



Optimisation du calcul du capital économique d'une compagnie d'assurance vie par la méthode des Simulations dans les Simulations

Anne-Claire Martial et Jean-Baptiste Garnier

ECOLE NATIONALE DE LA STATISTIQUE ET DE L'ADMINISTRATION ECONOMIQUE – ParisTech
MASTER MODELISATION ALEATOIRE (M2MO) – UNIVERSITE PARIS 7

Promotion 2013

Directeur de mémoire : Laurent Devineau (Milliman France)

Résumé

Dans le cadre de la nouvelle directive Solvabilité II, les assureurs pourront calculer leur capital économique suivant la Formule Standard ou suivant un modèle interne. Pour une assurance vie, la méthode donnant le plus d'informations sur la distribution des fonds propres à un an est la méthode des Simulations dans les Simulations (SdS). Cependant, sa mise en œuvre est limitée par le grand nombre de simulations nécessaires et les optionalités potentielles des contrats en portefeuille.

Nous proposons deux méthodes alternatives aux SdS, réduisant le budget de calcul nécessaire à l'obtention du capital économique. La première est une méthode probabiliste utilisant la décomposition d'un événement rare en événements moins rares, et la seconde se fonde sur les forêts aléatoires, qui sont une méthode d'apprentissage statistique.

Nous montrons que ces deux méthodes permettent d'obtenir une estimation du capital économique avec un budget de calcul moindre que celui des SdS. Aussi, nous proposons plusieurs variantes de ces méthodes, qui sont susceptibles de réduire davantage le budget de calcul. Enfin nous vérifions les résultats en utilisant ces méthodes d'abord dans un cadre simplifié, puis dans un cadre réel.

Ainsi, nous avons proposé deux nouvelles méthodes permettant de calculer efficacement le capital économique avec une mise en œuvre simple, ce qui pourrait faciliter la mise en place d'un modèle interne chez les assureurs vie.

Abstract

With the Solvency II framework, insurers will calculate economic capital using either the Standard Formula or an internal model. The method which yields the maximum information about the distribution of capital in a year of a life insurance portfolio is the Nested Simulations (NS) method. However, its implementation is severely limited by the large number of simulations needed and by the potential optionality of the products in the portfolio.

We propose two alternative methods, which reduce the computing budget, to obtain the economic capital. The first is a probabilistic method using the multilevel splitting method, and the second is based on random forests which are a statistical learning method.

We show that these two methods provide an estimate of the economic capital with a smaller computing budget than the NS method needs. Also, we propose several variants of these approaches, which are likely to further reduce the computing budget. Finally we check the results using these methods, first, in a simplified framework, then in a real setting.

Thus, we proposed two new methods to efficiently calculate the economic capital with a simple implementation, which could facilitate the establishment of an internal model in life insurance.

Remerciements

Tout d'abord, nous tenons à remercier nos encadrants :

M. Laurent Devineau, directeur du département de Recherche et Développement du cabinet de conseil en actuariat Milliman, pour nous avoir proposé ce sujet passionnant, ainsi que pour sa confiance et son accompagnement,

M. Julien Vedani, consultant de l'équipe de Recherche et Développement de Milliman et élève doctorant à Lyon 1, pour son investissement et ses conseils précieux,

M. Matthieu Chauvigny, ancien consultant de l'équipe de Recherche et Développement de Milliman et chargé d'étude actuarielle chez Axa France, pour sa disponibilité et ses conseils avisés.

Nous remercions aussi tous ceux qui nous ont aidés dans notre travail :

M. Arnaud Guyader, maître de conférences à l'Université Rennes 2, membre des équipes ASPI de l'INRIA et Processus Stochastiques de l'IRMAR, pour ses réponses précises à nos questions,

M. Arnak Dalalyan, professeur à l'ENSAE ParisTech, à l'Ecole des Ponts ParisTech, à l'ENS Cachan, et à l'Ecole Centrale Paris, pour avoir pris le temps de répondre à nos diverses interrogations.

Enfin, nous remercions **tous les collaborateurs du cabinet Milliman**, pour leur accueil et leur soutien.

Sommaire

Remerciements	4
Introduction.....	8
I. Cadre réglementaire.....	9
A. Evolution du cadre réglementaire.....	9
1. Solvabilité I	9
a) Historique de la directive Solvabilité I.....	9
b) Limites de Solvabilité I.....	10
2. Solvabilité II	11
a) Présentation générale de Solvabilité II.....	11
b) Zoom sur le pilier 1.....	12
c) Zoom sur le pilier 2.....	13
d) Zoom sur le pilier 3.....	13
B. Calcul du capital économique en général et pour une compagnie d'assurance-vie.....	13
a) La formule standard : une approche modulaire.....	14
b) Modèle interne.....	15
c) Description des portefeuilles d'assurance vie.....	16
II. Etat des lieux sur le calcul du capital économique pour une compagnie d'assurance vie	17
A. Modèle ALM	17
1. Scénarios primaires et secondaires.....	17
2. Gestion actif/passif.....	18
B. Méthode des Simulations dans les Simulations et quelques alternatives	19
1. La méthode des Simulations dans les Simulations.....	19
2. Zoom sur des méthodes alternatives proches	21
a) Réduction du nombre de simulations primaires : l'accélérateur SdS	21
b) Réduction du nombre de simulations secondaires : l'allocation optimale de scénarios..	22
3. Les méthodes utilisant un proxy	24
a) Méthodes paramétriques.....	24
b) Replicating Portfolio	26
c) Kriging stochastique	28
4. Récapitulatif des résultats et de la complexité des différentes méthodes.....	29
III. Deux nouvelles méthodes : la décomposition d'un événement rare en événements moins rares et l'estimation par forêts aléatoires.....	31
A. Décomposition d'un événement rare en événements moins rares.....	31

1.	Principe de la méthode	31
a)	Principe général de la méthode	31
b)	Littérature sur le sujet	32
c)	Zoom sur l'algorithme dont nous nous sommes inspirés	32
d)	Adaptation à notre problème.....	33
2.	Présentation d'un premier algorithme simple	35
a)	Inputs.....	35
b)	Explication de la règle d'arrêt des itérations verticales	35
c)	Algorithme.....	36
3.	Difficultés liées à l'implémentation pratique de l'algorithme : diffusion des scénarios selon une loi conditionnelle.....	37
a)	Algorithme de Metropolis Hastings.....	37
b)	Diffusion dans le cas gaussien	38
c)	Adaptation de l'algorithme de décomposition d'un événement rare	39
4.	Autre adaptation de l'algorithme.....	40
a)	Présentation des formes paramétriques locales.....	40
b)	Adaptation de l'algorithme grâce à l'utilisation de formes paramétriques locales.....	41
c)	Calcul de la complexité de l'algorithme	42
5.	Propriétés asymptotiques de l'estimateur du quantile	43
a)	Détermination de la loi asymptotique suivie par l'estimateur du quantile	43
b)	Intervalle de confiance asymptotique de l'estimateur du quantile	45
c)	Gain par rapport à un Monte Carlo classique	46
B.	Forêts aléatoires.....	47
1.	Présentation des arbres de décision	47
a)	Définition.....	47
b)	Construction d'un arbre de décision	49
c)	Limites et variantes	52
2.	Présentation des forêts aléatoires	52
a)	Préliminaires : le Bagging	52
b)	Les forêts aléatoires	54
3.	Premier algorithme : application directe de la méthode.....	55
4.	Optimisation de la base de calibrage	56
a)	Première méthode d'ajout de points de calibrage.....	56
b)	Deuxième méthode d'ajout de points de calibrage	57
5.	Optimisation du calcul des fonds propres de la zone quantile	57

6.	Calcul de la complexité.....	58
IV.	MISE EN OEUVRE ET RESULTATS	59
A.	Application dans le cas de modèles de gestion actif/passif simplifiés.....	59
1.	Description des formes paramétriques.....	59
2.	Premiers résultats obtenus	60
a)	Décomposition d'un événement rare en événements moins rares.....	61
b)	Forêts aléatoires.....	63
3.	Test de la robustesse des résultats	68
a)	Décomposition d'un événement rare en événements moins rares.....	68
b)	Forêts aléatoires.....	70
4.	Analyse de la sensibilité aux paramètres	71
a)	Décomposition d'un événement rare en événements moins rares.....	71
b)	Forêts aléatoires.....	75
5.	Analyse de la sensibilité par rapport à la forme paramétrique	80
a)	Décomposition d'un événement rare en événements moins rares.....	80
b)	Forêts aléatoires.....	82
6.	Comparaison des méthodes.....	85
B.	Application dans le cadre d'un modèle de gestion actif/passif réel	87
1.	Décomposition d'un événement rare en événements moins rares.....	87
2.	Forêts aléatoires.....	88
a)	Base de calibrage.....	89
b)	Etude des résultats obtenus.....	90
	Conclusion	93
	Bibliographie.....	95
V.	Annexes	98
A.	Annexe 1 : détermination de σ dans la diffusion des scénarios à l'aide d'un noyau de transition gaussien	98
B.	Annexe 2 : détermination des limites de $NM \ln 0,5 \% + MN$ et de MN	99
C.	Annexe 3 : détermination de la loi asymptotique d'un estimateur de Monte Carlo	100

Introduction

La directive Solvabilité II entrera en vigueur dans les prochaines années. Elle impose aux compagnies d'assurance de posséder un montant de fonds propres minimal, appelé capital économique, pour faire face à leurs obligations sans être insolvable. Plus précisément, le capital économique correspond aux réserves qui doivent être mises de côté pour réduire la probabilité de ruine à un an à 0,5 %. Nous nous intéressons, dans ce mémoire, à l'estimation du capital économique pour une compagnie d'assurance vie.

Pour calculer le capital économique, les compagnies d'assurance ont le choix entre l'utilisation de la formule standard proposée par la directive Solvabilité II ou la mise en place d'un modèle interne. Nous nous concentrons ici uniquement sur cette dernière approche. La méthode des Simulations dans les Simulations apparaît comme la méthode de détermination du capital économique la plus conforme à la définition donnée par Solvabilité II. Cependant, elle est très complexe d'un point de vue opérationnel, ce qui compromet fortement sa mise en œuvre. Dès lors, il est nécessaire de trouver des méthodes alternatives pour calculer le capital économique. Plusieurs ont déjà été proposées, mises en œuvre, et analysées par les acteurs du marché, mais les résultats obtenus dépendent de la complexité du portefeuille d'assurance vie.

L'un des enjeux actuels majeurs des compagnies d'assurance consiste donc à déterminer de nouvelles façons de calculer efficacement et rapidement le capital économique. Aussi, l'objectif de notre mémoire est de chercher de telles méthodes, de les mettre en œuvre, et de les comparer à celles déjà existantes. Nous nous intéressons dans la suite à deux techniques, déjà appliquées à d'autres secteurs que l'assurance, et que nous adaptons à notre problématique. La première méthode repose sur la décomposition d'un événement rare en événements moins rares, et a essentiellement été mise en œuvre pour calculer des probabilités faibles dans le domaine de la physique, tandis que la seconde est fondée sur les méthodes d'apprentissage statistique, qui se sont largement développées ces dernières années, en particulier en biostatistique et en génétique.

Ce mémoire est divisé en quatre parties. La première rappelle le cadre réglementaire dans lequel s'inscrit notre problématique. Nous y étudions l'évolution des réglementations auxquelles sont soumises les compagnies d'assurance, puis nous nous intéressons plus particulièrement à la définition du capital économique.

A partir de la deuxième partie, nous nous concentrons sur le calcul du capital économique dans le cadre d'une assurance vie. Après avoir mené une analyse de la méthode des Simulations dans les Simulations et mis en évidence ses limites, nous présentons différents algorithmes alternatifs.

Dans une troisième partie nous détaillons nos deux méthodes : la décomposition d'un événement rare en événements moins rares et les forêts aléatoires. Nous donnons d'abord leur principe général, puis nous expliquons comment nous les avons adaptées au calcul du capital économique dans le cadre d'une assurance vie.

Enfin, après avoir mené une analyse théorique des deux méthodes, nous les appliquons à des portefeuilles d'épargne. Nous commençons par les mettre en œuvre dans le cas de modèles de gestion actif/passif simplifiés, puis nous étudions une application pour un modèle de gestion réel. Cette dernière partie nous permet ainsi de comparer les deux méthodes entre elles, et avec celles déjà existantes.

I. Cadre réglementaire

Notre mémoire s'intéresse aux techniques permettant de déterminer la valeur du capital économique. Néanmoins, il est important de comprendre d'où vient la nécessité de calculer ce capital économique, c'est-à-dire de comprendre le contexte réglementaire. Aussi, nous allons, dans une première partie, présenter les deux directives relatives à la solvabilité des compagnies d'assurance, afin d'introduire, par la suite, la notion de capital économique.

A. Evolution du cadre réglementaire

Nous pouvons définir la solvabilité comme la capacité d'une entité à payer ses dettes. L'inversion du cycle de production des compagnies d'assurance rend cette notion très importante. En effet, un assureur reçoit d'abord des primes avant de rembourser des sinistres potentiels qui arriveront plus tard. Ainsi, pour une compagnie d'assurance, la solvabilité traduit la capacité de l'assureur à respecter les engagements qu'il a pris envers ses clients. Deux séries de directives la définissent : Solvabilité I et Solvabilité II. Nous commencerons par nous intéresser aux règles de solvabilité actuelles, puis, après avoir montré leurs limites, nous présenterons celles qui rentreront en vigueur prochainement, et qui sont définies par Solvabilité II.

1. Solvabilité I

Nous allons, dans cette partie, rappeler brièvement les caractéristiques principales de la directive Solvabilité I, puis nous mettrons en évidence les limites qui ont nécessité la création de Solvabilité II.

a) Historique de la directive Solvabilité I

Les premières directives en matière de solvabilité des compagnies d'assurance datent de 1973 pour l'assurance non-vie et de 1979 pour l'assurance vie. Ainsi, la réglementation Solvabilité I, à laquelle sont actuellement soumis les assureurs de l'Union Européenne et qui restera en vigueur au moins jusqu'au 1^{er} janvier 2014, est constituée de ces directives et d'une révision faite en 2002 pour la branche vie.

L'objectif de Solvabilité I est de définir un cadre prudentiel commun à tous les assureurs de l'Union Européenne, tout en laissant la liberté à chaque Etat de définir des règles plus strictes, afin de protéger les clients contre les risques d'insolvabilité de leur assureur.

Solvabilité I repose essentiellement sur trois piliers :

- Le caractère suffisant des provisions techniques (article R331) : les provisions techniques correspondent au montant permettant à un assureur d'honorer ses engagements vis-à-vis de ses assurés. Les principales sont les provisions pour sinistres à payer, dans le cas de l'assurance non-vie, et les provisions mathématiques, pour l'assurance vie. Ainsi, les compagnies d'assurance doivent savoir identifier les risques auxquelles elles sont soumises et mettre de côté des provisions au moins égales à l'espérance des sinistres.

- La qualité des actifs (article R332-R333) : comme l'assureur reçoit des primes avant de devoir payer des sinistres, il possède des sommes à placer sur les marchés. Ainsi, pour être sûr de pouvoir faire face à ses engagements, il est nécessaire que les placements effectués soient de bonne qualité, c'est-à-dire sûrs, liquides et rentables. Les placements sont donc soumis à certaines règles comme :
 - imposer qu'ils soient dans la même devise que les prestations (avec une tolérance de 20 %),
 - imposer qu'ils soient diversifiés entre différents secteurs et différentes classes d'actifs,
 - imposer qu'ils soient dispersés géographiquement.

- La marge de solvabilité (article R334) : elle est définie comme le capital que doivent posséder les compagnies d'assurance pour faire face à toutes dépenses supplémentaires dues à des événements inattendus. Le calcul de la marge de solvabilité diffère suivant la branche, vie ou non-vie, choisie :
 - En assurance vie, la marge de solvabilité est calculée à partir d'un pourcentage fixé des provisions mathématiques.
 - En assurance non-vie, deux calculs sont nécessaires pour déterminer la marge de solvabilité. Le premier s'appuie sur le montant annuel des primes tandis que le second est basé sur la charge annuelle des sinistres. La marge de solvabilité est définie comme étant le maximum des deux valeurs.

Notons que quelle que soit la branche (vie ou non-vie) choisie, la marge de solvabilité dépend également des dispositifs de réassurance.

b) Limites de Solvabilité I

A cause de sa simplicité, en particulier pour calculer la marge de solvabilité, la directive Solvabilité I présente certaines faiblesses.

Tout d'abord, nous pouvons observer deux cas aberrants, pour lesquels un assureur devra mettre plus de fonds propres de côté sans que cela ne soit justifié par une augmentation du risque auquel il est soumis. En effet, des provisions mathématiques trop prudentes en assurance vie et une augmentation des tarifs des primes d'assurance non-vie entraînent une augmentation de la marge de solvabilité.

Par ailleurs, la méthode de calcul de la marge de solvabilité ne distingue pas les différents risques. Par exemple, la marge de solvabilité associée aux actions est la même, quels que soient le type d'action et sa liquidité.

De plus, la formule permettant de calculer la marge de solvabilité se base uniquement sur le passé. Elle fait ainsi l'hypothèse que le passé permet d'estimer correctement le futur, ce qui lui donne un côté rétrospectif et statique.

Nous pouvons également noter que cette formule ne prend pas en compte tous les risques techniques, mais uniquement le risque de souscription. Ainsi, certains risques tels que le risque opérationnel ne rentrent pas dans le calcul de la marge de solvabilité.

Enfin, les niveaux de prudence requis dans le calcul des provisions techniques ne sont pas les mêmes pour tous les pays européens, ce qui peut entraîner des avantages concurrentiels entre différents pays. Par conséquent, une harmonisation européenne semble indispensable.

2. Solvabilité II

Comme nous l'avons vu, les exigences actuelles de marge de solvabilité dans les pays membres de l'Union Européenne sont régies par les directives Solvabilité I datant de 1973 et 1979 et de la révision faite en 2002. Pour pallier aux limites de Solvabilité I, abordées dans la partie précédente (harmonisation européenne, encourager les assureurs à mieux gérer leurs risques, prendre en compte certains risques négligés dans Solvabilité I...), un ensemble d'organismes se sont concertés pour définir la nouvelle directive de solvabilité des assurances : Solvabilité II. L'objet de cette partie est de présenter cette directive.

a) *Présentation générale de Solvabilité II*

La volonté d'augmenter l'intégration et l'harmonisation des normes prudentielles européennes fait suite aux crises de la fin des années 1990 et du début des années 2000. L'objectif est alors :

- d'améliorer la solvabilité des assureurs, pour garantir la sécurité des assurés,
- de continuer l'harmonisation européenne des normes actuarielles pour éviter les distorsions de concurrence entre les acteurs de différents pays.

La directive Solvabilité II est ainsi née de la concertation des organismes et personnes suivantes :

- la Commission européenne,
- le Comité Européen des Assureurs (CEA),
- le CEIOPS (Committee of European Insurances and Occupational Pensions Supervisors),
- des professionnels et associations du secteur (compagnies d'assurance, fédération professionnelles, actuaires).

Afin d'étudier l'impact de différents changements réglementaires le CEIOPS a demandé aux acteurs du marché de répondre à des vagues de consultations appelées QIS (Quantitative Impact Studies). Ces études quantitatives permettent d'estimer l'impact probable que pourraient avoir certaines modifications du cadre réglementaire sur le bilan des assureurs.

- La première étude effectuée concerne les méthodes de calcul des provisions techniques. Alors que des méthodes déterministes étaient proposées dans Solvabilité I, le CEIOPS a ici demandé aux assurances d'utiliser des méthodes stochastiques de provisionnement.
- La deuxième étude portait, elle aussi, sur la comparaison de plusieurs indicateurs de solvabilité et de calcul du SCR, dans le but de prendre en compte la diversification des risques.
- La troisième étude traitait du calibrage du SCR et du MCR ainsi que de l'agrégation des risques dans le cas des groupes d'assurances.
- La quatrième étude avait pour but d'affiner le calibrage du calcul du SCR par la formule standard.

- La cinquième étude concernait l'impact de Solvabilité II sur le bilan des assureurs et sur leur solvabilité.

Ces cinq consultations ont ainsi orienté le régulateur dans la construction de Solvabilité II. Cette nouvelle norme réglementaire conserve la structure à trois piliers de Solvabilité I. Ainsi, le premier pilier définit les règles quantitatives à suivre pour évaluer les provisions ainsi que le capital économique. Le deuxième pilier donne des directives concernant la gestion qualitative du risque, et sa gestion quantitative à moyen et long terme. Le troisième pilier porte sur la transparence des assurances et les informations qui doivent être transmises au marché et aux actionnaires.

Nous présentons ci-dessous un schéma récapitulatif des trois piliers constituant Solvabilité II :

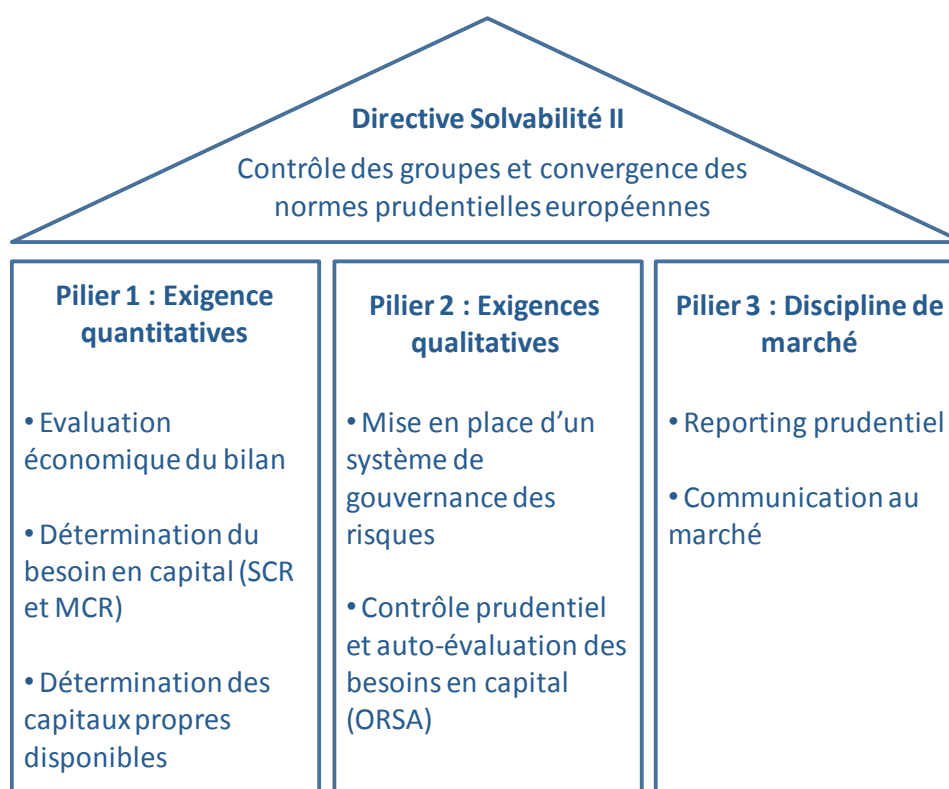


Figure 1 : Présentation des 3 piliers de Solvabilité II

Nous allons à présent nous intéresser de façon plus précise à chacun de ces trois piliers.

b) Zoom sur le pilier 1

Il peut arriver que les estimations faites ne soient pas exactes, et que l'assureur doive rembourser un montant supérieur à ce qu'il avait prévu. Pour éviter de devenir insolvable, les compagnies d'assurance provisionnent des fonds supplémentaires que nous appelons fonds propres.

Le pilier 1 définit la manière de calculer les provisions techniques et les fonds propres nécessaires aux compagnies.

Alors que dans Solvabilité I les provisions devaient être "suffisantes", il s'agit, dans Solvabilité II, de valoriser les actifs et les passifs des compagnies en valeur économique (méthode dite "Fair Value")

des normes IFRS). Ainsi les actifs sont évalués en valeur de marché, et les passifs le sont à leur "juste valeur". Nous parlons alors, dans le cas de passif, de valorisation *Best Estimate* (BE).

Le BE correspond donc à la valeur actualisée de l'espérance des flux de trésoreries futurs. Le taux d'actualisation utilisé est le taux "sans risque" du marché à la date de valorisation.

Le régulateur impose aux assureurs que leurs fonds propres soient supérieurs à deux montants nommés *Solvency Capital Requirement* (SCR) et *Minimum Capital Requirement* (MCR). Le maximum de ces deux montants est appelé le capital économique. Le MCR étant très simple à déterminer, nous allons nous concentrer sur le calcul du SCR. Par abus de langage, nous appellerons le SCR, capital économique.

Le MCR est la limite en dessous de laquelle le régulateur considère que le risque qu'une compagnie d'assurance ne puisse faire face à ses engagements est trop grand.

Le SCR est déterminé de telle sorte qu'il permette à l'assureur de remplir ses obligations à l'horizon d'une année avec un niveau de confiance de 99,5 %. Notons que l'ensemble des risques auxquels est soumis un assureur est pris en compte dans le calcul du SCR.

La relation considérée en générale pour déterminer le SCR est la suivante :

$$SCR(0) = FP(0) - Def(1) * q_{0.5\%}(FP(1)),$$

avec :

- $SCR(0)$: le SCR calculé à l'année 0,
- $FP(0)$: les fonds propres de la compagnie à l'année 0, qui sont, par définition, connus,
- $Def(1)$: le déflateur entre l'année 0 et l'année 1, calculé à la date 0,
- $q_{0.05\%}(FP(1))$: le quantile à 0,5 % des fonds propres à un an.

c) *Zoom sur le pilier 2*

Le deuxième pilier porte sur l'harmonisation européenne des méthodes de contrôle interne des risques. En particulier, les assureurs sont encouragés à améliorer la gouvernance de leurs risques de long terme grâce, notamment, au dispositif *Own Risk Solvency Assessment* (ORSA). Le régulateur pourra imposer aux assureurs une augmentation de leur marge de solvabilité (ajout d'un *add-in*) en fonction de la qualité de leur processus d'estimation et de gestion des risques.

d) *Zoom sur le pilier 3*

Le troisième pilier traite de la discipline de marché. Il impose aux assureurs une certaine transparence de leur méthode de calcul et de gestion de risque. En particulier, ils devront fournir au régulateur, aux actionnaires, et aux analystes, des informations concernant les performances financières, et les mesures de risque utilisées ainsi que les hypothèses sous-jacentes.

B. Calcul du capital économique en général et pour une compagnie d'assurance-vie

Nous avons vu que le capital économique est un montant placé en fonds propres dans le bilan des assureurs et calculé de telle manière qu'il permette aux compagnies de faire face à leurs obligations à un an, avec un seuil de confiance de 99,5 % .

Le CEIOPS propose deux méthodes de calcul de capital économique :

- la formule standard a été développée par le régulateur et permet un calcul simple et déterministe du capital économique. Cette approche reste suffisamment généraliste pour être applicable à tout portefeuille d'assurance,
- le modèle interne est une méthode personnalisée, et développée par les compagnies (mais nécessitant l'agrément des autorités de contrôle), du calcul du capital économique. Les modèles employés sont alors souvent stochastiques, et prennent en compte les spécificités du portefeuille de chaque assureur.

Remarque : Il est possible de mélanger ces deux approches, nous parlons alors de modèles internes partiels. Ces derniers ne seront pas développés dans la suite de notre mémoire.

Nous allons maintenant présenter ces deux approches, ainsi que leurs avantages et inconvénients.

a) La formule standard : une approche modulaire

Les assurances sont soumises à de très nombreux risques tels que des risques actions, de mortalité, de rachat... La multiplicité de ces risques rend la prise en compte de leur interdépendance très complexe. Le régulateur propose donc, dans la formule standard, un calcul simplifié du capital économique, en plusieurs étapes, et adapté aux profils de risque moyens des assureurs européens. Le capital économique obtenu par la formule standard approxime donc les fonds propres nécessaires à un assureur pour lui permettre de faire face à ses obligations. Il est calculé en appliquant des chocs instantanés à la date initiale.

Dans un premier temps, les assureurs doivent identifier leurs différents risques, appelés risques élémentaires, suivant la segmentation suivante :

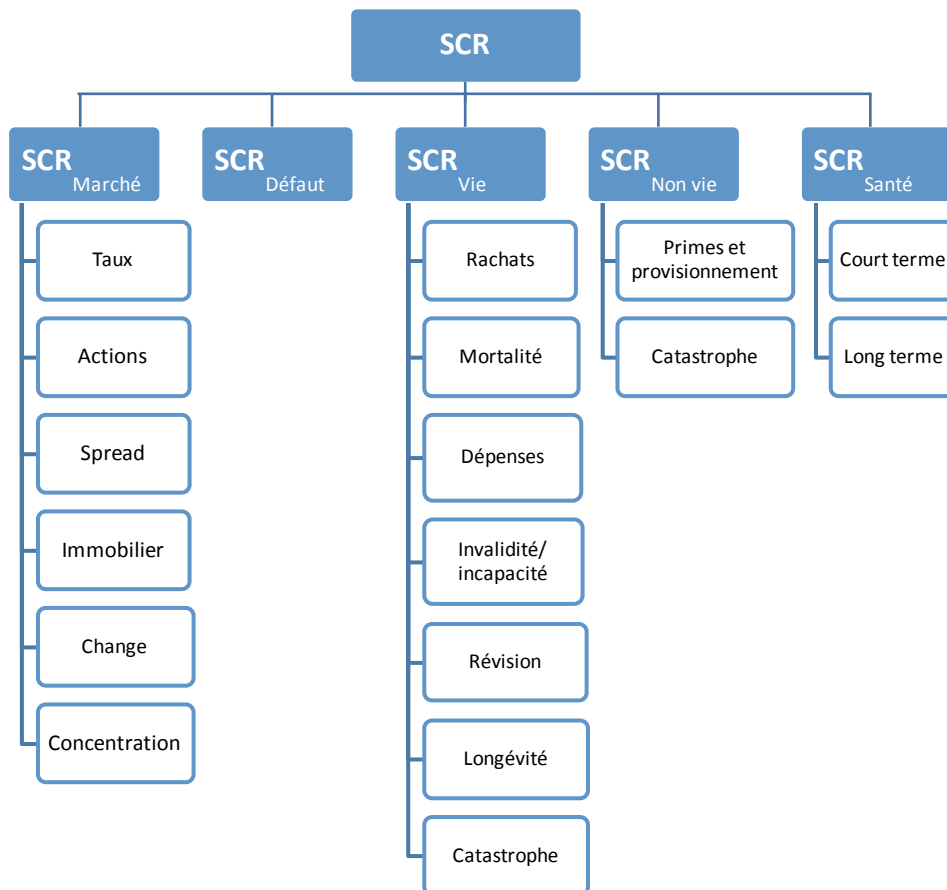


Figure 2 : Diagramme représentant la segmentation des risques de la formule standard pour le calcul du SCR

Puis, pour chaque risque élémentaire (correspondant aux petits rectangles dans le schéma ci-dessus), l'entreprise calcule le besoin en fonds propres suivant les règles dictées par le régulateur. Ces fonds propres sont en suite agrégés (grâce à une matrice de corrélation fournie, elle aussi, par le régulateur) une première fois pour donner les capitaux économiques associés aux risques de marché, santé, vie et non vie. Puis ils sont eux même agrégés pour donner le SCR.

Remarque : Nous avons ici négligé les risques opérationnels et incorporels, ainsi que les ajustements.

b) *Modèle interne*

Dans la directive Solvabilité II, le régulateur propose aux assureurs d'utiliser un modèle interne à la place de la formule standard. Ce modèle interne est alors un modèle d'évaluation du capital économique, construit par l'assureur, en fonction de ses profils de risque, dont le but est d'estimer le capital économique (nous gardons ici la même définition que dans la formule standard). L'inconvénient, pour les compagnies, d'employer une telle approche est qu'elle est très couteuse. En effet, en plus de son élaboration, chaque modèle interne doit être validé par le régulateur avant d'être utilisé. Cependant, cette approche peut être intéressante pour de gros assureurs ayant les moyens techniques et financiers d'entreprendre sa construction, puisqu'elle permet une prise en compte des risques spécifiques à l'assureur.

Un des principaux problèmes techniques soulevés par cette approche est le temps de calcul du capital économique. En effet, l'idéal serait de pouvoir calculer le risque sous-jacent à chaque produit de l'actif et du passif, ainsi que leur dépendance les uns avec les autres. Cependant la taille des portefeuilles (d'actifs et de passifs) des assureurs permet difficilement de réaliser un tel calcul.

Etudions maintenant le modèle interne dans le cas, plus spécifique, d'un portefeuille d'assurance vie.

De nombreux produits d'assurance vie comportent de l'optionnalité sur les actifs (notons que cette dernière peut être cachée). Il s'agit alors d'utiliser un modèle de gestion actif/passif, couplé avec un générateur de scénarios économiques, pour prendre ces optionnalités en compte dans le calcul du capital économique.

c) Description des portefeuilles d'assurance vie

L'assurance vie regroupe l'ensemble des risques dit "de vie" des assurances. Parmi les risques "de vie", nous comptons les risques d'invalidité, de mortalité, et de longévité.

Les encours d'assurance vie représentent près de 1 700¹ milliards d'euros fin 2011 dont une large majorité investie dans les entreprises. A cause de la taille des encours des contrats d'assurance vie, les assureurs sont susceptibles de préférer une approche de type modèle interne, à l'approche formule standard, puisqu'elle pourrait leur permettre de réduire leur besoin en fonds propres réglementaires.

Afin de mieux comprendre les problématiques liées à la présence d'optionnalité, nous allons présenter succinctement le risque de rachat.

Lors d'un emprunt, les particuliers français ont l'obligation de s'assurer contre leur décès. La durée de ces contrats est alors la même que celle de l'emprunt contracté. Cependant il peut arriver qu'après avoir souscrit à un emprunt (et à une assurance décès), l'emprunteur veuille racheter son emprunt, autrement dit, rembourser par anticipation l'intégralité du capital restant dû (cela peut être financé par un nouvel emprunt à un taux plus avantageux par exemple). N'ayant plus besoin d'être assuré, l'emprunteur clôture également son contrat d'assurance décès.

L'assureur ne peut prédire avec exactitude quels sont les contrats qui seront rachetés, ni leur date de rachat, mais il est possible de modéliser ces rachats, en fonction des caractéristiques des contrats et de la situation économique. Les rachats doivent donc être pris en compte dans les modèles de gestion actif/passif.

¹ Source : Rapport annuel FFSA 2012

² Les résultats sont donnés à titre indicatif et peuvent varier suivant la complexité du portefeuille.

II. Etat des lieux sur le calcul du capital économique pour une compagnie d'assurance vie

Nous allons maintenant nous placer uniquement dans le cadre d'une assurance vie. Nous avons vu que l'optionnalité de certains contrats rend nécessaire l'utilisation d'un modèle de gestion actif/passif (ALM). Nous allons donc présenter ce qu'est un modèle de gestion actif/passif, puis nous étudierons la méthode des Simulations dans les Simulations (SdS) ainsi que quelques alternatives.

A. Modèle ALM

Un modèle de gestion actif/passif permet de modéliser les interactions entre l'actif et le passif d'une compagnie d'assurance vie. Ce modèle utilise généralement, en *input*, des tables de scénarios économiques. Nous allons donc décrire, dans un premier temps, les deux types de scénarios utilisés dans la valorisation d'un portefeuille d'assurance vie à un an, puis nous présenterons le fonctionnement d'un modèle de gestion actif/passif.

1. Scénarios primaires et secondaires

Un scénario économique est un ensemble de variables décrivant une situation économique à une date donnée. Dans le cas d'un portefeuille d'assurance vie, les variables présentes dans les scénarios économiques sont souvent financières (telles que des prix d'actions ou d'obligations) mais représentent aussi des risques liés au passif de l'assureur (tels que la mortalité ou la longévité). Ces variables sont le plus souvent issues de modèles de projection d'actifs et de risque de passifs (mortalité par exemple), formant un générateur de scénarios économiques.

Toute l'information financière et technique portée par un scénario peut ainsi être synthétisée en un nombre restreint d'indicateurs, appelés facteurs de risque. Chaque facteur de risque caractérise entièrement l'aléa relatif à un risque donné. Nous pouvons par exemple définir le facteur de risque action comme un accroissement de browniens entre deux dates, si le risque action est modélisé par un modèle de type brownien géométrique. Notons qu'il est possible de considérer que les facteurs de risque sont gaussiens centrés réduits. Nous utiliserons en particulier cette hypothèse dans la méthode de décomposition d'un événement rare en événements moins rares. Nous verrons également dans la présentation des méthodes alternatives aux SdS qu'il est possible d'exprimer les fonds propres à un an en fonction des facteurs de risque.

Il existe deux grandes classes de tables de scénarios économiques : les tables dites risque-neutre (RN), et celles dites monde-réel (MR). Cette séparation est analogue à la différence entre la probabilité historique et risque-neutre en finance quantitative. En effet, une projection RN sert à calculer des prix d'actifs. Ainsi le calibrage de ces générateurs se fait sur des prix réels d'actifs à une date donnée (pour les variables financières), ou sur les prix de contrats d'assurance vie (pour le passif). Les générateurs MR servent, eux, à projeter des risques sous la probabilité historique. Ils se calibrent sur des historiques de prix d'actifs, et sur les réalisations passées des risques du passif.

Le capital économique étant le niveau de fonds propres nécessaire pour éviter la ruine économique à un an avec un seuil de confiance de 99,5 %, nous devons projeter les risques à un an de façon

monde-réel, avant de valoriser les fonds propres de la compagnie, sur chacune de ces simulations, en projetant les risques de façon risque-neutre. Les scénarios issus de la première projection, sont appelés scénarios primaires, les autres sont appelés scénarios secondaires.

Les scénarios secondaires sont donc générés conditionnellement aux scénarios primaires. Ces tables ont un objectif de valorisation, l'horizon de projection doit donc être supérieur ou égal à la durée d'extinction des passifs du portefeuille (en pratique l'horizon de projection est de l'ordre de 30 à 50 ans). Ainsi, à chaque scénario primaire est associée une table de scénarios secondaires.

C'est à l'aide de ces jeux de scénarios économiques primaires et secondaires qu'un modèle de gestion actif/passif est créé.

2. Gestion actif/passif

Pour étudier les modèles de gestion actif/passif, nous allons commencer par décrire le bilan d'une compagnie d'assurance, puis nous verrons comment valoriser chacun des postes.

Le capital économique défini dans la directive Solvabilité II repose sur une vision économique du bilan des compagnies d'assurance. Ce dernier peut se résumer, de manière simplifiée, par le schéma suivant :

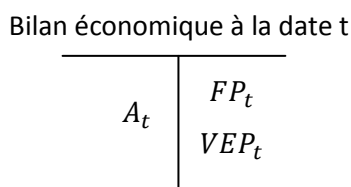


Figure 3 : Schéma simplifié du bilan économique d'une compagnie d'assurance à la date t

Nous observons ainsi que le bilan est constitué de trois éléments : la valeur de marché de l'actif A_t , les fonds propres économiques FP_t , et la valeur économique des passifs VEP_t . Il est alors important de noter que l'actif est calculé en valeur de marché, tandis que les fonds propres et le passif sont déterminés en valeur économique. Par ailleurs, comme le bilan est équilibré, nous obtenons, à toute date t , une relation entre les fonds propres, l'actif et le passif : $FP_t = A_t - VEP_t$.

Nous allons, à présent, nous intéresser plus précisément aux fonds propres. Ces derniers sont composés de deux éléments :

- L'actif net réévalué, noté ANR_t : il correspond à la richesse que la compagnie d'assurance a accumulée jusqu'à la date t .
- La « Value of In Force », notée VIF_t : elle représente les profits futurs qui seront générés par le stock de contrats que possède la compagnie à la date t , en supposant qu'aucun nouveau contrat ne sera vendu dans le futur.

Ainsi, la VIF est définie comme la valeur actuelle des résultats futurs. A la date 0, elle peut donc être calculée comme l'espérance, sous la probabilité risque-neutre Q , des résultats futurs, R_u , actualisés à l'aide du taux d'actualisation, $\delta_{0,u}$. Ce taux d'actualisation correspond au taux sans risque sous la probabilité risque-neutre. Les fonds propres à cette date sont donc donnés par la formule suivante :

$$FP_0 = ANR_0 + VIF_0 = ANR_0 + E_Q[\sum_{u \geq 1} \delta_{0,u} R_u | F_0],$$

où F_0 est la filtration correspondant à l'information financière disponible à la date 0.

De la même manière, nous pouvons calculer la valeur économique des passifs, en 0, à l'aide de l'espérance, sous la probabilité risque-neutre, des cash-flows futurs, P_u , actualisés au taux δ_u :

$$VEP_0 = E_Q \left[\sum_{u \geq 1} \delta_{0,u} P_u | F_0 \right].$$

Nous rappelons que pour calculer le capital économique, nous avons besoin de connaître la valeur des postes du bilan à un an. Pour cela, il est nécessaire que les aléas évoluent en adéquation avec les mouvements observés historiquement. Les projections se font donc sous la probabilité monde-réel sur la première période. Ainsi, en notant F_1^{MR} l'information monde-réel disponible à la fin de la première année, nous obtenons les expressions des fonds propres et de la valeur économique des passifs à un an :

$$FP_1 = ANR_1 + E_Q \left[\sum_{u \geq 1} \frac{\delta_{0,u}}{\delta_{0,1}} R_u | F_1^{MR} \right],$$

$$VEP_1 = E_Q \left[\sum_{u \geq 1} \frac{\delta_{0,u}}{\delta_{0,1}} P_u | F_1^{MR} \right].$$

Nous pouvons observer, dans les formules ci-dessus, que pour obtenir la valeur économique des fonds propres à un an, il est tout d'abord nécessaire d'obtenir la situation à un an en univers MR. Ensuite, à conditionnement monde-réel fixé, il faut effectuer un grand nombre de simulations RN puis calculer l'espérance actualisée des flux issus de ces simulations secondaires. Par conséquent, la valorisation économique des fonds propres nécessite beaucoup de simulations ALM, en particulier si nous voulons connaître la valeur des fonds propres pour plusieurs scénarios MR.

Maintenant que nous avons montré, de manière théorique, comment obtenir la valeur des fonds propres économiques à un an, nous allons expliquer comment en déduire, de manière pratique, le capital économique.

B. Méthode des Simulations dans les Simulations et quelques alternatives

Il existe plusieurs méthodes permettant d'estimer le capital économique défini dans les parties précédentes. Aussi, nous allons tout d'abord nous intéresser à celle qui permet d'obtenir le plus d'informations sur la distribution des fonds propres à un an, puis, après avoir mis en évidence ses limites, nous présenterons des méthodes alternatives déjà étudiées par certains acteurs du marché.

1. La méthode des Simulations dans les Simulations

La méthode des Simulations dans les Simulations (SdS) apparaît comme la façon la plus conforme à la directive Solvabilité II de calculer le capital économique. Elle consiste à effectuer deux séries de simulations imbriquées, d'où son nom de Simulations dans les Simulations.

Le premier niveau de simulations, appelées simulations primaires, repose sur la projection des facteurs de risques sur une période. Il est réalisé en univers MR, afin d'obtenir une évolution réaliste des différents risques sous-jacents sur cette période.

La méthode SdS consiste ensuite à lancer un nouveau jeu de simulations, les simulations secondaires. Celles-ci sont conditionnées par l'information issue des scénarios primaires et sont réalisées en univers RN afin d'avoir une évaluation du « bilan économique ». En calculant, pour chaque scénario primaire, la moyenne empirique des simulations secondaires associées, nous obtenons une estimation par méthode de Monte-Carlo de la valeur économique des bilans en fin de première période.

Nous présentons ci-dessous une illustration de cette méthode :

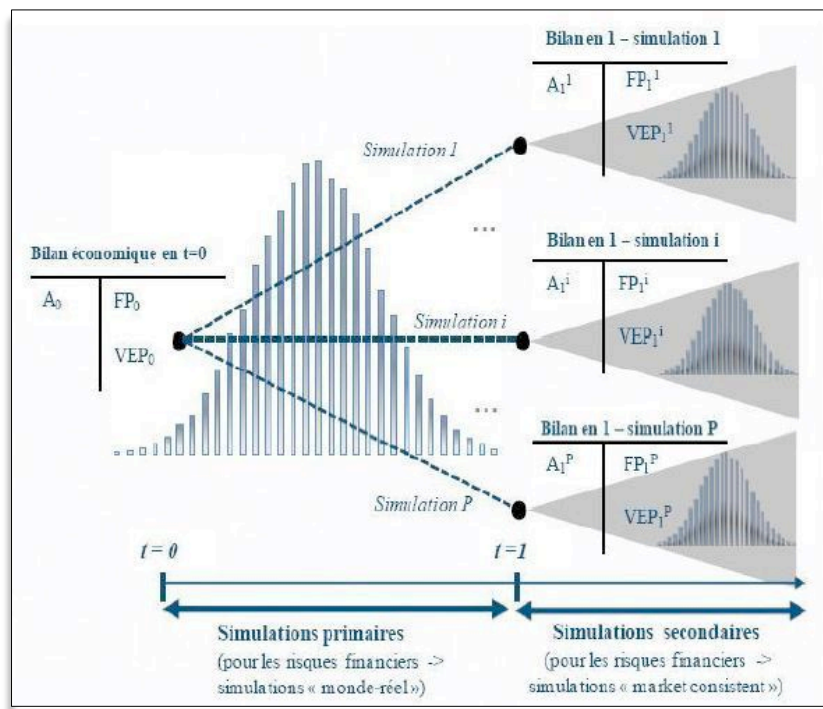


Figure 4 : Schéma de la méthode des Simulations dans les Simulations

La méthode SdS permet ainsi d'obtenir une distribution empirique de fonds propres à un an, dont nous pouvons déduire le capital économique.

Pour obtenir un résultat précis en utilisant cette méthode il est nécessaire d'effectuer un grand nombre de simulations primaires et secondaires. Nous appellerons, dans la suite du mémoire, « SdS complet » la méthode SdS utilisée avec 10 000 scénarios primaires et 5 000 simulations secondaires par scénario. Il est important de noter que, suivant la complexité du portefeuille utilisé, il peut être nécessaire d'effectuer un plus grand nombre de simulations.

Cette méthode apparaît comme étant très coûteuse, en tables d'une part, et en temps de projection d'autre part. En effet, elle nécessite l'utilisation d'une table de scénarios secondaires pour chacun des scénarios primaires. Par exemple, pour un SdS complet, il faudrait 10 000 tables de 5 000 simulations secondaires RN. Par ailleurs, le nombre total de simulations à effectuer est égal au produit du nombre de simulations primaires et du nombre de simulations secondaires, ce qui entraîne des temps de projection importants. Finalement, la complexité de cette méthode en termes de budget de calcul compromet très fortement sa mise en œuvre.

2. Zoom sur des méthodes alternatives proches

Comme il est très coûteux de mettre en place un SdS complet, il apparaît nécessaire de trouver des méthodes alternatives permettant d'obtenir le capital économique plus rapidement. En outre, nous pouvons remarquer que pour réaliser ce calcul, il n'est pas nécessaire de connaître précisément l'ensemble de la distribution des fonds propres à un an mais uniquement le quantile à 0,5 % de cette distribution. Il s'agit dès lors de réussir à estimer précisément ce quantile en utilisant le moins de calculs ALM possibles.

Nous pouvons déterminer plusieurs axes d'amélioration de la méthode SdS :

- diminuer le nombre de simulations primaires,
- diminuer le nombre de simulations secondaires,
- s'appuyer sur des fonctions proxy afin d'obtenir plus rapidement une distribution empirique (approchée efficacement) des fonds propres.

Nous allons tout d'abord nous intéresser à deux méthodes qui s'appuient sur la méthode SdS : l'accélérateur SdS et l'allocation optimale de scénarios. Nous étudierons ensuite des méthodes basées sur des fonctions proxy, telles que les méthodes paramétriques, le Replicating Portfolio, et le kriging stochastique.

a) Réduction du nombre de simulations primaires : l'accélérateur SdS

Nous présentons, dans cette partie, le principe de l'accélérateur SdS. Pour plus d'informations sur cette méthode et sa mise en œuvre, le lecteur pourra se référer à Devineau et Loisel (2009).

Nous observons que pour calculer le capital économique, il est uniquement nécessaire de connaître le quantile à 0,5 % des fonds propres. Ainsi, si nous nous basons sur une distribution composée de N réalisations de fonds propres, nous avons besoin de connaître précisément seulement les $[N \times 0,5 \ %]$ pires valeurs. Par conséquent, il est inutile de consommer une partie du budget de calcul pour estimer le reste de la distribution.

Cette méthode nécessite finalement de pouvoir identifier les $[N \times 0,5 \ %]$ scénarios les plus adverses. Néanmoins, nous pouvons observer, dans la pratique, que lorsque nous extrayons les facteurs de risque et que nous les représentons, les scénarios les plus adverses sont situés en périphérie du nuage. Il convient donc de définir une norme permettant de localiser ces scénarios. Cette norme, basée sur les facteurs de risque et sur la sensibilité des fonds propres à chacun des facteurs, peut être définie, pour deux facteurs de risque ϵ_1 et ϵ_2 , de la manière suivante : $\|\epsilon_1, \epsilon_2\|_{sensi} = \sqrt{s_1 \times \epsilon_1^2 + s_2 \times \epsilon_2^2}$, où s_1 est la sensibilité des fonds propres au facteur de risque ϵ_1 et s_2 la sensibilité par rapport à ϵ_2 . Nous pouvons alors appliquer la méthode de l'accélérateur SdS en nous basant sur cette norme.

L'algorithme est composé de plusieurs itérations détaillées ci-dessous :

- Lors de la première itération, nous sélectionnons les M scénarios les plus adverses en utilisant la norme définie précédemment. Notons que ce nombre M doit être choisi supérieur à $[N \times 0,5 \ %]$. Nous calculons alors, de manière approchée, les fonds propres associés à ces scénarios, puis nous les trions par ordre de fonds propres croissant.

- A la deuxième itération, nous prenons les M pires scénarios suivant, que nous déterminons grâce à la norme. De la même manière que lors de la première itération, nous estimons les fonds propres associés à ces scénarios, puis nous trions les $2M$ scénarios en fonction de leurs fonds propres.
- Si les $[N \times 0,5 \ %]$ pires scénarios obtenus sont les mêmes qu'à l'itération précédente, nous nous arrêtons. Sinon, nous répétons ce processus jusqu'à ce que les $[N \times 0,5 \ %]$ scénarios donnant les fonds propres les plus faibles soient les mêmes pour deux itérations successives.
- A la fin des itérations nous avons localisé les scénarios de la zone quantile. Afin d'obtenir une estimation plus précise du quantile, nous prenons les $2[N \times 0,5 \ %] + 1$ scénarios les plus adverses, d'après le tri effectué lors de la dernière itération, et nous calculons leurs fonds propres de manière exacte, à l'aide d'un grand nombre de simulations secondaires.
- Enfin, nous estimons le quantile à 0,5 % des fonds propres en prenant la médiane des fonds propres obtenus à l'issue de l'étape précédente.

Pour déterminer le budget de calcul utilisé par cette méthode, il convient de distinguer l'étape de localisation de la zone quantile à l'aide des itérations successives, pour laquelle les fonds propres sont calculés de manière approchée, à l'aide par exemple de 50 simulations secondaires, et l'étape dans laquelle nous recalculons, de manière exacte, les fonds propres et qui nécessitera un grand nombre de simulations secondaires, par exemple 1 000. Les résultats obtenus dans le cas de portefeuilles simples, pour lesquels peu d'itérations suffisent à localiser la zone quantile, sont excellents. En effet, le quantile à 0,5 % obtenu est quasiment égal à celui trouvé dans le cas d'un SdS complet et le budget de calcul est beaucoup plus faible. Cependant, lorsque la zone quantile devient plus difficile à localiser, le budget de calcul peut rapidement augmenter.

Ainsi, le principal avantage de cette méthode est qu'elle permet de diminuer grandement le budget de calcul utilisé, lorsque les portefeuilles étudiés sont simples. Par ailleurs, sa mise en œuvre ne pose pas de problème particulier. Cependant, il n'est pas possible de connaître à l'avance le budget qui sera nécessaire pour localiser la zone quantile, donc dans le cas de portefeuilles plus compliqués, l'intérêt de la méthode devient limité.

b) Réduction du nombre de simulations secondaires : l'allocation optimale de scénarios

Après avoir présenté une méthode basée sur la réduction des simulations primaires, nous pouvons nous intéresser à une technique basée sur une diminution du nombre de simulations secondaires total, l'allocation optimale de scénarios. Le principe de ce type de méthodes est de déterminer le nombre minimal de simulations secondaires à effectuer pour obtenir la meilleure estimation possible du quantile à 0,5 % des fonds propres.

Il existe plusieurs méthodes d'allocation optimale de scénarios pouvant être mises en œuvre. Elles se décomposent en deux types : les allocations uniformes et les allocations non-uniformes de scénarios. La première catégorie repose sur une optimisation préalable qui permet de déterminer quelle est la meilleure répartition du budget entre les simulations primaires et les simulations secondaires, tandis

que la seconde catégorie permet de déterminer le nombre de simulations secondaires à allouer en fonction du scénario primaire. Nous présenterons ici uniquement la méthode qui donne les meilleurs résultats et qui appartient à la deuxième catégorie : l'allocation optimale séquentielle de scénarios. Pour plus de précisions sur les autres méthodes existantes, se référer à l'article de Broadie et al. (2011).

La méthode d'allocation optimale séquentielle est fondée, comme la méthode précédente, sur le principe que nous avons uniquement besoin d'estimer correctement le quantile à 0,5 %, mais qu'il n'est pas nécessaire de connaître de façon précise l'ensemble de la distribution des fonds propres. Cette méthode alloue ainsi des simulations secondaires à un scénario primaire en fonction de sa proximité avec le scénario quantile.

Plus précisément, pour utiliser cette méthode, nous commençons par fixer un budget de calcul. Ensuite, l'allocation optimale séquentielle se décompose en deux étapes :

- Lors de la première étape, nous lançons un nombre de simulations secondaires fixé, m_0 , pour chaque scénario primaire. Nous obtenons ainsi une distribution de fonds propres et nous en déduisons une estimation du quantile à 0,5 %. Cette étape permet de repérer quels sont les scénarios de la zone quantile, c'est-à-dire les scénarios pour lesquels les fonds propres calculés à partir des m_0 simulations secondaires sont proches de l'estimateur du quantile.
- Dans une seconde étape, nous ajoutons des simulations secondaires aux scénarios de la zone quantile, afin de les calculer plus précisément. Nous répétons ce processus jusqu'à ce que tout le budget disponible ait été alloué.

A l'issue de l'algorithme, nous pouvons remarquer que, suivant la proximité entre le scénario primaire et le quantile, nous n'aurons pas alloué le même nombre de simulations secondaires. Ainsi, certains scénarios posséderont uniquement m_0 simulations secondaires, tandis que d'autres auront le maximum de simulations secondaires possible. Cette technique nous permet d'effectuer un zoom au voisinage du quantile et par conséquent d'obtenir une estimation précise de celui-ci avec un budget de calcul beaucoup moins important que dans le cas d'un SdS complet. Nous notons en effet que pour cette méthode, le budget de calcul, C , est donné par la formule suivante : $C = \sum_{i=1}^N m_i$, où N est le nombre de scénarios et m_i le nombre de simulations secondaires allouées au scénario primaire i .

Les résultats obtenus avec cet algorithme diffèrent selon le choix du paramètre m_0 , mais lorsque nous choisissons convenablement ce paramètre (en prenant par exemple $m_0 = 10$), nous parvenons à localiser correctement les scénarios de la zone quantile. Ensuite, en allouant un grand nombre de simulations secondaires à ces scénarios, l'algorithme permet de connaître de façon précise cette zone. Nous obtenons donc une très bonne estimation du quantile à 0,5 % des fonds propres.

Ainsi, les principaux avantages de cette méthode sont sa simplicité, à la fois mathématique et opérationnelle et ses bons résultats : elle aboutit à une estimation précise du quantile avec un budget de calcul fixé au départ et qui peut être choisi assez faible (par exemple égal à 300 000 pour 15 000 scénarios). Néanmoins, nous pouvons penser que, dans le cas de portefeuilles complexes, pour lesquels la zone quantile est très étendue, les résultats se dégradent. En effet, il sera alors nécessaire d'utiliser plus de simulations secondaires lors de la première étape pour localiser la zone quantile.

3. Les méthodes utilisant un proxy

Nous avons abordé, dans les parties précédentes, des méthodes assez proches du SdS, qui permettent d'obtenir une estimation précise du quantile à 0,5 % des fonds propres avec un budget de calcul réduit, grâce à une réduction des simulations primaires ou secondaires. Nous allons, à présent, exploiter le dernier axe d'optimisation de la méthode SdS, à l'aide de méthodes qui s'appuient sur la détermination d'une fonction proxy, permettant d'obtenir la valeur des fonds propres à partir des facteurs de risque.

Nous nous intéressons donc, dans cette partie, à quatre méthodes utilisant un proxy : les méthodes de type forme paramétrique incluant le curve fitting et le Least Square Monte Carlo, les Replicating Portfolios, et le kriging stochastique.

a) Méthodes paramétriques

Nous présentons tout d'abord des approches de type forme paramétrique. Celles-ci consistent à déterminer une formule donnant la valeur des fonds propres à un an en fonction des facteurs de risque. Nous pouvons distinguer plusieurs techniques de calibrage. En effet, la forme paramétrique peut être calibrée directement sur des réalisations de fonds propres obtenues par la mise en œuvre de la méthode SdS restreinte à quelques scénarios primaires, dans ce cas la méthode utilisée est appelée Curve Fitting. Le calibrage peut également être réalisé sur les Valeurs Actuelles Nettes de marges (VAN), qui sont des réalisations de fonds propres dégénérées, pour lesquelles une seule simulation secondaire RN a été utilisée. Dans ce dernier cas, nous appelons la méthode « Least Square Monte Carlo » (LSMC). Nous allons, dans cette partie, exposer le principe de ces deux méthodes. Nous renvoyons le lecteur à l'article de Barrie-Hibbert (2011) pour une présentation plus exhaustive.

(1) Curve Fitting

Le principe de la méthode du Curve Fitting est le calibrage d'une forme paramétrique sur des valeurs de fonds propres à un an. Nous pouvons distinguer trois étapes pour déterminer le capital économique à l'aide de cette technique :

- Nous sélectionnons, tout d'abord, un petit nombre de scénarios primaires (en pratique 100 à 150 sont suffisants pour obtenir de bons résultats). Pour chacun de ces scénarios, nous générons beaucoup de simulations secondaires, de façon à obtenir une estimation très précise des fonds propres associés.
- Dans une deuxième étape, nous calibrons la forme paramétrique, à partir des scénarios primaires sélectionnés et des fonds propres qui leur sont associés. Ce calibrage est effectué à l'aide d'une régression. A l'issue de cette étape, nous obtenons donc, dans le cas de deux facteurs de risque ϵ_1 et ϵ_2 , une fonction de la forme suivante : $FP(X) = a_0 + a_1\epsilon_1 + a_2\epsilon_1^2 + a_3\epsilon_1^3 + a_4\epsilon_2 + a_5\epsilon_2^2 + a_6\epsilon_2^3 + a_7\epsilon_1\epsilon_2 + a_8\epsilon_1^2\epsilon_2 + a_9\epsilon_1\epsilon_2^2$.

- Nous pouvons ainsi avoir une distribution de fonds propres, en appliquant directement la formule ci-dessus à un grand nombre de réalisations de facteurs de risque, et en déduire le capital économique.

Lorsque nous appliquons cette méthode, nous trouvons une estimation du quantile très proche de la valeur de référence avec une erreur relative souvent inférieure à 2 %. Par ailleurs, cette méthode permet de réduire très nettement le budget de calcul puisqu'elle nécessite d'effectuer un calcul précis avec beaucoup de simulations secondaires, pour peu de scénarios primaires, par exemple 150.

(2) Least Square Monte Carlo

La méthode LSMC est proche de celle du Curve Fitting. Cependant, le calibrage de la forme paramétrique n'est plus effectué sur la valeur des fonds propres mais sur les VAN de marge. En effet, nous pouvons remarquer, dans une approche de type SdS, que certaines simulations primaires proches conduisent à des simulations secondaires de lois quasi-identiques. Pour éviter cette redondance d'informations, le LSMC propose de n'effectuer qu'une seule simulation secondaire par scénario, pour un grand nombre de scénarios primaires. Nous notons alors, que même si la forme paramétrique est calibrée sur les VAN de marge, elle réplique convenablement les fonds propres.

Les étapes du LSMC sont les suivantes :

- Nous partons d'un grand nombre de scénarios primaires extrêmes, par exemple 50 000, et nous générons une unique simulation secondaire pour chacun de ces scénarios.
- Nous calibrons ensuite notre forme paramétrique de la même manière que précédemment.
- Nous en déduisons la distribution des fonds propres à un an et le capital économique.

Nous obtenons des résultats comparables à ceux du Curve Fitting : l'estimation du capital économique est proche de la valeur de référence, et le budget de calcul est beaucoup plus faible que celui d'un SdS complet. En effet, cette méthode nécessite uniquement la construction d'une simulation secondaire par scénario primaire.

Les schémas ci-dessous permettent d'illustrer les deux méthodes présentées dans cette partie, et de mettre en évidence leurs différences :

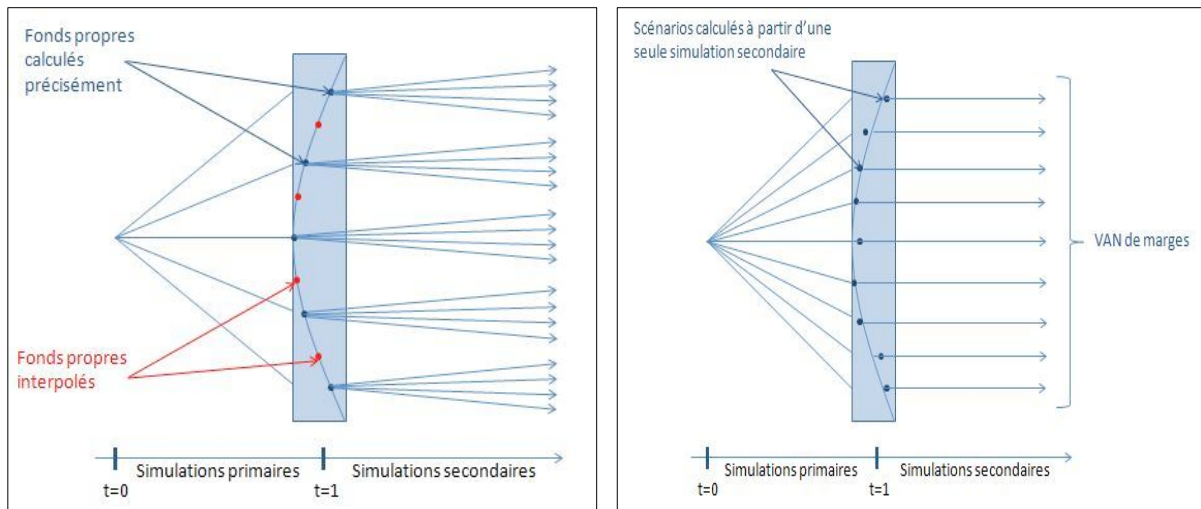


Figure 5 : Schéma des méthodes de types formes paramétriques : Curve Fitting (figure 5.a) et Least Square Monte Carlo (figure 5.b)

Les méthodes de type forme paramétrique permettent d'obtenir une bonne estimation du capital économique en réduisant considérablement le nombre de projections ALM par rapport à la méthode SdS. Par ailleurs, le calibrage de la forme paramétrique est relativement simple et rapide. Néanmoins, dans le cas de portefeuilles plus complexes, il peut être plus difficile d'obtenir une bonne estimation de la distribution des fonds propres à l'aide de formes paramétriques, où cela peut nécessiter l'utilisation de plus de points pour le calibrage.

b) *Replicating Portfolio*

Nous nous intéressons à présent à une méthode basée sur l'utilisation de portefeuilles répliquants. Un portefeuille répliquant est défini comme un portefeuille d'actifs reproduisant la valeur économique des fonds propres. Ainsi, cette méthode est fondée sur l'égalité suivante :

$$\text{Valeur du portefeuille répliquant} = \text{Valeur économique des fonds propres.}$$

Il existe plusieurs méthodes de type Replicating Portfolio basées sur différentes techniques de calibrage du portefeuille répliquant. Cependant, nous n'en présenterons ici qu'une seule, pour laquelle le calibrage du portefeuille répliquant est effectué par rapport à la valeur des fonds propres en $t = 1$. Le lecteur pourra se référer à Revelen (2009) pour obtenir des informations sur les autres techniques.

La méthode que nous nous proposons de présenter, est composée de cinq étapes :

- La première étape consiste à sélectionner des scénarios primaires adverses, grâce à l'utilisation d'une norme basée sur chacun des facteurs de risque et sur la corrélation entre ces facteurs.
- Dans une deuxième étape, nous calibrons une forme paramétrique à partir des N scénarios sélectionnés. Le calibrage est effectué à l'aide d'une approche de type MCO (Moindres Carrés Ordinaires). La forme paramétrique obtenue nous permet ainsi d'analyser la convexité des fonds propres.

- Nous déterminons ensuite les sous-portefeuilles répliquant chaque terme de la forme paramétrique, en nous appuyant sur le tableau ci-dessous, dans le cas de facteurs de risque action, ϵ_a , et zéro-coupon, ϵ_{ZC} :

Terme de la forme paramétrique	Actifs candidats pour le sous-portefeuille répliquant
ϵ_{ZC}	Cash, zéros-coupon
$\epsilon_{ZC}^2, \epsilon_{ZC}^3$	Cash, zéros-coupon, caplets, floorlets
$\epsilon_a, \epsilon_a^2, \epsilon_a^3$	Cash, indices actions, calls, puts

Tableau 1 : actifs candidats pour le sous-portefeuille répliquant chaque terme de la forme paramétrique

Notons que nous ne cherchons pas à répliquer les termes croisés de la forme paramétrique parce que nous postulons que certains actifs candidats les prennent implicitement en compte. Nous déterminons les paramètres associés à chaque sous-portefeuille (poids des actifs, maturité...) en minimisant l'écart quadratique entre le terme à répliquer et le sous-portefeuille. Le lecteur pourra se référer à Chauvigny et Devineau (2010) pour plus de détails sur la façon de choisir les actifs candidats en fonction du terme à répliquer.

- La quatrième étape consiste à déterminer la composition du portefeuille agrégé, c'est-à-dire que nous cherchons les poids associés à chaque actif candidat. Nous obtenons ainsi l'asset-mix du portefeuille répliquant final.
- Nous effectuons alors des travaux de validation afin de vérifier que le choix du portefeuille répliquant est correct.
- Enfin, nous projetons le portefeuille répliquant à un an, grâce à des techniques de Monte Carlo conditionnées par le réalisé de première période, et nous en déduisons la valeur des fonds propres et le capital économique.

Le schéma ci-dessous illustre les principales étapes de cette méthode :

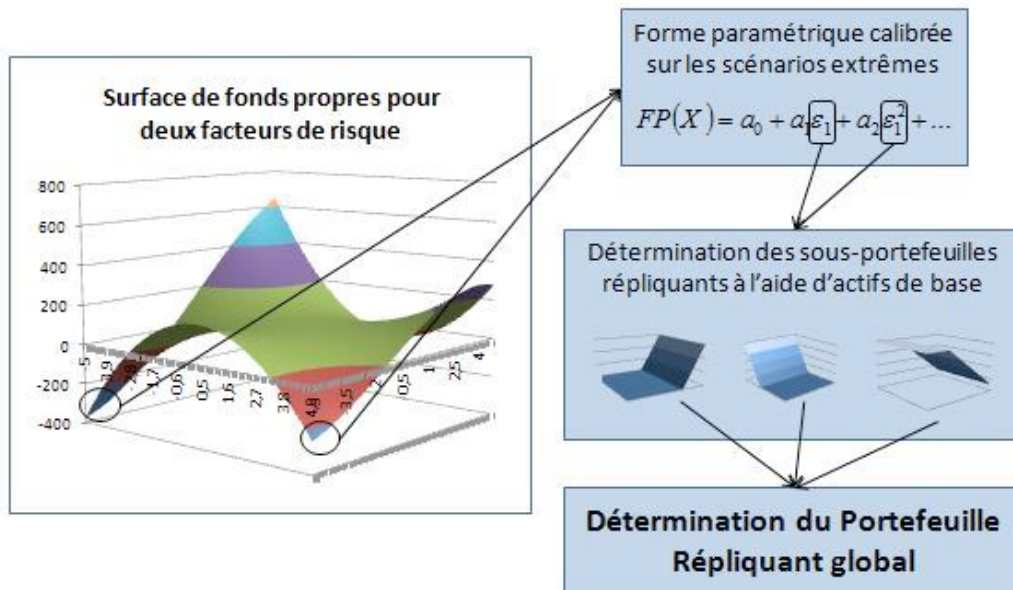


Figure 6 : Schéma de la méthode des Replicating Portfolios

La méthode des Replicating Portfolios, utilisée avec la technique de calibrage présentée ci-dessus, fournit des estimations robustes du capital économique et permet une réduction conséquente du budget de calcul par rapport à une approche de type SdS. Néanmoins cette méthode n'est pas la plus simple à mettre en œuvre, et nous pouvons regretter que le choix des actifs candidats soit basé sur des jugements d'experts.

c) *Kriging stochastique*

Comme les méthodes précédentes, l'algorithme de kriging stochastique est fondé sur l'utilisation d'une fonction proxy. En effet, il repose sur un calcul précis des fonds propres seulement pour quelques scénarios, appelés design points, et sur la construction, à l'aide de ces scénarios, d'une fonction donnant une estimation des fonds propres à partir des facteurs de risque. Cette dernière peut alors être appliquée à un grand nombre de scénarios, appelés points de prédiction, pour obtenir une distribution de fonds propres et en déduire le quantile à 0,5 %. La méthode présentée ci-après s'appuie sur les articles de Liu et Staum (2010) et Ankenman et al. (2010).

La méthode de kriging stochastique, adaptée au calcul du capital économique, est composée de trois étapes :

- Dans une première étape, nous nous appuyons sur les points de prédiction pour construire les design points. En effet, comme l'estimation des fonds propres aux points de prédiction est basée sur une interpolation des fonds propres aux design points, il est nécessaire que ces derniers soient répartis uniformément, dans tout l'espace, et que nous puissions en trouver un au voisinage de chaque point de prédiction. Pour choisir les design points, nous pouvons utiliser la méthode de maximisation des hypercubes latins. Celle-ci permet de placer N_1 points dans un convexe, de façon telle que ces points soient tous distincts et que la distance minimale entre eux soit maximisée.

Une fois que nous avons choisi les design points, nous pouvons générer un grand nombre de simulations secondaires pour ces scénarios et ainsi calculer les fonds propres qui leur sont associés. Nous construisons alors un métamodèle en calibrant une fonction proxy sur ces design points, ce qui nous permet de déterminer la valeur des fonds propres aux points de prédiction.

- A l'issue de la première étape, nous avons obtenu une distribution de fonds propres dont nous pouvons déduire le quantile à 0,5 %. La deuxième étape consiste alors à déterminer les scénarios les plus proches du quantile et à les rajouter à l'ensemble des design points. Ce processus permet ainsi de mieux calibrer la fonction proxy au voisinage du quantile. En effet, en générant des simulations secondaires pour les nouveaux design points, nous pouvons construire un nouveau métamodèle, plus précis que le précédent dans la zone quantile.
- Enfin, la dernière étape consiste à allouer de façon optimale le budget de calcul restant. Nous pouvons en effet calculer le nombre de simulations secondaires, m_i , à attribuer à chaque design point i , pour que la variance de l'estimateur du quantile à 0,5 % des fonds propres soit minimale. Après avoir générées ces simulations secondaires supplémentaires, nous construisons un dernier métamodèle, et nous en déduisons une estimation des fonds propres aux points de prédiction. Nous obtenons alors, à partir de la distribution des fonds propres obtenue, un estimateur du quantile à 0,5 % et donc une estimation du capital économique.

A la fin de l'algorithme, nous avons pu estimer le capital économique à l'aide de simulations secondaires générées pour les N_1 design points de départ et pour les N_2 design points qui ont été rajoutés à l'étape 2. Pour chacun de ces scénarios, nous avons généré un nombre m_i de simulations secondaires, qui a été déterminé à l'étape 3. Ainsi, le budget total utilisé pour cette méthode est égal à $C = \sum_{i=1}^{N_1+N_2} m_i$.

Même si les résultats obtenus dépendent du choix des paramètres N_1 et N_2 , cette méthode permet d'obtenir une estimation cohérente du quantile à 0,5 % des fonds propres avec un budget de calcul réduit. Cependant, elle peut paraître complexe d'un point de vue mathématique et les résultats obtenus ne sont pas meilleurs qu'avec une méthode plus simple comme l'allocation optimale séquentielle de scénarios. Par ailleurs, cette méthode semble plus complexe opérationnellement que la plupart de celles présentées précédemment puisque qu'il est nécessaire de construire plusieurs métamodèles.

4. Récapitulatif des résultats et de la complexité des différentes méthodes

Les méthodes présentées dans cette partie ont été testées sur des portefeuilles d'épargne standard. Aussi, nous pouvons récapituler les résultats obtenus, ainsi que la complexité, à la fois mathématique et opérationnelle, à l'aide du tableau² ci-dessous :

² Les résultats sont donnés à titre indicatif et peuvent varier suivant la complexité du portefeuille.

Méthode		Résultats	Budget de calcul (nombre de projections ALM)	Complexité ³ opérationnelle	Complexité ⁴ mathématique
Simulations dans les Simulations		Quantile de référence	50 000 000	++++	+
Accélérateur SdS		Quantile estimé quasiment égal au quantile de référence	300 000 – 400 000	+	+
Allocation optimale de scénarios		- Bons résultats pour des portefeuilles simples - Dégradation quand les zones quantile sont étendues	300 000	++	++
Méthodes paramétriques	Curve fitting	- Bons résultats pour des portefeuilles simples - Dégradation pour des zones quantile complexes	100 000	+	+
	LSMC		100 000	++	++
Replicating Portfolio		- Bons résultats pour des portefeuilles simples - Dégradation pour des portefeuilles plus complexes	100 000	+++	++
Kriging stochastique		Résultats moyens	300 000	+++	+++

Tableau 2 : Récapitulatif des méthodes permettant de calculer le capital économique

Nous avons ainsi présenté différentes méthodes permettant d'estimer précisément le capital économique. Cependant, nous avons pu noter que, pour la plupart des méthodes, la complexification du portefeuille risquait de détériorer la qualité des résultats.

Nous allons à présent proposer deux nouvelles méthodes pour évaluer le capital économique.

³ Plus il y a de signes « + », plus l'algorithme est complexe d'un point de vue opérationnelle.

⁴ Plus il y a de signes « + », plus l'algorithme est complexe d'un point de vue mathématique.

III. Deux nouvelles méthodes : la décomposition d'un événement rare en événements moins rares et l'estimation par forêts aléatoires

Dans cette partie, nous allons expliquer les fondements théoriques des deux méthodes que nous avons mises en œuvre dans la suite du mémoire pour calculer le capital économique. Il est important de noter que les techniques présentées ci-après ont déjà été appliquées à d'autres domaines, mais qu'elles sont adaptées au calcul du capital économique pour la première fois dans ce mémoire. Nous apportons en outre plusieurs modifications aux algorithmes de base. Nous commencerons par décrire la méthode de décomposition d'un événement rare en événements moins rares, puis nous nous intéresserons à la détermination des scénarios proches du quantile grâce aux forêts aléatoires.

A. Décomposition d'un événement rare en événements moins rares

Pour comprendre la théorie sur laquelle repose la méthode de décomposition d'un événement rare en événements moins rares, nous revenons tout d'abord sur les articles écrits sur ce sujet, puis nous proposons différentes adaptations des algorithmes déjà existant pour résoudre le problème de détermination du quantile à 0,5 % des fonds propres. Dans un second temps, nous reviendrons sur les difficultés liées à la mise en œuvre de l'algorithme. Enfin, dans une dernière partie, nous étudierons certaines propriétés de l'estimateur du quantile obtenu.

1. Principe de la méthode

La méthode que nous proposons ici est basée sur des itérations permettant d'obtenir un échantillon de réalisations d'événements de plus en plus rare. Nous pouvons donc l'utiliser pour estimer un quantile associé à une très faible probabilité.

a) Principe général de la méthode

L'estimation d'un quantile de manière précise peut devenir très difficile lorsque nous nous intéressons à des quantiles extrêmes. Par exemple, si nous utilisons une méthode de Monte Carlo pour estimer le quantile associé à une probabilité p , il est nécessaire de simuler un très grand nombre de variables, bien supérieur à la partie entière de $\frac{1}{p}$. Une telle approche consisterait à simuler toutes ces réalisations, à les ranger par ordre croissant puis à prendre la plus petite. Cependant, lorsque la probabilité p est très faible, le nombre de réalisations à simuler pour connaître le quantile devient très élevé. Par exemple, si nous recherchons le quantile associé à une probabilité de 10^{-6} , il est nécessaire de simuler plus de 10^6 réalisations. La méthode à laquelle nous nous intéressons dans cette partie permet de résoudre ce problème.

Notons A un événement rare, c'est-à-dire un événement tel que la probabilité $P(X \in A)$ soit très faible. La méthode que nous allons présenter dans cette partie consiste à décomposer l'événement A en événements moins rares contenant A et tels que le plus petit soit A .

Plus précisément, cette méthode repose sur la construction de $M+1$ événements de \mathbb{R}^d imbriqués : $\mathbb{R}^d = A_0 \supset A_1 \supset \dots \supset A_M = A$, tels que pour tout m compris entre 1 et M , la probabilité $P(X \in A_m | X \in A_{m-1})$ ne soit pas trop petite devant p . Estimer directement $P(X \in A_m)$ pourrait être compliqué, mais comme nous nous plaçons dans l'ensemble A_{m-1} , la probabilité conditionnelle n'est pas trop faible. Ainsi, nous pouvons estimer facilement la probabilité de passer d'un événement A_{m-1} à un événement A_m .

Cette méthode permet ainsi de construire des événements de plus en plus rares et de calculer les probabilités conditionnelles associées au passage d'un ensemble A_{m-1} vers un ensemble A_m . Nous obtenons finalement une estimation de la probabilité de l'événement rare A grâce au produit des probabilités conditionnelles :

$$P(X \in A) = \prod_{m=1}^M P(X \in A_m | X \in A_{m-1}).$$

b) Littérature sur le sujet

La méthode « Multilevel splitting », qui consiste à décomposer un événement de la forme $\Phi(X) > q$ en plusieurs événements imbriqués de la forme $\Phi(X) > L_m$ avec $-\infty = L_0 < L_1 < \dots < L_M = q$, a été introduite par Kahn et Harris (1951) puis par Rosenbluth et Rosenbluth (1955).

Cette méthode a été reprise pour la première fois par Au et Beck (2001) pour être appliquée à la simulation d'événements rares. Depuis, elle est utilisée pour estimer des probabilités très faibles dans différents domaines, comme la physique ou l'astronomie avec Morio et al. (2011), et des améliorations sont apportées régulièrement à l'algorithme de base. Par exemple, le choix des différents niveaux L_0, \dots, L_M est rapidement apparu comme l'une des faiblesses de la méthode initiale. Ainsi, Cérou et al. ont proposé en 2009 puis en 2012 une méthode adaptative pour déterminer les niveaux au fur et à mesure de l'algorithme, tout en minimisant la variance asymptotique de l'estimateur de la probabilité de l'événement rare. Par la suite, Guyader et al. (2011) et Caron et al. (2013) ont élaboré un algorithme alternatif pour calculer les quantiles associés à des probabilités faibles et ont proposé une nouvelle façon de choisir les niveaux L_0, \dots, L_M de manière optimale.

Parallèlement, ces auteurs ont également analysé les propriétés asymptotiques des estimateurs de la probabilité d'un événement rare et du quantile associé.

c) Zoom sur l'algorithme dont nous nous sommes inspirés

La méthode que nous proposons s'appuie essentiellement sur les articles de Guyader et al. (2011), Cérou et al. (2012) et Caron et al. (2013). Les auteurs y présentent des méthodes permettant d'estimer la probabilité d'un événement rare ou le quantile associé à une faible probabilité et les appliquent au tatouage numérique.

Un tatouage numérique est une signature ajoutée à des documents numériques comme les fichiers audios, les vidéos, les images... Cette signature donne souvent des informations de copyright qui permettent d'empêcher les copies. Ainsi, si un fichier contient ce type de tatouage numérique, il sera impossible de le copier. En principe, si un fichier n'a pas été tatoué numériquement, nous pouvons le copier. Cependant, il peut arriver qu'il soit impossible de copier un tel fichier. Cet événement est appelé une fausse alarme et il se produit avec une probabilité très faible, mais non nulle. Dans leurs

articles, les auteurs cherchent à évaluer cette probabilité. Ainsi, ils s'intéressent uniquement aux fichiers qui n'ont pas été tatoués numériquement et ils veulent estimer la probabilité qu'il ne soit pas possible de copier l'un de ces fichiers alors que cela devrait l'être.

Afin de mieux comprendre, nous pouvons établir un parallèle entre l'estimation de la probabilité de fausse alarme et l'erreur de première espèce d'un test statistique. En effet, nous pouvons poser un test statistique d'hypothèses :

- H_0 : le fichier n'est pas tatoué numériquement,
- contre H_1 : le fichier est tatoué numériquement.

Dans ce cas, l'erreur de première espèce est la probabilité de rejeter H_0 à tort, c'est-à-dire de considérer que le fichier est tatoué numériquement alors que ce n'est pas le cas. Autrement dit, c'est la probabilité de ne pas pouvoir copier un fichier qui n'a pas été tatoué numériquement. Cette erreur est celle qui doit être contrôlée et minimisée en priorité.

Dans les articles, les auteurs proposent deux algorithmes permettant d'estimer une probabilité ou un quantile. Ils définissent X comme un vecteur aléatoire de loi μ , et considèrent une fonction, notée $\Phi(X)$. Cette fonction peut être vue comme une boîte noire, c'est-à-dire qu'il est possible de connaître sa valeur pour un vecteur X donné mais que son expression exacte est inconnue. Si $\Phi(X)$ est supérieure à un niveau fixé q alors le fichier est considéré comme étant tatoué numériquement. Ainsi, si $\Phi(X) > q$, il est impossible de copier le fichier. L'objectif des algorithmes proposés est alors de considérer des fichiers qui n'ont pas été tatoués numériquement et d'estimer la probabilité $p = P(\Phi(X) > q)$ pour une valeur q fixée, ou d'estimer un quantile q pour une valeur de p fixée.

Dans le premier algorithme, qui permet uniquement de calculer une probabilité étant donné q , les auteurs fixent des niveaux $-\infty = L_0 < L_1 < \dots < L_M = q$, estiment séparément chaque probabilité $P(\Phi(X) > L_m | \Phi(X) > L_{m-1})$, puis en déduisent la probabilité estimée \hat{p} de la manière suivante :

$$\hat{p} = P(\Phi(X) > q) = \prod_{m=1}^M P(\Phi(X) > L_m | \Phi(X) > L_{m-1}).$$

Cependant, cette méthode possède des limites puisque les niveaux L_1, \dots, L_M sont choisis à l'avance, et puisqu'elle ne permet pas d'estimer un quantile.

Une adaptation de cette méthode a alors été proposée par les auteurs. Dans ce nouvel algorithme, les niveaux ne sont plus choisis avant de lancer l'algorithme, mais sont déterminés à chaque itération m , à partir de l'échantillon de N vecteurs, de la manière suivante : $L_m = \min(\Phi(X_1), \dots, \Phi(X_N))$. Après avoir déterminé L_m , nous re-générons le scénario X_{i^*} pour lequel L_m est atteint de façon à ce qu'il soit de loi $L(X | \Phi(X) > L_m)$.

Cette dernière adaptation permet, en outre, de déterminer la valeur d'un quantile, étant donné une probabilité p . En effet, à la fin de l'algorithme, nous obtenons : $p = P(\Phi(X) > L_M)$. Un estimateur naturel du quantile associé à la probabilité p est alors donné par L_M . Nous allons, à présent, proposer une adaptation de ce dernier algorithme au calcul du quantile à 0,5 % des fonds propres à un an, pour en déduire un capital économique.

d) Adaptation à notre problème

Précédemment, nous avons expliqué que la difficulté dans la détermination du capital économique venait uniquement du calcul du quantile à 0,5 % des fonds propres à un an. Nous allons ainsi adapter les algorithmes présentés ci-dessus au calcul de ce quantile.

Nous introduisons les notations suivantes :

- Nous notons X les vecteurs de facteurs de risque. Ainsi, X est un vecteur aléatoire de \mathbb{R}^d , défini par $X = (\epsilon_1, \dots, \epsilon_d)$, où $\epsilon_1, \dots, \epsilon_d$ sont les facteurs de risque. Nous pouvons supposer, sans perte de généralité, que les facteurs de risque suivent des lois normales centrées réduites indépendantes. X sera donc un vecteur gaussien de \mathbb{R}^d .
- FP est une fonction supposée continue (vérifié en pratique) de \mathbb{R}^d dans \mathbb{R} représentant la valorisation des fonds propres à un an. Cette fonction remplit les mêmes conditions que la fonction Φ des algorithmes de décomposition d'un événement rare en événements moins rares proposés dans la partie précédente. En effet, cette fonction est inconnue, mais nous pouvons calculer sa valeur pour chaque vecteur X en lançant un grand nombre de simulations secondaires.

Notre objectif est ainsi de déterminer la valeur de L , où L est défini par $P(FP(X) < L) = 0,5 \%$, c'est-à-dire que L représente le quantile à 0,5 % des fonds propres. Nous pouvons noter que le quantile que nous cherchons à déterminer est plus élevé que celui des algorithmes proposés dans la partie précédente, puisque les auteurs considèrent des quantiles associés à des probabilités de l'ordre de 10^{-12} . Cela pourra éventuellement limiter notre adaptation des algorithmes.

L'algorithme que nous proposons repose sur un procédé composé de deux niveaux d'itérations que nous appellerons dans la suite itérations verticales et itérations horizontales. Nous partons d'un échantillon de N scénarios, (X_1, \dots, X_N) . A chaque itération verticale m , le scénario (X_i^m) donnant les fonds propres les plus élevés au sein de l'échantillon est re-généré suivant la loi $L(X|FP(X) < L_m)$, où le niveau L_m est déterminé de la manière suivante : $L_m = \max(FP(X_1^{m-1}), \dots, FP(X_N^{m-1}))$. Pour simuler un scénario suivant cette loi, nous appliquerons plusieurs fois un noyau de transition. C'est ce que nous appellerons les itérations horizontales. Ce double procédé itératif peut être représenté par le schéma suivant :

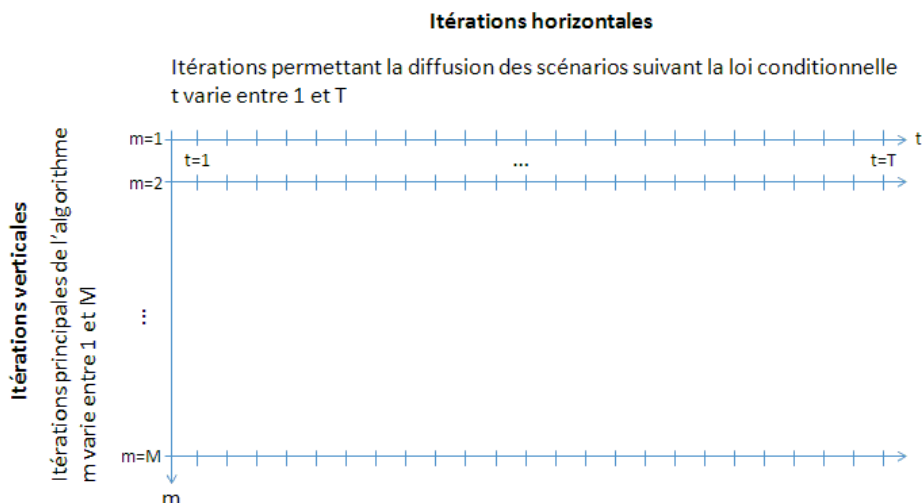


Figure 7 : Représentation schématique des deux procédés d'itérations de la méthode

A la fin de l'algorithme, nous avons donc construit des événements de \mathbb{R}^d , imbriqués les uns dans les autres : $A_0 \supset A_1 \supset \dots \supset A_M = A$, où A_m est défini par $A_m = \{X \in \mathbb{R}^d, FP(X) < L_m\}$.

2. Présentation d'un premier algorithme simple

Nous allons, à présent, expliquer concrètement comment mettre en œuvre notre adaptation de l'algorithme de décomposition d'un événement rare. Nous commençons par préciser les inputs, avant de détailler les différentes étapes à respecter pour appliquer cette méthode.

a) Inputs

Pour utiliser notre algorithme, nous avons tout d'abord besoin d'un échantillon de N scénarios, (X_1, \dots, X_N) . Nous rappelons que ces scénarios sont des vecteurs de facteurs de risque appartenant à \mathbb{R}^d , où d est le nombre de facteurs de risque considérés. Nous devons également connaître la valeur des fonds propres associés à ces scénarios, $(FP(X_1), \dots, FP(X_N))$. Ces scénarios, et leurs fonds propres, servent à initialiser l'algorithme.

Par ailleurs, il est nécessaire de choisir la valeur des deux paramètres N (nombre de scénarios) et T (nombre d'itérations horizontales utilisées pour diffuser les scénarios). Une étude des résultats suivant la valeur de ces paramètres sera effectuée dans la partie de mise en œuvre des algorithmes.

b) Explication de la règle d'arrêt des itérations verticales

Nous avons expliqué que la méthode reposait sur un procédé itératif permettant de construire des ensembles d'événements décroissants de la forme $A_m = \{X \in \mathbb{R}^d, FP(X) < L_m\}$. Cependant il est nécessaire de déterminer quand s'arrêter pour obtenir le quantile voulu, c'est-à-dire déterminer la valeur de M pour laquelle $A_M = \{X \in \mathbb{R}^d, FP(X) < q_{0,5\%}(FP(X))\}$.

Pour trouver cette valeur, nous allons tout d'abord chercher un estimateur de la probabilité, pour un scénario donné, d'être dans l'ensemble A_m , sachant qu'il appartient déjà à l'ensemble A_{m-1} . A chaque itération m , nous avons initialement $FP(X_i^{m-1}) < L_{m-1}$, c'est-à-dire que $X_i^{m-1} \in A_{m-1}$, pour tous les indices i compris entre 1 et N . Nous posons $L_m = \max(FP(X_1^{m-1}), \dots, FP(X_N^{m-1}))$ et nous notons i^* , le scénario pour lequel $FP(X_{i^*}^{m-1}) = L_m$. Nous avons alors $FP(X_i^m) < L_m$ pour tous les indices i compris entre 1 et N , excepté i^* . Ainsi, tous les scénarios qui sont dans A_{m-1} sont également dans A_m sauf $X_{i^*}^{m-1}$. Nous pouvons donc en conclure que :

$$P[X \in A_m | X \in A_{m-1}] = \left(1 - \frac{1}{N}\right).$$

Remarque : s'il existe deux scénarios i^* et i^{**} tels que $L_m = FP(X_{i^*}^{m-1}) = FP(X_{i^{**}}^{m-1})$, la formule donnée précédemment sera fautive. Néanmoins FP étant une fonction réelle supposée continue en les facteurs de risque, la probabilité d'un tel événement est de mesure nulle.

Comme nous effectuons M itérations indépendantes, nous pouvons estimer la probabilité d'être dans A_M par le produit des M probabilités conditionnelles : $\hat{P}(X \in A_M) = \hat{p} = \left(1 - \frac{1}{N}\right)^M$.

Inversement, si nous connaissons la probabilité d'être dans A_M , que nous la fixons égale à 0,5 %, et que nous recherchons le quantile associé, nous avons :

$$P(X \in A_M) = P(\text{FP}(X) < L_M) = 0,5 \% = P\left(\text{FP}(X) < q_{0,5\%}(\text{FP}(X))\right).$$

Nous pouvons donc estimer le quantile à 0,5 % des fonds propres par L_M . Nous cherchons ainsi la valeur de M telle que $P(X \in A_M) = 0,5 \% \approx \left(1 - \frac{1}{N}\right)^M$, ce qui nous amène à poser $M = \left\lceil \frac{\log(0,5 \%)}{\log(1-N^{-1})} \right\rceil$.

c) Algorithme

Notre algorithme se décompose en trois parties : l'initialisation, les itérations verticales au sein desquelles sont effectuées les itérations horizontales et l'arrêt des itérations déterminé grâce à la valeur de M trouvée précédemment. Nous présentons ci-dessous une première version d'un algorithme permettant de calculer le capital économique :

Initialisation :

Nous initialisons l'algorithme à l'aide d'un N -échantillon (X_1, X_2, \dots, X_N) de jeux de facteurs de risque indépendants et identiquement distribués de loi normale centrée réduite multivariée.

Nous posons : $X_1^0 = X_1, \dots, X_N^0 = X_N$ et $L_0 = +\infty$.

Notons que nous avons besoin en input des fonds propres associés à ces scénarios.

Itérations verticales :

Nous posons $M = \left\lceil \frac{\log(0,5 \%)}{\log(1-N^{-1})} \right\rceil$.

Pour $m = 1, \dots, M$:

- Soit $L_m = \max\left(\text{FP}(X_1^{m-1}), \dots, \text{FP}(X_N^{m-1})\right)$.
- Nous notons i^* la valeur de i pour laquelle $\text{FP}(X_i^{m-1})$ est maximal.
- $X_{i^*}^{m-1}$ n'appartient pas à l'événement $A_m = \{X \in \mathbb{R}^d, \text{FP}(X) < L_m\}$, or nous souhaitons avoir un échantillon de scénarios appartenant à cet ensemble. Nous expliquons exactement comment simuler un scénario suivant la loi $L(X|\text{FP}(X) < L_m)$ dans la partie III.A.3, mais nous donnons ici les étapes clés : pour obtenir un scénario $X_{i^*}^m$ appartenant à A_m , nous commençons par piocher aléatoirement un scénario X_j^{m-1} dans l'ensemble $\{X_1^{m-1}, \dots, X_N^{m-1}\} \setminus \{X_{i^*}^{m-1}\}$ et nous posons $X_{i^*}^m = X_j^{m-1}$. Comme X_j^{m-1} appartient à A_m , nous savons que $X_{i^*}^m$ est de loi $L(X|\text{FP}(X) < L_m)$. Cependant nous voulons que $X_{i^*}^m$ soit indépendant des autres scénarios et qu'il soit de loi normale. Nous allons donc diffuser $X_{i^*}^m$ à partir de X_j^{m-1} en utilisant un noyau de transition gaussien grâce aux itérations horizontales :
- Nous posons $t = 1$, puis nous appliquons la récurrence suivante tant que $t \leq T$:
 - Nous appliquons le noyau de transition à $X_{i^*}^m$, c'est-à-dire que nous obtenons un nouveau scénario X^* , pour lequel chaque facteur de risque ϵ_k^* donné par la formule
$$\epsilon_k^* = \frac{\epsilon_{i^*,k}^m + \sigma_k W_k}{\sqrt{1 + \sigma_k^2}},$$
 où $\epsilon_{i^*,k}^m$ est le $k^{\text{ième}}$ facteur de risque du scénario $X_{i^*}^m$.
 - Nous calculons $\text{FP}(X^*)$.
 - Si $\text{FP}(X^*) < L_m$, alors nous posons $X_{i^*}^m = X^*$ et nous augmentons t de 1.

- Pour tous les j différents de i^* , nous posons $X_j^m = X_j^{m-1}$, c'est-à-dire que nous ne changeons pas ces scénarios.
- Nous obtenons alors un nouvel échantillon de scénarios, $\{X_1^m, \dots, X_N^m\}$, pour lesquels les fonds propres sont tous inférieurs à L_m . Nous avons ainsi un échantillon de scénarios appartenant à A_m .

Arrêt des itérations verticales :

A la fin des $M = \left\lceil \frac{\log(0,5\%)}{\log(1-N^{-1})} \right\rceil$ itérations verticales, nous obtenons un échantillon de scénarios $\{X_1^M, \dots, X_N^M\}$ tels que les fonds propres associés à chacun de ces scénarios sont inférieurs à une valeur L_M déterminée par l'algorithme. Dans notre cas la probabilité d'appartenir à A_M est fixée et vaut 0,5 %. Nous avons donc $P(\text{FP}(X) < L_M) = P(X \in A_M) = 0,5\%$. Ainsi, nous pouvons estimer le quantile à 0,5 % des fonds propres en posant : $\hat{q}_{0,5\%}(\text{FP}) = L_M$.

Dans la suite du mémoire nous revenons plus en détail sur la façon de diffuser les scénarios suivant la loi conditionnelle $L(X|\text{FP}(X) < L_m)$.

3. Difficultés liées à l'implémentation pratique de l'algorithme : diffusion des scénarios selon une loi conditionnelle

Nous avons pu voir, dans l'algorithme ci-dessus, que la détermination du quantile à 0,5 % des fonds propres nécessitait de pouvoir simuler des scénarios X suivant la loi conditionnelle $L(X|\text{FP}(X) < L_m)$. Nous allons maintenant expliquer quels sont les problèmes posés par cette diffusion, et comment la réaliser.

a) Algorithme de Metropolis Hastings

L'une des difficultés principales de cette méthode réside dans la simulation des scénarios selon la loi $L(X|\text{FP}(X) < L_m)$.

Nous considérons que les facteurs de risque, associés aux scénarios (X_i^m), suivent une loi normale centrée réduite, que nous noterons μ par soucis de simplicité dans les expressions. Pour simuler un facteur de risque sachant que les fonds propres associés au scénario doivent être inférieurs à une valeur notée L_m , nous utilisons un noyau de transition K , μ -symétrique, c'est-à-dire tel que :

$$\forall (y, \epsilon) \in \mathbb{R}^{2d}, \mu(d\epsilon)K(y, d\epsilon) = \mu(dy)K(y, d\epsilon).$$

Ce noyau de transition joue un rôle très important, aussi il est important de bien le choisir.

Dans le cas général, nous ne connaissons pas la loi des variables que nous voulons diffuser. Nous noterons ici ces variables E , et leur réalisation ϵ . Nous pouvons alors utiliser l'algorithme de Metropolis-Hastings pour les simuler selon une matrice de transition K qui soit μ -symétrique.

Le principe de l'algorithme de Metropolis-Hastings est le suivant :

- Nous partons d'une matrice de transition Q et d'une variable de départ E_0 .

- Pour construire E_{t+1} à partir de E_t nous commençons par simuler deux variables aléatoires Y_t et U_t indépendantes de E_k , pour $k \leq t$, et indépendantes entre elles telles que :
 - Pour tout y , $P(Y_t = y) = Q(E_t, y)$,
 - U_t suit une loi uniforme sur l'intervalle $[0,1]$.

- Nous définissons $\alpha(\epsilon, y)$ par :

$$\begin{cases} \alpha(\epsilon, y) = \min\left(1, \frac{\mu(y)Q(y, \epsilon)}{\mu(\epsilon)Q(\epsilon, y)}\right) & \text{si } \mu(\epsilon)Q(\epsilon, y) \neq 0 \\ \alpha(\epsilon, y) = 1 & \text{si } \mu(\epsilon)Q(\epsilon, y) = 0 \end{cases}$$

Nous construisons ensuite E_{t+1} de la manière suivante :

- Si $U_t \leq \alpha(E_t, Y_t)$, nous posons $E_{t+1} = Y_t$. Dans ce cas, nous acceptons la transition.
- Si $U_t > \alpha(E_t, Y_t)$, nous posons $E_{t+1} = E_t$. Dans ce cas, nous rejetons la transition.

Nous pouvons alors montrer que $(E_t, t \geq 0)$ définit une chaîne de Markov de matrice de transition K et que μ est une mesure de probabilité invariante, où K est définie par :

$$\begin{cases} K(\epsilon, y) = Q(\epsilon, y) \times \alpha(\epsilon, y) & \text{si } \epsilon \neq y \\ K(\epsilon, \epsilon) = 1 - \sum_{\epsilon \neq y} P(\epsilon, y) & \text{sinon} \end{cases}$$

Pour adapter l'algorithme de Metropolis-Hastings à notre problème, nous pourrions considérer la matrice de transition K définie par : $\begin{cases} K(\epsilon, y) = Q(\epsilon, y) \times \alpha(\epsilon, y) & \text{si } \epsilon \neq y \\ K(\epsilon, \epsilon) = 1 - \sum_{\epsilon \neq y} K(\epsilon, y) & \text{sinon} \end{cases}$, où α et Q sont ceux définis dans la méthode de Metropolis-Hasting. Pour obtenir des facteurs de risque $\epsilon_{i^*,k}^m$ tels que le scénario associé soit de loi $L(X|FP(X) < L_m)$, et qu'il soit indépendant des autres scénarios de l'échantillon, nous commençons par diffuser chaque facteur de risque selon la matrice de transition K . Nous obtenons alors le scénario X^* , puis nous procédons de la manière suivante :

- Si $FP(X^*) < L_m$, nous acceptons la transition et nous remplaçons $X_{i^*}^m$ par X^* .
- Sinon nous rejetons la transition et $X_{i^*}^m$ reste à la même place.

Enfin, nous répétons cet algorithme jusqu'à avoir l'indépendance entre le scénario initial et le scénario obtenu à la suite des diffusions.

b) Diffusion dans le cas gaussien

Dans la partie précédente, nous avons expliqué comment simuler des scénarios selon la loi $L(X|FP(X) < L_m)$ dans le cas général, c'est-à-dire lorsque nous ne connaissons pas la loi des scénarios. Cependant, dans notre application au calcul du quantile à 0,5 % des fonds propres à un an, nous avons supposé que les facteurs de risque étaient gaussiens centrés réduits. Ainsi, nous connaissons directement un noyau de transition des facteurs de risque sans avoir à utiliser l'algorithme de Metropolis Hastings. En effet, il existe plusieurs noyaux de transition gaussiens connus de manière explicite :

- $K(\epsilon, d\epsilon') = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(\epsilon' - \epsilon)^2}{2\sigma^2}\right) \lambda(d\epsilon')$. La transition selon K est alors donnée par : $\epsilon^* = \epsilon + \sigma W$, où W suit une loi normale centrée réduite.

- $K(\epsilon, d\epsilon') = \sqrt{\frac{1+\sigma^2}{2\pi\sigma^2}} \exp\left(-\frac{1+\sigma^2}{2\sigma^2}\left(\epsilon' - \frac{\epsilon}{\sqrt{1+\sigma^2}}\right)^2\right) \lambda(d\epsilon')$. Dans ce cas, pour diffuser ϵ et pour que le facteur de risque ϵ^* obtenu suive une loi normale, nous utilisons l'expression : $\epsilon^* = \frac{\epsilon + \sigma W}{\sqrt{1+\sigma^2}}$.

Nous allons, dans la suite du mémoire, utiliser la seconde formule qui permet de renormaliser par rapport à l'écart-type du facteur de risque. Cette formule nous permet de diffuser les facteurs de risque selon une loi normale. Ainsi, pour tenir compte du conditionnement, nous rejetons le scénario X^* obtenu à l'issue de la diffusion, s'il ne respecte pas la condition $FP(X^*) < L_m$. Plus précisément, pour simuler un scénario X^m suivant la loi $L(X|FP(X) < L_m)$, nous commençons par diffuser chacun des facteurs de risque ϵ_k^m le composant, selon la matrice de transition K , c'est-à-dire que nous posons $\epsilon_k^* = \frac{\epsilon_k^m + \sigma W}{\sqrt{1+\sigma^2}}$.

- Si $FP(X^*) < L_m$ nous acceptons la transition et nous posons $X^m = X^*$ en égalisant chaque composante des deux vecteurs,
- Sinon nous rejetons la transition et X^m reste à la même place.

Remarque : Pour initialiser la diffusion avec un scénario remplissant la condition $FP(X) < L_m$, nous choisissons de tirer aléatoirement ce scénario parmi ceux suivant cette condition, c'est-à-dire que nous piochons aléatoirement un scénario dans l'ensemble $\{X_1^{m-1}, \dots, X_N^{m-1}\} \setminus \{X_{i^*}^{m-1}\}$, où $X_{i^*}^{m-1}$ est le scénario pour lequel $FP(X_{i^*}^{m-1}) = L_m$. Nous rappelons que ce scénario est noté X_j^{m-1} .

Remarque : Il est important de noter qu'une chaîne de Markov atteint sa loi stationnaire seulement au bout d'un certain nombre de transitions, aussi nous ne pouvons pas nous limiter à répéter une seule fois la procédure ci-dessus. Par ailleurs, il est également nécessaire de répéter la procédure plusieurs fois pour que le scénario $X_{i^*}^m$ obtenu soit indépendant du scénario de départ X_j^{m-1} . Par conséquent, cette procédure sera répétée jusqu'à ce que la loi stationnaire soit atteinte.

c) *Adaptation de l'algorithme de décomposition d'un événement rare*

Nous avons montré précédemment, que pour diffuser les scénarios selon la loi $L(X|FP(X) < L_m)$, nous appliquons, aux facteurs de risque, la transition $\epsilon_k^* = \frac{\epsilon_k^m + \sigma_k W_k}{\sqrt{1+\sigma_k^2}}$. Néanmoins, nous voyons que

celle-ci fait apparaître de nouveaux paramètres, les écart-types σ_k associés à chaque facteur de risque. Nous pouvons alors nous demander si le choix de ce paramètre influence les résultats.

Dans les articles étudiés, l'écart-type utilisé dans la diffusion des facteurs de risque, σ , est lié uniquement au nombre d'itérations nécessaires pour que la diffusion soit acceptée vingt fois. Cependant, nous pouvons penser que la valeur choisie pour ce paramètre joue un rôle important. En effet, si nous choisissons une valeur de σ_k très faible, les facteurs de risque seront peu diffusés, et X^* restera très proche de X^m . En revanche, si σ_k est élevé, X^* sera plus éloigné de X^m et le quantile trouvé pourrait alors être plus faible. Par ailleurs, le choix de ce paramètre peu apparaître comme étant arbitraire, et nous pouvons nous interroger sur la valeur à choisir pour obtenir de bons résultats.

Afin de résoudre ce problème, nous choisissons de relier σ_k à la variance observée dans l'échantillon avant la diffusion du scénario X_i^m . En effet, nous souhaitons qu'à la fin de la diffusion, la variance du $k^{\text{ième}}$ facteur de risque des scénarios de l'échantillon (X_1^m, \dots, X_N^m) soit égale à la variance d'une variable aléatoire $\epsilon_{i,k}$, de loi normale centrée réduite, et telle que $\text{FP}(X_i) < L_m$. Ainsi, pour chaque facteur de risque k , nous pouvons estimer cette variance par celle de l'échantillon $\{\epsilon_{1,k}^{m-1}, \dots, \epsilon_{N,k}^{m-1}\} \setminus \{\epsilon_{i^*,k}^{m-1}\}$, notée $\tilde{\sigma}_k^2$.

En effet, nous pouvons montrer (cf annexe 1) qu'en posant $\sigma_k = \sqrt{\left(\frac{1}{1-\tilde{\sigma}_k^2}\right)^{\frac{1}{T}} - 1}$, la variance obtenue à l'issue des T itérations horizontales est égale à notre estimation de la variance d'une variable aléatoire de loi normale centrée réduite sachant que $\text{FP}(X) < L_m$, c'est-à-dire à $\tilde{\sigma}_k^2$.

Ainsi, nous pouvons adapter l'algorithme présenté précédemment en appliquant la formule ci-dessus pour connaître la valeur de l'écart-type permettant le mieux de simuler les scénarios suivant la loi conditionnelle. Nous appellerons ce nouvel algorithme, « l'algorithme adapté », par opposition à « l'algorithme simple » qui considère que l'écart-type σ_k est constant au cours des itérations verticales.

Remarque : notons que les écarts-types utilisés pour diffuser chacun des facteurs de risque associés à un scénario sont obtenus indépendamment et que ces facteurs de risque sont diffusés de manière indépendante les uns des autres, sans introduire de corrélation entre eux.

4. Autre adaptation de l'algorithme

Jusqu'à maintenant, lorsque nous devons estimer les fonds propres associés à un scénario, nous utilisons un grand nombre de simulations secondaires. Cependant cela est très coûteux en projections ALM. Nous allons donc étudier une adaptation de notre algorithme permettant de résoudre ce problème.

En effet, nous remarquons, dans l'algorithme adapté présenté précédemment, qu'il n'est pas toujours nécessaire de connaître de manière exacte la valeur des fonds propres. En effet, lors de la diffusion des scénarios, nous appliquons le noyau de transition aux facteurs de risque $\epsilon_{i^*,k}^m$, pour obtenir $\epsilon_k^* = \frac{\epsilon_k^m + \sigma_k W_k}{\sqrt{1 + \sigma_k^2}}$. Nous comparons alors, à L_m , les fonds propres associés, pour savoir si la

transition est acceptée ou rejetée. Par conséquent, nous avons uniquement besoin de savoir si $\text{FP}(X^*)$ est inférieur ou supérieur à L_m , et non pas de connaître précisément la valeur de $\text{FP}(X^*)$. Nous pouvons ainsi optimiser notre algorithme en effectuant un calcul approché des fonds propres associés à X^* , à chaque itération horizontale. Avec cette technique, nous pouvons réduire le budget de calcul, en effectuant un calcul précis des fonds propres seulement lorsque cela est nécessaire.

a) Présentation des formes paramétriques locales

L'idée que nous avons choisie de développer pour effectuer un calcul approché des fonds propres sans consommer de budget de calcul, est, à chaque itération verticale m , de calibrer une forme paramétrique sur les scénarios obtenus à cette itération. Cette forme paramétrique sera dite locale, parce qu'au fur et à mesure des itérations verticales, elle sera calibrée sur des fonds propres de plus

en plus proches les uns des autres et de plus en plus faibles. Elle constitue ainsi une fonction proxy permettant d'obtenir les fonds propres à partir des facteurs de risque, et de fonctions simples des facteurs de risque.

Nous choisissons une forme paramétrique de degré maximal égal à 3. Nous considérons les puissances premières, secondes et troisièmes de chaque facteur de risque, et nous introduisons des termes croisés entre les facteurs de risque. Pour plus de simplicité dans les écritures, nous présentons ici la méthode dans le cas où nous avons deux facteurs de risque. Néanmoins, les résultats donnés sont également vérifiés lorsque nous avons plus de facteurs de risque. Ainsi, pour deux facteurs de risque, notre forme paramétrique est la suivante :

$$FP(X) = a_0 + a_1X_1 + a_2X_1^2 + a_3X_1^3 + a_4X_2 + a_5X_2^2 + a_6X_2^3 + a_7X_1X_2 + a_8X_1^2X_2 + a_9X_1X_2^2.$$

Nous définissons alors Y comme la matrice composée des vecteurs $Y_i = (1, X_{i,1}, X_{i,1}^2, X_{i,1}^3, X_{i,2}, X_{i,2}^2, X_{i,2}^3, X_{i,1}X_{i,2}, X_{i,1}^2X_{i,2}, X_{i,1}X_{i,2}^2)'$, et A comme le vecteur des coefficients a_0, \dots, a_9 . Nous notons également FP le vecteur formé à partir des fonds propres obtenus pour chaque scénario X_i . Nous pouvons alors réécrire la formule ci-dessus sous forme matricielle : $FP = Y'A$.

Notre objectif est, pour calibrer la forme paramétrique, de déterminer le vecteur A à partir des N scénarios composant notre échantillon à l'itération m et de leurs fonds propres associés. Nous pouvons déduire ce vecteur de l'égalité matricielle à l'aide des égalités suivantes :

$$YFP = YY'A$$

$$(YY')^{-1}YFP = A$$

Ainsi, nous pouvons calculer A à l'aide de Y et de FP , en posant : $A = (YY')^{-1}YFP$.

Une fois la forme paramétrique calibrée, nous pouvons l'utiliser à chaque itération horizontale associée à l'itération verticale m , pour avoir les fonds propres à chaque fois que nous appliquons le noyau de transition, sans effectuer de calculs ALM.

A la fin des T itérations horizontales, il est cependant nécessaire d'effectuer un calcul exact des fonds propres associés au scénario $X_{i^*}^m$ obtenu, pour être sûr que celui-ci soit calculé correctement et que nous n'introduisons pas d'erreur dans la détermination du maximum des fonds propres, lors des itérations verticales suivantes. En outre, cela permet également de vérifier que le scénario $X_{i^*}^m$ est bien inférieur au niveau L_m . Dans le cas contraire, nous recommençons la diffusion du scénario.

b) Adaptation de l'algorithme grâce à l'utilisation de formes paramétriques locales

Nous pouvons, à présent, donner un nouvel algorithme, utilisant la technique que nous venons de présenter, et qui permet de réduire le budget de calcul en limitant le nombre de simulations ALM. Pour distinguer facilement les fonds propres déterminés à partir de la forme paramétrique locale des fonds propres calculés à l'aide des simulations ALM, nous notons les premiers FP_{loc} . Nous pouvons ainsi reprendre l'algorithme présenté dans la partie III.A.2 et y ajouter les modifications dues à l'utilisation de la forme paramétrique locale :

Nous initialisons l'algorithme à l'aide d'un N -échantillon de scénarios, (X_1, X_2, \dots, X_N) , et nous posons : $X_1^0 = X_1, \dots, X_N^0 = X_N$ et $L_0 = +\infty$.

Nous posons ensuite $M = \left\lceil \frac{\log(0,5\%)}{\log(1-N^{-1})} \right\rceil$, puis, pour m allant de 1 à M , nous réalisons les calculs suivants :

- Nous calibrons notre forme paramétrique locale à l'aide des scénarios $X_1^{m-1}, \dots, X_N^{m-1}$ et des fonds propres associés $FP(X_1^{m-1}), \dots, FP(X_N^{m-1})$, en posant $A = (YY')^{-1}YFP$, où les notations sont celles introduites dans la partie précédente.
- Nous posons $L_m = \max\left(FP(X_1^{m-1}), \dots, FP(X_N^{m-1})\right)$.
- Nous notons i^* la valeur de i pour laquelle $FP(X_i^{m-1})$ est maximum.
- Nous piochons aléatoirement un scénario X_j^{m-1} dans l'ensemble $\{X_1^{m-1}, \dots, X_N^{m-1}\} \setminus \{X_{i^*}^{m-1}\}$, nous posons $X_{i^*}^m = X_j^{m-1}$, et nous diffusons ce scénario à l'aide d'un noyau de transition gaussien :
- Nous posons $t = 1$, puis nous appliquons la récurrence ci-dessous, tant que $t \leq T$:
 - Nous appliquons le noyau de transition aux facteurs de risque composant $X_{i^*}^m$. Nous obtenons alors un nouveau scénario, X^* .
 - Nous calculons $FP_{loc}(X^*)$ en utilisant la forme paramétrique locale.
 - Si $FP_{loc}(X^*) < L_m$, alors nous posons $X_{i^*}^m = X^*$ et nous augmentons t de 1.
 - A la fin des T itérations horizontales, nous calculons précisément $FP(X_{i^*}^m)$ en utilisant des simulations ALM, et nous vérifions la condition $FP(X_{i^*}^m) < L_m$. Si celle-ci n'est pas vérifiée, nous recommençons la diffusion du scénario $X_{i^*}^m$ à l'aide des T itérations horizontales et de la forme paramétrique locale.
- Pour tous les j différents de i^* , nous posons $X_j^m = X_j^{m-1}$.
- Nous obtenons ainsi un nouvel échantillon de scénarios, $\{X_1^m, \dots, X_N^m\}$, pour lesquels les fonds propres sont tous inférieurs à L_m .

A la fin des $M = \left\lceil \frac{\log(0,5\%)}{\log(1-N^{-1})} \right\rceil$ itérations verticales, nous estimons le quantile à 0,5 % des fonds propres en posant : $\hat{q}_{0,5\%}(FP) = L_M$.

c) Calcul de la complexité de l'algorithme

Comme les itérations horizontales ne nécessitent pas de calculs ALM, nous pouvons déterminer facilement le budget de calcul utilisé : pour initialiser l'algorithme, il est nécessaire d'avoir une estimation précise de la valeur des fonds propres associés aux scénarios de départ. Pour cela, nous générons $n_{secondaire}$ simulations secondaires pour chacun des N scénarios. Ensuite, nous effectuons $M = \left\lceil \frac{\log(0,5\%)}{\log(1-N^{-1})} \right\rceil$ itérations verticales et, pour chacune d'elles, nous devons effectuer un calcul exact de fonds propres, pour le scénario diffusé.

Ainsi le budget de calcul total, C , utilisé par cette méthode est le suivant :

$$C = N \times n_{secondaire} + \left\lceil \frac{\log(0,5\%)}{\log(1-N^{-1})} \right\rceil \times n_{secondaire}.$$

Nous verrons en pratique que les ordres de grandeurs des variables N et $n_{secondaire}$ sont respectivement égaux à 100 et 1 000. Ainsi, pour ces valeurs, nous obtenons un budget de calcul égal à 628 000, soit une réduction d'environ 99 % par rapport au SdS complet.

Remarque : La valeur de N est donnée à titre indicatif. Nous pourrions ainsi obtenir une valeur différente dans la partie de mise en œuvre de la méthode.

5. Propriétés asymptotiques de l'estimateur du quantile

Maintenant que nous avons présenté nos différentes méthodes pour estimer le quantile à 0,5 % des fonds propres, nous pouvons nous intéresser aux propriétés asymptotiques de cet estimateur. En particulier, nous cherchons ici à déterminer un intervalle de confiance, et nous comparons la précision de notre estimateur avec celle obtenue en utilisant la méthode de Monte Carlo.

a) Détermination de la loi asymptotique suivie par l'estimateur du quantile

Pour déterminer la loi asymptotique suivie par l'estimateur du quantile, nous procédons en plusieurs étapes : nous commençons par étudier la loi de la fonction de hasard cumulé des fonds propres, puis nous nous appuyons sur la création d'une nouvelle variable pour connaître la loi asymptotique de l'estimateur du quantile.

Loi suivie par la fonction de hasard cumulé des fonds propres à un an :

Pour étudier les propriétés de notre estimateur du quantile, nous allons nous intéresser à la fonction de hasard cumulé, notée $\Lambda(\cdot)$, et à la fonction de survie, $S(\cdot)$, de la variable aléatoire $-FP(X)$.

Pour commencer, nous allons montrer que la fonction de hasard cumulé appliquée à N variables aléatoires indépendantes, Y_1, \dots, Y_N , suit une loi exponentielle de paramètre 1, et plus précisément que :

$$(\Lambda(Y_1), \dots, \Lambda(Y_N)) \triangleq (E_1, \dots, E_N),$$

où E_1, \dots, E_N sont N variables aléatoires i.i.d. de loi exponentielle de paramètre 1.

Nous pouvons montrer que la fonction de hasard cumulé est une fonction croissante. En effet, la fonction de hasard $\lambda(\cdot)$ appliquée au point y est définie comme la probabilité instantanée de sortir d'un état à la date y sachant que le sujet est encore dans cet état en y , ce qui peut se traduire par la formule :

$$\lambda(y) = \lim_{\Delta_y \rightarrow 0} \frac{1}{\Delta_y} P(Y \in [y; y + \Delta_y] | Y \geq y).$$

Ainsi, cette fonction est positive, et par conséquent, la fonction de hasard cumulé, qui est égale à l'intégrale de la fonction de hasard, est croissante. Dans notre cas, Y sera la variable $-FP$ (opposé de la variable fonds propres), variable aléatoire à support réel dont nous pouvons supposer que la densité et la fonction de hasard sont toujours strictement positives. Nous pouvons alors en déduire les égalités suivantes :

$$P(Y_1 > y_1, \dots, Y_N > y_N) = P(Y_1 > y_1) \times \dots \times P(Y_N > y_N)$$

$$P(Y_1 > y_1, \dots, Y_N > y_N) = P(\Lambda(Y_1) > \Lambda(y_1)) \times \dots \times P(\Lambda(Y_N) > \Lambda(y_N))$$

Par ailleurs, nous savons que la fonction de survie et la fonction de hasard cumulé sont liées par la relation $S(y) = \exp(-\Lambda(y))$, ce qui nous permet d'écrire les égalités ci-dessous :

$$\begin{aligned} P(Y_1 > y_1, \dots, Y_N > y_N) &= P(Y_1 > y_1) \times \dots \times P(Y_N > y_N) \\ P(Y_1 > y_1, \dots, Y_N > y_N) &= S(y_1) \times \dots \times S(y_N) \\ P(Y_1 > y_1, \dots, Y_N > y_N) &= \exp(-\Lambda(y_1)) \times \dots \times \exp(-\Lambda(y_N)) \end{aligned}$$

En comparant les deux résultats ci-dessus, nous obtenons finalement l'égalité suivante :

$$P(\Lambda(Y_1) > \Lambda(y_1)) \times \dots \times P(\Lambda(Y_N) > \Lambda(y_N)) = \exp(-\Lambda(y_1)) \times \dots \times \exp(-\Lambda(y_N)).$$

Pour simplifier les notations, nous posons $\Lambda(Y_1) = E_1, \dots, \Lambda(Y_N) = E_N$ et $\Lambda(y_1) = e_1, \dots, \Lambda(y_N) = e_N$. Nous obtenons alors la relation ci-dessous :

$$P(E_1 > e_1) \times \dots \times P(E_N > e_N) = \exp(-e_1) \times \dots \times \exp(-e_N).$$

En particulier, E_1, \dots, E_N sont des variables aléatoires i.i.d. de loi exponentielle de paramètre 1. Finalement, la fonction de hasard cumulé suit une loi exponentielle.

Pour déterminer la loi suivie par les variables de la forme $\Lambda(-L_m)$, nous allons considérer l'itération verticale m et raisonner en sachant que les scénarios vérifient la condition $FP(X) < L_{m-1}$. Nous voulons ainsi déterminer la probabilité $P(FP(X) < L_m | FP(X) < L_{m-1})$, ou, écrit autrement, $P(-FP(X) > -L_m | -FP(X) > -L_{m-1})$. Pour cela, nous pouvons exprimer les fonctions de survie conditionnelles de la variable aléatoire $-FP(X)$ à l'aide de la fonction de survie $S()$:

$$P(-FP(X) > -L_m | -FP(X) > -L_{m-1}) = \frac{S(-L_m)}{S(-L_{m-1})}$$

$$P(-FP(X) > -L_m | -FP(X) > -L_{m-1}) = \frac{\exp(-\Lambda(-L_m))}{\exp(-\Lambda(-L_{m-1}))}$$

$$P(-FP(X) > -L_m | -FP(X) > -L_{m-1}) = \exp(-(\Lambda(-L_m) - \Lambda(-L_{m-1})))$$

Si nous notons $\Lambda_m(-L_m)$ la fonction de hasard cumulé conditionnelle pour l'itération verticale m , nous obtenons :

$$\Lambda_m(-L_m) = \Lambda(-L_m) - \Lambda(-L_{m-1}) \quad (1)$$

Par ailleurs, à chaque itération verticale, L_m est définie comme le maximum des fonds propres. Nous pouvons ainsi écrire les égalités suivantes :

$$\Lambda_m(-L_m) = \Lambda(-\max_{1 \leq i \leq N} FP(X_i^m)) = \Lambda(\min_{1 \leq i \leq N} -FP(X_i^m)) = \min_{1 \leq i \leq N} \Lambda(-FP(X_i^m)).$$

Or nous avons prouvé précédemment que les variables $\Lambda(-FP(X_i^m))$ suivaient une loi exponentielle de paramètre 1, et nous savons que le minimum de N variables aléatoires de telles lois exponentielles est une variable de loi exponentielle de paramètre N . Nous pouvons ainsi écrire $\Lambda_m(-L_m)$ sous la forme $\Lambda_m(-L_m) = \frac{E'_m}{N}$, où E'_m est une variable aléatoire de loi exponentielle de paramètre 1.

Nous déduisons alors de l'expression (1) et de la formule ci-dessus l'égalité en loi suivante :

$$\Lambda(-L_m) \triangleq \frac{1}{N} \sum_{j=1}^m E_j'.$$

Par conséquent, les variables $T_1 = \Lambda(-L_1), \dots, T_M = \Lambda(-L_M)$ peuvent être vues comme les temps d'arrivée d'un processus de poisson de paramètre N .

Loi asymptotique suivie par l'estimateur du quantile :

Un estimateur du quantile à 0,5 % des fonds propres est donné par $\hat{q} = -S^{-1}\left(\exp\left(-\frac{1}{N}G_M\right)\right)$, où G_M est défini par la formule $G_M = N(T_1 + \dots + T_M)$. Pour des détails sur cet estimateur, nous renvoyons le lecteur à l'article de Guyader et al. (2011).

Nous allons, à présent, déterminer la loi asymptotique suivie par notre estimateur. Nous pouvons remarquer que comme les variables T_1, \dots, T_M sont les temps d'arrivée d'un processus de poisson de paramètre N , G_M suit une loi Gamma de paramètres 1 et M (cf cours de Chabriac et cours de Lebarbier et Robin sur les processus de poisson). Ainsi, l'application du théorème central limite à la variable G_M donne le résultat suivant :

$$\sqrt{M}\left(\frac{G_M}{M} - 1\right) \xrightarrow{\mathcal{L}} \mathcal{N}(0,1), \text{ quand } M \text{ tend vers l'infini.}$$

Nous rappelons par ailleurs que $M = \left\lceil \frac{\ln(0,5\%)}{\ln(1-N^{-1})} \right\rceil$ et que par conséquent M tend vers l'infini lorsque N tend vers l'infini. Nous en déduisons ainsi que $\sqrt{M}\left(\frac{G_M}{M} - 1\right) \rightarrow \mathcal{N}(0,1)$, en loi quand N tend vers l'infini.

Nous pouvons alors nous servir de ce résultat pour connaître la limite, quand N tend vers l'infini, de $\sqrt{N} \times \left(\frac{G_M}{N} - (-\ln(0,5\%))\right)$. En effet, le terme $\sqrt{M}\left(\frac{G_M}{M} - 1\right)$ peut s'exprimer autrement :

$$\sqrt{M}\left(\frac{G_M}{M} - 1\right) = \sqrt{\frac{N}{M}} \times \sqrt{N} \times \left(\frac{G_M}{N} - \frac{M}{N}\right) = \sqrt{\frac{N}{M}} \left(\sqrt{N} \times \left(\frac{G_M}{N} - (-\ln(0,5\%))\right) \right) - \frac{N}{\sqrt{M}} \left(\ln(0,5\%) + \frac{M}{N} \right).$$

Nous pouvons démontrer (cf annexe 2) que nous obtenons les deux limites ci-dessous :

$$\lim_{N \rightarrow \infty} \frac{N}{\sqrt{M}} \left(\ln(0,5\%) + \frac{M}{N} \right) = 0 \text{ et } \lim_{N \rightarrow \infty} \sqrt{\frac{M}{N}} = \sqrt{-\ln(0,5\%)}. \text{}$$

Ainsi, nous en déduisons la convergence de $\sqrt{N} \times \left(\frac{G_M}{N} - (-\ln(0,5\%))\right)$ vers une variable aléatoire de loi $\sqrt{-\ln(0,5\%)} \times \mathcal{N}(0,1)$, ce qui peut se réécrire sous la forme suivante :

$$\sqrt{N} \times \left(\frac{G_M}{N} - (-\ln(0,5\%))\right) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, -\ln(0,5\%)\right), \text{ quand } N \text{ tend vers l'infini.}$$

Pour connaître la loi asymptotique de l'estimateur du quantile, nous appliquons alors la méthode delta à la fonction $g(x) = -S^{-1}(\exp(-x))$ au point $-\ln(0,5\%)$. En effet, nous rappelons que d'après la méthode delta, si Z est une variable aléatoire, et que nous avons la convergence $r_n \times (\hat{\theta} - \theta) \rightarrow Z$ en loi, alors $r_n \times (g(\hat{\theta}) - g(\theta)) \rightarrow g'(\theta) \times Z$. Dans notre cas, en notant f la densité de la variable aléatoire $-FP(X)$, nous obtenons la convergence suivante :

$$\sqrt{N} \times (\hat{q} - q) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \frac{-(0,5\%)^2 \times \ln(0,5\%)}{f(-q)^2}\right), \text{ en loi quand } N \text{ tend vers l'infini.}$$

Dès lors, il est possible de construire un intervalle de confiance asymptotique de l'estimateur \hat{q} .

b) Intervalle de confiance asymptotique de l'estimateur du quantile

Dans la partie précédente, nous avons déterminé la loi asymptotique suivie par l'estimateur du quantile à 0,5 % des fonds propres à un an. Nous pouvons utiliser ce résultat pour déterminer un intervalle de confiance asymptotique du quantile. Cependant, cet intervalle de confiance reposerait sur la densité de la variable aléatoire $-FP(X)$, qui nous est inconnue. Aussi, nous allons plutôt déterminer un intervalle de confiance de niveau α de M , puis en déduire un intervalle de confiance de $L_M = \hat{q}$.

Nous avons précédemment noté que les variables $T_1 = \Lambda(-L_1), \dots, T_M = \Lambda(-L_M)$ pouvaient être vues comme les temps d'arrivée d'un processus de poisson de paramètre N . Par ailleurs, nous avons également :

$$\begin{aligned} M &= \max\{m \mid -L_m < -q\} \\ M &= \max\{m \mid \Lambda(-L_m) < \Lambda(-q)\} \\ M &= \max\{m \mid T_m < -\ln S(-q)\} \\ M &= \max\{m \mid T_m < -\ln(0,5 \%) \} \end{aligned}$$

Nous en déduisons donc, plus précisément, que les variables $T_1 = \Lambda(-L_1), \dots, T_M = \Lambda(-L_M)$ sont les temps d'arrivée d'un processus de poisson de paramètre N jusqu'au temps $-\ln(0,5 \%)$. Ainsi, le nombre d'arrivées, M , peut s'interpréter comme une variable aléatoire de loi de poisson de paramètre $-N(\ln(0,5 \%))$. En outre, lorsque le paramètre d'une loi de poisson tend vers l'infini, nous pouvons approximer cette loi par une loi normale de moyenne et de variance égales au paramètre de la loi de poisson. Ainsi, nous obtenons asymptotiquement que la loi suivie par M est la suivante :

$$M \sim \mathcal{N}\left(-N(\ln(0,5 \%)), -N(\ln(0,5 \%))\right).$$

Nous pouvons alors en déduire un intervalle de confiance asymptotique de niveau α de M . En effet, en notant $z_{1-\frac{\alpha}{2}}$ le quantile d'ordre $1 - \frac{\alpha}{2}$ d'une loi normale centrée réduite, nous obtenons les égalités ci-dessous :

$$\begin{aligned} P\left(-z_{1-\frac{\alpha}{2}} \leq \frac{M - (-N(\ln(0,5 \%)))}{\sqrt{-N(\ln(0,5 \%))}} \leq z_{1-\frac{\alpha}{2}}\right) &= 1 - \alpha \\ P\left(-z_{1-\frac{\alpha}{2}} \sqrt{-N(\ln(0,5 \%))} - N(\ln(0,5 \%)) \leq M \leq z_{1-\frac{\alpha}{2}} \sqrt{-N(\ln(0,5 \%))} - N(\ln(0,5 \%))\right) &= 1 - \alpha \end{aligned}$$

Cette dernière égalité nous fournit ainsi un intervalle de confiance asymptotique de M de niveau α :

$$\left[-z_{1-\frac{\alpha}{2}} \sqrt{-N(\ln(0,5 \%))} - N(\ln(0,5 \%)); z_{1-\frac{\alpha}{2}} \sqrt{-N(\ln(0,5 \%))} - N(\ln(0,5 \%))\right].$$

Enfin, il nous suffit de poser $M_- = -z_{1-\frac{\alpha}{2}} \sqrt{-N(\ln(0,5 \%))} - N(\ln(0,5 \%))$ et $M_+ = z_{1-\frac{\alpha}{2}} \sqrt{-N(\ln(0,5 \%))} - N(\ln(0,5 \%))$, puis de prendre $[L_{M_-}; L_{M_+}]$ comme intervalle de confiance asymptotique de L_M .

c) Gain par rapport à un Monte Carlo classique

Nous allons à présent comparer ces résultats avec ceux obtenus lors de l'utilisation d'un Monte Carlo. Nous rappelons que si nous avons utilisé une méthode de Monte Carlo, nous aurions simulé N scénarios, puis nous aurions calculé les fonds propres associés. Ensuite, nous aurions classé ces fonds propres de la manière suivante : $FP(X)_{(1)} < \dots < FP(X)_{(N)}$, et enfin nous aurions estimé le quantile à 0,5 % en posant $\hat{q}_{MC} = FP(X)_{([0,5\% \times N])}$.

Nous pouvons montrer (cf annexe 3) que la loi suivie asymptotiquement par cet estimateur est donnée par $\sqrt{N}(\hat{q}_{MC} - q) \rightarrow \mathcal{N}\left(0, \frac{0,5\%(1-0,5\%)}{f(-q)^2}\right)$, où f est la densité de la variable aléatoire $-FP(X)$. Par ailleurs, nous rappelons que celle suivie par notre estimateur du quantile est de la forme $\mathcal{N}\left(0, \frac{-(0,5\%)^2 \times \ln(0,5\%)}{f(-q)^2}\right)$. Nous remarquons ainsi que notre estimateur possède une variance $-0,5\% \times \ln(0,5\%)$ fois plus faible que celle de l'estimateur obtenu par la méthode de Monte Carlo. Notre estimateur est donc plus précis.

Nous pouvons, en outre, noter que pour atteindre le même degré de précision avec une méthode de Monte Carlo, il serait nécessaire de simuler environ $\sqrt{\frac{1}{-0,5\% \times \ln(0,5\%)}} = 6,1$ fois plus de scénarios.

Dans cette partie, nous avons présenté le principe des méthodes de décomposition d'un événement rare en événements moins rares, et nous les avons adaptées au calcul du capital économique. Nous allons, maintenant, étudier une deuxième approche, fondée sur une méthode d'apprentissage statistique : les forêts aléatoires.

B. Forêts aléatoires

Dans cette partie, nous allons présenter une autre méthode d'estimation du capital économique. Alors que la méthode précédente estime le quantile réel des fonds propres, nous allons ici utiliser une méthode d'apprentissage statistique, pour optimiser le temps de calcul du quantile empirique à 0,5 % des fonds propres à un an associé à un jeu de scénarios économiques donnés. Cette méthode a donc pour but d'accélérer le calcul d'un estimateur du quantile empirique.

Nous présenterons, dans un premier temps, les arbres de décision puis les forêts aléatoires, et enfin des algorithmes de calcul de ce quantile.

1. Présentation des arbres de décision

L'objet de cette partie est de définir les arbres de décision, d'en présenter les principales méthodes de construction, ainsi que leurs limites.

a) Définition

Un arbre de décision est un outil non paramétrique d'aide à la décision qu'il est facile de représenter à l'aide d'un graphique sous la forme d'un arbre multinomial où apparaissent à l'extrémité de chaque

branche, les différents résultats possibles, en fonction des décisions prises à chaque étape ou nœud de l'arbre.

La construction d'un arbre se fait par segmentations successives de l'échantillon de départ, suivant un certain nombre de règles qui portent sur les variables explicatives. Chaque segmentation divise l'échantillon en plusieurs sous-ensembles. Dans la suite, nous n'étudierons que les arbres binaires : arbres pour lesquels chaque segmentation provient d'un test binaire.

Un exemple d'arbre est présenté dans l'encadré ci-dessous.

Exemple pratique :

Soit A un ensemble de points de $E = \mathbb{R}_+^2$. Notons les coordonnées des points de (ϵ_1, ϵ_2) .

Supposons que nous connaissions la position des éléments de A par rapport à une courbe \mathcal{C} inconnue séparant l'espace E en deux parties (s'ils sont au-dessus ou au-dessous de \mathcal{C}), et que nous souhaitions pouvoir prédire la position de tout point de E par rapport celle-ci.

Nous pourrions alors définir un arbre de décision et le représenter de la manière suivante :

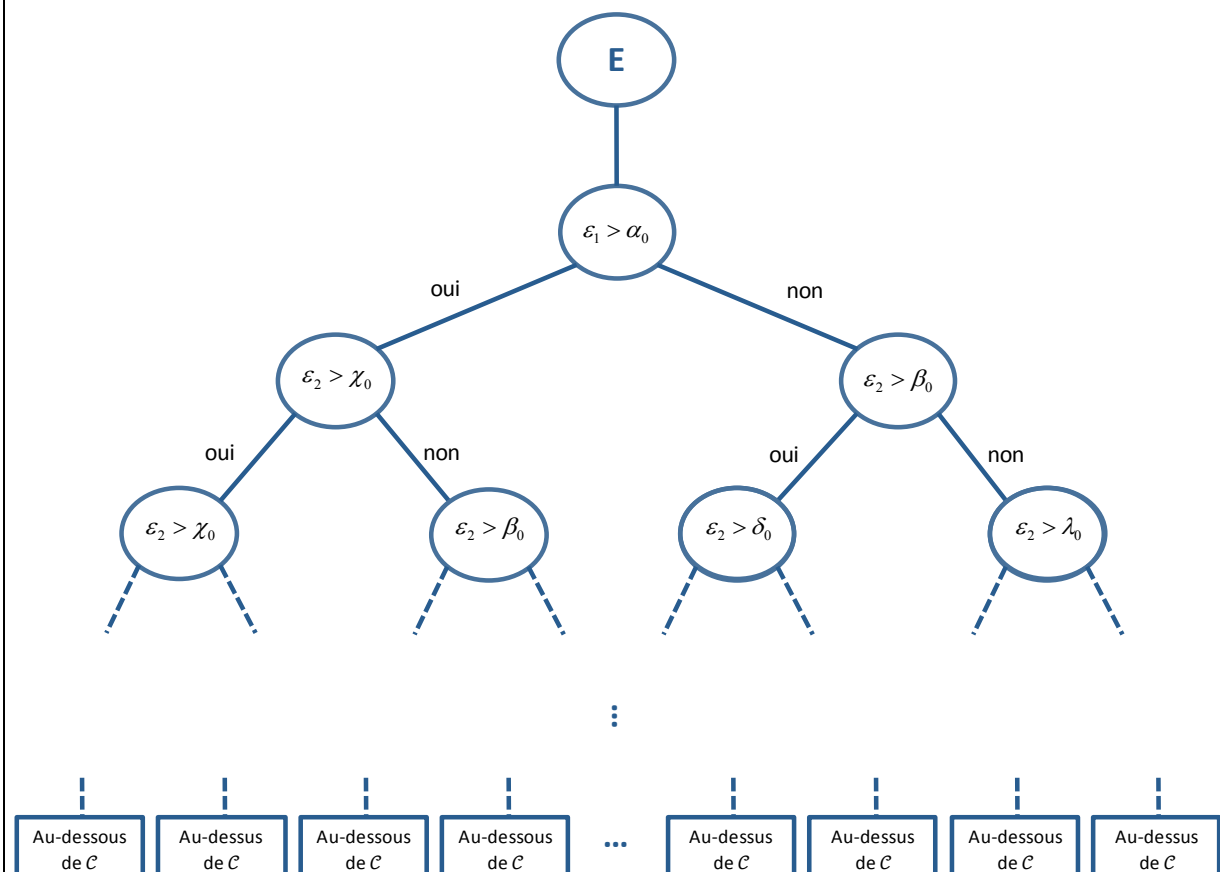


Figure 8 : Représentation graphique d'un arbre de décision

Sur le graphique ci-dessus, chaque nœud est une règle de décision permettant de séparer l'ensemble auquel elle s'applique en deux sous-ensembles. Les rectangles sont les sous-ensembles finaux, ou feuilles, de l'arbre de décision indiquant l'estimation de la position du point par rapport à \mathcal{C} . Ainsi pour tout élément de \mathbb{R}_+^2 , nous pouvons descendre l'arbre pour le placer dans une des feuilles de façon unique.

Deux types d'arbres binaires existent : ceux dont l'approche est dite orthogonale (dans l'encadré ci-dessus, et dans le schéma ci-dessous, à gauche), et ceux dont l'approche est dite oblique (dans le schéma ci-dessous, à droite). L'approche orthogonale consiste à utiliser une suite de nœuds ne portant que sur une variable explicative à la fois, tandis que l'approche oblique utilise des nœuds portant sur des combinaisons linéaires de variables explicatives.

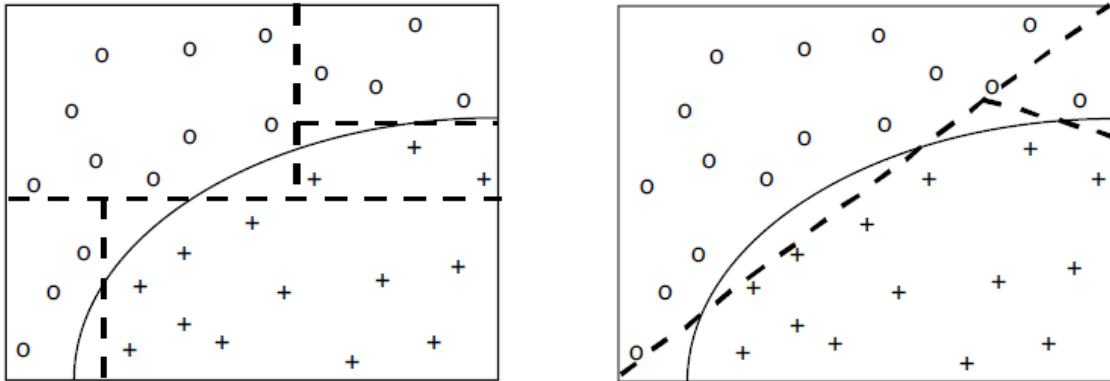


Figure 9 : Schéma représentant les segmentations d'un arbre orthogonal et d'un arbre oblique

Les figures ci-dessus représentent les règles de décision en pointillés, les points étant les croix et ronds, les variables explicatives étant l'abscisse et l'ordonnée, et la variable à expliquer étant la nature du point : rond ou croix. Ces arbres permettent d'estimer la nature d'un point à partir de son abscisse et de son ordonnée.

Les arbres obliques ont peu été étudiés dans la littérature scientifique publique, c'est pourquoi dans la suite nous ne nous intéresserons qu'aux arbres orthogonaux.

Nous allons maintenant développer les méthodes de construction d'un arbre de décision.

b) Construction d'un arbre de décision

Soit un vecteur aléatoire (X, FP) , avec X , vecteur de taille p , et FP , une variable aléatoire réelle.

Inputs de l'algorithme

L'algorithme de construction d'un arbre de décision nécessite les éléments suivants :

- Un critère permettant de sélectionner le meilleur test possible à chaque nœud, exemple : l'entropie de Shannon (algorithme C4.5), ou indice de Gini (algorithme CART).
- Des critères d'arrêt permettant de décider qu'un nœud est terminal (que c'est une feuille) :
 - le nombre minimal d'éléments par feuille,
 - la profondeur maximale de l'arbre,
 - le nombre maximal de feuilles.
- Un échantillon de variables aléatoires $(X_i, FP_i)_{i=1..N_{obs}}$, indépendantes et de même loi que (X, FP) . Cet échantillon est appelé base de calibrage.

Principe de l'algorithme

Nous allons présenter la construction d'un arbre de décision, dont l'objectif est prédire la variable réponse FP sachant $X = x$; autrement dit, nous allons estimer $FP|X = x$.

La construction d'un arbre de décision nécessite plusieurs étapes :

1. **Elaboration d'un arbre par méthode récursive,**
 - a. Pour construire un nœud nous recherchons le test le plus discriminant.
 - b. Le nœud ainsi défini, nous séparons l'échantillon mère en deux sous-échantillons filles.
 - c. Nous itérons la procédure tant qu'un critère d'arrêt n'est pas satisfait.
2. **Elagage (pruning)** : Etape facultative (mais très souvent nécessaire) consistant à supprimer des branches non significatives pour améliorer l'interprétation de l'arbre sans trop détériorer la performance prédictive. Cela permet aussi d'éviter le sur-apprentissage. L'élagage peut se faire aussi en testant l'arbre sur des données non-utilisées dans l'étape 1 : méthode *in sample/out of sample*.

Les deux principaux algorithmes de construction d'arbres sont les algorithmes CART (*Classification and Regression Tree*) et C5.0 (et versions antérieures). Leurs principales différences résident dans la détermination du test optimal à chaque nœud, et dans le critère d'élagage.

Que l'on se place dans le cadre de l'algorithme CART ou du C5.0, la recherche du test le plus discriminant à chaque nœud se fait de la même manière. Une fonction mesurant la qualité d'un test est définie, dépendant de la variable discriminante et du seuil choisi, et un algorithme de maximisation calcule le test ayant la meilleure qualité.

Nous allons maintenant détailler les principales fonctions mesurant la qualité.

L'entropie de Shannon : C4.5

Cet algorithme, développé par Quinlan et al. (1993), permet de faire des classifications.

Soient A_1 et A_2 les deux partitions de A issues du nœud étudié, et $q = \frac{N_{A_1}}{N_A}$, la proportion des

éléments de A_1 parmi ceux présents dans A . La qualité d'un test de classification se mesure par la

fonction suivante :

$$I_E(A_1, A_2) = E(A) - qE(A_1) - (1 - q)E(A_2),$$

avec $E()$ la fonction d'entropie de Shannon :

$$E(A) = \sum_{k, \text{classe}} -p_{k,A} * \log p_{k,A},$$

avec $p_{k,A}$ la proportion d'éléments de la classe k , dans A

L'algorithme C5.0 et les versions antérieures permettent, ainsi, de réaliser facilement des classifications.

L'indice de Gini : algorithme CART

L'algorithme CART (*Classification And Regression Tree*) a été inventé par Breiman et al. en 1984, cependant les fonctions associées aux classifications et aux régressions diffèrent légèrement.

Soient A_1 et A_2 les deux partitions d'un ensemble A , issues du nœud étudié, et $q = \frac{N_{A_1}}{N_A}$, la proportion des éléments de A_1 parmi ceux présents dans A .

Dans les deux cas, La qualité d'un test, se mesure par la fonction suivante :

$$I_G(A_1, A_2) = G(A) - qG(A_1) + (1 - q)G(A_2).$$

Les deux méthodes diffèrent dans la définition de la fonction G .

Dans le cas des classifications, G est définie par l'indice de Gini :

$$G(A) = \sum_{k \neq l} p(k|A)p(l|A),$$

avec k et l deux classes, et $p(k|A)$ la proportion d'éléments de la classe k dans A .

Si nous nous restreignons à deux classes, alors l'indice de Gini devient :

$$G(A) = 2 p(1|A) (1 - p(1|A)),$$

Ainsi $G(A) \in [0; 0,5]$, et plus $G(A)$ est faible, meilleur est le facteur discriminant du nœud.

Dans le cas de régressions, l'objectif est de minimiser la variance "intra-classe" de la variable réponse FP :

$$G(A) = V^A(FP)$$

$$\text{avec } V^A(FP) = \frac{1}{\#A} \sum_{i|x_i \in A} (FP_i - \overline{FP_A})^2, \text{ et } \overline{FP_A} = \frac{1}{\#A} \sum_{i|x_i \in A} FP_i$$

Les sous-ensembles obtenus de cette manière ont donc des valeurs de fonds propres homogènes.

Il n'est pas nécessaire d'étudier les méthodes d'élagage dans la mesure où les forêts aléatoires évitent le sur-apprentissage (nous étudierons ce phénomène plus loin).

Désormais nous nous restreignons aux cas où les variables explicatives et la variable réponse sont quantitatives. Nous utiliserons donc l'algorithme CART pour les régressions.

Remarque : Dans la mesure où nous travaillons sur des données quantitatives, nous pourrions penser que le nombre de couples (variable discriminante, seuil) à tester est infini, mais la base de calibrage étant fini, seul un nombre fini de tels couples est susceptibles d'être retenus. En effet, pour n observations d'une variable explicative fixée, $(\epsilon_i)_{i=1..n}$, nous pouvons restreindre les seuils à étudier à l'ensemble suivant :

$$\text{seuils à étudier} = \left\{ k \mid \min_{i, \epsilon_i \leq k} (k - \epsilon_i) = \min_{i, \epsilon_i > k} (\epsilon_i - k) \right\}$$

Cet ensemble est de taille au plus $n - 1$.

Nous remarquerons que l'algorithme CART permet de faire des régressions et des classifications tandis que C5.0 se restreint aux classifications. Nous allons donc nous intéresser à CART dans la suite.

c) *Limites et variantes*

Le premier risque inhérent aux arbres de décision est le sur-apprentissage (capture des informations présentes dans la base de calibrage dans leur intégralité et non dans leur généralité). Il devient alors plus difficile de généraliser la fonction réponse pour des valeurs autres que celles de la base de calibrage. Cependant il est possible d'éviter (en partie) ce phénomène en « élaguant » l'arbre (cette méthode est souvent appelée *Pruning*).

L'autre inconvénient majeur des méthodes d'apprentissage par arbres de décision est leur forte sensibilité à la base de calibrage. La conséquence est que ces algorithmes d'apprentissage par arbres de décision nuisent à la qualité de l'apprentissage.

Cependant des méthodes récentes permettent de considérablement améliorer la robustesse des classifications et régressions obtenues, ainsi que de réduire le risque de sur-apprentissage. Les forêts aléatoires sont une de ces méthodes.

2. *Présentation des forêts aléatoires*

L'objet de cette partie est de présenter les forêts aléatoires ainsi que leur méthode de construction. Nous verrons, dans un premier temps, la méthode du Bagging dont sont inspirées les forêts aléatoires, et dans un deuxième temps, nous étudierons les forêts aléatoires.

a) *Preliminaires : le Bagging*

Le *Bagging* (pour **Bootstrap Aggregating**) est un metamodelle introduit par Breiman (1994). Dans cette partie, nous allons présenter cette méthode, ainsi que les principaux résultats exposés dans l'article.

Le *Bagging* est une amélioration de l'algorithme de construction des arbres de décision qui consiste à faire un *Bootstrap* sur nos données, puis à agréger les arbres construits sur chaque échantillon *Bootstrap*.

Décrivons de manière plus formelle cet algorithme.

Notons \mathcal{L} l'échantillon d'apprentissage (ou base de calibrage) composé de n observations $(X_i, FP_i)_{i=1..n}$ indépendantes et identiquement distribuées, de loi (X, FP) inconnue.

Notons $\varphi(x, \mathcal{L})$ l'estimation de $E[FP|X = x]$ faite par l'arbre de décision calibré sur \mathcal{L} .

Il s'agit de construire des arbres légèrement décorrélés, mais dont l'espérance des estimations est égale aux estimations de $\varphi(\cdot, \mathcal{L})$. Pour cela, nous créons B échantillons Bootstrapp de \mathcal{L} , de taille n , et tiré avec remise.

Nous obtenons ainsi B échantillons, chacun composés de n observations $(X_i, FP_i)_{i=1..n}$ indépendantes et identiquement distribuées, de loi (X, FP) , notés $(\mathcal{L}_i)_{i=1..B}$.

L'agrégation des arbres obtenus sur chacune des bases de calibrage $(\mathcal{L}_i)_{i=1..B}$ se fait de la manière suivante :

$$\text{régression} : \varphi^B(x, (\mathcal{L}_i)_{i=1..B}) = \frac{1}{B} \sum_{i=1}^B \varphi(x, \mathcal{L}_i),$$

$$\text{classification} : \varphi^B(x, (\mathcal{L}_i)_{i=1..B}) = \operatorname{argmax}_{\text{classe } c} \sum_{i=1}^B 1_{\{\varphi(x, \mathcal{L}_i) \in c\}}.$$

Les arbres sont donc agrégés en calculant la moyenne de leur estimation, dans le cas d'une régression, et par décision majoritaire, dans le cas d'une classification.

Étudions formellement l'impact du *Bagging* sur les erreurs de prédiction dans le cas d'une régression.

En notant P la loi du vecteur (X, FP) , par la loi des grands nombres, nous avons l'égalité suivante :

$$\varphi^\infty(x, P) \triangleq \lim_{B \rightarrow \infty} \varphi^B(x, P) = E_{\mathcal{L}}[\varphi(x, \mathcal{L})] \text{ p.s.}$$

Soit (X_0, FP_0) de loi P et indépendant de \mathcal{L} . La moyenne des carrés des erreurs de prédiction de l'arbre de décision calibré sur \mathcal{L} s'écrit alors :

$$e \triangleq E_{\mathcal{L}} \left[E_{X_0, FP_0} \left[(FP_0 - \varphi(X_0, \mathcal{L}))^2 \right] \right].$$

De la même manière, nous pouvons écrire la moyenne des carrés des erreurs de prédiction de la méthode du *Bagging* e_B , ainsi que la limite de cette erreur e_∞ :

$$e_B \triangleq E_{X_0, FP_0} \left[(FP_0 - \varphi^B(X_0, P))^2 \right],$$

$$e_\infty \triangleq E_{X_0, FP_0} \left[(FP_0 - \varphi^\infty(X_0, P))^2 \right].$$

En développant l'expression de e , nous obtenons :

$$e = E_{\mathcal{L}} \left[E_{X_0, FP_0} [FP_0^2 - 2FP_0 \varphi(X_0, \mathcal{L}) + \varphi(X_0, \mathcal{L})^2] \right],$$

$$e = E_{X_0, FP_0} [FP_0^2] - 2E_{\mathcal{L}} \left[E_{X_0, FP_0} [FP_0 \varphi(X_0, \mathcal{L})] \right] + E_{\mathcal{L}} \left[E_{X_0, FP_0} [\varphi(X_0, \mathcal{L})^2] \right],$$

Nous pouvons appliquer le théorème de Fubini :

$$e = E_{X_0, FP_0} [FP_0^2] - 2E_{X_0, FP_0} \left[E_{\mathcal{L}} [FP_0 \varphi(X_0, \mathcal{L})] \right] + E_{X_0, FP_0} \left[E_{\mathcal{L}} [\varphi(X_0, \mathcal{L})^2] \right].$$

Puis, par l'inégalité de Jensen, nous obtenons :

$$e \geq E_{X_0, FP_0} [FP_0^2] - 2E_{X_0, FP_0} \left[E_{\mathcal{L}} [FP_0 \varphi(X_0, \mathcal{L})] \right] + E_{X_0, FP_0} \left[E_{\mathcal{L}} [\varphi(X_0, \mathcal{L})^2] \right],$$

$$e \geq E_{X_0, FP_0} [FP_0^2] - 2E_{X_0, FP_0} [FP_0 \varphi^\infty(X_0, P)] + E_{X_0, FP_0} [\varphi^\infty(x, P)^2],$$

$$e \geq E_{X_0, FP_0} \left[(FP_0 - \varphi^\infty(X_0, P))^2 \right],$$

$$e \geq e_\infty.$$

Ainsi, le *Bagging* permet de réduire la moyenne des carrés des erreurs de prédiction, l'amélioration est d'autant plus grande que $E_{\mathcal{L}}[\varphi(X_0, \mathcal{L})^2]$ est supérieur à $\varphi^\infty(x, P)^2$. L'estimation est, elle, en moyenne la même qu'avec des arbres de décision classiques. Cela est dû au caractère identiquement distribué des arbres agrégés dans le *Bagging*.

La principale limite du *Bagging* réside dans l'interprétation de celui-ci. En effet il est alors nécessaire d'agréger l'ensemble des arbres ainsi construits pour représenter de façon graphique les résultats obtenus. Cependant cette méthode bénéficie d'un plus faible risque de sur-apprentissage que les arbres de décision classiques.

Remarques :

- L'élagage n'est plus nécessaire : le sur-apprentissage est évité grâce à la décorrélation des arbres. Plus les arbres sont profonds et plus ils seront décorrélés.

- Il est toujours possible de jumeler cette technique avec une approche *in-sample/out-of-sample*, pouvant même être améliorée en donnant des poids plus importants aux meilleurs arbres lors de l'agrégation (méthode de type *Boosting*). Cette approche ne sera pas détaillée dans ce mémoire⁵.

b) Les forêts aléatoires

La méthode des forêts aléatoires a été introduite par Breiman en 2001. La différence avec le *Bagging*, présenté plus haut, se situe dans la construction de chaque arbre *Bootstrap* : à chaque nœud, nous choisissons aléatoirement m (avec $m \leq d$) variables explicatives (au lieu des d variables initiales) que nous mettons en compétition pour déterminer la règle de décision du nœud.

Les forêts aléatoires agrègent donc des arbres moins corrélés que dans le cas du *Bagging*, ce qui permet de réduire encore la variance du résultat obtenu.

Nous allons, maintenant, essayer de quantifier le gain de variance obtenu par l'utilisation des forêts aléatoires. Pour faciliter les calculs, nous ferons deux hypothèses :

- les arbres de la forêt ont tous la même variance σ^2 (cette variance est issue de la méthode du *Bagging* appliquée à chaque arbre),
- la corrélation entre deux arbres est invariante. Nous la noterons ρ .

Notons $\varphi_m^{RF}(x, (\mathcal{L}_i)_{i=1..B})$ la prédiction de x faite par la forêt aléatoire (avec m variables mises en concurrence), et $(\varphi_m^i(x, \mathcal{L}_i))_{i=1..B}$ les arbres la composant.

La variance empirique de la forêt aléatoire, composée de B arbres, s'écrit alors :

$$Var(\varphi_m^{RF}(x, (\mathcal{L}_i)_{i=1..B})) = Var\left(\frac{1}{B} \sum_{i=1}^B \varphi_m^i(x, \mathcal{L}_i)\right) = \frac{1}{B^2} \sum_{i=1}^B \sum_{j=1}^B Cov(\varphi_m^i(x, \mathcal{L}_i), \varphi_m^j(x, \mathcal{L}_j))$$

$$Var(\varphi_m^{RF}(x, (\mathcal{L}_i)_{i=1..B})) = \frac{1}{B^2} \sum_{i=1}^B Var(\varphi_m^i(x, \mathcal{L}_i)) + \frac{1}{B^2} \sum_{j \neq i}^B Cov(\varphi_m^i(x, \mathcal{L}_i), \varphi_m^j(x, \mathcal{L}_j))$$

$$Var(\varphi_m^{RF}(x, (\mathcal{L}_i)_{i=1..B})) = \frac{1}{B^2} \sum_{i=1}^B \sigma^2 + \frac{B(B-1)\sigma^2\rho}{B^2}$$

$$Var(\varphi_m^{RF}(x, (\mathcal{L}_i)_{i=1..B})) = \sigma^2\rho + \frac{\sigma^2(1-\rho)}{B}$$

Nous obtenons une expression similaire dans le cas du *Bagging* (avec une corrélation et une variance des arbres adéquates). En supposant que le paramètre B est suffisamment grand, seul le terme $\sigma^2\rho$ est significatif. La performance des forêts aléatoires sur la méthode du *Bagging*, dépend donc de la baisse de corrélation des arbres et de l'augmentation de la variance des arbres.

Un faible nombre de variables mises en concurrence m impliquera une faible corrélation ρ , mais une grande variance σ . Si m est égal au nombre total de variables explicatives, les forêts aléatoires sont équivalentes à la méthode du *Bagging*. Il s'agit donc de choisir m judicieusement. Breiman et al.

⁵ Pour plus d'informations, le lecteur pourra se reporter à Schapire et al. (1999).

(2001) conseille de fixer $m = \sqrt{d}$ pour les classifications et $m = d/3$ pour les régressions. Cependant la valeur optimale de ce paramètre dépend du problème étudié et des inputs.

Remarque : Il est possible de générer des forêts d'arbres orthogonaux ou obliques, cependant la littérature s'est principalement concentrée sur les forêts orthogonales. Nous utiliserons donc ces dernières.

Une utilisation possible des arbres de décision (ainsi que leurs améliorations : forêts aléatoires, etc.) est leur introduction dans des algorithmes déjà existants. Par exemple, nous pouvons envisager d'utiliser les forêts aléatoires comme outil de régression dans le cas du *Kriging*. Néanmoins, il peut être plus judicieux de construire un algorithme optimisant l'intégration des arbres de décision.

Trois algorithmes ont été étudiés. Le premier est une application directe des forêts aléatoires vues comme un simple outil de régression non-paramétrique. Le deuxième tend à améliorer les résultats en ajoutant des points de calibrage choisis de façon simple. Enfin le troisième algorithme ajoute des points de calibrage de manière plus judicieuse.

3. Premier algorithme : application directe de la méthode

Dans un premier temps nous avons étudié une application relativement simple des forêts aléatoires. Cette approche consiste à utiliser une forêt calibrée sur un petit nombre de points de calibrage pour estimer les valeurs des fonds propres pour une plus grande base, appelée « base totale », et à en déduire les scénarios les plus risqués. Une fois ces scénarios identifiés, il suffit de déterminer leurs fonds propres exacts en lançant un grand nombre de scénarios secondaires.

Voici une idée d'adaptation possible des forêts aléatoires à notre problématique de calcul de capital économique :

Input :

Supposons que nous avons à notre disposition :

- N scénarios économiques MR projetés à un an. Nous appelons cette base, la « base totale ».
- Un petit nombre, n (avec $n \leq N$), de scénarios économiques MR (ceux-ci peuvent être extraits de la base totale) ainsi que les fonds propres associés. Cette base sert à calibrer la forêt aléatoire, nous l'appellerons donc base de calibrage. Pour nous fixer les idées, nous pouvons supposer que cette base contient N scénarios.
- Un outil de projection RN permettant de construire des scénarios économiques RN conditionnellement au réalisé de première période.
- Un modèle de gestion actif/passif pour valoriser les fonds propres à partir de tables de scénarios économiques secondaires.

Algorithme de la méthode :

1. Construction de la forêt aléatoire : il s'agit d'expliquer les fonds propres en fonction des facteurs de risque. Nous calibrons donc la forêt sur la base de calibrage et les fonds propres associées.
2. Nous utilisons la forêt pour estimer les fonds propres de chaque scénario de la base totale. Nous pouvons alors définir une « zone quantile » : ensemble des scénarios de cette base dont les fonds propres estimés sont les plus faibles. Ces scénarios sont donc susceptibles de

contenir ceux dont les fonds propres réels seront parmi les 0,5 % les plus faibles de la base totale.

3. Nous calculons les fonds propres des scénarios de la zone quantile (en générant un grand nombre de scénarios secondaires puis en utilisant un modèle ALM de valorisation). Le cinquantième fond propre le plus faible est alors le quantile à 0,5 %.

Limites de la méthode :

Différents essais et tests seront effectués dans les parties suivantes, mais il est évident que le choix de la base de calibrage est crucial. En effet une base de calibrage trop petite ou négligeant certains phénomènes tels qu'une forte hausse des taux, ou une grande chute de la mortalité nuira à l'estimation des fonds propres. La partie suivante a pour but de remédier à ce problème en ajoutant des points à la base de calibrage.

4. Optimisation de la base de calibrage

Nous supposons, dans cette partie, que nous avons une base de calibrage initiale. Nous pouvons nous demander comment faire pour améliorer cette dernière, en particulier, pour savoir quels scénarios ajouter pour améliorer la détection des scénarios risqués de la base totale. Dans cette partie nous allons présenter deux améliorations du premier algorithme présenté *supra*, dont le but est de compléter judicieusement la base de calibrage, pour améliorer le caractère prédictif de la forêt, en particulier pour les scénarios les plus adverses.

a) Première méthode d'ajout de points de calibrage

Cet algorithme se fonde sur celui présenté à la partie précédente, à la différence près qu'un certain nombre de scénarios de calibrage sont ajoutés automatiquement par l'algorithme. En effet jusqu'à maintenant les n scénarios primaires de départ qui servent de points de calibrage sont fixés au lancement de l'algorithme. De manière à optimiser la répartition de ces scénarios, l'algorithme va déterminer, de façon itérative, quel est le point de calibrage dont le rang est mal estimé et pouvant contribuer à la détermination de la zone quantile. Pour chaque itération, ce point, appelé « point mal estimé », est déterminé et nous ajoutons aux points de calibrage celui de la base totale (composée de N points) qui lui est le plus proche (au sens de la norme 2) et n'étant pas déjà présent dans la base de calibrage.

Une fois que le nombre de points de calibrage voulu est atteint, nous lançons le premier algorithme sur la nouvelle base ainsi construite.

Remarque : D'après nos observations, le choix de la norme utilisée pour déterminer le point à ajouter a , de manière générale, peu d'influence sur les résultats. Choisir la norme 2, à la place d'une autre norme, n'est donc pas déterminant dans l'algorithme.

Reprenons l'algorithme de manière plus formelle et supposons que nous voulons ajouter n_{ajout} points de calibrage à une base de calibrage initiale donnée.

Pour i allant de 1 à n_{ajout} :

- Nous calibrons une forêt aléatoire sur la base de calibrage,
- Nous estimons chaque point de la base de calibrage à l'aide de la forêt aléatoire,
- Nous déterminons le « point mal estimé » :
 - Nous comparons les rangs réel et estimé de chaque point de calibrage,
 - Notons c le quantile d'ordre $q_{\text{risqué}}$ des estimés, et A l'ensemble des points dont les fonds propres estimés sont inférieurs à c ,
 - Notons B les éléments de A dont l'écart, entre les rangs réel et estimé, est le plus grand,
 - Le « point mal estimé » est l'élément de B le plus adverse,
- Nous déterminons le point à ajouter :
 - Nous calculons les distances entre le « point mal estimé » et l'ensemble des points de la base totale,
 - Le point le plus proche (selon la norme 2) et n'étant pas déjà présent dans la base de calibrage, est le point à ajouter,
 - Nous calculons les fonds propres réels du point à ajouter par la méthode SdS,
- Nous incluons le point à ajouter dans la base de calibrage.

Une base de calibrage complétée est ainsi construite, et elle contient $n' = n + n_{\text{ajout}}$ points distincts. Nous lançons alors le premier algorithme sur cette nouvelle base de calibrage.

b) Deuxième méthode d'ajout de points de calibrage

Cette méthode ressemble beaucoup à la précédente, la seule différence résidant dans l'approche utilisée pour déterminer le « point mal estimé ».

Après avoir trouvé le « point mal estimé », noté ici p_1 , nous déterminons son plus proche voisin, p_2 . Parmi les points de calibrage, nous définissons p_3 comme le milieu de p_1 et p_2 .

Le point à ajouter est alors le point de la base totale, n'étant pas déjà présent dans la base de calibrage et étant le plus proche du point p_3 .

L'intérêt de cet algorithme par rapport au précédent est qu'il permet de mieux répartir les points de calibrage et d'augmenter les chances de capter l'ensemble des mouvements de la forme paramétrique. En effet, avec la première méthode d'ajout de points de calibrage, nous risquons d'ajouter un point très proche du point mal estimé. Or, pour un nombre de points de calibrage fixé, les estimations bénéficient d'une grande distance entre les points de la base de calibrage. Ce dernier algorithme évite donc d'ajouter des points trop proches de ceux déjà présents dans la base de calibrage.

5. Optimisation du calcul des fonds propres de la zone quantile

L'objet de cette partie est d'étudier dans quelle mesure il est possible d'optimiser le calcul du quantile à 0,5 % de la base totale, la zone quantile étant connue. Nous supposons de plus que les $[0,5\% \times N]^{i\text{ème}}$ pires scénarios réellement les plus adverses sont dans la zone quantile.

Il peut arriver que la zone quantile contienne plus de $[0,5\% \times N]$ scénarios (nous verrons que c'est très souvent le cas). Il s'agit alors de pouvoir déterminer, de manière efficace, quel est le $[0,5\% \times N]^{\text{ième}}$ pire scénario de la zone quantile.

Une approche naïve serait de calculer de façon précise chaque scénario de la zone quantile en lançant un grand nombre de scénarios secondaires couplés au modèle de gestion actif/passif (comme dans une approche SdS). Cependant, il s'agit de ne calculer que les fonds propres du $[0,5\% \times N]^{\text{ième}}$ scénario le plus adverse, parmi ceux de la zone quantile. Nous sommes donc dans le cadre d'une application de l'algorithme d'allocation optimale de scénarios, qui pourra aider à réduire le coût de calcul.

6. Calcul de la complexité

Le calibrage et l'estimation par forêt aléatoire n'est pas très coûteux en temps de calcul. Nous nous intéresserons donc, dans cette partie, au nombre d'appels au modèle ALM à effectuer.

Rappelons les notations suivantes:

- n : La taille de la base de calibrage (n' si l'on se place dans un algorithme avec ajout de points à la base de calibrage),
- N : la taille de la base totale,
- $n_{\text{secondaire}}$: le nombre de simulations secondaires effectuées dans un calcul SdS classique.

Notons z la taille de la zone quantile définie à la fin de l'algorithme.

Alors le nombre d'appels à un logiciel de gestion actif/passif est le suivant :

$$C = (n + z) \times n_{\text{secondaire}}$$

Nous verrons *infra* qu'en pratique, les ordres de grandeur sont les suivants (pour un modèle à deux paramètres) :

$$n \sim 200,$$

$$N \sim 10\,000,$$

$$n_{\text{secondaire}} \sim 1\,000,$$

$$z \sim 3 \times [0,5\% \times N] \sim 150,$$

$$\text{Nombre d'appel ALM} \sim 350\,000,$$

$$\text{Nombre d'appels pour un SdS complet} \sim 50\,000\,000.$$

Remarque : ces valeurs sont données à titre illustratif et peuvent varier suivant le portefeuille utilisé.

Nous obtenons donc une réduction de l'ordre de 99 % du nombre d'appels au modèle de gestion actif/passif.

IV. MISE EN OEUVRE ET RESULTATS

Nous avons présentés deux nouvelles méthodes de calcul du capital économique : la méthode de décomposition des événements rares et la méthode des forêts aléatoires. Nous allons maintenant exposer les résultats obtenus en appliquant ces méthodes. Dans un premier temps nous allons les tester dans un cadre simplifié, puis nous étudierons les résultats dans des conditions réelles de gestion actif/passif.

A. Application dans le cas de modèles de gestion actif/passif simplifiés

Nous supposons, dans cette partie, qu'il existe une forme paramétrique polynomiale donnant les fonds propres à un an en fonction des facteurs de risque. Cette hypothèse va nous permettre de comparer la performance des algorithmes dans différentes conditions sans avoir besoin d'un portefeuille d'assurance vie ou d'un modèle de gestion actif/passif. Nous allons tout d'abord décrire les formes paramétriques utilisées, puis nous étudierons la robustesse des résultats, leurs sensibilités aux paramètres et aux formes paramétriques, et enfin nous comparerons les performances des méthodes.

1. Description des formes paramétriques

Nous avons appliqué nos différents algorithmes à deux formes paramétriques respectivement à deux et quatre facteurs de risque. Les formes paramétriques ont été calibrées en appliquant les méthodes utilisées pour le Curve Fitting et le Least Square Monte Carlo. L'intérêt de tester les méthodes présentées *supra*, sur des formes paramétriques, est que nous pouvons aisément calculer les fonds propres associés à un jeu de facteurs de risque. En effet, comme nous ne disposons pas d'un générateur de scénarios économiques ou d'un modèle ALM, les formes paramétriques sont un moyen simple et relativement efficace de définir un capital économique pour tout couple de facteur de risque.

Les deux formes paramétriques ont une forme polynomiale de degré 3, et les coefficients employés sont les suivants :

Puissances des facteurs de risque	Forme paramétrique à 4 facteurs
Constante	100,0
ϵ_1	4,8
ϵ_1^2	0,4
ϵ_1^3	-0,1
ϵ_2	0,2
ϵ_3	3,1
ϵ_3^2	0,8
ϵ_4	-14,5

ϵ_4^2	3,1
ϵ_4^3	-0,4
$\epsilon_1 \epsilon_4$	-0,2
$\epsilon_1 \epsilon_4^2$	0,0
$\epsilon_1 \epsilon_3$	0,1
$\epsilon_2 \epsilon_4$	-0,1
$\epsilon_3 \epsilon_4$	-2,6
$\epsilon_3 \epsilon_4^2$	0,5
$\epsilon_3^2 \epsilon_4$	-0,2

Puissances des facteurs de risque	Forme paramétrique à 2 facteurs
Constante	100
ϵ_1	2,3
ϵ_1^2	-1,0
ϵ_1^3	1,0
ϵ_2	-1,0
ϵ_2^2	0,01
ϵ_2^3	-0,1
$\epsilon_1\epsilon_2$	-0,02

$\epsilon_1^2\epsilon_2$	-0,1
$\epsilon_1\epsilon_2^2$	4,0

Tableau 3 : Tableau présentant les coefficients des deux formes paramétriques polynomiales

Le graphique ci-dessous représente la forme paramétrique à deux facteurs de risque, dans le plan (risque action, risque zéro-coupon). Nous constatons que les fonds propres sont les plus faibles pour des petites valeurs du facteur de risque action et des valeurs extrêmes pour le facteur de risque zéro-coupon. Nous observons également une plus grande sensibilité de la forme paramétrique, au risque action qu'au risque taux.

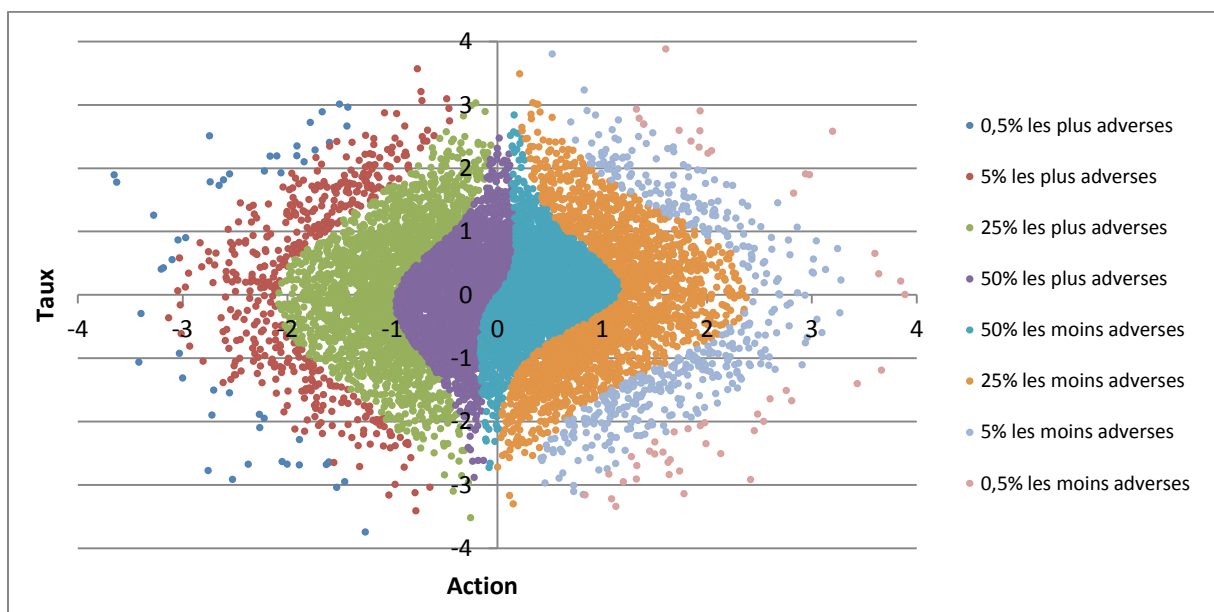


Figure 10 : Graphique représentant la répartition des scénarios par valeur de fonds propres

2. Premiers résultats obtenus

Nous allons étudier dans un premier temps les résultats obtenus à l'aide de l'algorithme de décomposition en événements rares, puis, dans un second temps, ceux donnés par la méthode des forêts aléatoires.

a) Décomposition d'un événement rare en événements moins rares

Comme expliqué précédemment, les algorithmes de décomposition d'un événement rare en événements moins rares ont été mis en œuvre pour deux formes paramétriques. Pour pouvoir analyser les différents résultats obtenus, nous avons effectué, dans chaque cas, un calcul de référence correspondant au quantile à 0,5 % obtenu en appliquant plus d'un million de fois la forme paramétrique choisie. Les valeurs de référence trouvées pour les formes paramétriques à deux et quatre facteurs de risque sont respectivement égales à 51,21 et à 74,14.

Dans la partie théorique sur la décomposition d'un événement rare en événements moins rares, nous avons tout d'abord présenté un premier algorithme simple, pour lequel l'écart-type utilisé dans la diffusion des scénarios était constant, puis nous avons adapté la méthode en déterminant, pour chaque facteur de risque, une valeur optimale de σ_k , à partir des scénarios qui n'étaient pas diffusés. Ainsi, nous allons commencer par comparer les quantiles obtenus en utilisant chacune des deux méthodes.

Pour cela, nous avons fixé la valeur des deux paramètres N et T à 100, et nous avons lancé, d'une part, vingt fois l'algorithme simple pour trois valeurs différentes de l'écart-type $\sigma_k = \sigma$, et, d'autre part, l'algorithme adapté. Ensuite, nous avons comparé les moyennes obtenues sur les vingt essais dans chacun des quatre cas aux valeurs de référence.

Nous présentons dans le tableau ci-dessous les résultats obtenus pour les deux formes paramétriques étudiées :

	Algorithme simple $\sigma = 0,01$	Algorithme simple $\sigma = 0,5$	Algorithme simple $\sigma = 0,99$	Algorithme adapté	Valeur de référence
Forme paramétrique à deux facteurs	65,91	34,46	31,39	50,07	51,21
Forme paramétrique à quatre facteurs	77,89	71,39	70,90	74,54	74,14

Tableau 4 : Quantile moyen obtenu en lançant vingt fois les algorithmes simples et adaptés pour chacune des formes paramétriques

Nous remarquons que les résultats obtenus avec l'algorithme adapté sont beaucoup plus proches de la valeur de référence que ceux obtenus lorsque l'écart-type utilisé dans la diffusion est constant, que ce soit avec la forme paramétrique à deux facteurs ou avec celle à quatre facteurs. Par ailleurs, nous observons que le quantile moyen obtenu sur les vingt lancers de l'algorithme converge vers des valeurs très différentes suivant la valeur de σ choisie. De plus, comme nous l'avions prévu, plus σ est élevé, plus les scénarios sont diffusés et le quantile trouvé à l'issue de l'algorithme est faible.

Ainsi, nous pouvons conclure des résultats ci-dessus que, pour notre portefeuille, il n'est pas judicieux de figer la valeur de σ_k . Dans la suite de la partie de mise en œuvre des algorithmes, nous nous intéresserons uniquement à la méthode de décomposition d'un événement rare en

événements moins rares adaptée, c'est-à-dire pour laquelle σ_k est déterminé à partir des scénarios non diffusés.

Nous pouvons analyser les premiers résultats obtenus avec cette méthode. Lorsque nous l'appliquons à la forme paramétrique à deux facteurs et que nous représentons les scénarios obtenus à la fin de l'algorithme dans le plan (facteur de risque action, facteur de risque zéro-coupon), nous obtenons le graphique ci-dessous :

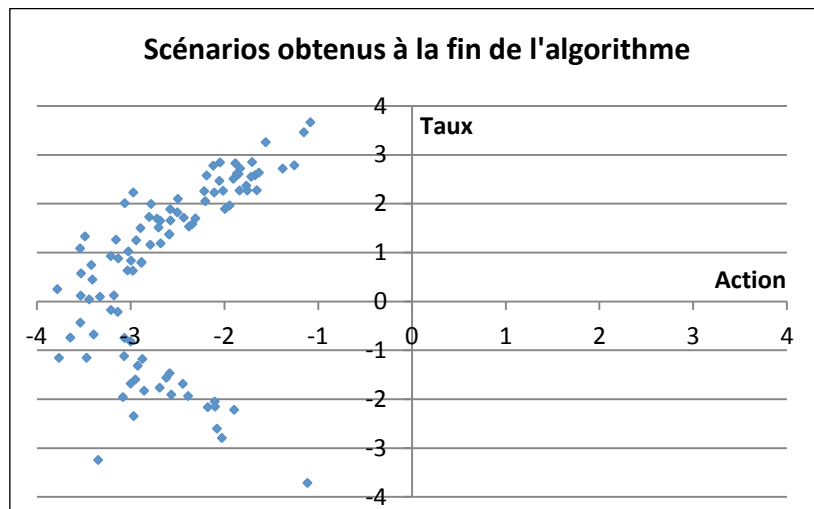


Figure 11 : Scénarios obtenus à la fin de l'algorithme, pour la forme paramétrique à deux facteurs, représentés dans le plan (action, zéro-coupon)

Les scénarios qui sont obtenus à la fin de l'algorithme et qui sont représentés dans le graphique ci-dessus sont des scénarios de queue. En effet, les fonds propres qui leur sont associés sont tous inférieurs au quantile à 0,5 % de la distribution des fonds propres à un an. Nous observons ainsi que les scénarios formant la zone quantile sont ceux pour lesquels il y a un choc à la hausse ou à la baisse des taux et un choc à la baisse des actions. Par ailleurs, nous notons que ces scénarios sont tous composés d'au moins un choc important (action ou taux).

Si nous nous intéressons au quantile moyen estimé avec l'algorithme adapté, nous remarquons que les résultats sont assez satisfaisants. En effet, le quantile moyen obtenu avec la forme paramétrique à deux facteurs est égal à 50,07, ce qui donne une erreur relative par rapport au quantile de référence de 2,23 %. Pour la forme paramétrique à quatre facteurs, le résultat obtenu est encore meilleur, puisque nous trouvons un quantile moyen égal à 74,54, soit une différence de seulement 0,54 % avec le calcul de référence.

Cependant, même si ces premiers résultats sont plutôt bons, il est important de noter qu'ils dépendent de la valeur des paramètres. En effet, notre algorithme repose sur le choix de deux paramètres : le nombre de scénarios, N , et le nombre d'itérations horizontales permettant la diffusion des scénarios, T . Dans l'exemple ci-dessus, nous avons choisi $N = 100$ et $T = 100$, mais nous pouvons penser que certains paramètres donnent de meilleurs résultats. Aussi une analyse de la sensibilité des résultats en fonction des paramètres sera effectuée dans la partie IV.B.4.A. Nous remarquons également que les résultats dépendent de la forme paramétrique utilisée. Nous

étudierons cette dépendance dans la partie V.B.4.A en nous intéressant à de nouvelles formes paramétriques.

b) Forêts aléatoires

Nous avons vu que la méthode des forêts aléatoires ne permet pas d'obtenir le quantile à 0,5 % réel, mais uniquement le quantile à 0,5 % de la base totale. Nous considérerons donc ce dernier comme le quantile réel. Cependant nous verrons, en dernière partie, qu'il est possible de réduire l'écart entre le quantile obtenu et le quantile réel.

Nous allons introduire une mesure de risque qui nous permettra d'analyser et de comparer facilement les résultats des algorithmes, puis d'analyser le caractère prédictif des forêts aléatoires. Enfin, nous présenterons les résultats obtenus par l'algorithme sans ajout de point, sur les deux formes paramétriques décrites *supra*.

Mesure de risque

Dans chacun des trois algorithmes présentés plus haut, nous utilisons les forêts aléatoires pour définir un ensemble contenant les scénarios les plus adverses, appelé zone quantile.

Soit ψ , la fonction suivante :

$$\psi(Z, q, N) = \frac{\#Z}{q \times N}$$

avec :

- Z : une zone quantile,
- $\#Z$: le cardinal d'un ensemble Z ,
- q : le quantile recherché,
- N : le nombre de scénarios de la base totale.

Nous avons vu que l'hypothèse clé des approches par forêts aléatoires est la conservation des rangs, notamment pour les scénarios les plus adverses. De manière à ce que le calcul du capital économique sachant la zone quantile soit valide, il est nécessaire que cette dernière inclue les scénarios dont les fonds propres réels sont inférieurs au quantile à 0,5 %.

Nous définirons donc la zone quantile, comme l'ensemble des scénarios de la base totale dont les fonds propres estimés sont inférieurs à un seuil quantile. Ce seuil quantile sera choisi le plus petit possible et tel que tous les scénarios dont les fonds propres réels sont inférieurs au quantile à 0,5 %, ont un estimé inférieur au seuil quantile.

ψ définit alors une mesure de risque.

Ainsi, la mesure de risque est le rapport du nombre de scénarios à sélectionner par ordre d'adversité estimé pour inclure les $[N \times \text{quantile souhaité}]$ scénarios les plus adverses réellement, sur le nombre de scénarios de la base totale.

Plus l'algorithme respecte l'ordre des scénarios, plus la mesure de risque tendra vers 1.

Etude préliminaire

Dans la mesure où notre objectif est l'estimation du quantile de la distribution de fonds propres à 0,5 %, deux approches sont envisageables. La première consiste à estimer directement le quantile à partir des résultats de la forêt. L'estimation du quantile à 0,5 % est donc la 50^{ième} plus petite valeur des fonds propres estimés (sur 10 000). L'hypothèse sous jacente est que les estimations de la forêt aléatoire conserve le niveau de fonds propres (et donc le rang). La deuxième méthode, moins conservatrice, suppose que les estimations de la forêt ne conservent que le rang. Il est alors possible de définir le scénario donnant le quantile comme le jeu des facteurs de risques associés à la 50^{ième} plus petite valeur de fonds propres estimés.

De manière à vérifier l'intérêt de l'utilisation des forêts aléatoires nous avons calibré une forêt aléatoire de taille 1 000 sur une base de 10 000 lignes composées de trois colonnes : deux colonnes de facteurs de risque (des lois normales) et une dernière colonne donnant les fonds propres (calculés grâce à la forme paramétrique). Enfin nous avons estimé les fonds propres à partir de ces mêmes facteurs de risque et de la forêt. Nous avons alors pu comparer les résultats obtenus avec les fonds propres réels.

Ci-dessous est représenté l'ensemble de la base, avec en abscisse la valeur réelle des fonds propres, et en ordonnée la valeur estimée :

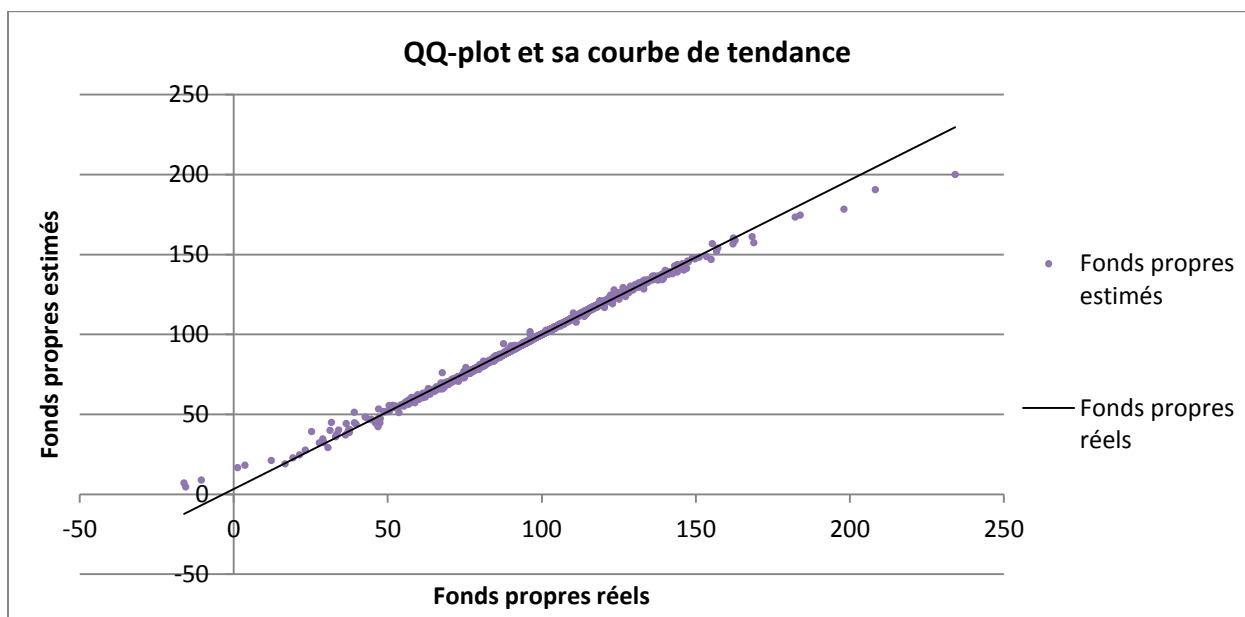


Figure 12 : Graphique représentant les fonds propres estimés en fonction des fonds propres réels

Nous pouvons constater une relativement bonne adéquation en niveau, principalement dans la zone centrale : pour des valeurs de fonds propres comprises entre -50 et 150. Cependant la qualité de l'estimation est moins bonne dans les extrêmes, en particulier pour les valeurs qui nous intéressent, celles inférieures à 50. Il paraît évident que si les niveaux ne sont pas bien conservés lors d'un test *in sample* (sur la base de calibrage), les valeurs estimées sur des bases *out of sample* ne pourront être considérées comme convenables. Cependant ce graphique nous laisse supposer que les forêts à défaut de conserver la valeur des fonds propres, peuvent conserver le rang de ceux-ci.

Le graphique ci-dessous compare les rangs des fonds propres réels avec ceux estimés, sur la même base que précédemment.

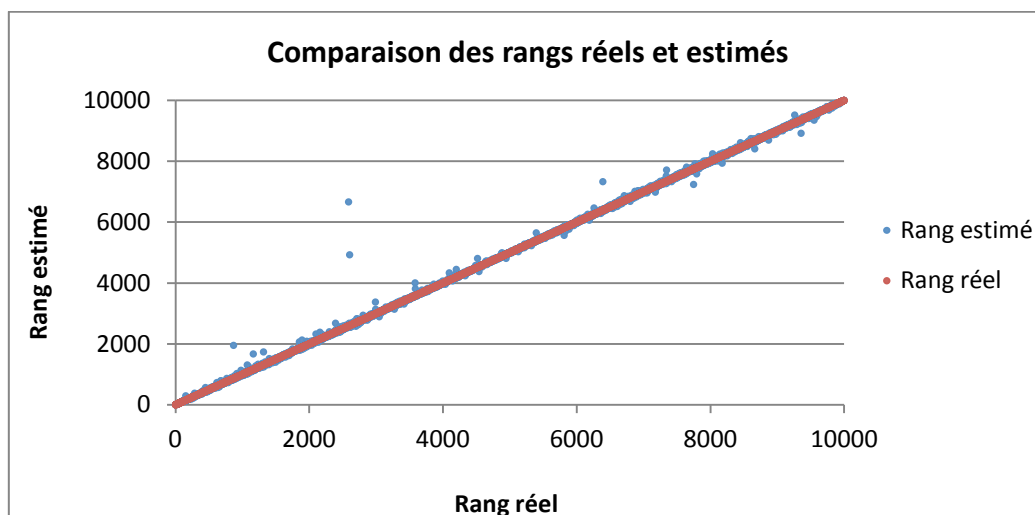


Figure 13 : Comparaison in sample des rangs réels et estimés par forêts aléatoires, avec 1 000 arbres et 10 000 scénarios.

Nous constatons une très bonne conservation des rangs, particulièrement dans les zones extrêmes (entre 0 et 500, et entre 9 500 et 10 000), ce qui est confirmé par le coefficient de corrélation de Spearman qui est égal à 99,99 %.

En effet, comme les fonds propres réels sont plus espacés dans les extrêmes, la conservation du rang est meilleure que la conservation des valeurs. Nous pouvons cependant émettre la réserve suivante : dans la mesure où la forêt doit être calibrée sur un minimum de points, une base de calibrage trop petite dégradera probablement la conservation des rangs. Comme abordé dans la partie précédente, nous devons définir une zone quantile, incluant les scénarios dont les fonds propres réels sont les plus adverses.

Pour mesurer la conservation du rang dans les extrêmes, pour différentes bases et méthodes, nous utiliserons donc la mesure de risque présentée précédemment.

Le tableau ci-dessous présente les valeurs des mesures de risque obtenues pour différents quantiles :

Quantile recherché :	0,5 %	1 %	2 %	3 %	4 %	5 %
Mesure de risque associée :	1,18	1,05	1,44	1,28	1,02	1,13

Tableau 5 : Présentation des mesures de risque par quantile recherché, pour un test in sample du premier algorithme (10 000 scénarios)

La mesure de risque est assez faible : de l'ordre de 1,2 (avec un maximum de 1,44 pour le quantile à 2 %). Ainsi, l'écart entre le rang des fonds propre estimés extrêmes et celui des fonds propres réels extrêmes est très faible et permet donc d'envisager l'utilisation des rangs donnés par les forêts. La mesure de risque associée au quantile à 3 % est égale à 1,28. Le calcul du quantile à 3 % nécessite alors de calculer précisément (par méthode SdS) les 384 scénarios ($1,28 \times 3 \% \times 10\,000 = 384$) donnant les fonds propres estimés les plus faibles, tandis que le quantile à 3 % correspond à la 300ème plus petite valeur de fonds propres de la base.

Evidemment, dans ce cas précis, le calcul du quantile est beaucoup plus simple, les fonds propres de la base totale étant connus (ce sont ceux de la base de calibrage, qui le sont par définition).

Nous allons donc nous placer en situation réelle, en supposant que les vrais fonds propres peuvent être calculés par la forme paramétrique décrite plus haut.

Application réelle : approche sans ajout de points de calibrage, sur données orthogonales

Nous allons appliquer les algorithmes présentés plus haut à la forme paramétrique polynomiale de degré 2 et en fonction de deux facteurs de risque. Cette forme paramétrique est décrite *supra*.

Il semble logique que les résultats varient beaucoup en fonction de la taille de la base de calibrage choisie. En effet plus la base est grande et plus la mesure de risque à 0,5 % est bonne. Nous pouvons aussi nous demander dans quelle mesure le choix des points de calibrage influe sur les résultats.

Nous avons lancé successivement l'algorithme 1 (sans amélioration) sur différentes bases de calibrage avec suffisamment d'arbres pour assurer la robustesse des valeurs obtenues (l'étude détaillée de la robustesse est faite dans la partie suivante).

Cinq bases de calibrage ont été étudiées pour la forme paramétrique à deux facteurs de risque. La première base, appelée « Loi normale *out of sample* », est constituée de facteurs de risque suivant une loi normale centrée réduite et des fonds propres associés. Les facteurs de risque de la deuxième base, « Loi normale *in sample* », ont été tirés aléatoirement parmi ceux de la base totale. Des quadrillages ont été utilisés pour répartir les facteurs de risque des trois bases suivantes. Les points extrêmes des quadrillages et le pas utilisé, diffèrent suivant les bases. Nous avons ainsi trois bases ayant un nombre de points différents, ce qui nous permettra d'étudier l'impact de celui-ci sur la mesure de risque.

Les résultats sont présentés dans le tableau ci-dessous :

Base de départ :					Mesure de risque par quantile recherché :		
Nom	Borne min	Borne max	Pas	Nombre de points	0,5 %	1 %	2 %
Loi normale <i>out of sample</i>	X	X	X	500	111	56.2	39.6
Loi normale <i>in sample</i>	X	X	X	500	40.1	50.3	39
Base de petite taille	-3.5 ⁶	3.5	0,7	121	3,29	4,62	5,44
Base de taille moyenne	-3.5	3.5	0,5	225	3,03	1,54	2,83
Base de	-3.4 ⁷	3.4	0,4	324	1,61	2,55	2,93

⁶ $P(N(0,1) \leq 3,5) = 0,02 \%$, ainsi une borne de 3,5 permet une bonne estimation des points les plus adverses.

⁷ $P(N(0,1) \leq 3,4) = 0,03 \%$, ainsi une borne de 3,4 permet une bonne estimation des points les plus adverses.

grande taille	
---------------	--

Tableau 6 : Mesures de risque obtenues pour la forme paramétrique à deux facteurs de risque

Nous constatons qu'il est plus intéressant de choisir une base de calibrage dont les points sont répartis de manière uniforme. Cela semble logique dans la mesure où les scénarios les plus risqués et donc les plus susceptibles de se situer dans la zone quantile, sont *a priori* loin de l'origine.

Par ailleurs, plus il y a de points répartis uniformément dans la base de calibrage et plus les résultats sont bons, ce qui confirme notre intuition.

Cependant, comme nous le verrons plus loin, (dans la partie consacrée à la sensibilité à la forme paramétrique), les résultats peuvent varier suivant la forme paramétrique utilisée, en particulier lorsque la base de calibrage est de petite taille, ou n'est pas adaptée à la forme paramétrique.

Les mesures de risque obtenues en calibrant sur des facteurs de risque suivant des lois normales ne sont pas satisfaisantes. Désormais nous nous restreignons aux bases de calibrage dont les points sont répartis uniformément.

De la même manière que pour la forme paramétrique à deux facteurs, nous allons construire une base de calibrage sous la forme d'un quadrillage de dimension 4 : les facteurs de risque de la base de calibrage que nous avons sélectionnée sont répartis uniformément sur les quatre axes, de -3,3 à 3,3⁸ avec un pas de 1,1.

Etudions maintenant les résultats obtenus sur la forme paramétrique à quatre facteurs.

Base de départ :					Mesure de risque par quantile recherché :		
Nom	Borne min	Borne max	Pas	Nombre de points	0,5 %	1 %	2 %
Base de grande taille	-3,3	3,3	1,1	2401	4,12	3,3	5,12

Tableau 7 : Premiers résultats obtenus pour la forme paramétrique à quatre facteurs, avec $m = 2$.

Les mesures de risque sont du même ordre de grandeur que pour la forme paramétrique à deux facteurs. L'approche des forêts aléatoires semble donc justifiée.

Remarque : Nous constatons que plus il y a de facteurs de risque, et plus le nombre de points de la base de calibrage doit être grand. En effet, dans le cas de la base de la forme paramétrique à quatre facteurs nous avons $\frac{3,3 - (-3,3)}{1,1} = 7$ valeurs distinctes par facteurs de risque. Le nombre total de points de la base est donc $7^4 = 2401$, ce qui est bien supérieur au nombre de points des bases testées sur la forme paramétrique à deux facteurs.

Que ce soit pour la méthode de décomposition d'un événement rare, ou celle des forêts aléatoires, les méthodes employées sont aléatoires, il s'agit donc de vérifier la robustesse des résultats obtenus.

⁸ $P(N(0,1) \leq 3,3) = 0,05 \%$, ainsi une borne de 3,3 permet une bonne estimation des points les plus adverses.

3. Test de la robustesse des résultats

Nous allons, dans cette partie, vérifier que les résultats sont robustes, c'est-à-dire peu volatils. Comme dans le paragraphe précédent, nous commencerons par nous intéresser à la méthode de décomposition d'un événement rare, puis nous étudierons la méthode basée sur les forêts aléatoires.

a) Décomposition d'un événement rare en événements moins rares

Pour tester si les résultats sont robustes, nous choisissons de lancer vingt fois de suite l'algorithme et d'observer comment varie l'estimateur du quantile entre ces vingt essais.

L'indicateur qui sera utilisé pour analyser la variation de l'estimateur du quantile est un écart-type relatif calculé sur ces vingt essais : $\sigma_r(\hat{q}) = \frac{\sigma(\hat{q})}{E[\hat{q}]}$. Nous sommes conscients que calculer un écart-type relatif à partir de vingt données ne permet pas de l'estimer correctement. Cependant, cet indicateur nous permet d'avoir une première idée de la façon dont varie le quantile fourni par notre algorithme.

Nous pouvons penser que la robustesse des résultats dépend des paramètres choisis. Aussi, pour vérifier cette conjecture, nous choisissons différents jeux de paramètres (N, T) et nous comparons l'écart-type relatif des quantiles à 0,5 % des fonds propres à un an obtenus. Pour repérer le rôle que jouent chacun des paramètres, nous fixons d'abord la valeur du nombre de scénarios N à 100 et nous faisons varier le nombre d'itérations horizontales T , puis nous choisissons un nombre d'itérations horizontales égal à 100 et nous testons différentes valeurs de N .

Cette étude est d'abord menée pour la forme paramétrique à deux facteurs de risque, puis pour la forme paramétrique à quatre facteurs.

Dans le cas de la forme paramétrique à deux facteurs, les écart-types relatifs obtenus pour $N = 100$ et pour des valeurs de T comprises entre 10 et 200 sont donnés dans le graphique ci-dessous :

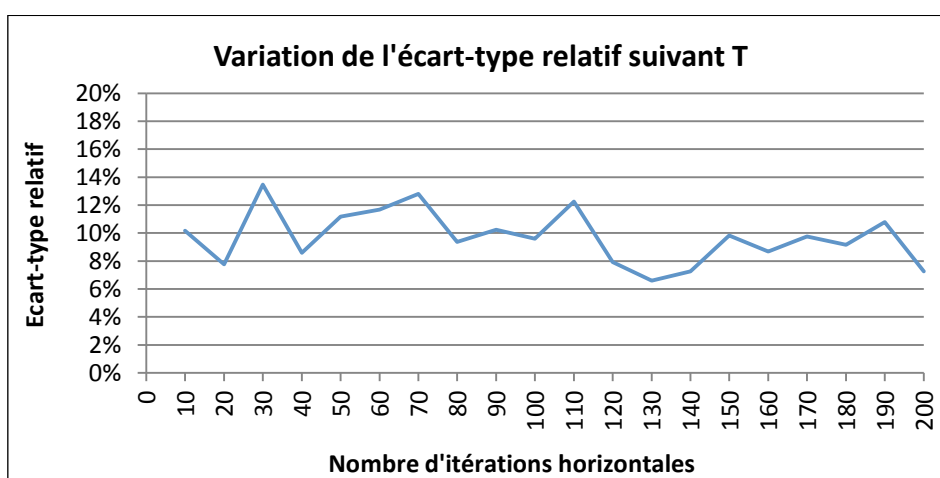


Figure 14 : Variation de l'écart-type relatif en fonction du nombre d'itérations horizontales T pour la forme paramétrique à deux facteurs

Nous observons que le paramètre T ne semble pas influencer beaucoup la robustesse du quantile, même si nous notons une très légère tendance à la baisse de l'écart-type relatif lorsque T augmente. Nous pouvons supposer qu'à partir d'une certaine valeur, ce paramètre n'influence pas la robustesse du résultat parce que le noyau de transition a été appliqué suffisamment de fois pour que la loi stationnaire soit atteinte.

Nous pouvons à présent fixer la valeur de T à 100 et étudier la variation de l'écart-type pour différentes valeurs de N , comprises entre 10 et 200. Nous obtenons alors le résultat suivant :

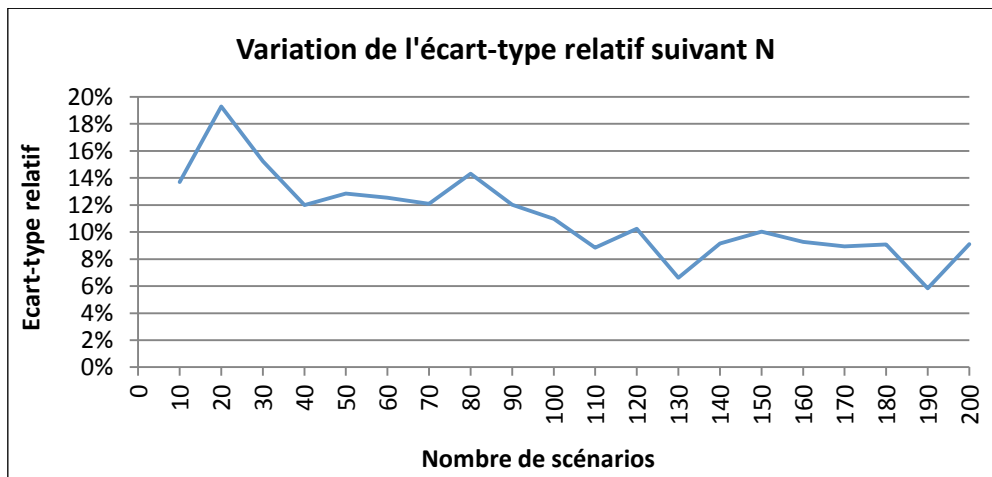


Figure 15 : Variation de l'écart-type relatif en fonction du nombre de scénarios N pour la forme paramétrique à deux facteurs

Nous remarquons que contrairement au paramètre T , N semble jouer un rôle dans la robustesse des résultats. Il semblerait en effet que plus N est grand, plus l'estimation du quantile est stable lorsque nous lançons plusieurs fois l'algorithme. En particulier, l'écart-type relatif devient inférieur à 12 % pour des valeurs de N supérieures à 90.

Nous rappelons que le nombre d'itérations verticales est calculé à partir de N . Aussi, plus le nombre de scénarios est grand, plus nous effectuons d'itérations verticales. Nous pouvons penser que c'est pour cette raison que les résultats obtenus sont plus robustes pour des valeurs de N plus élevées.

Nous estimons de la même manière les écart-types relatifs pour la forme paramétrique à quatre facteurs, en lançant vingt fois l'algorithme et en calculant le rapport de l'écart-type des quantiles estimés par leur moyenne. Nous remarquons tout d'abord que les écarts-type relatifs obtenus sont beaucoup plus faibles. En effet, les deux graphiques ci-dessous montrent que, quels que soient les paramètres (N, T) que nous avons testés, les écart-types relatifs sont toujours inférieurs à 4 % :

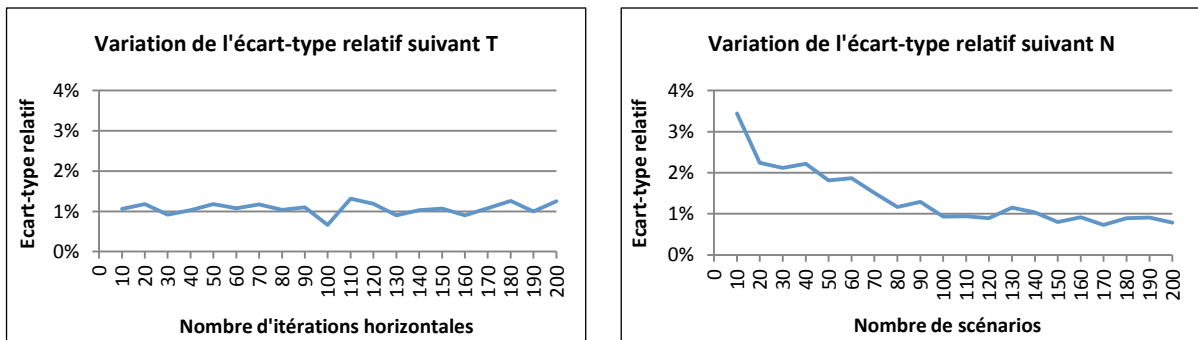


Figure 16 : Variation de l'écart-type relatif en fonction du nombre d'itérations horizontales T et du nombre de scénarios N pour la forme paramétrique à quatre facteurs

Nous observons cependant que la variation de l'écart-type relatif selon les paramètres choisis est la même que dans le cas de la forme paramétrique à deux facteurs. En effet, nous notons que le nombre d'itérations horizontales T ne semble pas réellement influencer la robustesse des résultats, tandis qu'augmenter le nombre de scénarios N permet d'obtenir un résultat plus précis.

b) Forêts aléatoires

Nous avons vu que la méthode des forêts aléatoires est une métaheuristique, les résultats sont donc aléatoires. L'étude des résultats d'un algorithme nécessite de connaître leur robustesse. En effet une faible robustesse demandera de lancer de nombreuses fois le programme pour réduire la volatilité des résultats. Nous avons donc lancé, pour différentes tailles de forêt, notre premier algorithme (sans ajout de point) vingt fois de manière à mesurer l'écart-type normalisé des estimations de la mesure de risque, en imposant un calibrage sur 121 points répartis uniformément dans le plan. Les résultats obtenus sont présentés dans les graphiques ci-dessous :

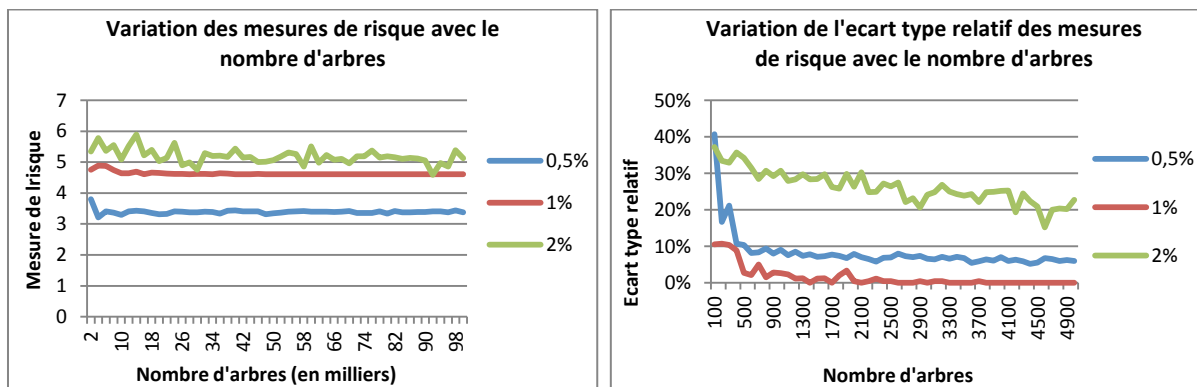


Figure 17 : Evolution de la mesure de risque et de son écart-type relatif pour différents quantiles pour la forme paramétrique à deux facteurs

Bien que les mesures de l'écart-type relatif ne soient pas très robustes (elles ne sont calculées que sur vingt points), elles montrent la relative stabilité des mesures de risque pour un nombre d'arbres supérieur à 1 000. De plus, les mesures de risques nous relativement insensibles au nombre d'arbres lorsque celui-ci est suffisamment grand (à partir de 5 000).

Remarque : Nous remarquons que les écarts types relatifs des mesures de risque pour le quantile à 1 % sont extrêmement faibles par rapport à ceux des quantiles à 0,5 % et 2 %. Cela est probablement dû aux aléas statistiques : la base de calibrage est particulièrement bien adaptée à la forme paramétrique.

Remarque : Dans la mesure où les forêts aléatoires sont utilisées pour faire des régressions, la valeur finale est la moyenne des valeurs obtenues sur chaque arbre. Il est donc équivalent de lancer une forêt de 10 000 arbres et de prendre la moyenne de 10 forêts de 1 000 arbres.

Nous obtenons des évolutions similaires avec la forme paramétrique à quatre facteurs.

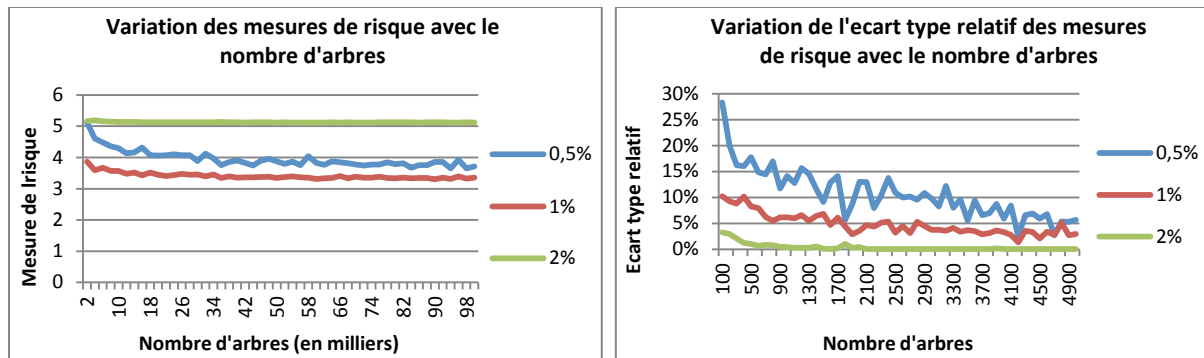


Figure 18 : Evolution des mesures de risque et de leur écart-type relatif en fonction du nombre d'arbres pour la forme paramétrique à quatre facteurs

Cela indique donc que le nombre d'arbres à construire doit être supérieur à 3 000 pour obtenir une précision à l'ordre 10^{-1} de la mesure de risque. Ainsi, en prenant 10 000 arbres, nous nous assurons de la robustesse des estimations.

4. Analyse de la sensibilité aux paramètres

Que ce soit pour la méthode de décomposition d'un événement rare en événement moins rares, ou pour celle des forêts aléatoires, nous pouvons noter que les algorithmes nécessitent de choisir un certain nombre de paramètres. Nous allons étudier, ici, l'influence de ces paramètres sur les résultats.

a) Décomposition d'un événement rare en événements moins rares

Nous avons montré précédemment que la robustesse des résultats dépendait du nombre de scénarios N . En effet, nous rappelons que pour de faibles valeurs de N , le quantile estimé n'est pas robuste. Ainsi, nous pouvons nous demander si la valeur obtenue pour l'estimateur du quantile dépend, elle aussi, de paramètres. Nous allons donc analyser quelle est la valeur moyenne du quantile lorsque le nombre d'itérations horizontales, T , et le nombre de scénarios, N , varient. Comme dans l'étude de la robustesse, nous lançons vingt fois l'algorithme pour chaque jeu de paramètres (N, T) , mais nous nous intéressons cette fois-ci à la moyenne obtenue.

i. Sensibilité au nombre d'itérations horizontales T

Nous allons tout d'abord nous intéresser au choix du nombre d'itérations horizontales, T . Ce paramètre correspond au nombre de fois où nous appliquons le noyau de transition lors de la diffusion des scénarios, aussi il doit respecter certaines contraintes. En effet, il doit permettre d'obtenir un scénario indépendant de celui à partir duquel il est diffusé.

Nous pouvons cependant penser que l'indépendance entre le scénario d'origine et le scénario diffusé dépend de la valeur de la variance σ_k^2 utilisée dans la diffusion, nous allons donc tout d'abord étudier comment varie la dépendance entre les scénarios avant et après la diffusion en fonction des valeurs de T et de σ_k .

(1) Etude de l'indépendance suivant T et σ_k

Nous rappelons tout d'abord que le noyau de transition utilisé lors de la diffusion des facteurs de risque est donné par la formule : $X_k^* = \frac{X_{i,k}^m + \sigma_k W_k}{\sqrt{1 + \sigma_k^2}}$. Nous avons choisi d'adapter l'algorithme initial pour lequel la variance σ_k^2 était constante en déterminant sa valeur à partir des scénarios qui n'étaient pas diffusés.

Cependant, avec notre adaptation, la variance peut prendre des valeurs très différentes. Nous allons donc, dans cette étude préliminaire, fixer à nouveau la valeur de $\sigma_k = \sigma$ et observer la dépendance entre le scénario initial et le scénario diffusé, afin de voir l'impact de σ plus facilement.

Pour analyser la dépendance, nous avons choisi de partir de 500 scénarios initiaux et de leur appliquer 150 fois le noyau de transition. Nous avons ensuite calculé le coefficient de corrélation entre le scénario initial et le scénario obtenu à l'issue de chaque nouvelle itération horizontale, pour pouvoir étudier comment il varie lorsque T augmente. Nous avons effectué cette analyse pour différentes valeurs de variance comprises entre 0,1 et 0,9. Nous présentons ci-dessous les graphiques donnant l'évolution du coefficient de corrélation en fonction du nombre d'itérations horizontales effectuées, pour $\sigma = 0,1$, $\sigma = 0,4$ et $\sigma = 0,9$:

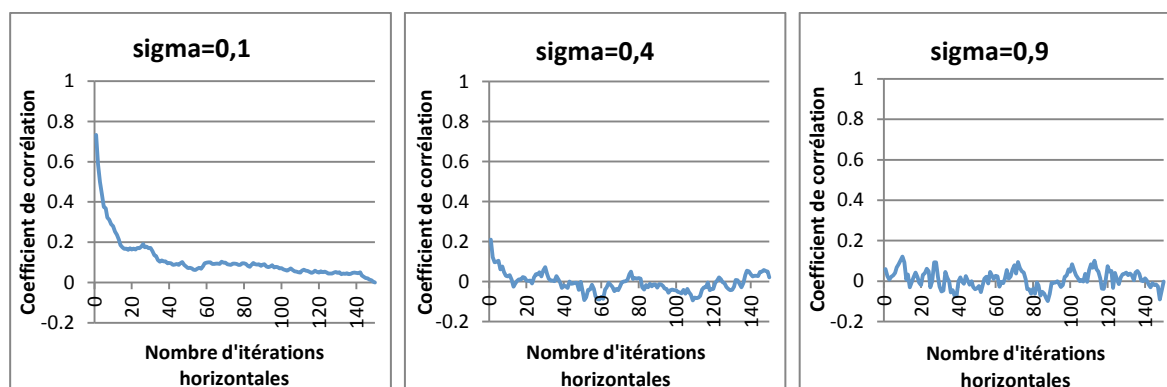


Figure 19 : Variation du coefficient de corrélation entre le scénario initial et le scénario obtenu après avoir appliqué T fois le noyau de transition, en fonction de T et pour différentes variances σ_k

Les résultats obtenus confirment notre intuition : plus la variance est faible, plus il faudra effectuer un grand nombre d'itérations pour obtenir l'indépendance. En effet, dans le cas où l'écart-type est égal à 0,1, le coefficient de corrélation au bout des premières itérations est supérieur à 0,6 tandis qu'il est inférieur à 0,2 lorsque σ_k est plus élevé.

En outre, nous remarquons que quelle que soit la valeur de σ , nous obtenons un coefficient de corrélation inférieur à 0,1 lorsque nous appliquons plus de 60 fois le noyau de transition. Nous pouvons donc considérer dans la suite du mémoire que le scénario initial et celui obtenu au bout de plus de 60 itérations sont indépendants.

Notons que cette étude est préalable à la mise en œuvre de l'algorithme et que nous n'avons pas tenu compte du rejet qui est parfois effectué dans l'algorithme, lorsque les fonds propres obtenus sont supérieurs au niveau fixé pour cette itération.

(2) Résultats pour différentes valeurs de T

Comme nous l'avons expliqué précédemment, le choix du nombre de fois où nous appliquons le noyau de transition pour diffuser le scénario peut jouer un rôle important. En effet, ce nombre doit être suffisamment élevé pour que le scénario obtenu à la fin de la diffusion soit indépendant du scénario à partir duquel il a été diffusé, mais appliquer le noyau de transition un nombre de fois trop important consommerait inutilement du budget.

Nous fixons la valeur de N à 100 et nous lançons vingt fois l'algorithme pour différents nombres d'itérations horizontales T , compris entre 10 et 150. Nous calculons ensuite le quantile moyen obtenu sur ces vingt lancers de l'algorithme pour chaque valeur de T .

Nous avons représenté ci-après ces quantiles moyens en fonction de T pour chacune des formes paramétriques. Nous avons également tracé en rouge la valeur de référence (51,21 pour la forme paramétrique à deux facteurs et 74,14 pour la forme paramétrique à quatre facteurs) afin de pouvoir visualiser rapidement pour quelles valeurs de T nous obtenons les meilleurs résultats.

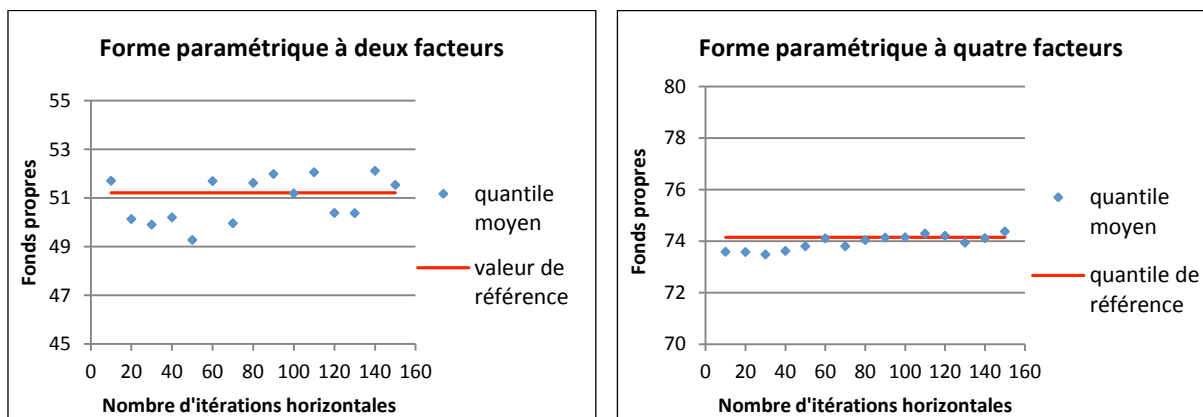


Figure 20 : Variation de la moyenne, sur vingt lancers de l'algorithme, de l'estimateur du quantile en fonction du nombre d'itérations horizontales T , pour les deux formes paramétriques

Nous remarquons tout d'abord que les quantiles moyens obtenus avec la forme paramétrique à quatre facteurs sont beaucoup plus proches de la valeur de référence qu'avec la forme paramétrique à deux facteurs. En effet, dans ce cas, quelle que soit la valeur de T choisie, l'écart relatif entre le quantile moyen et la valeur de référence est inférieur à 1 %. Nous déterminerons donc la valeur de T à choisir pour avoir les meilleurs résultats en nous basant sur la forme paramétrique à deux facteurs.

Pour cette dernière forme paramétrique, nous observons que les estimateurs du quantile sont les plus proches lorsque T est égale à 10, 60, 80, 100 et 150. En effet, dans ces quatre cas, l'écart relatif entre l'estimateur du quantile et la valeur de référence est inférieur à 1 %. Nous notons que pour les autres valeurs de T , l'écart entre l'estimateur du quantile et la valeur de référence est tout de même inférieur à 4 %. Finalement, il semblerait que le choix de ce paramètre influence peu les résultats.

Nous rappelons que dans l'analyse sur la dépendance entre le scénario initial et le scénario diffusé, nous avons mis en évidence qu'une soixantaine d'itérations horizontales étaient suffisantes pour obtenir l'indépendance. En outre, appliquer le noyau de transition plus de 60 fois n'améliore pas les résultats. Nous pouvons donc choisir une valeur de T égale à 60. Dans ce cas, la moyenne des estimateurs du quantile à 0,5 % des fonds propres est égale à 51,70, ce qui donne un écart relatif de 0,96 % par rapport au quantile de référence.

ii. Sensibilité au nombre de scénarios N

Nous allons à présent analyser les résultats obtenus suivant le nombre de scénarios N choisi. Intuitivement, nous pouvons penser que plus nous choisissons un N élevé, plus le quantile à 0,5 % des fonds propres sera calculé précisément. Cependant, choisir un nombre de scénarios élevé induit également un budget de calcul plus important.

Pour tester la sensibilité suivant le nombre de scénarios N , nous fixons la valeur du nombre d'itérations horizontales, T , à 100 et nous faisons varier N entre 10 et 150. Comme dans l'analyse de la sensibilité au paramètre T , nous lançons vingt fois l'algorithme pour chaque valeur de N et nous calculons la moyenne des estimateurs du quantile sur ces vingt lancers. Nous présentons ci-dessous les résultats obtenus pour la forme paramétrique à deux facteurs et pour celle à quatre facteurs :

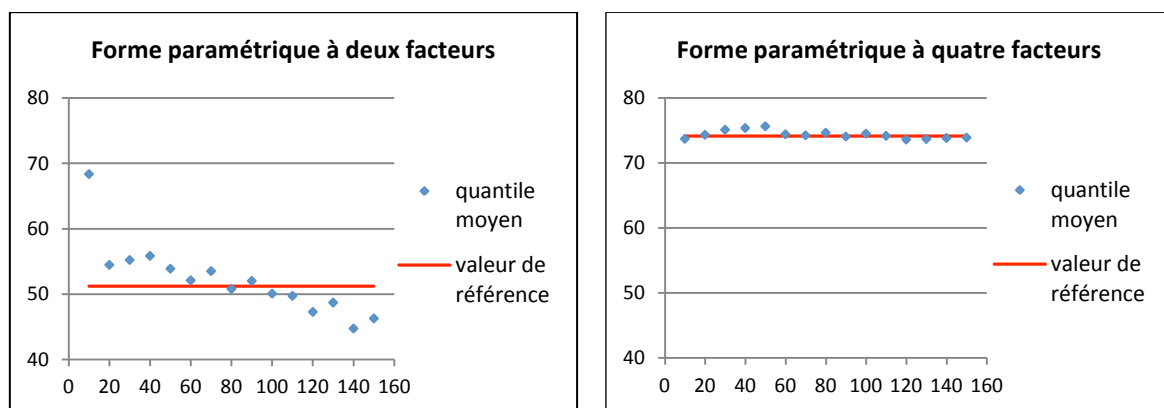


Figure 21 : Variation de la moyenne, sur vingt lancers de l'algorithme, de l'estimateur du quantile en fonction du nombre de scénarios N pour chaque forme paramétrique

Comme pour l'analyse de la sensibilité au nombre d'itérations horizontales T , le plus frappant dans les graphiques ci-dessus est que les résultats obtenus avec la forme paramétrique à quatre facteurs sont globalement beaucoup plus proches de la valeur de référence que ceux obtenus avec la forme paramétrique à deux facteurs.

Par ailleurs, nous observons sur le graphique de gauche que l'estimateur du quantile semble diminuer lorsque le nombre de scénarios N augmente. Nous rappelons que le nombre d'itérations verticales est lié à N . Ainsi, plus il y a de scénarios, plus nous effectuons d'itérations verticales. En outre, à chaque itération verticale m , la valeur de L_m telle que tous les scénarios appartiennent à $A_m = \{X|FP(X) < L_m\}$ diminue. Il n'est donc pas étonnant que pour des valeurs de N plus élevées l'estimateur du quantile soit plus faible que le quantile de référence. Ainsi, l'intuition que nous pouvions avoir initialement n'est pas vérifiée et lorsque nous choisissons un nombre de scénarios trop élevé, l'estimation du quantile est moins bonne.

Nous remarquons que, pour la forme paramétrique à deux facteurs, les meilleurs résultats sont obtenus pour des valeurs de N comprises entre 60 et 110. Lorsque le nombre de scénarios est inférieur à 60 les estimateurs du quantile sont plus de 5 % plus élevés que le quantile de référence. En particulier, lorsque $N = 10$, nous obtenons un quantile estimé égal à 68,37, soit un écart de 33,51 % avec la valeur de référence. Ainsi, il semblerait que prendre moins de 60 scénarios ne permette pas, dans cet exemple, de calculer de manière précise le quantile. De la même manière, lorsque le nombre de scénarios est supérieur à 110, le quantile à 0,5 % des fonds propres n'est pas bien estimé. Par ailleurs, pour la forme paramétrique à quatre facteurs, l'écart relatif entre le quantile estimé et le quantile de référence est inférieur à 1 % lorsque nous choisissons plus de 60 scénarios.

Ainsi, afin de ne pas consommer inutilement du budget de calcul, nous pouvons choisir de prendre un nombre de scénarios égal à 60. Cette valeur constitue en effet un bon compromis entre la proximité du résultat avec la valeur de référence et le budget de calcul. Dans ce cas, la moyenne de l'estimateur du quantile est égale à 52,11 pour la forme paramétrique à deux facteurs et à 74,43 pour la forme paramétrique à quatre facteurs.

b) Forêts aléatoires

Nous avons vu que les mesures de risque obtenues à l'aide de l'algorithme des forêts aléatoires étaient robustes lorsque le nombre d'arbres était suffisamment grand.

Nous allons maintenant étudier l'impact des autres paramètres sur la performance de cette approche. En particulier nous analyserons l'impact de faibles variations des bases de calibrage et totale sur les mesures de risque, leur sensibilité aux paramètres, ainsi qu'aux variables explicatives et à la forme paramétrique.

(1) Sensibilité aux faibles perturbations de la base totale ou de calibrages

Maintenant que nous sommes assurés de la robustesse des résultats, il est intéressant d'étudier quel est l'impact de petites variations sur les données en *input*. Nous allons ainsi mesurer la sensibilité des résultats aux perturbations de la base de calibrage en faisant :

- des variations aléatoires suivant une loi normale centrée de faible variance, sur les valeurs des facteurs de risque de la base de calibrage,
- des variations aléatoires suivant une loi normale centrée de faible variance, sur les valeurs des fonds propres de la base de calibrage,
- des translations ou dilatations des valeurs des fonds propres des bases de calibrage et de la base totale,
- en appliquant l'inverse de la fonction de répartition d'une normale centrée réduite à l'ensemble des facteurs de risque, qui suivent ainsi une loi uniforme.

Pour chacune des modifications listées ci-dessous, les résultats ne varient pas ou varient très peu : la modification de la mesure de risque à 0,5 %, 1 % et 2 % est au maximum de l'ordre de 10^{-2} . Nous pouvons donc faire l'hypothèse que les mesures de risque sont insensibles aux légères variations de la base totale et de la base de calibrage. Cette hypothèse est importante pour notre étude. En effet, dans le cas réel (lorsque les fonds propres de la base de calibrage sont obtenus par SdS), les fonds propres associés à chaque scénario ne sont que des estimations des fonds propres réels. Puisqu'il est impossible, dans la pratique, d'utiliser les SdS avec une infinité de scénarios secondaires, les fonds propres de la base de calibrage sont légèrement différents de leur contrepartie réelle.

(2) Sensibilité aux nombres de variables, m , mises en concurrence

Nous avons vu que la différence entre la méthode du Bagging et les forêts aléatoires réside dans le nombre de variables explicatives mises en concurrence lors de la construction des nœuds. Dans le cas des forêts, ce nombre est inférieur au nombre total de variables explicatives, d .

Cette pratique permet de décorréler les arbres de la forêt, et ainsi de diminuer le risque de sur-apprentissage. Breiman conseille de ne mettre en concurrence qu'un tiers des variables explicatives pour des régressions (la racine du nombre de variables explicatives pour les classifications). Le choix de ce paramètre est laissé à l'utilisateur puisque la valeur optimale dépend principalement de la fonction sous-jacente étudiée (forme paramétrique dans notre étude) et de la corrélation des variables explicatives. Nous avons cependant voulu tester la sensibilité des forêts à ce paramètre. Nous avons constaté que le paramètre optimal se situait entre les deux tiers et les trois quarts du nombre de variables explicatives.

Le tableau ci-dessous présente les valeurs des mesures de risque obtenues pour différentes valeurs de m sur la forme paramétrique à quatre facteurs, calibrée sur la même base que précédemment :

m	Mesure de risque par quantile recherché :		
	0,5 %	1 %	2 %
1	15,82	16,97	15,46
2	4,12	3,3	5,12
3	3,87	3,51	4,96
4	3,88	3,58	5

Figure 22 : Variation des mesures de risque pour différentes valeurs du paramètre m , avec la forme paramétrique à quatre facteurs

Nous constatons, que la valeur optimale de m se situe, ici, au trois quarts du nombre de variables explicatives. De plus, nous observons une faible sensibilité à ce paramètre pour des valeurs de m strictement supérieures à 1. Nous garderons ce paramètre fixé à deux tiers du nombre de variables explicatives dans la suite de notre étude.

(3) Sensibilité aux variables explicatives

Jusqu'à maintenant, nous n'avons utilisé que des approches dites orthogonales. Cela signifie que les divisions faites à chaque nœud, ne portent que sur un seul facteur de risque, et sont bilatérales.

En effet, comme nous l'avons évoqué dans la présentation des forêts aléatoires, la littérature actuelle (et les fonctions implémentées dans le logiciel R) s'est concentrée, jusqu'à aujourd'hui, sur les arbres dit « orthogonaux ». Ces arbres ont la particularité d'être construits avec des divisions de la forme : « $param_i > k$ » avec k et i des valeurs spécifiques à chaque nœud de chaque arbre.

Cependant, d'autre type d'arbres existent, notamment les arbres binaires obliques. Les segmentations, à chaque nœud, de ces derniers sont des tests bilatéraux sur les combinaisons linéaires des variables explicatives.

En créant, à partir des facteurs de risque initiaux (des lois normales centrées réduites indépendantes), de nouvelles variables, comme la somme ou la différence des facteurs de risque de départ, nous pouvons alors créer non plus des arbres orthogonaux mais semi-obliques. Le caractère oblique ne peut être obtenu dans la mesure où il faudrait créer autant de variables qu'il y a de directions possibles dans le plan (pour la dimension 2) autrement dit il faudrait créer une infinité de nouvelles variables de calibrage.

Nous avons donc testé cette méthode sur la forme paramétrique à deux facteurs, en ajoutant deux variables explicatives, ϵ_3 et ϵ_4 , construites de la manière suivante :

$$\epsilon_3 = \epsilon_1 - \epsilon_2$$

$$\epsilon_4 = \epsilon_1 + \epsilon_2$$

Les résultats obtenus sont les suivants :

Base de départ :		Mesure de risque par quantile recherché :			Rappels des résultats avec l'approche orthogonale :		
Nom	Nombre de points	0,5 %	1 %	2 %	0,5 %	1 %	2 %
Base de petite taille	121	1,18	2,3	1,54	3,29	4,62	5,44
Base de taille	225	1,48	1,63	1,64	3,03	1,54	2,83

moyenne							
Base de grande taille	324	1,2	1,36	1,34	1,61	2,55	2,93

Tableau 8 : Comparaison des mesures de risque entre les approches orthogonale et semi-oblique sur la forme paramétrique à deux facteurs de risque

Nous observons une nette amélioration des mesures de risque, particulièrement marquée pour la base de calibrage de petite taille. De plus, l'écart-type relatif des mesures de risque chute : ce dernier est divisé par plus de 2 avec l'approche semi-oblique.

Etudions l'impact de l'ajout de variables explicatives dans le cas de la forme paramétrique à quatre facteurs de risque.

Afin de garder une certaine symétrie de traitement, nous ajoutons le même type de variables explicatives :

$$\begin{aligned} \epsilon_5 &= \epsilon_1 - \epsilon_2 \\ \epsilon_6 &= \epsilon_1 + \epsilon_2 \\ \epsilon_7 &= \epsilon_1 - \epsilon_3 \\ \epsilon_8 &= \epsilon_1 + \epsilon_3 \\ \epsilon_9 &= \epsilon_1 - \epsilon_4 \\ \epsilon_{10} &= \epsilon_1 + \epsilon_4 \\ \epsilon_{11} &= \epsilon_2 - \epsilon_3 \\ \epsilon_{12} &= \epsilon_2 + \epsilon_3 \\ \epsilon_{13} &= \epsilon_2 - \epsilon_4 \\ \epsilon_{14} &= \epsilon_2 + \epsilon_4 \\ \epsilon_{15} &= \epsilon_3 - \epsilon_4 \\ \epsilon_{16} &= \epsilon_3 + \epsilon_4 \end{aligned}$$

Les résultats obtenus avec cette approche sont les suivants :

Base de départ :					Mesure de risque par quantile recherché :			Rappels des résultats avec l'approche orthogonale :		
Nom	Borne min	Borne max	Pas	Nombre de points	0,5 %	1 %	2 %	0,5 %	1 %	2 %
Base de grande taille	-3,3	3,3	1,1	2401	2,31	1,84	2,11	4,12	3,3	5,12

Tableau 9 : Comparaison des mesures de risque entre l'approche orthogonale et semi-oblique sur la forme paramétrique à quatre facteurs de risque

Remarque : le paramètre, m , qui est le nombre de variables explicatives mises en concurrence lors de la construction des nœuds, a été fixé à 10.

De la même manière que pour la forme paramétrique à deux facteurs de risque, nous observons une nette diminution des mesures de risque.

Nous pouvons donc considérer que l'approche semi-oblique est plus pertinente que l'approche orthogonale.

(4) Ajout de points de calibrage

Nous avons vu que l'ajout de variables explicatives améliore la mesure de risque. Cependant les différences de résultats, entre les trois bases de calibrage testées, demeurent significatives. Dans cette partie, nous allons nous intéresser à l'optimisation de la base de calibrage. Pour cela, nous allons mettre en œuvre les méthodes d'ajout de points à la base de calibrage, présentées *supra*.

Base de calibrage :			Mesure de risque avec le premier algorithme d'ajout de points			Mesure de risque avec le deuxième algorithme d'ajout de points		
Nom	Ajout de points	Nombre de points	0,5 %	1 %	2 %	0,5 %	1 %	2 %
Base de petite taille	OUI	121+104 = 225	1,08	1,77	1,10	1,29	1,5	1,56
Base de taille moyenne	NON	225	1,48	1,63	1,64	1,48	1,63	1,64
Base de petite taille	OUI	121+203 = 324	1,07	1,95	1,22	1,08	1,08	1,15
Base de taille moyenne	OUI	225+99 = 324	1,25	1,21	1,30	1,14	1,19	1,30
Base de grande taille	NON	324	1,2	1,36	1,34	1,2	1,36	1,34

Tableau 10 : Comparaison des mesures de risque entre les deux approches d'ajout de points avec des arbres semi-obliques, sur la forme paramétrique à deux facteurs de risque

Nous constatons que les forêts aléatoires sont globalement plus performantes sur des bases de calibrage complétées par un algorithme d'ajout de points, que sur les bases initiales de même taille. De plus, ces deux algorithmes donnent des résultats du même ordre de grandeur, même si le premier algorithme semble mieux marcher sur la petite base de calibrage. Cela semble logique, puisque dans la deuxième approche le point ajouté sera d'autant plus éloigné du point mal estimé que la base est de petite taille.

Nous n'avons étudié qu'une seule base de calibrage pour la forme paramétrique à quatre facteurs. Nous rappelons que nous avons fait ce choix à cause du trop grand nombre de facteurs de risque. En effet, rajouter une segmentation par axe dans notre maillage, aurait conduit à une base de taille 4096 ($8^4 = 4096$), ce qui est trop grand pour rendre l'approche des forêts aléatoires efficace. Nous avons donc étudié l'évolution des mesures de risque (calculées avec des arbres semi-obliques) en fonction du nombre de points ajoutés à la base initiale (composée de 2401 points). Les résultats sont présentés dans le graphique ci-dessous :

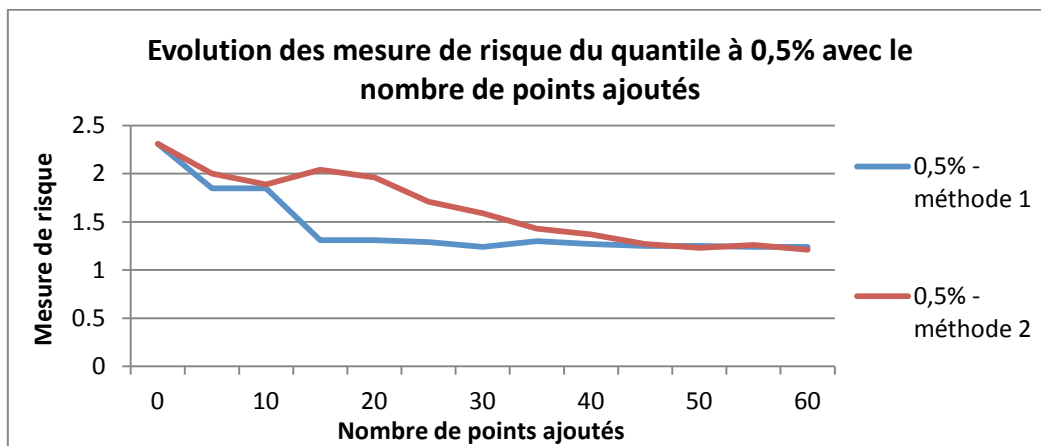


Figure 23 : Evolution des mesures de risque pour le quantile à 0,5 % en fonction du nombre de points ajoutés, calculé avec les deux méthodes d'ajouts de points et des arbres semi-obliques sur la forme paramétrique à quatre facteurs

Nous constatons une diminution significative de la mesure de risque, principalement pour les 40 premiers points ajoutés. Un seuil apparaît à partir de 40 points ajoutés, ce qui semble montrer les limites des méthodes d'ajouts de points. La première méthode d'ajout de points semble un peu plus performante que la deuxième sur cet exemple, dans la mesure où les résultats sont meilleurs lors de l'ajout de 30 à 40 points.

5. Analyse de la sensibilité par rapport à la forme paramétrique

Dans les parties précédentes, nous avons appliqué nos méthodes à deux formes paramétriques, basées respectivement sur deux et quatre facteurs de risque. Nous avons ainsi remarqué que les résultats obtenus pouvaient être différents. Nous allons, à présent, nous intéresser uniquement à des formes paramétriques à deux facteurs de risque. Nous comparons, dans cette partie, trois formes paramétriques, dont celle que nous avons utilisée dans les parties précédentes.

a) Décomposition d'un événement rare en événements moins rares

Nous pouvons commencer par comparer les résultats obtenus avec les trois formes paramétriques en analysant l'allure de leur zone quantile. En effet, l'algorithme de décomposition d'un événement rare en événements moins rares aboutit à l'obtention d'un échantillon de scénarios situés au voisinage du quantile. Nous pouvons représenter cet échantillon dans le plan (facteur de risque action, facteur de risque zéro-coupon), comme le montrent les graphiques ci-dessous :

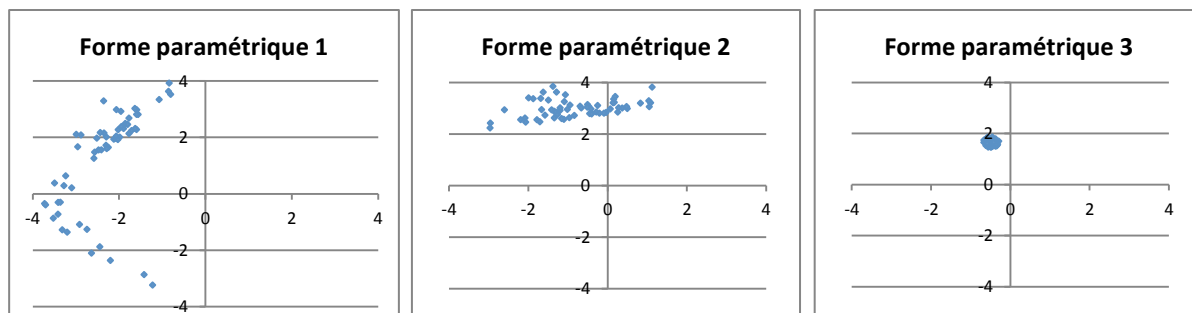


Figure 24 : Représentation de la zone quantile pour chacune des formes paramétriques

Nous remarquons que les zones quantiles sont complètement différentes pour chacune des formes paramétriques. Pour la première forme paramétrique, les scénarios de queue forment un croissant de lune, avec des facteurs de risque action négatifs. Dans la deuxième forme paramétrique, le facteur de risque action ne semble pas jouer un rôle important dans la détermination des fonds propres, contrairement au facteur de risque zéro-coupon. Le voisinage du quantile forme ainsi une droite oblique, avec des facteurs de risque zéro-coupon compris entre 2 et 4. Enfin, la troisième forme paramétrique présente une zone quantile très ciblée, puisque le voisinage du quantile peut être représenté par un point de coordonnées environ égaux à -0,4 et 1,7.

Nous avons vu que les zones quantiles des trois formes paramétriques étudiées étaient complètement différentes, et nous pouvons nous demander s'il en est de même des résultats sur le quantile estimé. Ainsi, nous avons calculé dans les trois cas un quantile de référence. Par ailleurs, comme nous avons déterminé précédemment que les paramètres optimaux étaient $T = 60$ et $N = 60$, nous avons lancé vingt fois l'algorithme avec ces paramètres, pour chacune des formes paramétrique, et nous avons comparé la moyenne obtenue avec le quantile de référence. Les résultats sont présentés dans le tableau ci-dessous :

	Quantile moyen	Valeur de référence	Ecart relatif entre le quantile moyen et la valeur de référence
Forme paramétrique 1	51,26	51,21	0,10 %
Forme paramétrique 2	-480,9	-451,7	6,46 %
Forme paramétrique 3	90,71	90,68	0,03 %

Tableau 11 : Quantile moyen obtenu sur vingt lancers de l'algorithme pour les trois formes paramétriques étudiées

Nous observons tout d'abord que les meilleurs résultats sont ceux obtenus avec la première et la troisième forme paramétrique. En effet, nous obtenons des écarts relatifs entre l'estimateur du quantile et le résultat de référence respectivement égaux à 0,10 % et 0,03 %. Nous pouvons penser, pour la forme paramétrique 3, que ce résultat est dû au fait que la zone quantile est très localisée. En effet, nous rappelons que, pour cette forme paramétrique, les scénarios de queue forment un point. A l'opposé, les résultats obtenus pour la forme paramétrique 2 sont beaucoup moins bons.

Nous pouvons conclure de cette analyse que les résultats obtenus par l'algorithme de décomposition d'un événement rare en événements moins rares dépendent fortement de la forme paramétrique, et

plus précisément de l'allure de la zone quantile. Ainsi, les résultats seront excellents dans le cas de scénarios de queue très localisés, mais pourront se dégrader lorsque le voisinage du quantile est plus étalé.

b) Forêts aléatoires

De la même manière que pour les événements rares, nous allons étudier la sensibilité de nos résultats à la forme paramétrique. Différents algorithmes, abordés précédemment, seront donc testés sur les deux nouvelles formes paramétriques.

(1) Comparaison des approches orthogonales et semi-obliques

Nous pouvons comparer les résultats obtenus avec des bases orthogonales et semi-obliques :

Résultats pour la forme paramétrique numéro 2 :

Base de calibrage :			Mesure de risque par quantile recherché :		
Nom	Semi-oblique?	Nombre de points	0,5 %	1 %	2 %
Base de petite taille	NON	121	2,48	2,08	1,80
Base de petite taille	OUI	121	1,78	1,75	1,91
Base de taille moyenne	NON	225	2,5	1,54	1,51
Base de taille moyenne	OUI	225	1,63	1,41	1,35
Base de grande taille	NON	324	1,69	1,64	1,68
Base de grande taille	OUI	324	1,59	1,24	1,33

Tableau 12 : Comparaison des mesures de risque entre l'approche orthogonale et semi-oblique, pour la forme paramétrique numéro 2

Résultats obtenus pour la forme paramétrique numéro 3 :

Base de calibrage :			Mesure de risque par quantile recherché :		
Nom	Semi-	Nombre de points	0,5 %	1 %	2 %

oblique?					
Base de petite taille	NON	121	10,96	5,47	2,89
Base de petite taille	OUI	121	3,86	2,62	2,51
Base de taille moyenne	NON	225	2,94	2,96	2,74
Base de taille moyenne	OUI	225	2,05	1,63	2,30
Base de grande taille	NON	324	2,14	2,72	1,54
Base de grande taille	OUI	324	1,90	1,73	1,32

Tableau 13 : Comparaison des mesures de risque entre l'approche orthogonale et semi-oblique, pour la forme paramétrique numéro 3

De la même manière que pour la forme paramétrique numéro 1, les mesures de risques sont bien meilleures lorsque les arbres sont semi-obliques, bien que les forêts aléatoires soient un peu moins performantes sur la troisième forme paramétrique que sur les deux autres.

Par ailleurs, les zones quantiles de la forme paramétrique numéro 3 sont plus grandes (mesures de risque plus élevées pour une base totale de la même taille) que dans le cas de la forme paramétrique à quatre facteurs. Ainsi, malgré le doublement du nombre de facteurs de risque, notre algorithme est plus efficace sur la forme paramétrique à quatre facteurs que sur la forme paramétrique numéro 3. Cela est probablement dû à la répartition des scénarios les plus adverses de la forme paramétrique numéro 3. En effet, plus les scénarios les plus risqués sont proches les uns des autres, et plus il sera difficile pour la forêt aléatoires de les définir.

(1) Etude de l'ajout de points à la base de calibrage

Il s'agit maintenant de nous intéresser à l'ajout de points à la base de calibrage. Nous avons présenté deux méthodes dont la différence se situe dans le calcul du « point mal estimé ». Comme dans la partie précédente, nous allons comparer les performances des différentes méthodes sur les deux nouvelles formes paramétriques, à budget de calcul (ou nombre de points dans la base de calibrage finale) fixé.

Résultats pour la forme paramétrique numéro 2 :

Base de calibrage :	Mesure de risque avec le premier algorithme d'ajout de points :	Mesure de risque avec le deuxième algorithme d'ajout de points :

Nom	Ajout de points	Nombre de points	0,5 % 1 % 2 %			0,5 % 1 % 2 %		
			0,5 %	1 %	2 %	0,5 %	1 %	2 %
Base de petite taille	OUI	121+104 = 225	1,13	1,03	1,62	1,22	1,36	2,22
Base de taille moyenne	NON	225	1,63	1,41	1,35	1,63	1,41	1,35
Base de petite taille	OUI	121+203 = 324	1,07	1,01	1,99	1,35	1,18	1,37
Base de taille moyenne	OUI	225+99 = 324	1,18	1,08	1,33	1,12	1,17	1,28
Base de grande taille	NON	324	1,59	1,24	1,33	1,59	1,24	1,33

Tableau 14 : Comparaison des méthodes d'ajout de points avec des arbres semi-obliques, pour la forme paramétrique numéro 2

Résultats pour la forme paramétrique numéro 3 :

Base de calibrage :			Mesure de risque avec le premier algorithme d'ajout de points :			Mesure de risque avec le deuxième algorithme d'ajout de points :		
Nom	Ajout de points	Nombre de points	0,5 %	1 %	2 %	0,5 %	1 %	2 %
Base de petite taille	OUI	121+104 = 225	2,56	1,73	1,76	1,02	1,27	1,67
Base de taille moyenne	NON	225	2,05	1,63	2,30	2,05	1,63	2,30
Base de petite taille	OUI	121+203 = 324	1,03	1,25	1,55	1,02	1,16	1,08
Base de taille moyenne	OUI	225+99 = 324	1,84	1,33	1,74	1,10	1,42	1,27
Base de grande taille	NON	324	1,90	1,73	1,32	1,90	1,73	1,32

Tableau 15 : Comparaison des méthodes d'ajout de points avec des arbres semi-obliques, pour la forme paramétrique numéro 3

Nous constatons qu'il est généralement préférable d'utiliser une méthode d'ajout de points plutôt que d'utiliser une base de calibrage de même taille et uniformément répartie. En effet, l'algorithme ajoute les points de la base totale qui ont des fonds propres estimés faibles, et dont l'estimation du rang n'est pas bonne dans la base de calibrage. Les points ainsi ajoutés sont mieux estimés par la

forêt aléatoire, or comme ils sont, par définition, risqués (ils font partie des 5 % pires scénarios estimés), leur présence dans la base de calibrage a tendance à diminuer la mesure de risque.

Sur la deuxième forme paramétrique, l'amélioration des résultats est du même ordre de grandeur pour les deux algorithmes. Cependant, le deuxième algorithme d'ajout de points performe bien mieux que le premier, sur la forme paramétrique numéro 3. Cela est sans doute dû à la base de calibrage. En effet la première approche tend à ajouter des points proches de ceux déjà présents dans la base de calibrage. Or la base de départ ne fait que 121 points ce qui rend cette méthode non optimale. Dans la mesure où le deuxième algorithme d'ajout de points évite d'introduire des points trop proches les uns des autres dans la base de calibrage, nous pouvons le considérer comme plus prudent, et généralement plus performant.

Cependant, dans le cas des mesures de risque du quantile à 2 %, il peut être préférable de ne pas ajouter de points. Ce phénomène est probablement dû au critère de rang dans le calcul du « point mal estimé » (paramètre $q_{risqué}$). Effectivement, par ce critère, nous imposons que le point mal estimé fasse partie des 2 % de scénarios de calibrage les plus adverses. La probabilité de sélectionner ainsi un scénario appartenant réellement à la zone quantile à 2 % est plus faible que si $q_{risqué}$ était fixé à 5 %. Il s'agit donc de régler ce paramètre avec minutie.

Remarque : les mesures de risque obtenues pour les quantiles à 0,5 % et 1 % varient peu si nous modifions $q_{risqué}$ de 2 % à 5 %. Ce n'est qu'à partir de 10 % que les mesures de risques se dégradent.

6. Comparaison des méthodes

Dans les parties précédentes, nous avons mis en œuvre nos deux techniques de détermination du capital économique, et nous avons étudié les caractéristiques des résultats (robustesse et sensibilité aux paramètres et à la forme paramétrique). Nous pouvons à présent les comparer entre elles.

La principale différence entre les deux approches repose sur la nature même du résultat obtenu. En effet, la méthode de décomposition d'un événement rare en événements moins rares aboutit à l'estimation du quantile réel de la distribution des fonds propres, tandis que dans le cas des forêts aléatoires il s'agit d'un quantile empirique, égal au quantile à 0,5 % de la base totale étudiée. Cependant, cette différence induit aussi que le quantile obtenu en utilisant les forêts aléatoires est égal à la valeur de référence (calculée avec un SdS complet), puisque dans les deux cas nous le déterminons en générant un grand nombre de simulations secondaires à partir du scénario quantile.

Nous avons pu montrer que dans le cas de la méthode de décomposition d'un événement rare en événements moins rares l'algorithme adapté donnait des meilleurs résultats que l'algorithme simple. En effet, en déterminant la valeur de l'écart-type utilisé dans la diffusion à partir des scénarios non diffusés, nous n'avons pas besoin de choisir de manière arbitraire ce paramètre, et nous obtenons un quantile à 0,5 % estimé beaucoup plus proche de la valeur de référence. Pour ce dernier algorithme, les paramètres optimaux, c'est-à-dire ceux aboutissant au meilleur résultat avec le budget de calcul le plus faible, sont $T = 60$ et $N = 60$. Nous pouvons ainsi calculer le nombre de projections ALM lorsque nous utilisons ces paramètres :

$$C = N \times 1000 + \left\lceil \frac{\log(0,5\%)}{\log(1-N^{-1})} \right\rceil \times 1000 = 376\,000.$$

Pour l'algorithme utilisant les forêts aléatoires, nous avons pu voir que le calibrage était meilleur si nous utilisions des points répartis uniformément plutôt que selon une loi normale centrée réduite. En effet, cela permettait d'avoir plus de points de calibrage dans les extrêmes. Par ailleurs, nous avons mis en évidence que la méthode optimale était celle utilisant des arbres semi-obliques, avec l'ajout d'une cinquantaine de points à la base de calibrage. Dans ce dernier cas, nous rappelons qu'en utilisant $n = 255$ points de calibrage, la mesure de risque obtenue est de l'ordre de 1,30 pour les formes paramétriques à deux facteurs, excepté pour la troisième, pour laquelle la mesure de risque vaut 2,56. Pour être certain que la zone quantile déterminée par l'algorithme contient le scénario quantile, nous pouvons calculer précisément les fonds propres pour $z = 3 \times [N \times 0,5 \text{ \%}]$, c'est-à-dire que dans le cas d'une base totale contenant 10 000 scénarios nous effectuons 150 calculs précis de fonds propres. Le nombre de projections ALM utilisées par cet algorithme est donc le suivant :

$$C = (n + z) \times 1000 = 405\,000.$$

Nous remarquons que les budgets de calcul des deux algorithmes proposés sont proches, et correspondent à une réduction de plus de 99 % par rapport au SdS complet. Nous notons également que le nombre de projections ALM est du même ordre de grandeur que pour les méthodes alternatives présentées dans la partie II.B. En outre, dans le cas de formes paramétriques à deux facteurs de risque, les deux méthodes donnent une bonne estimation du quantile : cette valeur est égale à celle de référence dans le cas des forêts aléatoires et à 51,26 pour la méthode de décomposition d'un événement rare, soit une différence de seulement 0,10 %. Ainsi, pour cette forme paramétrique, les deux méthodes peuvent être utilisées et donnent des résultats tout à fait comparables.

En revanche, pour des formes paramétriques à quatre facteurs de risque, il est préférable d'utiliser la méthode de décomposition d'un événement rare. En effet, lorsque le nombre de facteurs de risque augmente, le nombre de points nécessaires pour calibrer les forêts aléatoires explose, ce qui empêche de diminuer le temps de calcul par rapport au SdS complet. Au contraire, la méthode de décomposition d'un événement rare peut facilement être utilisée, et nous remarquons même que les résultats sont meilleurs que pour la forme paramétrique à deux facteurs. Néanmoins, il convient de noter que les résultats obtenus pour ce dernier algorithme semblent dépendre davantage de la forme de la zone quantile que du nombre de facteurs de risque utilisés.

Nous pouvons conclure de cette analyse que les deux méthodes permettent d'estimer correctement le quantile à 0,5 % dans des cas simples (zone quantile peu étendue pour la méthode de décomposition d'un événement rare et nombre de facteurs de risque inférieur à quatre pour les forêts aléatoire) avec une grande diminution du nombre de projections ALM par rapport au SdS complet. Cependant, il semblerait que nous soyons confrontés au même problème que pour les méthodes alternatives de la partie II.B., et qu'il soit difficile d'obtenir une bonne estimation du quantile pour des portefeuilles plus complexes.

Nous allons à présent étudier dans quelle mesure s'affranchir d'une forme paramétrique peut nous permettre d'améliorer davantage les résultats, grâce à l'utilisation d'une forme paramétrique locale pour la méthode de décomposition d'un événement rare en événements moins rares, et à l'ajout de nouvelles variables comme des puissances de facteurs de risque ou des termes croisés pour les forêts aléatoires.

B. Application dans le cadre d'un modèle de gestion actif/passif réel

Dans cette partie, nous allons présenter les résultats obtenus dans un cas réel, en supposant que nous ne disposons pas directement des facteurs de risque et d'une forme paramétrique mais seulement de scénarios économiques ainsi que des facteurs de risques extraits de ceux-ci.

Le modèle étudié est un modèle à trois facteurs. Les risques associés sont les suivants :

- un risque action,
- un risque de taux,
- un risque de mortalité.

Nous avons à notre disposition une base de données issue d'un portefeuille « classique » d'assurance vie, comportant les facteurs de risque de chaque scénario ainsi que les fonds propres associés. Cette base contient 50 000 couples (facteurs de risque, fonds propres).

1. Décomposition d'un événement rare en événements moins rares

Nous avons proposé, dans la partie théorique sur la décomposition d'un événement rare en événements moins rares, un algorithme permettant de calculer le quantile à 0,5 % des fonds propres à un an pour en déduire le capital économique avec un nombre de projection ALM réduit. Nous allons donc, dans cette partie, mettre en œuvre cette méthode afin de vérifier que la diminution du budget de calcul ne détériore pas la qualité des résultats.

Nous nous intéressons tout d'abord à la partie de l'algorithme utilisant des formes paramétriques locales. Celles-ci doivent permettre de dire si les fonds propres sont supérieurs ou inférieurs au niveau L_m pour chaque itération m . Nous allons donc commencer par vérifier que les formes paramétriques locales conservent le rang des fonds propres. Pour cela, nous calibrons une forme paramétrique polynomiale de degré 3 à l'aide de 100 scénarios primaires, comme nous avons proposé de le faire dans l'algorithme. Nous appliquons ensuite cette forme paramétrique à 30 000 scénarios, et nous comparons les rangs obtenus aux rangs réels. Nous obtenons alors le graphique ci-dessous :

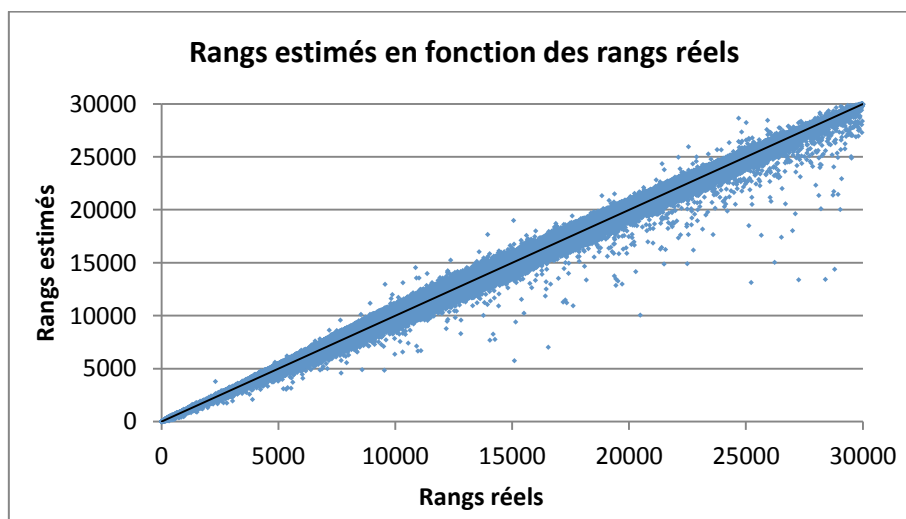


Figure 25 : Représentation des rangs des fonds propres calculés à partir de la forme paramétrique locale en fonction des rangs des fonds propres réels

Ce graphique laisse supposer que la forme paramétrique calibrée sur 100 scénarios conserve bien le rang des fonds propres, en particulier pour les fonds propres les plus faibles. Cela est confirmé par le taux de corrélation des rangs de Spearman, qui vaut 0,997. Cependant, cette conservation n'est pas parfaite. Aussi, la sécurité que nous avons placée dans l'algorithme et qui consiste à vérifier, à l'aide de calculs ALM, que le scénario obtenu à la fin de la diffusion correspond bien à des fonds propres inférieurs à L_m , et à relancer la diffusion dans le cas contraire, apparaît comme étant nécessaire.

Cette étude ayant montré qu'il est possible d'utiliser une forme paramétrique locale pour estimer les fonds propres lors des diffusions horizontales et les comparer à L_m , nous pouvons à présent appliquer notre méthode. Dans les parties précédentes, nous avons montré que les meilleurs résultats étaient obtenus lorsque le nombre d'itérations horizontales T , et le nombre de scénarios N , étaient égaux à 60. Nous avons donc fixé la valeur des paramètres à 60, et lancé vingt fois l'algorithme utilisant la forme paramétrique locale et celui ne l'utilisant pas, afin de comparer les quantiles moyens obtenus. Les résultats sont présentés dans le tableau ci-dessous :

	Algorithme utilisant la forme paramétrique locale	Algorithme n'utilisant pas la forme paramétrique locale	Valeur de référence
Quantile à 0,5 % des fonds propres	3 720	3 518	3 675
Ecart relatif par rapport à la valeur de référence	1,24 %	-4,27 %	

Tableau 16 : Quantile moyen obtenu, sur vingt lancers, avec un algorithme utilisant la forme paramétrique locale et un algorithme ne l'utilisant pas

Nous pouvons remarquer que l'utilisation de la forme paramétrique locale semble améliorer les résultats. En effet, dans cet exemple, nous observons que l'estimation du quantile est plus proche de la valeur de référence lorsque nous avons utilisé la forme paramétrique locale.

Néanmoins, nous notons que l'algorithme de décomposition d'un événement rare ne permet pas d'estimer correctement le quantile à 0,5 % des fonds propres à l'aide de trois facteurs de risque pour ce portefeuille. En effet, dans les exemples précédents, excepté pour la forme paramétrique 2, l'erreur relative sur l'estimation du quantile est inférieure à 1 % pour un bon choix de paramètres, tandis que pour ce portefeuille l'erreur vaut -4,27 %.

2. Forêts aléatoires

Nous avons pu constater, dans la partie précédente, qu'ajouter des variables explicatives telles que des combinaisons linéaires des facteurs de risque apportait un gain en termes de mesures de risque : elles pouvaient être divisées par un facteur 2. Cependant nous ne pouvions ajouter des puissances et termes croisés des facteurs de risque aux variables explicatives, puisque les formes paramétriques en dépendaient directement. Nous risquons de biaiser notre étude. Puisque nous ne disposons plus

d'une forme paramétrique, il est désormais possible de tester cette approche. Nous allons donc vérifier si l'ajout de puissances et de termes croisés de facteurs de risque, améliore les résultats.

Nous ne disposons pas de base de calibrage, il s'agit donc d'en créer une. Pour cela nous allons utiliser une technique analogue à celle employée dans les algorithmes d'ajout de points à la base de calibrage.

a) Base de calibrage

Il est évident qu'il n'existe pas de base de calibrage idéale quelle que soit la table totale. Étant donné que les facteurs de risques suivent des lois normales centrées réduites et qu'il faut tenir compte de la taille de la base de calibrage, nous avons défini comme objectif de base de calibrage, une grille (de dimension 3) dont les coefficients varient de -3,85 à 3,85 avec un pas de 1,1. Ces valeurs ont été définies de manière arbitraire de manière à limiter la taille de cette base tout en ayant des points extrêmes.

La base de calibrage est constituée des points de la base totale les plus proches (pour la norme 2) des points de la base objectif.

Ainsi nous avons une base de calibrage ayant un nombre de lignes n_{cali} :

$$n_{cali} = \left(\frac{3,85 - (-3,85)}{1,1} \right)^3 = 512$$

La base de calibrage fait donc 512 lignes. Après calibrage sur les 50 000 points à notre disposition, il est possible de représenter les projections des points de calibrage sur trois plans (chacun étant généré par deux facteurs de risque) ainsi que la base de calibrage objectif.

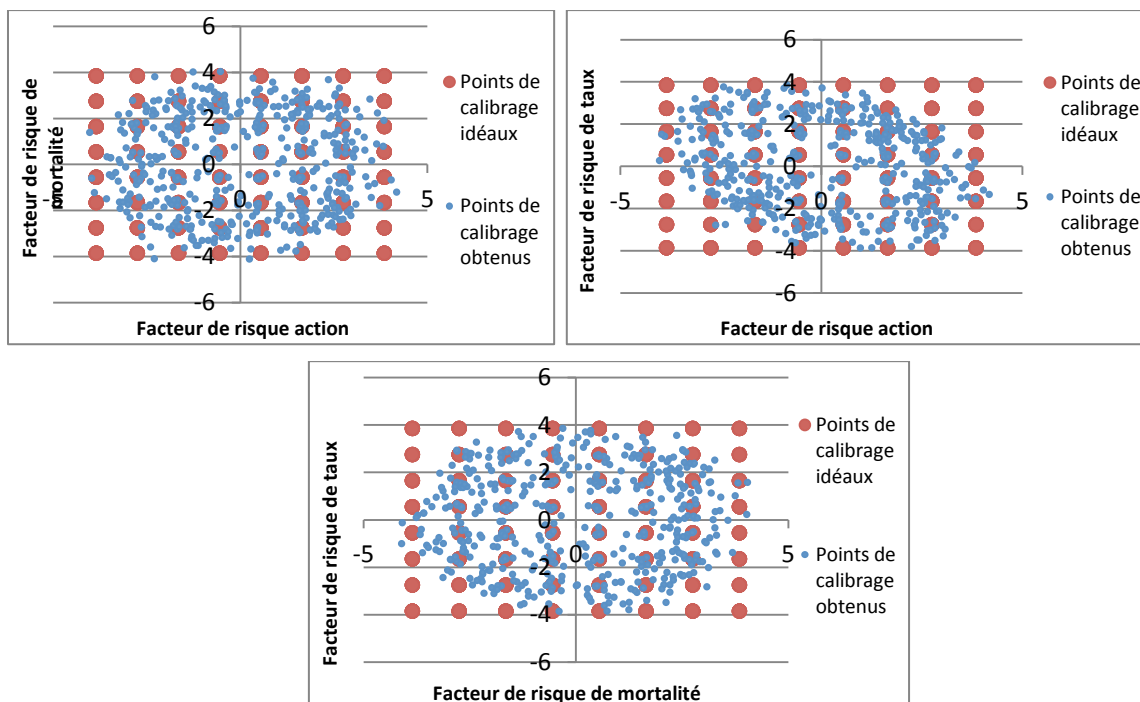


Figure 26 : Graphique représentant les points de calibrage idéaux et obtenus

Nous pouvons constater que les points de la base de calibrage objectif, les plus éloignés de l'origine (se situant aux 4 coins sur les graphiques ci-dessus) ne sont pas bien représentés dans la base. En effet, les points bleus, représentant les points de la base de calibrage réelle, sont éloignés des coins. Ceci a pour effet de limiter le pouvoir prédictif de notre modèle dans ces zones extrêmes. Cependant, dans la mesure où l'objectif final est l'estimation du quantile à 0,5 % de la distribution de fonds propres de la base totale, ce phénomène est marginal, les points les plus adverses de cette dernière étant tout de même inclus dans la base de calibrage (ce sont souvent ceux dont les coordonnées sont les plus extrêmes).

b) Etude des résultats obtenus

La base totale faisant 50 000 lignes, plusieurs options s'offrent à nous. Nous pouvons, par exemple, étudier plusieurs lots de 10 000 lignes chacun ou lancer les algorithmes sur la base totale toute entière.

Nous avons vu que les méthodes semi-obliques sont plus performantes que les méthodes orthogonales. Nous n'allons donc étudier que les méthodes semi-obliques. Comme évoqué *supra*, nous porterons une attention particulière au gain d'efficacité résultant de l'utilisation des puissances et termes croisés des facteurs de risque comme variables explicatives.

Nous avons donc testé trois approches différentes, sur l'algorithme qui n'ajoute pas de points de calibrage pour faciliter l'interprétation des résultats :

- Première approche : les variables explicatives sont composées des trois facteurs de risque ($\epsilon_1, \epsilon_2, \epsilon_3$) ainsi que des bissectrices : ($\epsilon_1 - \epsilon_2, \epsilon_1 + \epsilon_2, \epsilon_1 - \epsilon_3, \epsilon_1 + \epsilon_3, \epsilon_2 - \epsilon_3, \epsilon_2 + \epsilon_3$). Nous obtenons 9 variables explicatives.
- Deuxième approche : nous ajoutons aux variables explicatives de la première approche les puissances et termes croisés de facteurs de risque de degré 2 : ($\epsilon_1^2, \epsilon_2^2, \epsilon_3^2, \epsilon_1\epsilon_2, \epsilon_1\epsilon_3, \epsilon_2\epsilon_3$). Le nombre total de variables explicatives augmente à 15.
- Troisième approche : nous ajoutons aux variables explicatives de la deuxième approche les puissances et termes croisés de facteurs de risque de degré 3 : ($\epsilon_1^3, \epsilon_2^3, \epsilon_3^3, \epsilon_1^2\epsilon_2, \epsilon_1^2\epsilon_3, \epsilon_2^2\epsilon_3, \epsilon_2^2\epsilon_1, \epsilon_3^2\epsilon_1, \epsilon_3^2\epsilon_2, \epsilon_1\epsilon_2\epsilon_3$). Le nombre total de variables explicatives augmente finalement à 25.

Les résultats obtenus avec la première approche (semi-oblique sans terme croisé) sont présentés dans le tableau ci-dessous :

Base totale			Mesure de risque par quantile en approche semi- oblique		
Première ligne de la base totale	Dernière ligne de la base totale	Nombre de lignes	0,5 %	1 %	2 %
1	10 000	10 000	1,20	1,31	1,54
10 001	20 000	10 000	1,48	1,71	1,42
20 001	30 000	10 000	2,25	1,40	1,34

30 001	40 000	10 000	1,48	1,19	1,50
40 001	50 000	10 000	1,95	1,12	1,68
1	50 000	50 000	1,98	1,28	1,52

Tableau 17 : Comparaison des mesures de risques obtenues par l'approche semi-oblique sur différentes parties de la base mise à notre disposition

Les mesures de risque sont généralement inférieures à 2 sauf lorsque nous définissons comme base les scénarios compris entre la 20 001 ième et la 30 000 ième ligne. Ainsi en dépit d'une certaine variance dans les résultats, ces derniers justifient l'utilisation de la méthode des forêts aléatoires. Le tableau suivant détaille les résultats obtenus avec la deuxième approche qui est un peu plus précise : des termes croisés d'ordre 2 étant ajoutés aux variables explicatives.

Les résultats obtenus avec la deuxième approche (semi-oblique avec termes croisés de degré 2) sont présentés ci-dessous :

Base totale			Mesure de risque par quantile en approche semi-oblique		
Première ligne de la base totale	Dernière ligne de la base totale	Nombre de lignes	0,5 %	1 %	2 %
1	10 000	10 000	1,08	1,25	1,20
10 001	20 000	10 000	1,23	1,36	1,16
20 001	30 000	10 000	1,80	1,20	1,24
30 001	40 000	10 000	1,48	1,09	1,34
40 001	50 000	10 000	1,85	1,18	1,10
1	50 000	50 000	1,64	1,28	1,28

Tableau 18 : Comparaison des mesures de risques obtenues par l'approche semi-oblique en incluant des termes croisés d'ordre 2, sur différentes parties de la base mise à notre disposition

Nous constatons que les résultats sont meilleurs avec les termes croisés d'ordre 2 que sans ceux-ci. En effet, les mesures de risque sont toujours inférieures à 2. Nous pouvons dès lors nous demander si l'ajout de nouvelles variables explicatives (d'autres puissances et termes croisés des facteurs de risque par exemple) est bénéfique.

Les résultats obtenus avec la troisième approche (semi-oblique avec termes croisés de degré 3) sont présentés ci-dessous:

Base totale	Mesure de risque par quantile en approche semi-oblique
-------------	--------------------------------------------------------

Première ligne de la base totale	Dernière ligne de la base totale	nombre de lignes	0,5 %	1 %	2 %
1	10 000	10 000	1,08	1,27	1,16
10 001	20 000	10 000	1,33	1,29	1,09
20 001	30 000	10 000	1,38	1,26	1,27
30 001	40 000	10 000	1,30	1,12	1,43
40 001	50 000	10 000	2,12	1,29	1,17
1	50 000	50 000	1,57	1,27	1,28

Tableau 19 : Comparaison des mesures de risques obtenues par l'approche semi-oblique en incluant des termes croisés d'ordre 3, sur différentes parties de la base mise à notre disposition

Les mesures de risque de cette dernière approche sont sensiblement équivalentes à celles de la méthode précédente. Nous pouvons cependant remarquer que même si la dernière méthode semble être plus performante de façon générale, elle ne l'est pas lorsque la base utilisée correspond aux lignes comprises entre 40 001 et 50 000.

Bien que la méthode des forêts aléatoires sélectionne, d'elle-même, les variables ayant un caractère explicatif significatif, l'ajout de variables explicatives peut diminuer les mesures de risque. Ce phénomène est dû soit à un trop grand nombre de variables explicatives non significatives, soit à un paramètre m trop faible. Une étude complémentaire devrait être menée pour retirer automatiquement les variables non significatives.

Ainsi, cette partie nous a permis de voir comment nous pouvons adapter nos deux méthodes dans le cas d'un modèle de gestion actif/passif réel. Pour la méthode de décomposition d'un événement rare en événements moins rares, l'utilisation d'une forme paramétrique locale lors de la diffusion des scénarios nous permet de diminuer le budget de calcul, et aboutit à de meilleures estimations du capital économique. Dans le cas des forêts aléatoires, nous avons pu voir qu'en utilisant des puissances de facteurs de risque et des termes croisés nous obtenions une mesure de risque plus faible.

Conclusion

Dans ce mémoire, nous avons tout d'abord rappelé que, dans le cadre de la directive Solvabilité II, les compagnies d'assurance doivent mettre de côté des fonds propres, appelés capital économique, afin d'éviter la ruine à horizon un an et avec un niveau de confiance de 99,5 %. La méthode la plus conforme à cette définition du capital économique est la méthode SdS, mais celle-ci est compliquée à mettre en œuvre. Nous nous sommes donc intéressés à des méthodes alternatives, fondées sur l'observation qu'il est uniquement nécessaire de connaître de façon précise le quantile à 0,5 % des fonds propres à un an pour estimer correctement le capital économique.

La méthode de décomposition d'un événement rare en événements moins rares et les forêts aléatoires sont des techniques qui avaient déjà été appliquées à d'autres domaines, mais jamais à l'assurance. Aussi, nous avons dû les adapter à notre problématique, afin de déterminer de façon précise le capital économique en utilisant le minimum de projections ALM. Plusieurs adaptations et améliorations ont alors été apportées aux algorithmes dont nous nous étions inspirés. Nous nous sommes également servis de certaines méthodes alternatives qui existaient déjà pour les optimiser.

Après avoir expliqué la théorie relative à nos deux méthodes, nous les avons mises en œuvre dans le but de les comparer aux autres méthodes alternatives aux SdS. Dans un premier temps, l'application à un portefeuille d'épargne des algorithmes que nous avons proposés nous a ensuite permis de justifier nos adaptations. Nous avons en effet pu observer que l'algorithme de décomposition d'un événement rare en événements moins rares pour lequel la valeur de l'écart-type utilisé dans la diffusion n'était pas figée aboutissait à une estimation du quantile à 0,5 % des fonds propres plus proche de la valeur de référence que l'algorithme simple. Par ailleurs, nous avons remarqué que l'utilisation d'une forme paramétrique locale permettait d'améliorer encore davantage les résultats, tout en diminuant le nombre de projections ALM. Concernant les forêts aléatoires, nous avons observé que le calibrage était meilleur lorsqu'il était effectué sur des points répartis uniformément, et nous avons mis en évidence que la méthode donnant les meilleurs résultats était celle utilisant des arbres semi-obliques avec ajout de points. En appliquant directement ce dernier algorithme à des modèles de gestion actif/passif réels, nous avons encore pu améliorer les résultats en prenant en compte des puissances et des termes croisés de facteurs de risque dans la forêt aléatoire.

Dans un second temps, la mise en œuvre des algorithmes a montré que nous pouvions obtenir des estimations précises du quantile à 0,5 % des fonds propres, en diminuant le nombre de projections ALM de plus de 99 % par rapport au SdS complet. Cependant, nous obtenons le même type de problème que pour les méthodes alternatives déjà existantes puisqu'il semble difficile d'obtenir une bonne estimation du quantile lorsque le portefeuille devient plus complexe. En effet, les résultats obtenus avec la méthode de décomposition d'un événement rare se détériorent lorsque la zone quantile devient plus étendue, et la taille de la base de calibrage des forêts aléatoires pose problème quand nous considérons plus de trois facteurs de risque.

Nous pouvons imaginer certaines modifications de nos algorithmes dans le but de résoudre ces problèmes. D'une part, pour la première méthode, nous pouvons supposer qu'en choisissant de manière optimale les points de calibrage de la forme paramétrique locale au lieu de prendre tous les points composant l'échantillon de scénarios, nous obtiendrons une estimation des fonds propres plus précise dans la zone voulue. En outre, lors de la diffusion des scénarios, il serait intéressant d'utiliser une matrice de variance-covariance non diagonale, c'est-à-dire d'introduire des corrélations entre les

facteurs de risque. Celles-ci pourraient, comme les variances, être calculées à partir des scénarios non diffusés.

D'autre part, pour les forêts aléatoires, nous avons remarqué que le choix de la base de calibrage jouait un rôle très important, en particulier lorsque le nombre de facteurs de risque augmentait. Ainsi, il semblerait possible de diminuer les projections ALM en optimisant ce choix. Nous pourrions par exemple observer que certains points constituant la base telle qu'elle est définie actuellement, ont une probabilité quasiment nulle d'être générés et pourraient être supprimés. Ensuite, nous rappelons qu'en pratique nous ne connaissons pas la mesure de risque. Aussi, au lieu de prendre une grande zone quantile pour être sûr de contenir le scénario quantile, nous pourrions intégrer à notre algorithme une approche de type accélérateur SdS. Celle-ci itérerait le nombre de scénarios pour lequel nous devons lancer un grand nombre de simulations secondaires jusqu'à être certain d'avoir le scénario quantile, ce qui permettrait encore de diminuer le nombre de projections ALM. Enfin, nous avons vu qu'ajouter des puissances de facteurs de risque et des termes croisés permettait de diminuer la mesure de risque. Nous pouvons donc penser qu'en utilisant d'autres fonctions des facteurs de risque nous obtiendrons de meilleurs résultats.

Bibliographie

- Ankenman B, Nelson B.L., Staum J. (2010), *Stochastic Kriging for Simulation Metamodeling*, Operations Research, Vol. 58, No. 2, 371-382.
- Arnold B.C., Balakrishnan N., Nagaraja H.N. (2008), *A First Course in Order Statistics*, Classics in Applied Mathematics.
- Au S. K., Beck J. L. (2001), *Estimation of small failure probabilities in high dimensions by subset simulation*, Probabilistic Engineering Mechanics, Vol.16, 263-277.
- Au S. K., Beck J. L. (2003), *Subset Simulation and its Application to Seismic Risk Based on Dynamic Analysis*, Journal of Engineering Mechanics, 901-917.
- Breiman L. (1994), *Bagging Predictors*, Technical Report No. 412, Berkeley.
- Breiman, L. (2001), *Random Forests*, Machine Learning, 45 :5-32.
- Broadie M., Du Y., Moallemi C.C. (2011), *Efficient Risk Estimation via Nested Sequential Simulation*, Management Science, Vol. 57, No. 6, 1172-1194.
- Caron V., Guyader A., Munoz Zuniga M., Tuffin B. (2013), *Some recent results in rare event estimation*, ESAIM: Proceedings, 1-10.
- Cérou F., Guyader A. (2005), *Adaptative multilevel splitting for rare event analysis*, IRISA, Publication interne, No. 1747.
- Cérou F., Guyader A. (2007), *Adaptative particle techniques and rare event estimation*, ESAIM: Proceedings, Vol. 19, 65-72.
- Cérou F., Del Moral P., Furon T., Guyader A. (2009), *Rare event simulation for a static distribution*, INRIA, Rapport de recherche, No. 6792.
- Cérou F., Del Moral P., Furon T., Guyader A. (2012), *Sequential Monte Carlo for Rare Event Estimation*, Statistics and Computing, Vol. 22, Issue 3, 795-808.
- Chabriac C. (2013), *Processus de Poisson*, Polycopié de cours du Master ISMAG, Université de Toulouse le Mirail.
- David H. A., Nagaraja H. N. (2003), *Order Statistics (Third Edition)*, Wiley Series in Probability and Statistics.
- Decupère S. (2011), *Agrégation des risques et Allocation de capital sous Solvabilité II*, Mémoire d'actuariat, ENSAE ParisTech.
- Del Moral P., Doucet A., Jasra A. (2006), *Sequential Monte Carlo Samplers*, Journal of the Royal, Statistical Society: Series B, Statistical Methodology, Vol. 68, Issue 3, 411-436.
- Devineau L., Loisel S. (2009), *Construction d'un algorithme d'accélération de la méthode des simulations dans les simulations pour le calcul du capital économique Solvabilité II*, Bulletin Français d'Actuariat, Vol. 10, No. 17, 188-221.
- Devineau L., Chauvigny M. (2011), *Replicating portfolios : techniques de calibrage pour le calcul du capital économique Solvabilité II*, Bulletin Français d'Actuariat, Vol. 11, No. 21, 59-97.
- Fermanian J.D. (2006), *Modèles de durées*, Polycopié de cours, ENSAE.
- FFSA (2012), *Rapport annuel*.

- Gadat S. (2009), *Analyse de données et apprentissage statistique*, cours du M2SID, Université de Toulouse.
- Genuer R. (2010), *Forêts aléatoires : aspects théoriques, sélection de variables et applications*, thèse, Université Paris-Sud 11.
- Genuer R., Poggi J.M., Tuleau C. (2008), *Random Forest: some methodological insights*, Rapport de recherche N°6729, INRIA.
- Graham C. (2008), *Chaînes de Markov : cours et exercices corrigés*, Dunod, Sciences Sup.
- Guyader A., Hengartner N., Matzner-Lober E. (2011), *Simulation and estimation of Extreme Quantiles and extreme Probabilities*, Applied Mathematics & Optimization, Vol. 64, Issue 2, 171-196.
- Hastie T., Tibshirani R., Friedman J. (2009), *The Elements of Statistical Learning*, Springer Series in Statistics.
- Jian-Feng Y. (2004), *limites et méthode delta*, Polycopié de cours du Master Statistique, Econométrie, Université de Rennes 1.
- Kahn H., Harris T. (1951), *Estimation of particle transmission by random sampling*, National Bureau of Standards, Appl. Math., Ser. 12, 27-30.
- Koursaris A. (2011), *A comparison between curve fitting and least squares Monte Carlo techniques*, Barrie & Hibbert.
- Lebarbier E., Robin S. (2007), *Processus de Poisson Processus de Naissances et Morts*, Polycopié de cours, AgroParisTech.
- L'Ecuyer P., Demers V., Tuffin B. (2007), *Rare Events, Splitting, and Quasi-Monte Carlo*, ACM Transactions on Modeling and Computer Simulation, Vol. 17, Issue 2, No. 9.
- Le Mer S (2010), *Calcul du capital économique en assurance vie*, Mémoire d'actuariat, Université Paris Dauphine.
- Liu M. (2010), *Efficient Simulation in Financial Risk Management*.
- Liu M., Staum J. (2010), *Stochastic Kriging for Efficient Nested Simulation of Expected Shortfall*, The Journal of Risk, Vol. 12, No. 3, 3-27.
- Millet A. (2012), *Méthodes de Monte-Carlo*, Polycopié de cours du Master M2Mo, Université Paris 7.
- Morio J., Pastel R., Le Gland F. (2011), *Estimation de probabilités et de quantiles rares pour la caractérisation d'une zone de retombée d'un engin*, Journal de la Société Française de Statistique, Vol. 152, No. 4.
- Planchet F. (2012), *Modèles de durée*, Polycopié de cours, ISFA.
- Quinlan R., Kohavi R. (1999), *Decision Tree discovery*, Data Mining, 10-18.
- Ramanampisoa T. (2012), *Application de la méthode Least-Square Monte-Carlo pour la mise en place de l'ORSA en Assurance vie*, Mémoire d'actuariat, ISUP.
- Revelen J. (2009), *« Replicating Portfolio » et capital économique en assurance vie*, Mémoire de fin d'études, Ecole Centrale de Lyon.
- Robert C.P., Casella G. (2004), *Monte Carlo Statistical Methods*, Springer Texts in Statistics.
- Rosenbluth M., Rosenbluth A. (1955), *Monte Carlo calculation of the average extension of molecular chains*, The Journal of Chemical Physics, Vol. 23, Issue 2.

- Sauvet C. (2006), *Solvency II – Quelle modélisation stochastique des provisions techniques prévoyance et non vie ?*, Mémoire d'actuariat, ISFA.
- Schapire E. et Freund Y., *A Short Introduction to Boosting*, Journal of Japanese Society for Artificial Intelligence, 14(5):771-780.
- Stempf G. (2009), *Fouilles de données, Arbres de décision*, Cours de M2 MASS, Université d'Aix-Marseille.
- Therond P. (2005), *Contrôle de la solvabilité des compagnies d'assurance : évolutions récentes*, Slides Séminaire « Gestion des risques » - DIAF.
- Van Dam E.R., Husslage B., den Hertog D., Melissen H. (2007), *Maximin Latin Hypercube Designs in Two Dimensions*, Operations Research, Vol. 55, No. 1, 158-169.
- Vedani J., Virepinte P.A. (2011), *Modèle structurel de crédit dans une optique assurantielle*, Mémoire d'actuariat, ENSAE Paristech.
- Vedani J., Devineau L. (2013), *Solvency assessment within the ORSA framework: issues and quantitative methodologies*, Milliman.
- Wolinski P. (2012), *Simulation et estimation de quantiles extrêmes et de probabilités d'événements rares*, Mémoire de L3, ENS.

V. Annexes

A. Annexe 1 : détermination de σ dans la diffusion des scénarios à l'aide d'un noyau de transition gaussien

Nous rappelons que pour diffuser un scénario suivant la loi $L(X|FP(X) < L_m)$, nous piochons aléatoirement un scénario dans l'ensemble $\{X_1^{m-1}, \dots, X_N^{m-1}\} \setminus \{X_i^{m-1}\}$. Nous notons X_j^{m-1} ce scénario et $\epsilon_{j,k}^{m-1}$ le $k^{\text{ième}}$ facteur de risque de ce scénario. Ensuite, nous appliquons plusieurs fois un noyau de transition gaussien à $\epsilon_{j,k}^{m-1}$, c'est-à-dire que nous initialisons notre récurrence en posant $\epsilon_{0,k} = \epsilon_{j,k}^{m-1}$ et que nous utilisons, pour chaque facteur de risque indicé par k , la relation de récurrence suivante :

$$\epsilon_{t,k} = \frac{\epsilon_{t-1,k} + \sigma_k W_k}{\sqrt{1 + \sigma_k^2}},$$

Dans l'expression ci-dessus, W_k est une variable aléatoire gaussienne indépendante de $\epsilon_{t-1,k}$, et σ_k est l'écart-type utilisé pour le noyau de transition associé au $k^{\text{ième}}$ facteur de risque.

Afin de ne pas avoir à choisir la valeur de σ_k pour chaque facteur de risque, nous allons chercher comment la définir de manière adéquate. Nous notons $\tilde{\sigma}_k^2$ la variance de l'échantillon $\{\epsilon_{1,k}^{m-1}, \dots, \epsilon_{N,k}^{m-1}\} \setminus \{\epsilon_{i^*,k}^{m-1}\}$. Cette variance est celle du $k^{\text{ième}}$ facteur de risque des scénarios de loi $L(X|FP(X) < L_m)$. Nous voulons que la variance du scénario diffusé (vu en tant que variable aléatoire) soit égale à $\tilde{\sigma}_k^2$ à la fin des itérations horizontales.

A chaque fois que nous appliquons le noyau de transition à un facteur de risque k , nous avons :

$$\epsilon_{t,k} \times \sqrt{1 + \sigma_k^2} = \epsilon_{t-1,k} + \sigma_k W_k.$$

Ainsi, en prenant la variance, nous obtenons :

$$\text{Var}(\epsilon_{t,k} | X_{j,k}^{m-1}) \times (1 + \sigma_k^2) = \text{Var}(\epsilon_{t-1,k} | \epsilon_{j,k}^{m-1}) + \sigma_k^2.$$

Nous définissons la suite (u_t) par $u_t = \text{Var}(\epsilon_{t,k} | X_{j,k}^{m-1})$. Ainsi, l'égalité ci-dessus se traduit par la relation de récurrence $u_t = \frac{1}{1 + \sigma_k^2} u_{t-1} + \frac{\sigma_k^2}{1 + \sigma_k^2}$. Nous reconnaissons alors une suite arithmético-géométrique, c'est-à-dire de la forme $u_t = a \times u_{t-1} + b$, avec $a = \frac{1}{1 + \sigma_k^2}$ et $b = \frac{\sigma_k^2}{1 + \sigma_k^2}$.

Nous allons à présent déterminer l'expression de la suite arithmético-géométrique (u_t) en fonction de t . Nous posons $r = \frac{b}{1-a} = 1$.

Nous pouvons alors calculer $u_t - r$ en fonction de u_{t-1} :

$$u_t - r = u_t - \frac{b}{1-a} = a \times u_{t-1} + b - \frac{b}{1-a} = a \times u_{t-1} - \frac{ab}{1-a} = a \left(u_{t-1} - \frac{b}{1-a} \right)$$

Nous obtenons finalement : $u_t - r = a(u_{t-1} - r)$, ce qui nous permet de remarquer que la suite $(u_t - r)$ est une suite géométrique de raison a . Nous pouvons donc exprimer cette nouvelle suite en fonction de t : $u_t - r = a^t(u_0 - r)$.

Par ailleurs, $u_0 = Var(\epsilon_{0,k}|X_{j,k}^{m-1}) = Var(\epsilon_{j,k}^{m-1}|X_{j,k}^{m-1}) = 0$. Nous obtenons alors l'expression de (u_t) en fonction de t : $u_t = a^t(u_0 - r) + r$, et nous en déduisons l'expression de $Var(\epsilon_{t,k}|\epsilon_{j,k}^{m-1})$:

$$Var(\epsilon_{t,k}|\epsilon_{j,k}^{m-1}) = 1 - \left(\frac{1}{1+\sigma_k^2}\right)^t.$$

Nous souhaitons qu'au bout de T itérations horizontales, $Var(\epsilon_T|X_j^{m-1}) = \tilde{\sigma}_k^2$. Nous obtenons ainsi l'équation suivante : $\tilde{\sigma}_k^2 = 1 - \left(\frac{1}{1+\sigma_k^2}\right)^T$ dont nous pouvons déduire l'expression de σ_k en fonction de $\tilde{\sigma}_k^2$:

$$\sigma_k = \sqrt{\left(\frac{1}{1-\tilde{\sigma}_k^2}\right)^{\frac{1}{T}} - 1}.$$

Remarque : nous rappelons que nous considérons que les facteurs de risque sont indépendants les uns des autres et que l'écart-type utilisé dans la diffusion est calculé pour chaque facteur de risque.

B. Annexe 2 : détermination des limites de $\frac{N}{\sqrt{M}} \left(\ln(0,5 \%) + \frac{M}{N} \right)$ et de $\sqrt{\frac{M}{N}}$

Calcul de la limite de $\frac{N}{\sqrt{M}} \left(\ln(0,5 \%) + \frac{M}{N} \right)$:

Nous rappelons que $M = \left\lceil \frac{\ln(0,5 \%) }{\ln(1-N^{-1})} \right\rceil$. Par conséquent, nous pouvons encadrer M de la façon suivante :

$$\frac{\ln(0,5 \%) }{\ln(1-N^{-1})} \leq M < \frac{\ln(0,5 \%) }{\ln(1-N^{-1})} + 1.$$

Nous en déduisons alors les deux encadrements ci-dessous :

$$0 < \ln(0,5 \%) + \frac{1}{N} \times \frac{\ln(0,5 \%) }{\ln(1-N^{-1})} \leq \ln(0,5 \%) + \frac{M}{N} < \ln(0,5 \%) + \frac{1}{N} \times \frac{\ln(0,5 \%) }{\ln(1-N^{-1})} + \frac{1}{N}$$

$$0 < \frac{N}{\sqrt{\frac{\ln(0,5 \%) }{\ln(1-N^{-1})} + 1}} < \frac{N}{\sqrt{M}} \leq \frac{N}{\sqrt{\frac{\ln(0,5 \%) }{\ln(1-N^{-1})}}}$$

En multipliant membre à membre, nous montrons finalement que nous pouvons encadrer $\frac{N}{\sqrt{M}} \left(\ln(0,5 \%) + \frac{M}{N} \right)$ de la manière suivante :

$$\left(\ln(0,5 \%) + \frac{1}{N} \times \frac{\ln(0,5 \%) }{\ln(1-N^{-1})} \right) \frac{N}{\sqrt{\frac{\ln(0,5 \%) }{\ln(1-N^{-1})} + 1}} < \frac{N}{\sqrt{M}} \left(\ln(0,5 \%) + \frac{M}{N} \right)$$

$$\text{et } \frac{N}{\sqrt{M}} \left(\ln(0,5 \%) + \frac{M}{N} \right) < \left(\ln(0,5 \%) + \frac{1}{N} \times \frac{\ln(0,5 \%) }{\ln(1-N^{-1})} + \frac{1}{N} \right) \frac{N}{\sqrt{\frac{\ln(0,5 \%) }{\ln(1-N^{-1})}}}$$

Pour déterminer la limite du terme de gauche et celle du terme de droite, nous allons calculer leurs équivalents. Pour cela, nous devons tout d'abord déterminer un équivalent de $\frac{\ln(0,5 \%) }{\ln(1-N^{-1})}$:

$$\frac{\ln(0,5 \%) }{\ln(1-N^{-1})} = \frac{\ln(0,5 \%) }{-\left(\frac{1}{N} + \frac{1}{N^2} + o\left(\frac{1}{N^2}\right)\right)} = \frac{-N \ln(0,5 \%) }{1 + \frac{1}{N} + o\left(\frac{1}{N}\right)} = -N \ln(0,5 \%) \left(1 + \frac{1}{N} + o\left(\frac{1}{N}\right) \right)$$

$$\frac{\ln(0,5\%)}{\ln(1-N^{-1})} = -N(\ln(0,5\%)) \left(1 + \frac{1}{N}\right) + o(1) = -N(\ln(0,5\%)) - \ln(0,5\%) + o(1)$$

Nous pouvons ainsi déduire de ce calcul préliminaire un équivalent du terme de gauche :

$$\begin{aligned} \left(\ln(0,5\%) + \frac{1}{N} \times \frac{\ln(0,5\%)}{\ln(1-N^{-1})}\right) \frac{N}{\sqrt{\frac{\ln(0,5\%)}{\ln(1-N^{-1})} + 1}} &= \left(-\frac{\ln(0,5\%)}{N} + o\left(\frac{1}{N}\right)\right) \times \frac{N}{(1-N(\ln(0,5\%)) + o(1))^{1/2}} \\ \left(\ln(0,5\%) + \frac{1}{N} \times \frac{\ln(0,5\%)}{\ln(1-N^{-1})}\right) \frac{N}{\sqrt{\frac{\ln(0,5\%)}{\ln(1-N^{-1})} + 1}} &= \\ (-\ln(0,5\%) + o(1)) \times \frac{1}{\sqrt{-N(\ln(0,5\%)) \left(1 - \frac{1}{N(\ln(0,5\%))} + o\left(\frac{1}{N}\right)\right)^{1/2}} & \\ \left(\ln(0,5\%) + \frac{1}{N} \times \frac{\ln(0,5\%)}{\ln(1-N^{-1})}\right) \frac{N}{\sqrt{\frac{\ln(0,5\%)}{\ln(1-N^{-1})} + 1}} &= \sqrt{\frac{-\ln(0,5\%)}{N}} \times \left(1 + \frac{1}{2N(\ln(0,5\%))}\right) + o\left(\frac{1}{N^2}\right) \end{aligned}$$

Nous obtenons alors finalement la convergence du terme de gauche vers 0.

Nous pouvons procéder de la même manière et déterminer également un équivalent du terme de droite :

$$\begin{aligned} \left(\ln(0,5\%) + \frac{1}{N} \times \frac{\ln(0,5\%)}{\ln(1-N^{-1})} + \frac{1}{N}\right) \frac{N}{\sqrt{\frac{\ln(0,5\%)}{\ln(1-N^{-1})} + 1}} &= \left(-\frac{\ln(0,5\%)}{N} + \frac{1}{N} + o\left(\frac{1}{N}\right)\right) \frac{N}{(-N(\ln(0,5\%)) + o(1))^{1/2}} \\ \left(\ln(0,5\%) + \frac{1}{N} \times \frac{\ln(0,5\%)}{\ln(1-N^{-1})} + \frac{1}{N}\right) \frac{N}{\sqrt{\frac{\ln(0,5\%)}{\ln(1-N^{-1})} + 1}} &= (-\ln(0,5\%) + 1 + o(1)) \frac{1}{\sqrt{-N(\ln(0,5\%)) + o(1)}} \end{aligned}$$

Ainsi, nous en déduisons également la convergence de ce terme vers 0, ce qui nous donne la limite voulue :

$$\lim_{N \rightarrow \infty} \frac{N}{\sqrt{M}} \left(\ln(0,5\%) + \frac{M}{N}\right) = 0.$$

Calcul de la limite de $\sqrt{\frac{M}{N}}$:

A partir des encadrements sur M , nous pouvons déduire un encadrement de $\sqrt{\frac{M}{N}}$:

$$\sqrt{\frac{1}{N} \times \frac{\ln(0,5\%)}{\ln(1-N^{-1})}} \leq \sqrt{\frac{M}{N}} < \sqrt{\frac{1}{N} \times \frac{\ln(0,5\%)}{\ln(1-N^{-1})} + \frac{1}{N}}$$

Par ailleurs, nous avons déterminé précédemment que $\frac{\ln(0,5\%)}{\ln(1-N^{-1})}$ était équivalent à $-N(\ln(0,5\%))$ en l'infini. Nous pouvons alors en déduire que les deux termes encadrant $\sqrt{\frac{M}{N}}$ convergent vers $-(\ln(0,5\%))$, ce qui nous donne la limite ci-dessous :

$$\lim_{N \rightarrow \infty} \sqrt{\frac{M}{N}} = -(\ln(0,5\%)).$$

C. Annexe 3 : détermination de la loi asymptotique d'un estimateur de Monte Carlo

Nous allons, dans cette annexe, montrer que la loi suivie asymptotiquement par un estimateur de Monte Carlo est donnée par $\sqrt{N}(\hat{q}_{MC} - q) \rightarrow \mathcal{N}\left(0, \frac{p(1-p)}{f(-q)^2}\right)$, où f est la densité du quantile.

Nous notons F la fonction de répartition de la variable $FP(X)$ et nous posons, pour plus de simplicité dans les écritures, $r = 0,5 \% \times N$. Nous pouvons alors exprimer $FP(X)_{(r)}$ en fonction de la $r^{\text{ième}}$ plus petite réalisation d'une variable aléatoire de loi uniforme sur l'intervalle $[0,1]$:

$$FP(X)_{(r)} = F^{-1}(U_{(r)}).$$

Ainsi, nous allons tout d'abord nous intéresser à la variable aléatoire $U_{(r)}$, et plus précisément, nous allons commencer par déterminer sa densité, sa moyenne et sa variance.

Calcul de la densité de $U_{(r)}$:

Nous pouvons définir la densité de la variable $U_{(r)}$ de la manière suivante :

$$f_{(r)}(u) = \lim_{du \rightarrow 0} \frac{1}{du} P(u < U_{(r)} \leq u + du).$$

Nous remarquons que l'événement $u < U_{(r)} \leq u + du$ signifie que sur un ensemble de N variables aléatoires de loi uniforme, $r-1$ sont inférieures à u , une est comprise entre u et $u+du$ et $N-r$ sont supérieures à $u+du$. Nous pouvons ainsi en déduire la probabilité de cet événement :

$$P(u < U_{(r)} \leq u + du) = \binom{N}{r-1} u^{r-1} \times \binom{N-(r-1)}{1} ((u+du) - u) \times (1 - (u+du))^{N-r}.$$

Par ailleurs le produit des deux combinaisons peut se simplifier et s'écrire sous la forme $\frac{N!}{(r-1)!(N-r)!}$.

Nous en déduisons ainsi que la densité de la variable aléatoire $U_{(r)}$ est donnée par la formule ci-dessous :

$$f_{(r)}(u) = \frac{N!}{(r-1)!(N-r)!} u^{r-1} \times (1-u)^{N-r}.$$

Calcul de la moyenne de $U_{(r)}$:

A partir de la densité calculée précédemment, nous pouvons déterminer l'expression de la moyenne de $U_{(r)}$, notée p_r :

$$p_r = E[U_{(r)}] = \int_0^1 u \times \frac{N!}{(r-1)!(N-r)!} u^{r-1} \times (1-u)^{N-r} du$$

$$p_r = \int_0^1 \frac{N!}{(r-1)!(N-r)!} u^r \times (1-u)^{N-r} du$$

Pour calculer cette intégrale, nous effectuons r intégrations par parties en dérivant à chaque fois u^r et en intégrant $(1-u)^{N-r}$. A chaque intégration par partie, le premier terme de l'intégration s'annule puisque nous prenons la valeur d'une fonction de la forme $u^\alpha(1-u)^\beta$ en 0 et en 1. Ainsi, au bout des r intégrations par parties, nous obtenons les expressions ci-dessous :

$$p_r = \int_0^1 \frac{N!}{(r-1)!(N-r)!} \times (r! \times 1) \times \left(\frac{1}{(N-r+1) \times \dots \times N}\right) (1-u)^N du$$

$$p_r = \frac{N!}{(r-1)!(N-r)!} \times r! \times \frac{(N-r)!}{N!} \left[-\frac{(1-u)^{N+1}}{N+1} \right]_0^1$$

$$p_r = \frac{N!}{(r-1)!(N-r)!} \times r! \times \frac{(N-r)!}{N!} \times \frac{1}{N+1}$$

$$p_r = \frac{r}{N+1}$$

Nous avons donc déterminé l'expression de la moyenne de $U_{(r)}$, notée p_r : $p_r = \frac{r}{N+1}$.

Calcul de la variance de $U_{(r)}$:

Nous procédons de la même manière que précédemment pour déterminer la variance de $U_{(r)}$. En effet, nous commençons par calculer $E[U_{(r)}^2]$ à l'aide de $r+1$ intégrations par parties :

$$E[U_{(r)}^2] = \int_0^1 u^2 \times \frac{N!}{(r-1)!(N-r)!} u^{r-1} \times (1-u)^{N-r} du$$

$$E[U_{(r)}^2] = \int_0^1 \frac{N!}{(r-1)!(N-r)!} u^{r+1} \times (1-u)^{N-r} du$$

$$E[U_{(r)}^2] = \int_0^1 \frac{N!}{(r-1)!(N-r)!} \times (r+1)! \times \left(\frac{1}{(N-r+1) \times \dots \times (N+1)} \right) (1-u)^{N+1} du$$

$$E[U_{(r)}^2] = \frac{N!}{(r-1)!(N-r)!} \times (r+1)! \times \frac{(N-r)!}{(N+1)!} \left[-\frac{(1-u)^{N+2}}{N+2} \right]_0^1$$

$$E[U_{(r)}^2] = \frac{r(r+1)}{(N+1)(N+2)}$$

Nous pouvons alors calculer la variance de $U_{(r)}$:

$$Var[U_{(r)}] = E[U_{(r)}^2] - E[U_{(r)}]^2 = \frac{r(r+1)}{(N+1)(N+2)} - \frac{r^2}{(N+1)^2} = \frac{r((r+1)(N+1) - r(N+2))}{(N+1)^2(N+2)} = \frac{r(N-r+1)}{(N+1)^2(N+2)}$$

A l'issue de ces calculs, nous obtenons finalement que la variance de $U_{(r)}$ est donc donnée par la formule suivante : $Var[U_{(r)}] = \frac{p_r(1-p_r)}{N+2}$.

Calcul de la moyenne asymptotique de \hat{q}_{MC} :

Nous rappelons que $U_{(r)}$ et $FP(X)_{(r)}$ sont liés par la relation $FP(X)_{(r)} = F^{-1}(U_{(r)})$. Nous pouvons alors effectuer un développement limité de $F^{-1}(U_{(r)})$ au deuxième ordre, au voisinage de $p_r = \frac{r}{N+1}$, ce qui nous donne l'expression ci-dessous :

$$FP(X)_{(r)} = F^{-1}(p_r) + (F^{-1})'(p_r)(U_{(r)} - p_r) + \frac{1}{2}(F^{-1})''(p_r)(U_{(r)} - p_r)^2 + o\left((U_{(r)} - p_r)^2\right) \quad (3)$$

En utilisant les relations concernant $U_{(r)}$ que nous venons de démontrer, nous pouvons en déduire que l'espérance de $FP(X)_{(r)}$ est donnée par la formule suivante :

$$E[FP(X)_{(r)}] = F^{-1}(p_r) + \frac{1}{2}(F^{-1})''(p_r) \frac{p_r(1-p_r)}{N+2} + o\left(\frac{1}{N+2}\right).$$

Nous obtenons alors la limite $\lim_{N \rightarrow \infty} E[FP(X)_{(r)}] = q$, où écrit autrement $\lim_{N \rightarrow \infty} E[\hat{q}_{MC}] = q$.

Calcul de la variance asymptotique de \hat{q}_{MC} :

Nous pouvons effectuer le même type de calcul pour calculer la variance de $FP(X)_{(r)}$. En effet, un développement limité de $FP(X)_{(r)}^2 = \left(F^{-1}(U_{(r)})\right)^2$ donne le résultat suivant :

$$FP(X)_{(r)}^2 = \left(F^{-1}(p_r)\right)^2 + ((F^{-1})^2)'(p_r)(U_{(r)} - p_r) + \frac{1}{2}((F^{-1})^2)''(p_r)(U_{(r)} - p_r)^2 + o\left((U_{(r)} - p_r)^2\right)$$

Nous pouvons alors calculer l'espérance de $FP(X)_{(r)}^2$:

$$E[FP(X)_{(r)}^2] = (F^{-1}(p_r))^2 + \frac{1}{2}((F^{-1})^2)''(p_r) \frac{p_r(1-p_r)}{N+2} + o\left(\frac{1}{N+2}\right).$$

Nous en déduisons alors l'expression de la variance de $FP(X)_{(r)}$:

$$Var[FP(X)_{(r)}] = (F^{-1}(p_r))^2 + \frac{1}{2}((F^{-1})^2)''(p_r) \frac{p_r(1-p_r)}{(N+2)} - \left[F^{-1}(p_r) + \frac{1}{2}(F^{-1})''(p_r) \frac{p_r(1-p_r)}{(N+2)} \right]^2 + o\left(\frac{1}{N+2}\right)$$

$$Var[FP(X)_{(r)}] = \frac{1}{2}((F^{-1})^2)''(p_r) \frac{p_r(1-p_r)}{(N+2)} - F^{-1}(p_r) \times (F^{-1})''(p_r) \frac{p_r(1-p_r)}{(N+2)} + o\left(\frac{1}{N+2}\right) \quad (4)$$

Parallèlement, nous calculons la dérivée seconde de $(F^{-1})^2$:

$$((F^{-1})^2)'(u) = 2 \times (F^{-1})'(x) \times (F^{-1})(x)$$

$$((F^{-1})^2)''(u) = 2 \times (F^{-1})''(x) \times (F^{-1})(x) + 2 \times ((F^{-1})'(x))^2$$

Nous remplaçons alors, dans l'égalité (4), $((F^{-1})^2)''(p_r)$ par l'expression que nous venons de trouver et nous obtenons le résultat suivant :

$$Var[FP(X)_{(r)}] = ((F^{-1})'(p_r))^2 \frac{p_r(1-p_r)}{(N+2)} + o\left(\frac{1}{N+2}\right)$$

Ainsi, en notant f la densité de la variable aléatoire $-FP(X)$, nous en déduisons un développement limité de la variance de $FP(X)_{(r)}$ en fonction de f :

$$Var[FP(X)_{(r)}] = \frac{1}{f^2(-q)} \frac{p_r(1-p_r)}{(N+2)} + o\left(\frac{1}{N+2}\right).$$

Nous pouvons alors déduire de l'ensemble des calculs précédents que la loi asymptotique de l'estimateur de Monte Carlo du quantile est la suivante :

$$\sqrt{N}(\hat{q}_{MC} - q) \rightarrow \mathcal{N}\left(0, \frac{p(1-p)}{f(-q)^2}\right).$$