Liability Management une approche Value at Risk, Mémoire d'actuariat, ISFA, décembre 1998
- [35] KWIATKOWSKI D., PHILIPPS P., SCHMIDT P., SHIN Y., Testing the Null Hypothesis of stationnarity Against the Alternative of a unit Root, Journal of Econometrics, 54, 1992.
- [36] SAPORTA G., Probabilité, analyse des données et statistique, Édition Technip
- [37] TOPSACALIAN P., Les indices boursiers, Banque & Marchés, 47, juillet-août 2000
- [38] TOKPAVI S., La Value-at-risk comme mesure du risque de marché : cas de la catastrophe financière du Comté d'Orange, mars 2004
- [39] URYASEV S., Conditional Value-at-Risk, methodology and applications : overview, juin 2004

Annexes

Annexe 1. Histogrammes des rendements à 1 et 10 jours

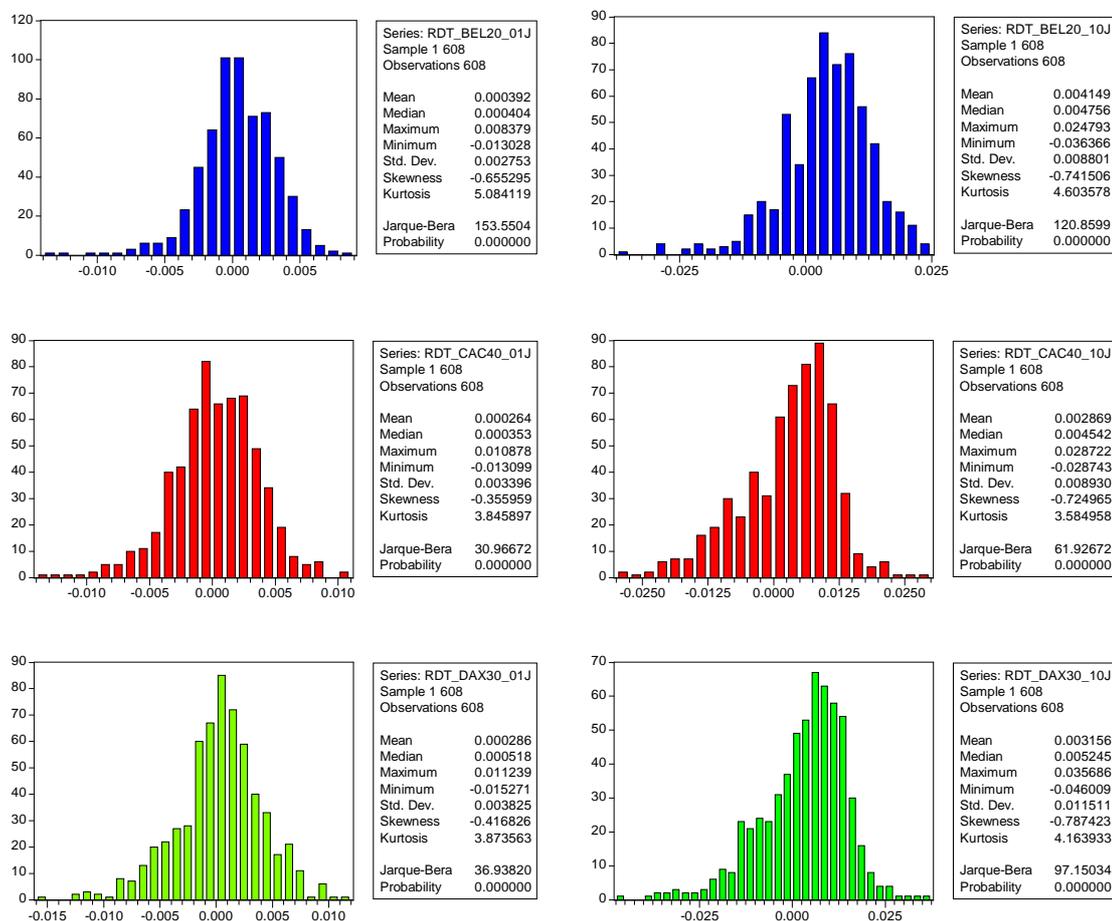


FIGURE 2 – Histogramme des rendements à 1 et 10 jours

Annexe 2. Séries des rendements à 1 et 10 jours

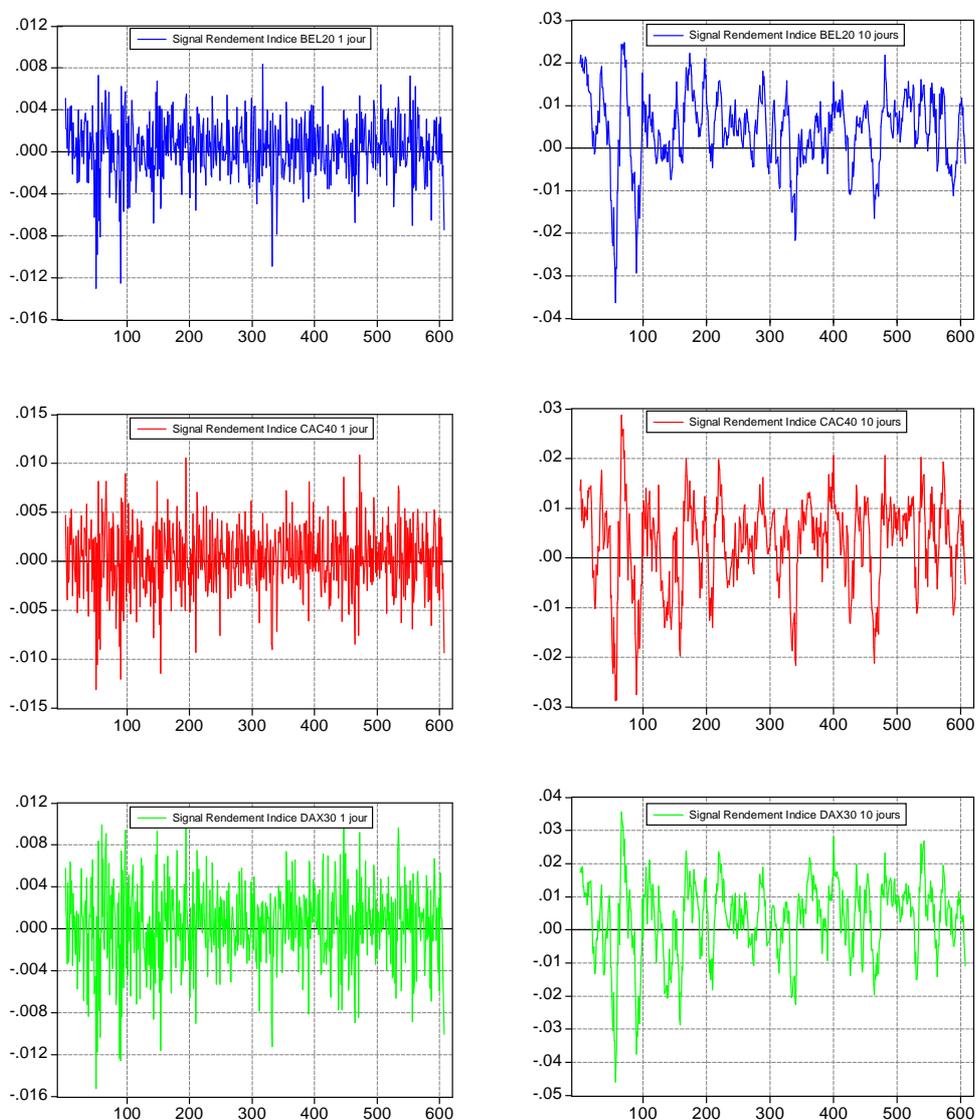


FIGURE 3 – Séries des rendements à 1 et 10 jours

Annexe 3. Tests de stationnarité à 1 et 10 jours

Null Hypothesis: RDT_BEL20_01J has a unit root
Exogenous: Constant, Linear Trend
Bandwidth: 1 (Newey-West using Bartlett kernel)

	Adj. t-Stat	Prob.*
Phillips-Perron test statistic	-24.71083	0.0000
Test critical values:		
1% level	-3.973245	
5% level	-3.417239	
10% level	-3.131010	

*Mackinnon (1996) one-sided p-values.

FIGURE 4 – Test PP des rendements à 1 jour de l'indice BEL20

Null Hypothesis: RDT_BEL20_01J is stationary
Exogenous: Constant, Linear Trend
Bandwidth: 1 (Newey-West using Bartlett kernel)

	LM-Stat.
Kwiatkowski-Phillips-Schmidt-Shin test statistic	0.043909
Asymptotic critical values*:	
1% level	0.216000
5% level	0.146000
10% level	0.119000

*Kwiatkowski-Phillips-Schmidt-Shin (1992, Table 1)

FIGURE 5 – Test KPSS des rendements à 1 jour de l'indice BEL20

Null Hypothesis: RDT_BEL20_10J has a unit root
Exogenous: Constant, Linear Trend
Bandwidth: 9 (Newey-West using Bartlett kernel)

	Adj. t-Stat	Prob.*
Phillips-Perron test statistic	-6.550958	0.0000
Test critical values:		
1% level	-3.973245	
5% level	-3.417239	
10% level	-3.131010	

*Mackinnon (1996) one-sided p-values.

FIGURE 6 – Test PP des rendements à 10 jours de l'indice BEL20

Null Hypothesis: RDT_BEL20_10J is stationary
 Exogenous: Constant, Linear Trend
 Bandwidth: 17 (Newey-West using Bartlett kernel)

	LM-Stat.
Kwiatkowski-Phillips-Schmidt-Shin test statistic	0.060457
Asymptotic critical values*:	
1% level	0.216000
5% level	0.146000
10% level	0.119000

*Kwiatkowski-Phillips-Schmidt-Shin (1992, Table 1)

FIGURE 7 – Test KPSS des rendements à 10 jours de l'indice BEL20

Null Hypothesis: RDT_CAC40_01J has a unit root
 Exogenous: Constant, Linear Trend
 Bandwidth: 12 (Newey-West using Bartlett kernel)

	Adj. t-Stat	Prob.*
Phillips-Perron test statistic	-27.53818	0.0000
Test critical values:		
1% level	-3.973245	
5% level	-3.417239	
10% level	-3.131010	

*MacKinnon (1996) one-sided p-values.

FIGURE 8 – Test PP des rendements à 1 jour de l'indice CAC40

Null Hypothesis: RDT_CAC40_01J is stationary
 Exogenous: Constant, Linear Trend
 Bandwidth: 14 (Newey-West using Bartlett kernel)

	LM-Stat.
Kwiatkowski-Phillips-Schmidt-Shin test statistic	0.026543
Asymptotic critical values*:	
1% level	0.216000
5% level	0.146000
10% level	0.119000

*Kwiatkowski-Phillips-Schmidt-Shin (1992, Table 1)

FIGURE 9 – Test KPSS des rendements à 1 jour de l'indice CAC40

Null Hypothesis: RDT_CAC40_05J has a unit root
 Exogenous: Constant, Linear Trend
 Bandwidth: 39 (Newey-West using Bartlett kernel)

	Adj. t-Stat	Prob.*
Phillips-Perron test statistic	-6.975186	0.0000
Test critical values:		
1% level	-3.973245	
5% level	-3.417239	
10% level	-3.131010	

*MacKinnon (1996) one-sided p-values.

FIGURE 10 – Test PP des rendements à 5 jours de l'indice CAC40

Null Hypothesis: RDT_CAC40_05J is stationary
 Exogenous: Constant, Linear Trend
 Bandwidth: 10 (Newey-West using Bartlett kernel)

	LM-Stat.
Kwiatkowski-Phillips-Schmidt-Shin test statistic	0.024733
Asymptotic critical values*:	
1% level	0.216000
5% level	0.146000
10% level	0.119000

*Kwiatkowski-Phillips-Schmidt-Shin (1992, Table 1)

FIGURE 11 – Test KPSS des rendements à 5 jours de l'indice CAC40

Null Hypothesis: RDT_CAC40_10J has a unit root
 Exogenous: Constant, Linear Trend
 Bandwidth: 6 (Newey-West using Bartlett kernel)

	Adj. t-Stat	Prob.*
Phillips-Perron test statistic	-7.343350	0.0000
Test critical values:		
1% level	-3.973245	
5% level	-3.417239	
10% level	-3.131010	

*MacKinnon (1996) one-sided p-values.

FIGURE 12 – Test PP des rendements à 10 jours de l'indice CAC40

Null Hypothesis: RDT_CAC40_10J is stationary
 Exogenous: Constant, Linear Trend
 Bandwidth: 16 (Newey-West using Bartlett kernel)

	LM-Stat.
Kwiatkowski-Phillips-Schmidt-Shin test statistic	0.031177
Asymptotic critical values*:	
1% level	0.216000
5% level	0.146000
10% level	0.119000

*Kwiatkowski-Phillips-Schmidt-Shin (1992, Table 1)

FIGURE 13 – Test KPSS des rendements à 10 jours de l'indice CAC40

Null Hypothesis: RDT_DAX30_01J has a unit root
 Exogenous: Constant, Linear Trend
 Bandwidth: 11 (Newey-West using Bartlett kernel)

	Adj. t-Stat	Prob.*
Phillips-Perron test statistic	-25.09412	0.0000
Test critical values:		
1% level	-3.973245	
5% level	-3.417239	
10% level	-3.131010	

*MacKinnon (1996) one-sided p-values.

FIGURE 14 – Test PP des rendements à 1 jour de l'indice DAX30

Null Hypothesis: RDT_DAX30_01J is stationary
 Exogenous: Constant, Linear Trend
 Bandwidth: 11 (Newey-West using Bartlett kernel)

	LM-Stat.
Kwiatkowski-Phillips-Schmidt-Shin test statistic	0.027864
Asymptotic critical values*:	
1% level	0.216000
5% level	0.146000
10% level	0.119000

*Kwiatkowski-Phillips-Schmidt-Shin (1992, Table 1)

FIGURE 15 – Test KPSS des rendements à 1 jour de l'indice DAX30

Null Hypothesis: RDT_DAX30_05J has a unit root
 Exogenous: Constant, Linear Trend
 Bandwidth: 36 (Newey-West using Bartlett kernel)

	Adj. t-Stat	Prob.*
Phillips-Perron test statistic	-5.959049	0.0000
Test critical values:		
1% level	-3.973245	
5% level	-3.417239	
10% level	-3.131010	

*MacKinnon (1996) one-sided p-values.

FIGURE 16 – Test PP des rendements à 5 jours de l'indice DAX30

Null Hypothesis: RDT_DAX30_05J is stationary
 Exogenous: Constant, Linear Trend
 Bandwidth: 11 (Newey-West using Bartlett kernel)

	LM-Stat.
Kwiatkowski-Phillips-Schmidt-Shin test statistic	0.022906
Asymptotic critical values*:	
1% level	0.216000
5% level	0.146000
10% level	0.119000

*Kwiatkowski-Phillips-Schmidt-Shin (1992, Table 1)

FIGURE 17 – Test KPSS des rendements à 5 jours de l'indice DAX30

Null Hypothesis: RDT_DAX30_10J has a unit root
 Exogenous: Constant, Linear Trend
 Bandwidth: 5 (Newey-West using Bartlett kernel)

	Adj. t-Stat	Prob.*
Phillips-Perron test statistic	-6.835692	0.0000
Test critical values:		
1% level	-3.973245	
5% level	-3.417239	
10% level	-3.131010	

*MacKinnon (1996) one-sided p-values.

FIGURE 18 – Test PP des rendements à 10 jours de l'indice DAX30

Null Hypothesis: RDT_DAX30_10J is stationary
 Exogenous: Constant, Linear Trend
 Bandwidth: 16 (Newey-West using Bartlett kernel)

	LM-Stat.
Kwiatkowski-Phillips-Schmidt-Shin test statistic	0.027687
Asymptotic critical values*:	
1% level	0.216000
5% level	0.146000
10% level	0.119000

*Kwiatkowski-Phillips-Schmidt-Shin (1992, Table 1)

FIGURE 19 – Test KPSS des rendements à 10 jours de l'indice DAX30

Annexe 4. Signaux sources à 1 et 10 jours

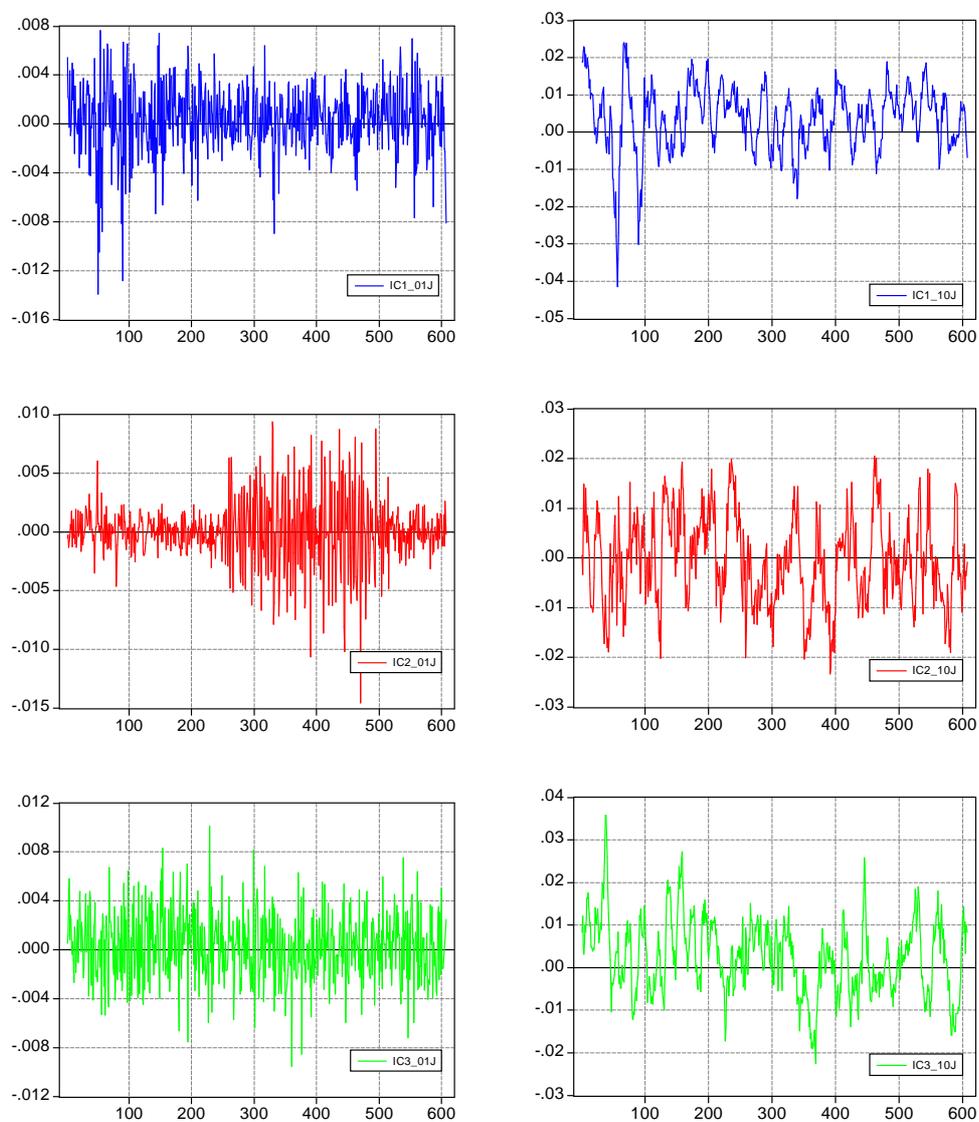


FIGURE 20 – Séries des signaux sources à 1 et 10 jours

Annexe 5. Histogrammes des sources à 1 et 10 jours

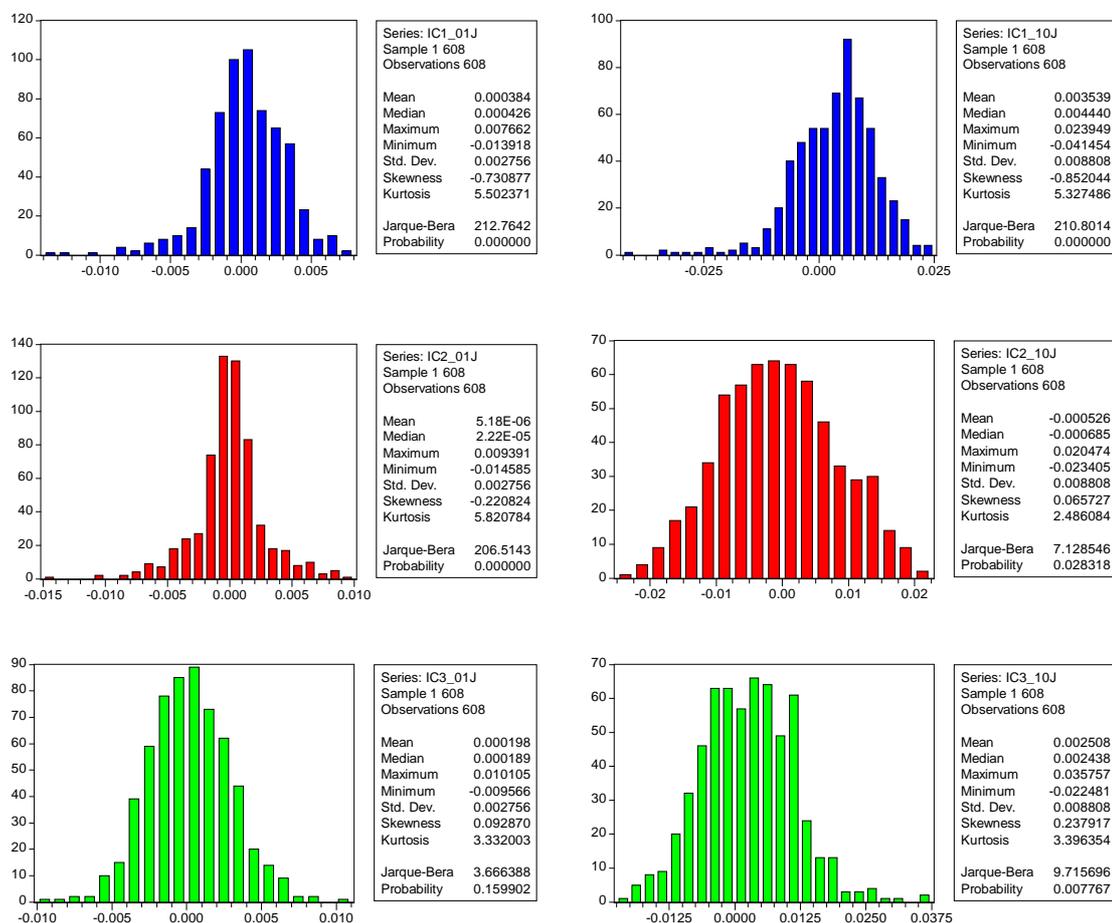


FIGURE 21 – Histogrammes des signaux sources à 1 et 10 jours