



N° d'ordre NNT : 2016LYSE1245

THÈSE DE DOCTORAT DE L'UNIVERSITÉ DE LYON

opérée au sein de
l'Université Claude Bernard Lyon 1

École Doctorale ED 512
InfoMaths

Spécialité de doctorat : Mathématiques
Discipline : Mathématiques Appliquées

Soutenue publiquement le 02/12/2016, par :
Khalil Said

Mesures de risque multivariées et applications en science actuarielle

Devant le jury composé de :

Girard Stéphane, Directeur de Recherche, INRIA Grenoble Rhône-Alpes

Président

Cossette Hélène, Professeur, Université Laval

Rapporteure

Guillou Armelle, Professeur, Université de Strasbourg

Rapporteure

Bellini Fabio, Professeur, Università di Milano-Bicocca

Examineur

Vanduffel Steven, Professeur, Vrije Universiteit Brussel

Examineur

Maume-Deschamps Véronique, Professeur, Université Claude Bernard Lyon 1

Directrice de thèse

Rullière Didier, Maître de conférences, Université Claude Bernard Lyon 1

Co-directeur de thèse

Remerciements

Je voudrais remercier en tout premier lieu mes deux directeurs de thèse :

- Madame Véronique MAUME-DESCHAMPS pour l’honneur qu’elle m’a fait en acceptant de diriger ce travail. Son soutien, sa disponibilité, ses qualités humaines, et ses précieuses directives tout au long de la réalisation de ce travail m’ont beaucoup impressionné.
- M. Didier RULLIERE pour le privilège qu’il m’a fait en acceptant de codiriger ce travail. Sa gentillesse, sa modestie, sa riche expérience et l’accueil cordial qu’il m’a toujours réservé m’ont inspiré une grande admiration à son égard.

Qu’ils puissent trouver dans ce modeste rapport le témoignage de ma sincère gratitude et de mon profond respect.

Je suis très honoré que :

- Madame Hélène COSSETTE, Professeur à l’Université Laval,
 - Madame Armelle GUILLOU, Professeur à Université de Strasbourg
- aient accepté d’être rapporteurs de cette thèse. Je suis très reconnaissant pour l’intérêt et le temps qu’elles ont consacré à ce travail.

Je tiens à remercier chaleureusement :

- Monsieur Fabio BELLINI, Professeur à l’Università di Milano-Bicocca,
 - Monsieur Stéphane GIRARD, Directeur de Recherche à l’INRIA Grenoble Rhône-Alpes,
 - Monsieur Steven VANDUFFEL, Professeur à Vrije Universiteit Brussel (VUB)
- de m’avoir fait l’honneur d’accepter de faire partie du jury de ma thèse.

Je voulais témoigner ma gratitude à Madame Sylvie BENZONI, Madame Sylvie CALABRETTO, Monsieur Christophe EVERAERE, et Monsieur Luca ZAMBONI. Sans leur aide et soutien, cette soutenance ne pouvait pas avoir lieu.

Mes remerciements s’adressent également à Monsieur Jean-Yves LESUEUR qui a décidé de m’offrir la chance de réaliser cette thèse avec un contrat doctoral.

Je voudrais aussi remercier du fond du cœur Areski COUSIN et Esterina MASIELLO, sans qui je n’aurais jamais eu la chance de préparer cette thèse.

Un grand merci à tous les membres du laboratoire SAF, je garderais un excellent souvenir de la bonne ambiance qui régnait au sein du laboratoire.

Enfin, j’exprime ma plus grande et humble gratitude pour la France. Je suis très reconnaissant envers cette grande nation, qui m’a offert la chance de suivre mes études dans ses établissements d’enseignement publics.

*À ma mère qui n'a jamais fréquenté une école
mais qui était toujours une école à ses enfants
À mon père qui n'a jamais obtenu des diplômes
mais qui a sacrifié toute sa vie pour que ses enfants en obtiennent
À ma sœur et mes trois frères qui me considèrent beaucoup plus qu'un simple frère aîné
À tous les jeunes égyptiens qui ont investi leurs âmes dans un rêve de changement qui deviendra sûrement
un jour une réalité.*

Résumé

Mesures de risque multivariées et applications en science actuarielle

L'entrée en application depuis le 1^{er} Janvier 2016 de la réforme réglementaire européenne du secteur des assurances Solvabilité 2 est un événement historique qui va changer radicalement les pratiques en matière de gestion des risques. Elle repose sur une prise en compte importante du profil et de la vision du risque, via la possibilité d'utiliser des modèles internes pour calculer les capitaux de solvabilité et l'approche ORSA (Own Risk and Solvency Assessment) pour la gestion interne du risque. La modélisation mathématique est ainsi un outil indispensable pour réussir un exercice réglementaire. La théorie du risque doit être en mesure d'accompagner ce développement en proposant des réponses à des problématiques pratiques, liées notamment à la modélisation des dépendances et aux choix des mesures de risques. Dans ce contexte, cette thèse présente une contribution à l'amélioration de la gestion des risques actuariels. En quatre chapitres nous présentons des mesures multivariées de risque et leurs applications à l'allocation du capital de solvabilité.

La première partie de cette thèse est consacrée à l'introduction et l'étude d'une nouvelle famille de mesures multivariées *élicitables* de risque qu'on appellera *des expectiles multivariés*. Son premier chapitre présente ces mesures et explique les différentes approches utilisées pour les construire. Les expectiles multivariés vérifient un ensemble de propriétés de cohérence que nous abordons aussi dans ce chapitre avant de proposer un outil d'approximation stochastique de ces mesures de risque. Les performances de cette méthode étant insuffisantes au voisinage des niveaux asymptotiques des seuils des expectiles, l'analyse théorique du comportement asymptotique est nécessaire, et fera le sujet du deuxième chapitre de cette partie. L'analyse asymptotique est effectuée dans un environnement à variations régulières multivariées, elle permet d'obtenir des résultats dans le cas des queues marginales équivalentes. Nous présentons aussi dans le deuxième chapitre le comportement asymptotique des expectiles multivariés sous les hypothèses précédentes en présence d'une dépendance parfaite, ou d'une indépendance asymptotique, et nous proposons à l'aide des statistiques des valeurs extrêmes des estimateurs de l'expectile asymptotique dans ces cas.

La deuxième partie de la thèse est focalisée sur la problématique de l'allocation du capital de solvabilité en assurance. Elle est composée de deux chapitres sous forme d'articles publiés. Le premier présente une axiomatisation de la cohérence d'une méthode d'allocation du capital dans le cadre le plus général possible, puis étudie les propriétés de cohérence d'une approche d'allocation basée sur la minimisation d'indicateurs multivariés de risque. Le deuxième article est une analyse probabiliste du comportement de cette dernière approche d'allocation en fonction de la nature des distributions marginales des risques et de la structure de la dépendance. Le comportement asymptotique de l'allocation est aussi étudié et l'impact de la dépendance est illustré par différents modèles marginaux et différentes copules.

La présence de la dépendance entre les différents risques supportés par les compagnies d'assurance fait de l'approche multivariée une réponse plus appropriée aux différentes problématiques de la gestion des risques. Cette thèse est fondée sur une vision multidimensionnelle du risque et propose des mesures de nature multivariée qui peuvent être appliquées pour différentes problématiques actuarielles de cette nature.

Mots-clés: *Mesures de risque multivariées, Solvabilité 2, Gestion des risques, Modèles internes, Capitaux de solvabilité, Approche ORSA, La théorie du risque, Modélisation de la dépendance, Allocation du capital, Expectiles multivariés, Elicitabilité, Propriétés de cohérence, Approximation stochastique, Indépendance asymptotique, Valeurs extrêmes, Copules.*

Abstract

Multivariate risk measures and their applications in actuarial science

The entry into force since January 1st, 2016 of Solvency 2, the European regulatory reform of insurance industry, is a historic event that will radically change the practices in risk management. It is based on taking into account the own risk profile and the internal view of risk through the ability to use internal models for calculating solvency capital requirement and ORSA (Own Risk and Solvency Assessment) approach for internal risk management. It makes the mathematical modeling an essential tool for a successful regulatory exercise. The risk theory must allow to support this development by providing answers to practical problems, especially those related to the dependence modeling and the choice of risk measures. In the same context, this thesis presents a contribution to improving the management of insurance risks. In four chapters we present multivariate risk measures and their application to the allocation of solvency capital.

The first part of this thesis is devoted to the introduction and study of a new family of multivariate *elicitable* risk measures that we will call *multivariate expectiles*. The first chapter presents these measures and explains the different construction approaches. The multivariate expectiles verify a set of coherence properties that we also discuss in this chapter before proposing a stochastic approximation tool of these risk measures. The performance of this method is insufficient in the asymptotic levels of the expectiles thresholds. That makes the theoretical analysis of the asymptotic behavior necessary. The asymptotic behavior of multivariate expectiles is then the subject of the second chapter of this part. It is studied in a multivariate regular variations framework, and some results are given in the case of equivalent marginal tails. We also study in the second chapter of the first part the asymptotic behavior of multivariate expectiles under previous assumptions in the presence of a perfect dependence, or in the case of asymptotic independence. Finally, we propose using extreme values statistics some estimators of the asymptotic expectile in these cases.

The second part of the thesis is focused on the issue of solvency capital allocation in insurance. It is divided into two chapters; each chapter consists of a published paper. The first one presents an axiomatic characterization of the coherence of a capital allocation method in a general framework. Then it studies the coherence properties of an allocation approach based on the minimization of some multivariate risk indicators. The second paper is a probabilistic analysis of the behavior of this capital allocation method based on the nature of the marginal distributions of risks and the dependence structure. The asymptotic behavior of the optimal allocation is also studied and the impact of dependence is illustrated using some selected models and copulas.

Faced to the significant presence of dependence between the various risks taken by insurance companies, a multivariate approach seems more appropriate to build responses to the various issues of risk management. This thesis is based on a multidimensional vision of risk and proposes some multivariate risk measures that can be applied to several actuarial issues of a multivariate nature.

Keywords: *Multivariate risk measures, Solvency 2, Risk management, Internal models, Solvency capital requirement, ORSA approach, Risk theory, Dependence modeling, Capital allocation, Multivariate expectiles, Elicitability, Coherence properties, Stochastic approximation, Asymptotic independence, Extreme values, Copulas*

Table des matières

Remerciements	iii
Résumé	vii
Tables des matières	xi
Introduction générale	1
Introduction	3
La théorie des mesures de risque	3
La modélisation de la dépendance	7
La théorie des valeurs extrêmes	9
La problématique de l'allocation du capital en assurance	15
Objectifs et organisation de la thèse	17
Notations de la thèse	18
Principaux résultats	21
Construction de mesures expectiles multivariées	21
Comportement asymptotique des expectiles multivariés	24
Allocation du capital par minimisation du risque	27
Impact de la dépendance sur l'allocation du capital par minimisation du risque	30
Des mesures de risque expectiles multivariées	37
Chapitre 1 Des extensions multivariées des mesures de risque expectiles	39
1.1 Élicitabilité des mesures de risque	41
1.2 Constructions de mesures expectiles multivariées	48
1.3 Propriétés de cohérence	54
1.4 Approximation stochastique	64
1.5 Comportement asymptotique	71
Bibliographie	75
1.6 Annexe : Illustrations numériques	77

Chapitre 2 Comportement asymptotique des expectiles multivariés	87
2.1 Variations régulières multivariées et dépendance extrême	87
2.2 Modèle Fréchet à queues équivalentes	93
2.3 Modèle Fréchet à queue dominante	104
2.4 Modèle par mélange d'échelle	110
2.5 Estimation de l'expectile multivarié asymptotique	113
Bibliographie	121
Allocation du capital par minimisation du risque	123
Chapitre 3 Allocation du capital par minimisation d'indicateurs multivariés de risque	125
3.1 Optimal allocation presentation	127
3.2 Coherence of a capital allocation in a multivariate context	130
3.3 Coherence of the optimal allocation	132
3.4 Discussion: What could be the best choice for a capital allocation?	138
Bibliography	141
Chapitre 4 Impact de la dépendance	143
4.1 Optimal allocation	145
4.2 Independence models	148
4.3 Asymptotic independence models	157
4.4 Asymptotic dependence models	160
4.5 Perfect dependence	167
Bibliography	171
4.6 Appendix: Proofs	172
Conclusion	175
Conclusion et perspectives	177
Bibliographie	179

Introduction générale

Introduction

Dans cette thèse, nous nous sommes intéressés aux mesures de risque multivariées et à leurs possibles applications en science actuarielle. Les mesures doivent être à même de prendre en compte la nature différente des risques ainsi que leurs relations de dépendance. Le parfait exemple de l'intérêt de ces mesures est l'allocation du capital sous risque. Il s'agit d'un important exercice actuariel qui concerne toutes les compagnies d'assurance possédant plusieurs activités, et qui peut avoir un impact crucial sur les résultats financiers de l'entreprise. Cette problématique est de nature multivariée et nécessite une prise en compte de la dépendance qui existe entre l'ensemble des risques couverts par l'assureur.

L'objectif de cette thèse est la construction de mesures de risque multivariées qui vérifient un certain nombre de propriétés de cohérence, et qui peuvent répondre à des problématiques actuarielles précises. Pour cela, la première partie est consacrée à la présentation des mesures construites et à l'analyse de leur comportement asymptotique dans le contexte des variations régulières. La seconde partie aborde la méthode d'allocation du capital par minimisation d'indicateurs multivariés de risque. Elle donne une axiomatisation des critères de cohérence pour cette technique d'allocation et analyse l'impact de la dépendance sur le comportement de l'allocation optimale.

Ce travail s'appuie sur plusieurs concepts développés dans le cadre de la théorie du risque. La théorie des mesures de risque avec sa riche littérature constitue le point de départ de cette thèse qui cible la contribution au développement de la vision multivariée de cette théorie. Toute approche multivariée de gestion des risques est construite sur une modélisation de la dépendance, ce travail utilise dans l'ensemble de ses parties les outils de modélisation de la dépendance présents dans la littérature, notamment les copules et les différentes mesures de dépendance et de dépendance asymptotique. Nous faisons appel aussi dans le cadre de cette thèse à la théorie des valeurs extrêmes nécessaire pour l'étude du comportement asymptotique des mesures. Nous utilisons sa relation avec les fonctions à variations régulières pour caractériser le comportement des mesures que nous allons introduire en fonction de la nature du domaine d'attraction des marginales, et nous proposons quelques estimateurs statistiques à l'aide de la littérature des statistiques des extrêmes. Enfin, l'ensemble des mesures définies dans le cadre de ce travail est utilisable pour plusieurs applications en assurance, elles offrent une réponse multivariée à plusieurs problématiques notamment celle de l'allocation du capital.

La théorie des mesures de risque

Quantifier les risques est la préoccupation essentielle des actuaires dans l'exercice de leur métier. Les mesures de risque sont des outils présents dans la quasi-totalité de la chaîne actuarielle. Calculer une prime, déterminer et allouer un capital de solvabilité, évaluer un contrat de réassurance, et estimer une perte ou une probabilité de ruine, les mesures de risque sont au cœur de toute opération visant la maîtrise du risque.

Nous adopterons dans cette thèse la convention actuarielle selon laquelle les variables aléatoires représentent des risques ou des pertes lorsqu'elles sont positives. Une mesure de risque ρ est une application sur l'ensemble des variables aléatoires, qui prend ses valeurs dans \mathbb{R} , sa valeur représente le montant d'argent nécessaire pour que le risque devienne acceptable. Cette définition est traduite mathématiquement à l'aide de la notion d'ensemble d'acceptation (en anglais *acceptance set*). En dimension 1,

les mesures de risque invariantes par translation sont totalement caractérisées par leurs ensembles d'acceptation \mathcal{A}_ρ définis par

$$\mathcal{A}_\rho := \{X : \rho(X) \leq 0\},$$

où ρ désigne la mesure de risque univariée concernée. \mathcal{A}_ρ représente l'ensemble des risques qui peuvent être considérés acceptables de point de vue économique, ou par le régulateur dans un contexte actuariel. La mesure de risque est définie à l'aide de son ensemble d'acceptation par

$$\rho(X) = \inf\{r : X - r \in \mathcal{A}_\rho\}.$$

Il s'agit de la représentation duale des mesures de risques introduite dans Artzner et al. (1999) [3]. Les mesures de risque sont caractérisées par des propriétés qui reflètent des comportements économiques souhaitables. L'axiomatisation de la cohérence des mesures de risque représente sans doute un point de repère dans le développement de la théorie du risque. Cette caractérisation apparue dans Artzner et al. (1999) [3], est considérée aujourd'hui comme la référence en matière d'analyse des propriétés des mesures. Une généralisation multidimensionnelle a été proposée par Jouini et al. (2004) [71] pour définir la cohérence des mesures de risque sous forme d'ensemble de vecteurs. En dimension 1, une mesure de risque est qualifiée de cohérente si elle vérifie les quatre axiomes suivants :

1. *L'invariance par translation* : Pour toute constante c ,

$$\rho(X + c) = \rho(X) + c;$$

2. *L'homogénéité positive* : Pour toute constante positive c ,

$$\rho(cX) = c\rho(X),$$

en particulier $\rho(0) = 0$;

3. *La monotonie* : Pour tous risques X et Y définis sur le même espace de probabilité,

$$\mathbb{P}(X \leq Y) = 1 \Rightarrow \rho(X) \leq \rho(Y);$$

4. *La sous-additivité* : Pour tous risques X et Y définis sur le même espace de probabilité,

$$\rho(X + Y) \leq \rho(X) + \rho(Y).$$

La première propriété traduit aussi la neutralité de la mesure car

$$\rho(X - \rho(X)) = 0.$$

La quatrième propriété est une traduction mathématique de la prise en compte du gain de la diversification. D'autres propriétés non considérées comme des axiomes de cohérences sont aussi souhaitables et parfois naturelles. Nous donnons à titre d'exemples :

- *L'invariance par loi* : Pour un risque X , la mesure $\rho(X)$ ne dépend que de sa loi de probabilités. Pour tous risques X et Y de même loi,

$$\rho(X) = \rho(Y);$$

- *L'additivité comonotone* : Pour tout risque X et toutes fonctions croissantes f et g ,

$$\rho(f(X) + g(X)) = \rho(f(X)) + \rho(g(X));$$

- *La convexité* : Pour tous risques X et Y définis sur le même espace de probabilité, pour toute constante $\alpha \in [0, 1]$,

$$\rho(\alpha X + (1 - \alpha)Y) \leq \alpha\rho(X) + (1 - \alpha)\rho(Y).$$

Plusieurs familles de mesures de risque sont présentées dans la littérature de la théorie du risque. Les mesures usuelles les plus utilisées par les praticiens sont :

- La *Value at Risk* de seuil α est le quantile d'ordre α de la distribution de X

$$VaR_\alpha(X) = \inf\{x \in \mathbb{R} : \mathbb{P}(X \leq x) \geq \alpha\},$$

si la X est de distribution continue alors $VaR_\alpha(X) = \inf\{x \in \mathbb{R} : F(x) \geq \alpha\} = F^{-1}(\alpha)$, où F est la fonction de répartition de X .

La VaR n'est pas une mesure cohérente car elle ne satisfait pas la propriété de la sous-additivité. Elle est comonotoniquement additive et vérifie, pour toute fonction croissante et continue, φ

$$VaR_\alpha(\varphi(X)) = \varphi(VaR_\alpha(X)).$$

En pratique, le principal inconvénient de la VaR est son incapacité à fournir l'information sur le comportement du risque au-delà de sa valeur.

- La *Tail Value at Risk* (TVaR) avec un niveau de confiance α est définie comme une moyenne des VaR dépassant la $VaR_\alpha(X)$:

$$TVaR_\alpha(X) = \frac{1}{1-\alpha} \int_\alpha^1 VaR_\kappa(X) d\kappa.$$

La TVaR est une mesure cohérente pour tous les risques de distribution continue. Elle représente la plus petite mesure cohérente majorant la VaR, et elle fournit l'information sur le comportement du risque et l'épaisseur de la queue de sa distribution.

- La *Conditional Tail Expectation* (CTE) est définie par

$$CTE_\alpha(X) = \mathbb{E}[X|X > VaR_\alpha(X)],$$

la CTE et la TVaR coïncident pour les risques de distribution continue.

- La *Conditional VaR* (CVaR) est définie par

$$CVaR_\alpha(X) = \mathbb{E}[X - VaR_\alpha(X)|X > VaR_\alpha(X)] = CTE_\alpha(X) - VaR_\alpha(X),$$

- L'*Expected Shortfall* (ES) est définie par

$$ES_\alpha(X) = \mathbb{E}[(X - VaR_\alpha(X))_+] = CVaR_\alpha(X)\mathbb{P}(X > VaR_\alpha(X)).$$

La variance est aussi considérée comme une mesure de risque importante, utilisée en particulier pour la construction de la frontière efficiente de Markovitz en allocation des actifs financiers. En assurance, l'espérance et la variance sont des mesures de risques limitées dans leur utilisation mais restent néanmoins utiles.

La théorie des mesures de risque est en relation étroite avec la théorie de la décision économique. Une mesure performante permet d'expliquer les comportements observés des risques. Toutefois, beaucoup de paramètres d'aversion au risque restent difficilement mesurables. Deux grandes familles de mesures illustrent cette relation. La première est constituée par les mesures de risque basées sur la théorie de l'*espérance d'utilité*, cette famille de mesures est construite sur la notion de fonction d'utilité, qui traduit l'aversion au risque d'un agent économique. Considérons un assureur possédant une richesse initiale R_0 , et une fonction d'utilité notée u croissante et concave, il accepte d'assurer un risque X au prix $\rho(X)$ s'il est solution de l'équation suivante :

$$\mathbb{E}[u(R_0 - X + \rho(X))] = u(R_0),$$

$\rho(X)$ est donc la mesure associée au risque X . La deuxième est la famille des *mesures de distorsion* composée des mesures de risque issues de la théorie de l'*espérance déformée*. Elle suppose l'existence d'une fonction $g : [0, 1] \rightarrow [0, 1]$ non croissante telle que $g(0) = 0$ et $g(1) = 1$, appelée fonction de déformation ou de distorsion. Le critère du choix est la maximisation de l'espérance déformée :

$$\mathbb{H}_g[X] = - \int_{-\infty}^0 (1 - g(\bar{F}_X(x))) dx + \int_0^\infty g(\bar{F}_X(x)) dx = \int_0^1 VaR_\alpha(X) dg(1 - \alpha).$$

Wang (2000) [110] a étudié la famille des mesures de distorsion appelées *mesures de Wang*. À toute fonction de déformation convexe g , il associe la mesure donnée par :

$$\rho(X) = \mathbb{H}_{\bar{g}}[X] = \int_0^1 VaR_\alpha(X) dg(\alpha),$$

où $\bar{g}(\alpha) = 1 - g(1 - \alpha)$ est une fonction de déformation concave. Toute mesure de Wang s'exprime donc sous forme d'un mélange de VaR. Les mesures de Wang sont positivement homogènes, invariantes par translation et monotones.

Les mesures de Wang associées à des fonctions de déformation \bar{g} concaves sont cohérentes. La VaR est une mesure de Wang associée à la fonction de déformation $\bar{g}_\alpha(x) = \mathbb{1}_{[x \geq 1 - \alpha]}$. La TVaR est aussi une mesure de Wang obtenue avec $\bar{g}_\alpha(x) = \min\{\frac{x}{1 - \alpha}, 1\}$ comme fonction de distorsion pour le seuil α . Acerbi (2002) [2] a introduit la classe des mesures de risques spectrales. Cette famille est composée, selon Tasche (2002) [105], des mesures de risques cohérentes qui sont invariantes par loi et comonotoniquement additives.

Nous nous appuyons dans cette thèse sur une littérature récente, qui se focalise sur une nouvelle propriété désirable statistiquement pour toute mesure de risque. Une mesure de risque ρ est dite *élicitable* si elle peut être écrite, pour toute variable aléatoire X , sous la forme d'une solution de minimisation d'une fonction d'erreur

$$\rho(X) = \arg \min_x \mathbb{E}_{\mathbb{P}}[s(x, X)], \quad X \sim \mathbb{P}.$$

Cette propriété introduite dans la littérature de la théorie du risque par Gneiting (2011) [60] est importante pour comparer les différentes méthodes d'estimations des mesures de risque. Elle donne un sens aux procédures de *backtesting* nécessaires dans la gestion pratique des risques financiers et actuariels.

La VaR est une mesure élicitable pour tout seuil $\alpha \in [0, 1]$, elle s'écrit sous la forme

$$VaR_\alpha(X) = \arg \min_{x \in \mathbb{R}} \{\mathbb{E}[\alpha(X - x)_+ + (1 - \alpha)(x - X)_+]\},$$

tandis que la TVaR n'est pas une mesure élicitable. Ce constat a remis en question la notion de cohérence des mesures de risque, et a lancé une nouvelle discussion académique pour définir le choix de mesure qui répond le mieux aux besoins de la modélisation des risques. Une des réponses proposées est fondée sur la caractérisation des mesures de risque qui sont à la fois cohérentes et élicitables. Dans ce sens, les *mesures expectiles* sont apparues récemment comme une alternative souhaitable et défendable de la VaR et la TVaR.

Les expectiles ont été introduits par Newey et Powell (1987) [72]. Ils sont définis pour une variable aléatoire X comme suit

$$e_\alpha(X) = \arg \min_{x \in \mathbb{R}} \mathbb{E}[\alpha(X - x)_+^2 + (1 - \alpha)(x - X)_+^2],$$

où $(x)_+ = \max(x, 0)$, et α le seuil de confiance. Bellini et al. (2014) [12] ont démontré que ces mesures sont cohérentes et élicitables pour un niveau de confiance $\alpha \geq 1/2$, elles sont, selon Bellini and Bigozzi (2015) [10], les seules mesures de risques, invariantes par loi, qui vérifient à la fois ces deux propriétés.

En assurance, les normes prudentielles *Solvabilité 2* invitent les assureurs à considérer leur propre vision des risques qu'ils supportent. Elles offrent la possibilité à toute compagnie d'implémenter un modèle interne de gestion des risques fondé sur des hypothèses réalistes, notamment en termes de nature des risques et des structures de dépendances. Le choix des mesures de risque à utiliser dans le calcul des capitaux de solvabilité est aussi libre dans le cadre d'un modèle interne. Ce constat a rendu nécessaire le choix d'une mesure de risque adaptée au portefeuille des risques assurés. La comparaison des mesures entre ainsi dans les priorités des travaux de la recherche en gestion de risques.

Les compagnies d'assurance, notamment les réassureurs et les grands assureurs, sont généralement des structures mutli-branches. Elles sont composées de plusieurs lignes d'affaires, d'activités différentes ou implantées dans différents pays ou zones géographiques. Dans ce cas, les mesures de risque univariées sont capables de fournir une quantification du degré de risque au niveau local ou, par agrégation, au niveau global. Cependant, elles perdent leur efficacité dans toute analyse multivariée qui considère le vecteur des risques, comme la détermination de la contribution locale dans l'état de la solvabilité globale. Cela justifie une *approche multivariée* de la gestion du risque.

Dans le cas univarié, la présence d'un ordre naturel donne un sens simple à la définition des mesures de risque. En dimension supérieure, la perte de cet ordre rend le concept de mesure de risque plus difficile à définir et admet différentes interprétations.

Les nouveaux enjeux de la théorie du risque sont surtout la formalisation des propriétés désirées, pour répondre à des besoins pratiques, et la prise en compte de la dépendance.

L'ensemble de cette thèse adopte une vision multivariée de la gestion du risque. Nous consacrons ainsi les deux chapitres de la première partie à la présentation et l'étude d'extensions multidimensionnelles des mesures expectiles. Ces extensions héritent d'une grande partie des propriétés de cohérence de la version univariée, notamment l'élicitabilité. Le choix d'une construction parmi les méthodes proposées dépendra de l'utilisation envisagée et du choix de modélisation de la dépendance. La deuxième partie de cette thèse est aussi fondée sur une approche multivariée qui répond à une problématique de même nature.

Nous présentons maintenant les principaux outils utilisés dans cette thèse.

La modélisation de la dépendance

En assurance, l'hypothèse d'indépendance des risques n'est pas réaliste. La prise en compte de la dépendance est indispensable. La théorie des copules représente aujourd'hui un outil fondamental dans la gestion des risques.

Une copule C est une fonction de répartition multivariée ayant pour marginales des lois uniformes sur $[0, 1]$. Elle représente la structure de dépendance qui relie les distributions marginales pour former la distribution jointe.

Théorème de Sklar (1959) [103] Pour tout vecteur aléatoire $\mathbf{X} = (X_1, \dots, X_d)^T$ dans \mathbb{R}^d dont les lois marginales F_{X_1}, \dots, F_{X_d} sont continues, il existe une unique copule C telle que

$$F_{\mathbf{X}}(x_1, \dots, x_d) = \mathbb{P}(X_1 < x_1, \dots, X_d < x_d) = C(F_{X_1}(x_1), \dots, F_{X_d}(x_d)).$$

L'importance de ce théorème réside dans la possibilité de séparer l'analyse multivariée des risques en deux parties, une univariée qui traite uniquement des distributions marginales, et une deuxième qui concerne la structure de la dépendance. La densité jointe de la distribution multivariée est aussi caractérisée à l'aide de la densité de sa copule si cette dernière est de classe C^d

$$f_{\mathbf{X}}(x_1, \dots, x_d) = c(F_{X_1}(x_1), \dots, F_{X_d}(x_d)) \prod_{i=1}^d f_{X_i}(x_i),$$

avec $c = \frac{\partial^d C}{\partial x_1 \dots \partial x_d}$.

La copule de survie \bar{C} est la structure de dépendance qui relie les fonctions de survie marginales $\bar{F}_{X_1}, \dots, \bar{F}_{X_d}$.

Le problème de Fréchet occupe une place importante dans la littérature de la modélisation de la dépendance. Il s'intéresse aux bornes de la distribution jointe en fonction des marginales. Ces bornes

portent le nom *bornes de Fréchet-Hoeffding* et sont données par l'inégalité suivante

$$\max \left(\sum_{i=1}^d F_{X_i}(x_i) - d + 1 \right) \leq F_{\mathbf{X}}(x_1, \dots, x_d) \leq \min (F_{X_1}(x_1), \dots, F_{X_d}(x_d)).$$

Les mesures de dépendance entre variables aléatoires ont aussi connu un développement important grâce à l'utilisation des copules pour définir les distributions multivariées.

Parmi ces outils de mesure de la dépendance, nous utilisons dans le cadre de cette thèse, le *rho de Spearman* qui représente une mesure de concordance définie pour les distributions bivariées continues à l'aide de leurs copules C par

$$\rho_S = 12 \int \int_{[0,1]^2} uv dC(u, v) - 3 = 12 \int \int_{[0,1]^2} C(u, v) dudv - 3,$$

et le *tau de Kendall* qui s'exprime en fonction de la copule associée à la distribution jointe bivariée comme suit

$$\tau = 4 \int \int_{[0,1]^2} C(u, v) dC(u, v) - 1.$$

Nous utilisons aussi des mesures de dépendance caudale des distributions que nous appellerons la *dépendance asymptotique*. Le coefficient de dépendance des queues supérieures a été introduit dans Joe (1997) [69]. Il est défini pour une distribution bivariée (X, Y) de marginales continues par

$$\lambda_U = \lim_{u \rightarrow 1^-} \mathbb{P}(Y > F_Y^{-1}(u) | X > F_X^{-1}(u)),$$

et peut être exprimé à l'aide de la copule C correspondante au couple (X, Y) comme suit

$$\lambda_U = \lim_{u \rightarrow 1^-} \frac{1 - 2u + C(u, u)}{1 - u},$$

si cette limite existe.

De la même façon, le coefficient de dépendance des queues inférieures est défini par

$$\lambda_L = \lim_{u \rightarrow 0^+} \mathbb{P}(Y \leq F_Y^{-1}(u) | X \leq F_X^{-1}(u)) = \lim_{u \rightarrow 0^+} \frac{C(u, u)}{u},$$

quand la limite existe.

L'analyse profonde de la dépendance asymptotique nécessite l'utilisation d'une définition plus générale que ces coefficients constants. Les fonctions de dépendances des queues répondent à ce besoin. Considérons un vecteur aléatoire \mathbf{X} dans \mathbb{R}^d de distributions marginales continues. La fonction de dépendance extrême supérieure est définie par Klüppelberg et al (2008) [77] comme suit

$$\lambda_U^{\mathbf{X}}(x_1, \dots, x_d) = \lim_{t \rightarrow 0} t^{-1} \mathbb{P}(1 - F_{X_1}(X_1) \leq tx_1, \dots, 1 - F_{X_d}(X_d) \leq tx_d).$$

À l'aide de la copule de survie \tilde{C}

$$\lambda_U^{\mathbf{X}}(u_1, \dots, u_d) = \lim_{t \rightarrow 0^+} \frac{\tilde{C}_{\mathbf{X}}(tu_1, \dots, tu_d)}{t}.$$

Dans le cas bivarié, les coefficients de dépendance asymptotique sont uniquement les valeurs des fonctions de dépendance extrême en $(1, 1)$.

Ordres stochastiques

Pour comparer les risques, on fait appel à la notion des ordres stochastiques. Nous rappelons la définition, présentée dans Shaked et Shanthikumar (2007)[102], de la dominance d'ordre stochastique.

Pour deux risques représentés par les variables aléatoires X et Y , Y est dit dominé par X selon la dominance stochastique de premier ordre si et seulement si :

$$\bar{F}_X(x) \leq \bar{F}_Y(x), \quad \forall x \in \mathbb{R}^+,$$

dans ce cas, on note cette dominance par : $X \leq_{st} Y$.

Cette définition est équivalente à

$$X \leq_{st} Y \Leftrightarrow \mathbb{E}[u(X)] \leq \mathbb{E}[u(Y)], \text{ pour toute fonction croissante } u.$$

Pour comparer des copules bivariées nous utilisons l'ordre des copules noté \preceq . Pour deux copules bivariées C et C'

$$C \preceq C' \Leftrightarrow C(u, v) \leq C'(u, v) \quad \forall (u, v) \in [0, 1]^2.$$

Risques comonotones

Dans un contexte de gestion des risques, la dépendance forte entre les variables aléatoires est un cas extrêmement important qui doit être pris en compte dans l'étude des mesures de risque. Ce cas est appelé la comonotonie. Le concept des variables aléatoires comonotones est apparu dans les travaux de Hoeffding et Fréchet (1940) [63] et (1951) [56]. Dans cette thèse nous utilisons la définition des risques comonotones dans sa formulation actuarielle, introduite dans la littérature dans Borch (1962) [15].

Les variables aléatoires (X_1, X_2, \dots, X_n) sont considérées comonotones si et seulement si il existe une variable aléatoire Y et des fonctions non décroissantes $\varphi_1, \dots, \varphi_n$ telles que :

$$(X_1, \dots, X_n) =_d (\varphi_1(Y), \dots, \varphi_n(Y)).$$

La théorie des valeurs extrêmes

Les seuils de solvabilité réglementaires étant élevés en assurance, le comportement asymptotique occupe une place importante dans l'analyse des mesures de risque. La théorie des valeurs extrêmes est un outil important pour étudier ce comportement. Elle permet d'appréhender les volumes de pertes extrêmes et de quantifier la dangerosité asymptotique des risques en fonction de la nature de leurs distributions.

Pour des variables aléatoires $X_{1,n}, X_{2,n}, \dots, X_{n,n}$ ordonnées et identiquement distribuées, le théorème des valeurs extrêmes annoncé dans Gnedenko (1943) [59] permet de caractériser la distribution limite du maximum $X_{n,n}$ à l'aide d'une opération de normalisation. Il annonce que s'il existe deux suites, une positive d'échelle (b_n) et une de position (a_n) telles que la loi limite de $\frac{X_{n,n} - a_n}{b_n}$ soit non-dégénérée représentée par une distribution H , alors H ne peut prendre que l'une des trois formes suivantes :

Fréchet

$$\Phi_\theta(x) = \begin{cases} 0 & \text{si } x \leq 0 \\ \exp(-x^{-\theta}) & \text{si } x > 0 \end{cases}$$

Weibull

$$\Psi_\theta(x) = \begin{cases} \exp(-(-x)^{-\theta}) & \text{si } x \leq 0 \\ 1 & \text{si } x > 0 \end{cases}$$

Gumbel

$$\Lambda(x) = \exp(-e^{-x}), \quad x \in \mathbb{R}$$

Ces trois distributions limites forment naturellement les trois domaines d'attractions qui portent leurs noms. Elles sont regroupées dans la forme générale introduite dans Jenkinson (1955) [68]

$$\lim_{n \rightarrow \infty} \mathbb{P}\left(\frac{X_{n,n} - a_n}{b_n} \leq x\right) = H_\gamma(x),$$

avec

$$H_\gamma(x) = \begin{cases} \exp\left(-(\mathbf{1} + \gamma x)_+^{-1/\gamma}\right) & \text{si } \gamma \neq \mathbf{0} \\ \exp(-e^{-x}) & \text{si } \gamma = \mathbf{0} \end{cases}.$$

γ est l'indice des valeurs extrêmes.

Pour identifier le domaine d'attraction de la distribution des valeurs extrêmes en pratique, on peut utiliser le test d'exponentialité pour les excès, présenté en détail avec des illustrations numériques et graphiques dans la thèse de Garrido (2002) [58]. Il s'agit d'une méthode de type *QQ-plot* qui consiste à représenter graphiquement le quantile de la distribution exponentielle comme représentation du domaine d'attraction de Gumbel contre les excès dans l'ordre pour un échantillon d'extrema.

Pour estimer la distribution des valeurs extrêmes, on peut aussi utiliser l'approche POT (*Peaks Over Threshold*). Cette méthodologie utilise la distribution des excès et nécessite le choix d'un seuil u , puis l'étude de la variable aléatoire $X - u$ quand $X \geq u$. La fonction de répartition de cette variable aléatoire est

$$\bar{F}_u(y) = \mathbb{P}(X \geq y + u \mid X \geq u) = \frac{\mathbb{P}(X \geq y + u, X \geq u)}{\mathbb{P}(X \geq u)} = \frac{\bar{F}(y + u)}{\bar{F}(u)},$$

lorsque le seuil est assez grand, $\bar{F}_u(y)$ peut être approchée par la fonction de survie de la loi de Pareto Généralisée

$$\bar{G}_{\gamma, \sigma(u)}(y) = \begin{cases} \left(1 + \gamma \frac{y}{\sigma}\right)^{-1/\gamma} & \text{si } \gamma \neq \mathbf{0} \\ \exp\left(-\frac{y}{\sigma}\right) & \text{si } \gamma = \mathbf{0} \end{cases},$$

définie sur \mathbb{R}^+ si $\gamma \geq 0$, et sur $[0, -\sigma/\gamma[$ si $\gamma < 0$. Le paramètre d'échelle σ doit être strictement positif. Le théorème de Pickands (1975) permet de relier la distribution des excès et la loi des valeurs extrêmes.

Théorème Pickands (1975) [93] - Balkema et De Haan (1974) [5]

$$\lim_{n \rightarrow \infty} \mathbb{P}\left(\frac{X_{n,n} - a_n}{b_n} \leq x\right) = H_\gamma(x) \Leftrightarrow \lim_{u \rightarrow x_F} \sup_{y \in [0, x_F - u]} |\bar{F}_u(y) - \bar{G}_{\gamma, \sigma(u)}(y)| = 0.$$

Un ensemble de résultats et d'applications de la convergence de la distribution des excès est présenté dans le livre d'Embrechts (2000) [47]. Denuit et Charpentier (2005)[42] ont démontré que le nombre de dépassements du seuil des excès suit asymptotiquement une loi de Poisson de paramètre $\lambda = \lim_{n \rightarrow +\infty} n\bar{F}_X(u_n)$, avec u_n le seuil choisi pour un échantillon de taille n , si cette limite existe.

Par un changement de variable simple ($x = y + u$), la fonction de survie de X au-delà de u peut s'écrire sous la forme

$$\bar{F}(x) = \bar{F}(u)\bar{F}_u(x - u).$$

Le théorème de Pickands permet donc de l'approximer par

$$\bar{F}(x) \simeq \bar{F}(u)\bar{G}_{\gamma, \sigma}(x - u) = \begin{cases} \bar{F}(u) \left(1 + \gamma \frac{x-u}{\sigma}\right)^{-1/\gamma} & \text{si } \gamma \neq 0 \\ \bar{F}(u) \exp\left(-\frac{x-u}{\sigma}\right) & \text{si } \gamma = 0 \end{cases}.$$

Plusieurs méthodes d'estimation des paramètres de la distribution des valeurs extrêmes sont présentes dans la littérature des statistiques des extrêmes.

Statistique des extrêmes

La distribution marginale asymptotique est entièrement définie par son indice des valeurs extrêmes, et ses deux suites de normalisation. En pratique, et sans connaissance de la forme exacte de la distribution,

il faut estimer ses trois paramètres. La fonction de répartition de la loi des valeurs extrêmes s'écrit sous la forme

$$H_{\gamma,a,b}(x) = H_{\gamma} \left(\frac{x-a}{b} \right) = \exp \left(- \left(1 + \gamma \left(\frac{x-a}{b} \right) \right)_+^{-1/\gamma} \right).$$

Pour estimer les paramètres de la loi des valeurs extrêmes, la méthode du maximum de vraisemblance ne permet pas d'obtenir des estimateurs explicites. On peut utiliser la méthode des moments pondérés proposée par Hosking et al. (1985) [64], et qui consiste à utiliser un échantillon de maxima $X_{i,n-k+1,n}, \dots, X_{i,n,n}$ de fonction de répartition $H_{\gamma,a,b}$, pour déterminer les moments pondérés. Ces moments sont définis pour un ordre m et pour une variable aléatoire Y de fonction de répartition $H_{\gamma,a,b}$ par

$$\mu_m = \mathbb{E}[Y H_{\gamma,a,b}^m(Y)].$$

Ces moments existent pour un indice $\gamma < 1$, et ont la forme explicite suivante

$$\mu_m = \frac{1}{m+1} \left(a - \frac{b}{\gamma} (1 - (m+1)^{\gamma} \Gamma(1-\gamma)) \right),$$

avec $\Gamma(x) = \int_0^{+\infty} t^{x-1} \exp(-t) dt$. Les trois premiers moments pondérés sont donc suffisants pour estimer les paramètres de la loi des extrêmes.

D'un autre côté, on peut utiliser des estimateurs linéaires de ces moments

$$\hat{\mu}_m = \frac{1}{k} \sum_{\ell=n-k+1}^n X_{\ell,n} \left(\frac{\ell - n + k - 1}{k} \right)^m.$$

Ces estimateurs sont obtenus directement en remplaçant l'espérance par la moyenne empirique et la fonction de répartition $H_{\gamma,a,b}$ par la fonction de répartition empirique. Les estimateurs des paramètres de la loi des extrêmes sont déduits ensuite des moments

– $\hat{\gamma}$ est solution de l'équation

$$\frac{1 - 3^{\hat{\gamma}}}{1 - 2^{\hat{\gamma}}} = \frac{3\hat{\mu}_2 - \hat{\mu}_0}{2\hat{\mu}_1 - \hat{\mu}_0};$$

–

$$\hat{b} = \frac{\hat{\gamma}(\hat{\mu}_0 - 2\hat{\mu}_1)}{(1 - 2^{\hat{\gamma}}) \Gamma(1 - \hat{\gamma})};$$

–

$$\hat{a} = \hat{\mu}_0 + \frac{(\hat{\mu}_0 - 2\hat{\mu}_1)(1 - \Gamma(1 - \hat{\gamma}))}{(1 - 2^{\hat{\gamma}}) \Gamma(1 - \hat{\gamma})}.$$

L'utilisation d'un échantillon d'extrema, composé de la statistique d'ordre de $k = k(n)$ plus grandes valeurs, est nécessaire pour toutes les estimations liées aux valeurs extrêmes. $k(n)$ doit être grand ($k(n) \rightarrow +\infty$), mais négligeable devant la taille de l'échantillon initial ($k(n)/n \rightarrow 0$). Le choix optimal de k a fait le sujet de plusieurs travaux. Nous citons à titre d'exemple la méthode de détermination des données de queue proposée dans Danielsson et de Vries (1997) [32].

On peut aussi utiliser l'approche POT (*Peaks Over Threshold*), présentée également dans l'introduction de la thèse, et qui permet d'approximer la loi des des valeurs extrêmes à l'aide de la distribution des excès par application du Théorème de Pickands (1975) [93]

$$\bar{F}(x) \simeq \bar{F}(u) \bar{G}_{\gamma,\sigma}(x-u) = \begin{cases} \bar{F}(u) (1 + \gamma \frac{x-u}{\sigma})^{-1/\gamma} & \text{si } \gamma \neq 0 \\ \bar{F}(u) \exp(-\frac{x-u}{\sigma}) & \text{si } \gamma = 0 \end{cases}.$$

L'estimation des paramètres d'échelle σ et de forme γ peut se faire à l'aide de la méthode des moments pondérés présentée précédemment, en utilisant un échantillon des excès au lieu d'un échantillon des extrema. En effet, pour un échantillon $\{Y_1, \dots, Y_k\}$ de k excès indépendants, Hosking et Wallis (1987) [65] ont défini le moment pondéré d'ordre m par

$$\mu_m = \mathbb{E}[Y \bar{G}_{\gamma,\sigma}^m(Y)],$$

qui existe si $\gamma < 1$ et qui est donné par

$$\mu_m = \frac{\sigma}{(m+1)(m+1-\gamma)}.$$

Les deux premiers moments sont donc suffisants pour obtenir γ et σ :

$$\gamma = \frac{4\mu_1 - \mu_0}{2\mu_1 - \mu_0} \quad \text{et} \quad \sigma = \frac{2\mu_1\mu_0}{\mu_0 - 2\mu_1}.$$

Il ne reste alors qu'à estimer les moments pondérés μ_0 et μ_1 . Pour cela, il faut choisir un seuil des excès. En pratique, on choisit $u = X_{n-k+1,n}$, et dans ce cas on estime $F_X(u)$ par $\frac{k}{n}$ et on obtient les estimateurs des moments pondérés

$$\hat{\mu}_m = \frac{1}{k} \sum_{\ell=n-k+1}^n (X_{\ell,n} - X_{n-k+1,n}) \left(\frac{\ell - n + k - 1}{k} \right)^m.$$

Finalement,

$$\hat{\gamma} = \frac{4\hat{\mu}_1 - \hat{\mu}_0}{2\hat{\mu}_1 - \hat{\mu}_0} \quad \text{et} \quad \hat{\sigma} = \frac{2\hat{\mu}_1\hat{\mu}_0}{\hat{\mu}_0 - 2\hat{\mu}_1}.$$

Estimation non paramétrique

Plusieurs estimateurs de l'indice des valeurs extrêmes ont été présentés dans la littérature. On choisit k très grand ($k \rightarrow +\infty$) tel que $k/n \rightarrow 0$.

Estimateur de Hill Le plus simple des estimateurs de l'indice des valeurs extrêmes est l'estimateur de Hill (1975) [62] défini par

$$\hat{\gamma}_{\text{Hill}} = \frac{1}{k(n)} \sum_{\ell=1}^{k(n)} \log(X_{n-\ell+1,n}) - \log(X_{n-k(n)+1,n}).$$

Le principal inconvénient de cet estimateur est qu'il n'est utilisable que pour les distributions de Fréchet ($\gamma > 0$). Sa construction est basée sur l'idée de la linéarité du graphique

$$(\log(k/\ell), \log(X_{n-\ell+1,n}) - \log(X_{n-k+1,n})),$$

et l'estimateur de Hill est donc un estimateur de la pente. En effet, une distribution du domaine d'attraction de Fréchet F se caractérise à l'aide de l'indice des valeurs extrêmes par l'existence d'une fonction à variation lente L telle que

$$\bar{F}(x) = x^{-\frac{1}{\gamma}} L(x),$$

donc

$$\lim_{u \rightarrow \infty} \frac{\bar{F}(tu)}{\bar{F}(u)} = t^{-\frac{1}{\gamma}},$$

pour tout $t > 1$. On peut donc approximer $\bar{F}(x)$ au-delà de u par

$$\bar{F}(x) \simeq \bar{F}(u) \left(\frac{x}{u} \right)^{-\frac{1}{\gamma}}. \quad (1)$$

Cette approximation peut aussi s'écrire sous la forme

$$\gamma \log \left(\frac{\bar{F}(u)}{\bar{F}(x)} \right) \simeq \log(x) - \log(u),$$

et en utilisant l'approche GPD, avec $u = X_{n-k+1,n}$, on obtient pour $\ell \in \{1, \dots, k\}$

$$\gamma \log \left(\frac{k/n}{\ell/n} \right) = \gamma \log \left(\frac{k}{\ell} \right) \simeq \log(X_{n-\ell+1,n}) - \log(X_{n-k+1,n}).$$

L'indice γ est par suite approximé par

$$\gamma \simeq \frac{\sum_{\ell=1}^k \log(X_{n-\ell+1,n}) - \log(X_{n-k+1,n})}{\sum_{\ell=1}^k \frac{k}{\ell}} = \frac{\sum_{\ell=1}^k \log(X_{n-\ell+1,n}) - \log(X_{n-k+1,n})}{k \log(k) - \log(k!)},$$

et son estimateur est basé sur l'équivalence

$$k \log(k) - \log(k!) \stackrel{k \uparrow +\infty}{\sim} k.$$

Cette dernière équivalence découle de la Formule de Stirling $k! \stackrel{k \uparrow +\infty}{\sim} \sqrt{2\pi k} (k/e)^k$.

L'estimateur de Hill $\hat{\gamma}_H(k)$ converge en probabilité vers γ si $k \rightarrow \infty$ et $k/n \rightarrow 0$ quand $n \rightarrow \infty$. La convergence presque sûre nécessite la condition supplémentaire $k/\log \log n \rightarrow \infty$. La normalité asymptotique est obtenue si on impose une condition de second ordre supplémentaire sur la fonction à variations lentes qui donne sa vitesse de convergence vers 1. Cette condition est l'existence d'un indice réel $\rho > 0$ et une fonction auxiliaire $A(\cdot)$ qui vérifie $\lim_{x \rightarrow +\infty} A(x) = 0$, tels que

$$\lim_{x \rightarrow +\infty} \frac{\log(tx) - \log(x)}{A(x)} = \frac{t^\rho - 1}{\rho},$$

pour tout $t > 1$. Elle est très proche de la définition des $2RV_{\gamma,\rho}$ les variations régulières de second ordre. La normalité asymptotique est obtenue si cette condition est satisfaite et si $\sqrt{k}A(n/k) \rightarrow 0$ quand $n \rightarrow +\infty$. Sous les hypothèses précédentes, l'indicateur vérifie la convergence

$$\lim_{k \rightarrow +\infty} \hat{\gamma}_{\text{Hill}} = \gamma \quad p.s.,$$

et la normalité asymptotique

$$\sqrt{k} \frac{\hat{\gamma}_{\text{Hill}} - \gamma}{\gamma} \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1).$$

Estimateur de Pickands L'estimateur de Pickands (1975) [93] est défini par la statistique suivante

$$\hat{\gamma}_{\text{Pickands}} = \frac{1}{\log(2)} \log \left(\frac{X_{n-\lfloor \frac{k(n)}{4} \rfloor, n} - X_{n-\lfloor \frac{k(n)}{2} \rfloor, n}}{X_{n-\lfloor \frac{k(n)}{2} \rfloor, n} - X_{n-k(n), n}} \right).$$

Il converge en probabilité. La convergence presque sûre est obtenue si de plus $k/\log \log n \rightarrow \infty$ quand $n \rightarrow \infty$. La normalité asymptotique est obtenue sous des conditions supplémentaire détaillées dans Dekkers et de Haan (1989) [37]. Sous ces conditions, cet estimateur est asymptotiquement normal,

$$\sqrt{k} \frac{\hat{\gamma}_{\text{Pickands}} - \gamma}{v(\gamma)} \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1).$$

et sa variance asymptotique est

$$v(\gamma) = \frac{\gamma \sqrt{2^{\gamma+1} - 1}}{2 \log(2) (2^\gamma - 1)}.$$

Il est utilisable pour tous les domaines d'attraction, d'où son intérêt pratique. Pour les distributions du domaine d'attraction de Fréchet, l'estimateur de Pickands est moins efficace que l'estimateur de Hill, et plus sensible à la taille k des observations utilisées.

Estimateur des moments Afin de pallier l'inconvénient du signe de l'indice des valeurs extrêmes de l'estimateur de Hill, une généralisation a été proposée par Dekkers et al. (1989) [38]. Il s'agit de l'estimateur des moments défini par la statistique suivante

$$\hat{\gamma}_M(k) = \hat{\gamma}_{\text{Hill}}(k) + 1 - \frac{1}{2} \left(1 - \frac{\hat{\gamma}_{\text{Hill}}(k)}{S(k)} \right)^{-1},$$

avec

$$S(k) = \frac{1}{k} \sum_{\ell=1}^k (\log(X_{n-\ell+1,n}) - \log(X_{n-k+1,n}))^2.$$

L'estimateur des moments est convergent et asymptotiquement normal sous des conditions de régularité simples

$$\sqrt{k} \frac{\hat{\gamma}_M - \gamma}{\sqrt{\gamma^2 + 1}} \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1),$$

quel que soit le signe de γ , mais il n'a pas d'interprétation graphique.

D'autres extensions de l'estimateur de Hill ont été introduites dans la littérature. Nous citons à titre d'exemple l'estimateur UH proposé par Beirlant et al. (1996) [9], l'estimateur Hill Alternatif proposé par Resnick et Stărică (1997) [96], et l'estimateur connu sous le nom de Zipf, proposé par Kratz et Resnick (1996) [78] et qui utilisent la méthode des moindres carrés ordinaires pour estimer l'indice des valeurs extrêmes par

$$\hat{\gamma}_Z(k) = \frac{\frac{1}{k} \sum_{\ell=1}^k \log\left(\frac{k+1}{\ell}\right) \left(\log(X_{n-\ell+1,n}) - \frac{1}{k} \sum_{\ell=1}^k \log(X_{n-\ell+1,n})\right)}{\frac{1}{k} \sum_{\ell=1}^k \left(\log\left(\frac{k+1}{\ell}\right)\right)^2 - \left(\sum_{\ell=1}^k \log\left(\frac{k+1}{\ell}\right)\right)^2}.$$

Cet estimateur est consistant et asymptotiquement normal mais avec une variance deux fois celle de l'estimateur de Hill. Une liste détaillée des estimateurs proposés dans la littérature est présentée par Csörgő et Viharos (1998) [31].

La théorie des valeurs extrêmes est en relation étroite avec le concept des variations régulières. Les domaines d'attractions caractérisent la nature de la décroissance des queues de distributions, la nature des variations des fonctions de survie déterminera le domaine d'attraction de la distribution.

Le concept des variations régulières

Une fonction mesurable positive f est dite à variations régulières en $+\infty$ d'indice θ , si pour tout $t > 0$,

$$\lim_{x \rightarrow +\infty} \frac{f(tx)}{f(x)} = t^\theta,$$

et on note dans ce cas $f \in \text{RV}_\theta(+\infty)$.

Une variable aléatoire X et sa distribution sont considérées à variations régulières d'indice $\theta \geq 0$ si la queue droite de la distribution $\bar{F}_X \in \text{RV}_{-\theta}(+\infty)$ et on note dans ce cas $\bar{F}_X \in \text{RV}_{-\theta}$. Nous présentons en détail cette caractérisation et ses propriétés dans la première section du chapitre 2. Pour illustrer son importance, Embrechts et al. (1997) [48] montrent

$$F \in \text{MDA}(\Phi_\theta) \iff x_F = \sup\{x : F(x) < 1\} = +\infty \text{ et } \bar{F} \in \text{RV}_{-\theta}.$$

Le domaine d'attraction de Fréchet est donc complètement caractérisé par les distributions à variations régulières.

Les variations régulières multivariées

En dimension supérieure à 1 plusieurs définitions des variations régulières sont proposées dans la littérature. Pour introduire ce concept, un vecteur aléatoire \mathbf{X} dans \mathbb{R}^d est dit à queue à variations régulières d'exposant θ , si et seulement si il existe une mesure finie μ sur la sphère unité \mathbb{S}^{d-1} , une fonction à variations lentes* L , et un réel strictement positif $\theta > 0$, tels que

$$\lim_{x \rightarrow +\infty} \frac{x^\theta}{L(x)} \mathbb{P}\left(\|\mathbf{X}\| > x, \frac{\mathbf{X}}{\|\mathbf{X}\|} \in B\right) = \mu_{\mathbf{X}, \|\cdot\|}(B),$$

pour tout $B \in \mathcal{B}(\mathbb{S}^{d-1})$ avec $\mu_{\mathbf{X}, \|\cdot\|}(\partial B) = 0$. La mesure μ dépend de la norme utilisée et du vecteur \mathbf{X} , on l'appelle la *mesure spectrale* de \mathbf{X} .

Nous revenons dans la première section du chapitre 2 de la thèse sur cette notion. Nous présentons d'autres versions de cette définition, et nous expliquerons son interprétation.

Nous présentons maintenant une introduction à l'allocation du capital en assurance.

*. Fonction dans $\text{RV}_0(+\infty)$

La problématique de l'allocation du capital en assurance

Les applications des mesures de risque en actuariat sont multiples. Les normes *Solvabilité 2* apportent un changement radical en matière de gestion des risques dans le secteur actuariel. Elles sont basées sur une vision de renforcement du contrôle de risque encouru et de minimisation de la probabilité d'insolvabilité. La détermination du capital économique réglementaire connaît sous ce dispositif prudentiel une révolution méthodologique. Elle nécessite une modélisation de la dépendance entre les différents risques et le choix d'une méthodologie pour leur agrégation, cela se traduit dans le cas de l'utilisation de la Formule Standard par exemple par des matrices de corrélations qui relient les familles et les sous-familles de risques. Une fois le *capital de solvabilité requis SCR* calculé, son allocation entre les différentes branches d'activités s'impose comme un exercice interne, certes non contrôlé par le premier pilier de *Solvabilité 2*, mais qui joue un rôle crucial dans la détermination des performances de toute l'activité de l'assureur. Le cas des groupes d'assurance nécessite un traitement particulier dans le cadre de l'approche *Own Risk and Solvency Assessment ORSA*, une analyse multivariée des risques semble donc inévitable.

La détermination du capital économique se construit avec une vision globale du risque, basée sur une approche *bottom-up* tel que c'est le cas dans la formule standard, qui propose l'étude de chaque risque séparément au niveau local, puis l'agrégation de l'ensemble des risques à l'aide des matrices de corrélations entre les modules et les sous-modules de risques. L'allocation de ce capital nécessite une approche *top-down* pour identifier le poids réel de la participation de chaque risque dans le risque global. Il s'agit dans un sens de déterminer des seuils locaux de solvabilité qui prennent en compte à la fois le comportement marginal de l'activité et sa structure de dépendance avec les autres risques.

Afin de formaliser la problématique de l'allocation du capital, qui représente finalement l'opération complémentaire de l'agrégation des risques, on considère un portefeuille de n risques, qui représentent les coûts ou les pertes d'une compagnie d'assurance. Ces risques peuvent représenter les pertes des différentes lignes d'affaires ou branches. La même modélisation peut être généralisée pour une problématique d'allocation de fonds pour un portefeuille financier ou portefeuille de contrats. Nous représentons ces risques à l'aide d'un vecteur de variables aléatoires (X_1, \dots, X_n) .

Le calcul du capital économique nécessite une agrégation des risques après avoir déterminé leur structure de dépendance. Le risque global peut être représenté par la somme $S = \sum_{i=1}^n X_i$. Pour une mesure de risque ρ choisie, le capital économique est calculé par $\rho(S)$.

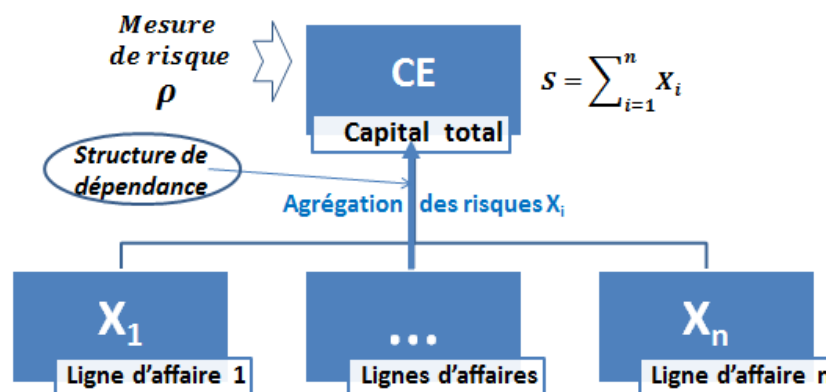


FIGURE 1 – L'agrégation des risques et le calcul du capital économique

L'allocation du capital est l'opération qui consiste à distribuer ce montant $\rho(S) = \rho(\sum_{i=1}^n X_i)$ entre les différents risques $X_i, (i = 1, \dots, n)$, et donc la détermination de la participation de chaque ligne d'affaires dans l'immobilisation des capitaux propres de solvabilité.

Nous présentons graphiquement dans la figure 2 le principe de l'allocation du capital, et sa relation

avec l'opération de l'agrégation des risques.

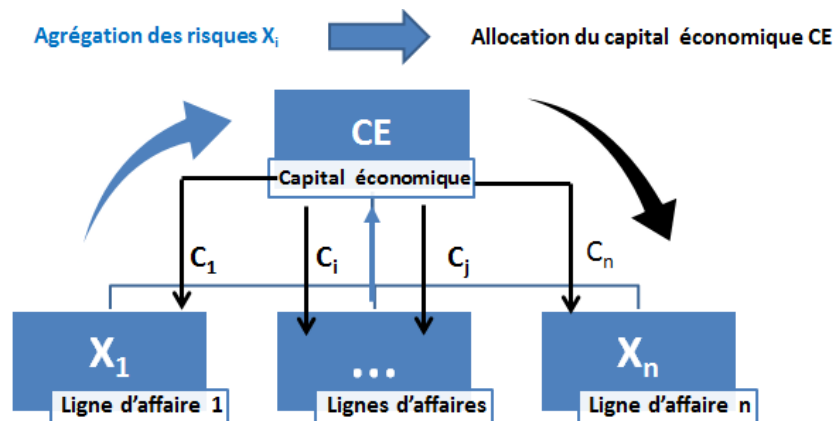


FIGURE 2 – De l'agrégation des risques à l'allocation du capital

Cette problématique touche en premier lieu les groupes d'assurances, constitués de plusieurs lignes d'affaires ou plusieurs branches, parfois implémentées dans des pays dont les règles de solvabilités sont différentes. Le cas de l'AIG* et sa mise sous tutelle publique le 16 septembre 2008, suite à la crise des *subprimes*, est un parfait exemple pour illustrer l'importance d'une bonne allocation du capital. En effet, la filiale américaine du quatrième assureur mondial a connu des fortes difficultés alors qu'il était impossible de transférer ses capitaux des filiales européennes en bonne santé financière.

Dans la deuxième partie de cette thèse, nous nous focalisons sur la problématique de l'allocation du capital, et nous la traitons indépendamment de la méthodologie de la détermination du capital de solvabilité.

La littérature propose un ensemble de méthodes pour répondre à la problématique de l'allocation du capital. Les plus présentes dans les pratiques des compagnies d'assurance, sont la méthode de Shapley et la méthode d'Euler. La première est issue de la théorie des jeux coopératifs. La deuxième utilise l'homogénéité des mesures de risque univariées pour quantifier l'impact marginal de chaque branche sur le groupe.

Nous pouvons résumer les étapes des méthodologies classiques d'allocation de capital dans le processus suivant :

- Établir la distribution multivariée des pertes $X = (X_1, \dots, X_n)$;
- Choisir une mesure de risque univariée ρ qui sera utilisée pour calculer le capital économique ;
- Calculer le capital économique $CE = \rho(S)$, on peut aussi s'intéresser au dépassement $\rho(S) - \mathbb{E}[S]$;
- Choisir un principe d'allocation pour allouer le capital économique entre les branches ;
- Choisir une mesure de risque pour l'allocation, qui n'est pas forcément celle utilisée pour le calcul du capital économique ;
- Déterminer C_i la part de la contribution au risque allouée au risque X_i .

*. **American International Group, Inc.** (AIG) est un des leaders mondiaux de l'assurance, classé quatrième en 2012, derrière AXA, Allianz et Generali, avec un réseau de plus de 135 pays, plus de 74 millions clients). Basé à New York, AIG est listé sur le NYSE, ainsi qu'aux bourses de Zurich, de Paris et de Tokyo.

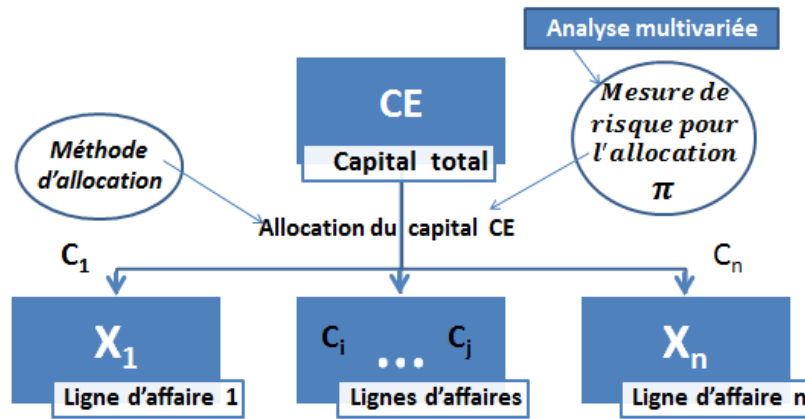


FIGURE 3 – Principe de l'allocation du capital

Ces méthodologies sont donc fondées sur le choix d'une mesure de risque univariée pour l'allocation, et ont pour unique objectif la détermination du poids du risque local par rapport au global. Dans un objectif de minimisation du risque, nous pensons que cette approche peut être améliorée à l'aide d'une méthodologie basée sur la minimisation d'indicateurs multivariés de risque. Cette nouvelle approche dont l'idée a été présentée dans Cénac et al. (2014) [23], et que nous formalisons dans un cadre plus général dans cette thèse, représente une réponse multivariée à une problématique de la même nature. Nous consacrons la deuxième partie de ce manuscrit à cette méthode d'allocation du capital en assurance par minimisation du risque global.

L'allocation doit vérifier quelques propriétés souhaitables, comme l'additivité $CE = \sum_{i=1}^n C_i$. Cela introduit la nécessité d'une axiomatisation de la cohérence économique de toute méthode d'allocation. Dans le premier chapitre de la deuxième partie de cette thèse nous proposons une caractérisation d'une *allocation optimale cohérente* à l'aide d'un ensemble d'axiomes de cohérence. Nous étudions ensuite ces propriétés de cohérence dans le cas d'une allocation optimale par minimisation des indicateurs proposée dans les deux articles Cénac et al. (2014) [23] et Dhaene et al. (2012) [43].

Le deuxième chapitre de cette partie est une analyse du comportement de l'allocation optimale en fonction de la nature des distributions marginales des risques et de la structure de dépendance. Nous étudions aussi dans cet article le comportement asymptotique de l'allocation pour les grands niveaux du capital.

Objectifs et organisation de la thèse

L'objectif général de la thèse est la participation au développement de la théorie multivariée du risque à travers :

- La construction d'une nouvelle famille de mesures de risque multivariées, l'analyse de ses propriétés de cohérence et l'étude de son comportement asymptotique.
- L'axiomatisation de la cohérence d'une méthode d'allocation du capital fondée sur une vision multivariée.
- La proposition de méthodes d'allocation du capital basées sur les mesures de risque introduites.
- L'analyse de l'impact de la dépendance sur la composition d'une allocation optimale par minimisation d'indicateurs multivariés de risque.

Nous nous sommes efforcés dans le cadre de cette thèse de proposer des méthodes applicables en assurance en prenant en compte les règles imposées par les autorités de contrôle, et les objectifs principaux des dispositifs prudentiels.

La thèse est organisée en deux parties. La première partie est dédiée à la présentation de mesures de risque expectiles multivariées. Cette partie s'organise en deux chapitres. Le premier chapitre explique

en détail les idées des constructions proposées. Plusieurs méthodologies de construction de mesures de risque expectiles multivariées sont introduites. Ces mesures vérifient un ensemble de propriétés de cohérence que nous présentons également dans ce chapitre. Nous proposons un outil pratique pour l'approximation stochastique de ces mesures fondé sur la version multivariée de l'algorithme de Robbin-Monro (1951). Le deuxième chapitre est une analyse du comportement asymptotique des expectiles multivariés. Nous nous appuyons sur le concept des variations régulières et sur les outils de la théorie des valeurs extrêmes pour évaluer ce comportement dans les cas de l'équivalence et de la dominance des queues. Cette étude est focalisée sur le cas du domaine d'attraction de Fréchet qui contient les distributions à queues lourdes. Les résultats théoriques obtenus sont ensuite utilisés pour proposer des estimateurs statistiques dans des cas particuliers. Nous avons décidé de rédiger cette partie en français pour pouvoir justifier clairement les constructions proposées.

La deuxième partie est totalement consacrée à la problématique de l'allocation du capital en assurance. Ses deux chapitres sont deux articles sur le sujet. Le premier introduit une axiomatisation de critères de cohérence d'une méthode d'allocation du capital dans un cadre multivarié général, puis démontre que l'approche de l'allocation par minimisation d'indicateurs multivariés de risque présente bien une réponse cohérente à cette problématique. Le deuxième chapitre de cette partie est un article consacré à l'étude de l'impact de dépendance et de la nature des lois marginales sur la composition de l'allocation optimale. Le comportement asymptotique de cette allocation est aussi étudié pour différents cas de modèles.

Un ensemble d'illustrations numériques est présenté pour chaque partie de la thèse.

Notations de la thèse

Dans ce manuscrit, nous allons utiliser des notations et conventions unifiées pour désigner les notions usuelles des sciences actuarielles. Nous présentons dans cette sous-section les notations qui seront les plus utilisées. Nous rappellerons ces notations dans les autres chapitres quand cela sera nécessaire.

\emptyset L'ensemble vide.

\mathbb{R} L'ensemble des réels.

\mathbb{R}^+ L'ensemble des réels positifs.

$(\cdot)_+$ La partie positive définie par $(x)_+ = \max\{x, 0\}$.

$(\cdot)_-$ La partie négative définie par $(x)_- = \max\{-x, 0\}$.

$[\cdot]$ Partie entière.

f^{\leftarrow} La fonction inverse continue à gauche d'une fonction croissante $f : \mathbb{R} \rightarrow \mathbb{R}$ définie par $f^{\leftarrow}(y) = \inf\{x : f(x) \geq y\}$.

$f \stackrel{x \rightarrow x_0}{\sim} g$ Équivalence des fonctions, valide si $\lim_{x \rightarrow x_0} f(x)/g(x) = 1$.

$o(1)$ Négligeable dans le sens $\liminf o(1) = 0$.

$O(1)$ Borné dans le sens $\limsup |O(1)| < +\infty$.

\xrightarrow{v} La convergence vague.

$\xrightarrow{\mathcal{L}}$ La convergence en loi.

$\xrightarrow{p.s.}$ La convergence presque sûre.

$\overline{\lim}$ La limite supérieure.

$\underline{\lim}$ La limite inférieure.

$F_X(\cdot)$ La fonction de répartition de la variable aléatoire X .

$\bar{F}_X(\cdot)$ La fonction de survie de la variable aléatoire X définie par $\bar{F}_X(\cdot) = 1 - F_X(\cdot)$.

$f_X(\cdot)$ La densité de la variable aléatoire X .

x_F Point terminal supérieur de la fonction de répartition F , $x_F = \sup\{x : F_X(x) < 1\}$.

$\mathbb{1}_E$ La fonction indicatrice de l'événement E .

\mathbb{E} L'espérance mathématique.

$\mathbb{P}(\cdot)$ Probabilité de l'événement \cdot .

$\pi_X(x)$ La fonction Stop-Loss définie par $\pi_X(x) = \mathbb{E}[(X - x)_+]$.

$m_X(x)$ La fonction de la moyenne des excès.

-
- $\stackrel{\mathcal{L}}{=}$ Égalité de distributions : $X \stackrel{\mathcal{L}}{=} Y \Leftrightarrow \mathbb{P}(X \leq x) = \mathbb{P}(Y \leq x)$ pour tout x .
 RV_ρ Fonctions à variations régulières d'indice ρ .
 ERV_γ Fonctions à variations régulières prolongées d'indice ρ .
 MRV Variations régulières multivariées.
 MDA Domaine d'attraction des valeurs extrêmes (*max-domain of attraction*).
 $X_{k,n}$ Statistique d'ordre k dans un échantillon de taille n .
 \hat{x} Un estimateur statistique de x .
 $\text{MDA}(\Phi_\gamma)$ Domaine d'attraction de Fréchet pour les valeurs extrêmes.
 $\text{MDA}(\Lambda)$ Domaine d'attraction de Gumbel pour les valeurs extrêmes.
 $\text{MDA}(\Psi_\gamma)$ Domaine d'attraction de Weibull pour les valeurs extrêmes.
 $\mathcal{N}(0, 1)$ La loi normale standard.
 GPD La distribution de Pareto généralisée.
 $\mathcal{E}(\lambda)$ La distribution exponentielle de paramètres λ .
 $\text{Pa}(\alpha, \lambda)$ La distribution de Pareto de paramètres α et λ .
 FGM La copule Farlie-Gumbel-Morgenstern. Copulas
 $C_{X,Y}$ La copule représentant la structure de dépendance du couple aléatoire (X, Y) .
 \bar{C} La copule de survie associé à la copule C .
 λ_U Fonction de dépendance extrême supérieure.
 λ_L Fonction de dépendance extrême inférieure.
 \preceq_{st} Le premier ordre de dominance stochastique.
 \preceq_{sl} L'ordre de dominance stochastique Stop-Loss.

Principaux résultats

L'objectif de la première partie de cette thèse est de présenter une approche de construction de mesures multivariées de risque basée sur la propriété d'élicitabilité. L'allocation de capital et la réassurance optimale représentent des champs pratiques d'application de ce type de mesures en assurance. Nous proposons un outil numérique pour la détermination des mesures présentées dans le cas le plus général possible. Nous étudions le comportement asymptotique, qui représente le cas du risque extrême, et nous proposons des estimateurs statistiques pour quelques cas de modèles.

Construction de mesures expectiles multivariées

Le point de départ de ce chapitre est la propriété de l'élicitabilité des mesures de risque univariées. L'objectif étant de proposer des mesures multivariées élicitable par construction, nous utilisons la caractérisation suivante de cette propriété dans un contexte multidimensionnel

Élicitabilité multivariée Une mesure de risque vectorielle $\rho : \mathcal{P} \rightarrow \mathbb{R}^k$ est élicitable s'il existe une fonction score $s : \mathbb{R}^d \times \mathbb{R}^k \rightarrow \mathbb{R}$, telle que

$$\rho(\mathbf{X}) \in \arg \inf_{\mathbf{x} \in U \subset \mathbb{R}^k} \mathbb{E}[s(\mathbf{X}, \mathbf{x})],$$

où \mathcal{P} désigne l'ensemble des vecteurs aléatoires.

Dans le cas des mesures vectorielles de risque construites dans cette thèse $d = k$. Les expectiles étant les seules mesures de risque univariées invariantes par lois à la fois cohérentes et élicitable, l'idée est de construire des extensions multivariées de ces mesures dans l'objectif de conserver leurs propriétés en dimensions supérieure à 1 semble naturelle.

Construction à l'aide des normes

La construction la plus simple sera une généralisation à l'aide des normes. Ainsi, pour tout vecteur aléatoire $\mathbf{X} = (X_1, \dots, X_d)^T$ dans \mathbb{R}^d , et toute norme $\| \cdot \|$ sur \mathbb{R}^d , on définit la fonction de score suivante

$$s_\alpha(\mathbf{X}, \mathbf{x}) = \alpha \| (\mathbf{X} - \mathbf{x})_+ \|^2 + (1 - \alpha) \| (\mathbf{x} - \mathbf{X})_+ \|^2, \quad \forall \mathbf{x} \in \mathbb{R}^d,$$

avec $(\mathbf{X})_+ = ((X_1)_+, \dots, (X_d)_+)$. Sa version normalisée est

$$s_\alpha(\mathbf{X}, \mathbf{x}) = \alpha \left(\| (\mathbf{X} - \mathbf{x})_+ \|^2 - \| \mathbf{X}_+ \|^2 \right) + (1 - \alpha) \left(\| (\mathbf{x} - \mathbf{X})_+ \|^2 - \| \mathbf{X}_- \|^2 \right).$$

Sous certaines conditions sur la norme pour garantir la stricte convexité de la fonction s_α , nous définissons l'expectile multivarié à seuil commun par

$$\mathbf{e}_\alpha(\mathbf{X}) = \arg \min_{\mathbf{x} \in \mathbb{R}^d} \mathbb{E}[\alpha \| (\mathbf{X} - \mathbf{x})_+ \|^2 + (1 - \alpha) \| (\mathbf{x} - \mathbf{X})_+ \|^2].$$

Cette construction donne naissance à une large famille de mesures expectiles multidimensionnelles. Nous définissons ainsi des L_p -expectiles à l'aide des normes L_p

$$\mathbf{e}_\alpha^{L_p}(\mathbf{X}) = \arg \min_{\mathbf{x} \in \mathbb{R}^d} \mathbb{E}[\alpha \| (\mathbf{X} - \mathbf{x})_+ \|_p^2 + (1 - \alpha) \| (\mathbf{x} - \mathbf{X})_+ \|_p^2],$$

et des expectiles matriciels à l'aide d'une matrice semi-définie positive $\Sigma = (\pi_{ij})_{1 \leq i, j \leq d}$

$$\mathbf{e}_\alpha^\Sigma(\mathbf{X}) = \arg \min_{\mathbf{x} \in \mathbb{R}^d} \mathbb{E}[\alpha(\mathbf{X} - \mathbf{x})_+^T \Sigma (\mathbf{X} - \mathbf{x})_+ + (1 - \alpha)(\mathbf{X} - \mathbf{x})_-^T \Sigma (\mathbf{X} - \mathbf{x})_-].$$

La matrice Σ peut être statique à coefficients fixes, ou dynamique dont les coefficients dépendent du vecteur aléatoires \mathbf{X} . Plusieurs exemples sont présentés dans le premier chapitre de cette thèse. Les fonctions score de cette construction sont, sous des conditions que nous donnerons, strictement convexes ce qui garantit l'existence et l'unicité des mesures définies. Le vecteur expectile vérifie dans ce cas la condition d'optimalité de premier ordre, donnée par exemple dans le cas matriciel par

$$\alpha \sum_{i=1}^d \pi_{ki} \mathbb{E}[(X_i - x_i)_+ \mathbb{1}_{\{X_k > x_k\}}] = (1 - \alpha) \sum_{i=1}^d \pi_{ki} \mathbb{E}[(x_i - X_i)_+ \mathbb{1}_{\{x_k > X_k\}}],$$

l'interprétation économique de cette mesure est plus claire dans l'écriture suivante de ce système d'optimalité

$$\frac{\sum_{i=1}^d \pi_{ki} \mathbb{E}[(X_i - x_i)_+ \mathbb{1}_{\{X_k > x_k\}}]}{\sum_{i=1}^d \pi_{ki} \mathbb{E}[(x_i - X_i)_+ \mathbb{1}_{\{x_k > X_k\}}]} = \frac{1 - \alpha}{\alpha}, \quad \forall k \in \{1, \dots, d\},$$

qui la présente comme un vecteur de seuils qui permet de stabiliser un ratio multivarié pertes/gains entre toutes les composantes du vecteur \mathbf{X} . Le principal avantage de cette construction est la simplicité de l'agrégation des risques à l'aide des normes, mais son inconvénient est de se limiter à la modélisation d'un unique type de structure de dépendance. Ainsi, les expectiles matriciels ne sont capables de prendre en compte que la dépendance bivariée entre les couples de composantes de \mathbf{X} .

Construction à l'aide de l'optimalité

L'idée de cette méthode est de partir d'une interprétation économique sous forme d'un système d'optimalité afin de définir une mesure multidimensionnelle. Nous donnons ici deux exemples, le premier qu'on appellera *And-expectile* et qui permet dans le cas d'un portefeuille d'assurance de plusieurs activités, d'obtenir un vecteur de seuils qui stabilise le ratio des gains en cas de solvabilité de toutes les activités et pertes en cas d'insolvabilité globale, pour toutes les activités. Il est défini comme solution du système suivant

$$\frac{\mathbb{E}[(X_k - x_k)_+ \mathbb{1}_{\{X_1 > x_1, \dots, X_d > x_d\}}]}{\mathbb{E}[(X_k - x_k)_- \mathbb{1}_{\{X_1 < x_1, \dots, X_d < x_d\}}]} = \frac{1 - \alpha}{\alpha}, \quad \forall k \in \{1, \dots, d\}.$$

Le deuxième exemple sera appelé *Sum-expectile*, il permet d'obtenir le vecteur de l'allocation qui stabilise pour toutes les activités le ratio des gains en cas de solvabilité du groupe et pertes en cas d'insolvabilité du groupe. Il est défini comme solution du système suivant

$$\frac{\mathbb{E} \left[(X_k - x_k)_+ \mathbb{1}_{\left\{ \sum_{k=1}^d X_k > \sum_{k=1}^d x_k \right\}} \right]}{\mathbb{E} \left[(X_k - x_k)_- \mathbb{1}_{\left\{ \sum_{k=1}^d X_k < \sum_{k=1}^d x_k \right\}} \right]} = \frac{1 - \alpha}{\alpha}, \quad \forall k \in \{1, \dots, d\}.$$

Cette construction permet de répondre directement à une problématique précise, et prend en compte toute la structure de dépendance. Ces principaux inconvénients sont la possibilité de la perte de l'unicité qui donnera naissance à des mesures de risque sous forme d'ensembles, et la difficulté de construire la fonction score de la forme élicitable. Pour pallier ces problèmes tout en répondant au besoin économique, nous avons proposé d'autres formes de ces expectiles par minimisation d'erreurs asymétriques. L'ensemble des mesures de risque définies dans cette partie sont des mesures vectorielles élicitables par construction à l'exception des expectiles *And* et *Sum*. Plusieurs applications de ces mesures en assurance et en finance sont possibles, nous citons notamment l'allocation du capital, la détermination des contrats optimaux de réassurance, et le contrôle du risque systémique.

Propriétés de cohérence

Nous nous focalisons dans la suite sur les expectiles matriciels. Ces expectiles vérifient, en plus de l'élicitabilité, et sous des conditions sur les coefficients de la matrice de construction, un ensemble de propriétés de cohérence. Nous présentons ici quelques-unes des plus importantes

– *Homogénéité positive* : Pour tout vecteur aléatoire $\mathbf{X} = (X_1, \dots, X_d)$ sur \mathbb{R}^d , et toute constante réelle positive a ,

$$\mathbf{e}_\alpha(a\mathbf{X}) = a\mathbf{e}_\alpha(\mathbf{X}).$$

– *Invariance par translation* : Pour tout vecteur aléatoire $\mathbf{X} = (X_1, \dots, X_d)$ sur \mathbb{R}^d et tout vecteur $\mathbf{a} = (a_1, \dots, a_d)$ dans \mathbb{R}^d ,

$$\mathbf{e}_\alpha(\mathbf{X} + \mathbf{a}) = \mathbf{e}_\alpha(\mathbf{X}) + \mathbf{a}.$$

– *Invariance par loi* : Pour tous vecteurs aléatoires $X = (X_1, \dots, X_d)$ et $Y = (Y_1, \dots, Y_d)$ dans \mathbb{R}^d tels que pour tout $(i, j) \in \{1, \dots, d\}^2$, on a $(X_i, X_j) \sim (Y_i, Y_j)$, alors

$$\forall \alpha \in [0, 1], \quad \mathbf{e}_\alpha(X) = \mathbf{e}_\alpha(Y).$$

– *Symétrie par rapport au seuil* : Pour tout vecteur aléatoire $\mathbf{X} = (X_1, \dots, X_d)$ sur \mathbb{R}^d , et pour tout seuil $\alpha \in [0, 1]$,

$$\mathbf{e}_\alpha(-\mathbf{X}) = -\mathbf{e}_{1-\alpha}(\mathbf{X}).$$

– *Monotonie interne* : Pour tout vecteur aléatoire $\mathbf{X} = (X_1, \dots, X_d)^T$ dans \mathbb{R}^d , s'il existe un couple $(i, j) \in \{1, \dots, d\}^2$ tel que $X_i \leq X_j$, p.s et les couples (X_i, X_k) et (X_j, X_k) ont la même copule pour tout $k \in \{1, \dots, d\} \setminus \{i, j\}$, alors sous des conditions supplémentaires sur les coefficients de la matrice de construction

$$\mathbf{e}_\alpha^i(\mathbf{X}) \leq \mathbf{e}_\alpha^j(\mathbf{X}).$$

Ces mesures sont entièrement déterminées par la condition d'optimalité du premier ordre. Une approximation des vecteurs expectiles sera donc possible à l'aide de l'application de l'algorithme Robbins-Monro. Cette procédure est expliquée dans la quatrième section du premier chapitre. Des illustrations numériques sont également présentées. L'analyse numérique de la convergence permet de remarquer la faible performance de cet outil pour les niveaux asymptotiques du seuil.

Expectile asymptotique

Soit $(X_1, \dots, X_d)^T$ un vecteur aléatoire dans \mathbb{R}^d . Le support de chaque variable aléatoire X_i est noté $[x_i^l, x_i^r]$, avec $x_i^l \in \mathbb{R} \cup \{-\infty\}$ et $x_i^r \in \mathbb{R} \cup \{+\infty\}$. On note \mathbf{X}_F le vecteur $(x_F^1, \dots, x_F^d)^T$, et \mathbf{X}_I le vecteur $(x_I^1, \dots, x_I^d)^T$. Nous considérons le cas d'un expectile matriciel associé à une matrice statique semi-définie positive à coefficients positifs et strictement positifs sur la diagonale. Nous avons démontré dans ce cas que pour tout vecteur aléatoire $\mathbf{X} = (X_1, \dots, X_d)^T$ dans \mathbb{R}^d , tel que ses distributions marginales sont toutes continues sur leurs supports respectifs et $\mathbb{E}[|X_i|] < +\infty$

$$\lim_{\alpha \rightarrow 1} \mathbf{e}_\alpha(\mathbf{X}) = \mathbf{X}_F, \quad \text{et} \quad \lim_{\alpha \rightarrow 0} \mathbf{e}_\alpha(\mathbf{X}) = \mathbf{X}_I,$$

pour tout $i \in \{1, \dots, d\}$. Si de plus tous les supports sont infinis, ces limites vérifient l'équivalence, c.-à-d., pour tout $k \in \{1, \dots, d\}$

$$\lim_{x_k \rightarrow +\infty} \alpha(x_1, \dots, x_d) = 1, \quad \text{et} \quad \lim_{x_k \rightarrow -\infty} \alpha(x_1, \dots, x_d) = 0,$$

avec

$$\alpha(x_1, \dots, x_d) = \frac{\sum_{i=1}^d \pi_{ki} \mathbb{E}[(x_i - X_i)_+ \mathbb{1}_{\{x_k > X_k\}}]}{\sum_{i=1}^d \pi_{ki} \left(\mathbb{E}[(X_i - x_i)_+ \mathbb{1}_{\{X_k > x_k\}}] + \mathbb{E}[(x_i - X_i)_+ \mathbb{1}_{\{x_k > X_k\}}] \right)},$$

où $\alpha(\cdot)$ est la fonction inverse de $\alpha \rightarrow \mathbf{e}_\alpha(\mathbf{X})$.

Ce comportement de convergence vers le vecteur des points terminaux des supports des distributions marginales ne fournit pas de précisions sur la composition du vecteur expectile au voisinage du niveau

asymptotique. Afin de répondre à ce besoin, le chapitre suivant sera consacré à une analyse plus détaillée du comportement asymptotique des expectiles multivariées en fonction de la nature des queues marginales.

Une série d'illustrations numériques pour des expectiles matriciels, dans quelques cas de modèles, est présentée dans l'annexe du chapitre 1.

Comportement asymptotique des expectiles multivariés

La convergence de l'algorithme de Robbins-Monro (1951) pour l'approximation stochastique du vecteur expectiles n'est pas satisfaisante pour les niveaux asymptotiques du seuil. Pour pallier ce problème nous consacrons ce chapitre à l'étude du comportement asymptotique des expectiles multivariés. Nous considérons le cas des L_1 -expectiles. Une grande partie des résultats obtenus est généralisable pour les expectiles matriciels construits à l'aide de matrices semi-définies positives à coefficients positifs. L'objectif principal de ce chapitre sera de caractériser le comportement asymptotique des L_1 -expectiles en fonction du domaine d'attraction des valeurs extrêmes des distributions marginales, afin d'obtenir des estimations des composantes du vecteur expectile.

Nous utilisons dans le cadre de cette étude asymptotique le concept des variations régulières et des variations régulières multivariées. L'importance de cette notion réside essentiellement dans sa relation étroite avec la caractérisation des domaines d'attraction des valeurs extrêmes.

Nous considérons dans ce chapitre le modèle à queues équivalentes. Pour tout vecteur aléatoire $\mathbf{X} = (X_1, \dots, X_d)^T$, ce modèle se traduit par deux hypothèses :

H1 X_1 est à distribution à queue à variations régulières

$$\bar{F}_{X_1} \in \text{RV}_{-\theta},$$

avec $\theta > 1$.

H2 Les queues des $X_i, i = 1, \dots, d$ sont équivalentes. C-à-d, pour tout $i \in \{2, \dots, d\}$, il existe une constante strictement positive c_i telle que

$$\lim_{x \rightarrow +\infty} \frac{\bar{F}_{X_i}(x)}{\bar{F}_{X_1}(x)} = c_i.$$

L'étude du comportement asymptotique des expectiles sous ces hypothèses permet d'établir des relations entre les composantes du vecteur expectiles et leur comparaison avec les VaR univariées.

Dans un contexte de variations régulières multivariées, et sous les hypothèses H1 et H2, la mesure spectrale du vecteur aléatoire \mathbf{X} est complètement caractérisée par la fonction de la dépendance extrême supérieure. Nous nous focalisons sur les distributions du domaine d'attraction de Fréchet des valeurs extrêmes. Nous étudions le cas général, et nous déterminons le comportement de l'expectile dans les cas de l'indépendance asymptotique, la comonotonie et la dominance d'une composante. Nous traitons dans la suite le modèle à mélange commun. Ce chapitre se termine par des estimations statistiques dans les cas de l'indépendance asymptotique et de la dépendance parfaite. Nous nous basons sur les outils des statistiques des extrêmes pour proposer des estimateurs de l'expectile asymptotique.

Modèle Fréchet à queues équivalentes

Dans ce cas, les distributions marginales appartiennent au domaine d'attraction des valeurs extrêmes de Fréchet $F_{X_1} \in \text{MDA}(\Phi_\theta)$, et l'indice des variations régulières est supposé supérieur strictement à 1 ($\theta > 1$) pour garantir $\mathbb{E}[\mathbf{X}] < +\infty$. Nous avons démontré que sous l'hypothèse de l'équivalence des queues, les composantes de l'expectile multivarié asymptotique sont comparables et vérifient

$$0 < \lim_{\alpha \rightarrow 1} \frac{\mathbf{e}_\alpha^i(\mathbf{X})}{\mathbf{e}_\alpha^1(\mathbf{X})} \leq \overline{\lim}_{\alpha \rightarrow 1} \frac{\mathbf{e}_\alpha^i(\mathbf{X})}{\mathbf{e}_\alpha^1(\mathbf{X})} < +\infty, \forall i \in \{2, \dots, d\},$$

et

$$0 < \lim_{\alpha \rightarrow 1} \bar{F}_{X_i}(\mathbf{e}_\alpha^i(\mathbf{X})) \leq \overline{\lim}_{\alpha \rightarrow 1} \frac{1 - \alpha}{\bar{F}_{X_i}(\mathbf{e}_\alpha^i(\mathbf{X}))} < +\infty, \forall i \in \{2, \dots, d\}.$$

Si de plus les limites

$$\lim_{\alpha \rightarrow 1} \frac{1 - \alpha}{\bar{F}_{X_1}(\mathbf{e}_\alpha^1(\mathbf{X}))} = \eta, \text{ et } \lim_{\alpha \rightarrow 1} \frac{\mathbf{e}_\alpha^i(\mathbf{X})}{\mathbf{e}_\alpha^1(\mathbf{X})} = \beta_i, \forall i \in \{2, \dots, d\},$$

existent alors elles sont déterminées par la solution unique du système d'équations

$$\frac{1}{\theta - 1} - \eta \frac{(\beta_k)^\theta}{c_k} = - \sum_{i=1, i \neq k}^d \left(\int_{\frac{\beta_i}{\beta_k}}^{+\infty} \lambda_U^{ik} \left(\frac{c_i}{c_k} t^{-\theta}, 1 \right) dt - \eta \frac{\beta_k^{\theta-1}}{c_k} \beta_i \right), \forall k \in \{1, \dots, d\}.$$

La résolution de ce système est équivalente à celle du système d'équations suivant

$$\sum_{i=1}^d \int_{\frac{\beta_i}{\beta_k}}^{+\infty} \lambda_U^{ik} \left(c_i t^{-\theta}, c_k \beta_k^{-\theta} \right) dt = \sum_{i=1}^d \int_{\beta_i}^{+\infty} \lambda_U^{i1} \left(c_i t^{-\theta}, 1 \right) dt, \forall k \in \{2, \dots, d\},$$

qui permet de visualiser clairement l'impact de la nature des dépendances asymptotiques bivariées entre les composantes du vecteur \mathbf{X} sur les limites β_i .

Si le vecteur $\mathbf{X} = (X_1, \dots, X_d)$ est comonotone, alors

$$\lim_{\alpha \rightarrow 1} \frac{1 - \alpha}{\bar{F}_{X_k}(\mathbf{e}_\alpha^k(\mathbf{X}))} = \frac{1}{\theta - 1} \text{ et } \beta_k = c_k^{1/\theta}, \forall k \in \{1, \dots, d\}.$$

Dans le cas de l'indépendance asymptotique ($\lambda_U = 0$) entre les variables aléatoires $X_i, i = 1, \dots, d$

$$\lim_{\alpha \rightarrow 1} \frac{\mathbf{e}_\alpha^k(\mathbf{X})}{\mathbf{e}_\alpha^1(\mathbf{X})} = c_k^{\frac{1}{\theta-1}} \text{ et } \lim_{\alpha \rightarrow 1} \frac{1 - \alpha}{\bar{F}_{X_k}(\mathbf{e}_\alpha^k(\mathbf{X}))} = \frac{c_k^{\frac{1}{\theta-1}}}{(\theta - 1) \left(1 + \sum_{j=2}^d c_j^{\frac{1}{\theta-1}} \right)},$$

pour tout $k \in \{1, \dots, d\}$. Ce résultat permet de comparer la composante expectile correspondante à chaque risque et sa VaR du même seuil

$$\mathbf{e}_\alpha^k(\mathbf{X}) \underset{\alpha \rightarrow 1}{\sim} \text{VaR}_\alpha(X_k) (\theta - 1)^{-\frac{1}{\theta}} \left(\frac{1 + \sum_{i=2}^d c_i^{\frac{1}{\theta-1}}}{c_i^{\frac{1}{\theta-1}}} \right)^{-\frac{1}{\theta}}, \quad k \in \{1, \dots, d\}.$$

Modèle Fréchet à queue dominante

Nous avons examiné aussi le cas de la dominance de la queue d'un risque. Cette hypothèse se traduit par des coefficients d'équivalence nuls $c_i = 0$ pour tout $i \in \{2, \dots, d\}$

$$\lim_{x \uparrow +\infty} \frac{\bar{F}_{X_i}(x)}{\bar{F}_{X_1}(x)} = 0, \quad \forall i \in \{2, \dots, d\}.$$

Nous avons montré que sous l'hypothèse de la régularité des variations de tous les risques, les composantes de l'expectile asymptotique héritent la propriété de la dominance

$$\beta_i = \lim_{\alpha \uparrow 1} \frac{\mathbf{e}_\alpha^i(\mathbf{X})}{\mathbf{e}_\alpha^1(\mathbf{X})} = 0, \quad \lim_{\alpha \uparrow 1} \frac{1 - \alpha}{\bar{F}_{X_k}(\mathbf{e}_\alpha^i(\mathbf{X}))} = 0, \quad \forall i \in \{2, \dots, d\},$$

et

$$\lim_{\alpha \uparrow 1} \frac{1 - \alpha}{\bar{F}_{X_1}(\mathbf{e}_\alpha^1(\mathbf{X}))} = \frac{1}{\theta - 1}.$$

Modèle à mélange commun

Nous avons dans une deuxième étape considéré le modèle à mélange commun qui consiste à construire un vecteur aléatoire \mathbf{X} par une variable de mélange R

$$\mathbf{X} = (X_1, \dots, X_d)^T = (RT_1, \dots, RT_d)^T,$$

où le vecteur aléatoire (T_1, \dots, T_d) est composé de distributions marginales à supports positifs et indépendantes de R . La variable aléatoire de mélange R est supposée positive et possède une queue droite à variation régulière $\bar{F}_R \in \text{RV}_{-\theta}$. Nous avons démontré que s'il existe $0 < \delta < 1$ telle que les variables aléatoires $T_i, i = 1, \dots, d$ admettent toutes des moments d'ordre $\theta + \delta$, strictement positifs, $0 < \mathbb{E}[T_i^{\theta+\delta}] < +\infty$, alors le vecteur aléatoire $\mathbf{X} = (X_1, \dots, X_d)^T = (RT_1, \dots, RT_d)^T$ satisfait les hypothèses H1 et H2. Cela découle du fait que

$$\lim_{x \rightarrow +\infty} \frac{\bar{F}_{X_i}(x)}{\bar{F}_R(x)} = \mathbb{E}[T_i^\theta],$$

pour tout $i \in \{1, \dots, d\}$. Nous avons démontré ensuite que sous les mêmes conditions, les composantes de l'expectile multivarié asymptotique vérifient, sous l'hypothèse de l'existence des limites, le système d'équations suivant

$$\eta \beta_k^{\theta-1} \sum_{i=1}^d \beta_i = \frac{[T_k^\theta]}{\theta - 1} + \sum_{i=1, i \neq k}^d A_{ik}, \quad \forall k \in \{1, \dots, d\},$$

avec

$$A_{ik} = \mathbb{E} \left[\left(\frac{T_i}{T_k} - \frac{\beta_i}{\beta_k} \right)_+ T_k^\theta \right] + \frac{\left(\frac{\beta_i}{\beta_k} \right)^{-\theta+1} \mathbb{E} \left[T_i^\theta \bar{F}_{T_k} \left(\frac{\beta_k}{\beta_i} T_i \right) \right] + \mathbb{E} \left[T_k^{\theta-1} \left(T_i - T_k \frac{\beta_i}{\beta_k} \right)_+ + \frac{\beta_i}{\beta_k} T_k^\theta \bar{F}_{T_i} \left(\frac{\beta_i}{\beta_k} T_k \right) \right]}{\theta - 1}.$$

Estimation statistique

Pour clôturer ce chapitre, nous avons construit des estimateurs statistiques des expectiles asymptotiques à l'aide des statistiques des extrêmes. L'estimation des intégrales des fonctions de dépendance des queues étant compliquée, nous nous sommes limités aux cas de l'indépendance asymptotique et de la dépendance parfaite. À l'aide des estimateurs de l'indice des valeurs extrêmes présents dans la littérature nous avons proposé d'estimer les coefficients d'équivalence des queues $c_i, i = 2, \dots, d$ par

$$\hat{c}_i = \left(\frac{X_{i,n-k+1,n}}{X_{1,n-k+1,n}} \right)^{\frac{1}{\hat{\gamma}(k)}},$$

où $\hat{\gamma}(k)$ est un estimateur de l'indice des valeurs extrêmes, comme, par exemple, l'estimateur de Hill (1975) [62] défini par

$$\hat{\gamma}_{\text{Hill}} = \frac{1}{k(n)} \sum_{\ell=1}^{k(n)} \log(X_{n-\ell+1,n}) - \log(X_{n-k(n)+1,n}),$$

construit à l'aide des k plus grandes observations où k est très grand ($k \rightarrow +\infty$) tel que $k/n \rightarrow 0$. Sous les hypothèses du modèle Fréchet à variations régulières multivariées, dans le cas de l'indépendance asymptotique ($\lambda_U = 0$) deux à deux entre les composantes du vecteur aléatoire $\mathbf{X} = (X_1, \dots, X_d)^T$, le L_1 -expectile peut être estimé par

$$\hat{\mathbf{e}}_\alpha^\perp(\mathbf{X}) = X_{1,n-k(n)+1,n} \left(\frac{k(n)}{(1-\alpha)n} \right)^{\hat{\gamma}} \left(\frac{\hat{\gamma}}{1-\hat{\gamma}} \right)^{\hat{\gamma}} \left(\frac{1}{1 + \sum_{\ell=2}^d \hat{c}_\ell^{\frac{\hat{\gamma}}{1-\hat{\gamma}}}} \right)^{\hat{\gamma}} (1, \hat{c}_2^{\frac{\hat{\gamma}}{1-\hat{\gamma}}}, \dots, \hat{c}_d^{\frac{\hat{\gamma}}{1-\hat{\gamma}}})^T.$$

Dans le cas d'un vecteur de risque comonotone, nous avons proposé l'estimation du L_1 -expectile par

$$\hat{\mathbf{e}}_\alpha^+(\mathbf{X}) = X_{1,n-k(n)+1,n} \left(\frac{k(n)}{(1-\alpha)n} \right)^{\hat{\gamma}} \left(\frac{\hat{\gamma}}{1-\hat{\gamma}} \right)^{\hat{\gamma}} (1, \hat{c}_2^{\hat{\gamma}}, \dots, \hat{c}_d^{\hat{\gamma}})^T.$$

La convergence des estimateurs découle de celle des estimateurs utilisés dans la construction.

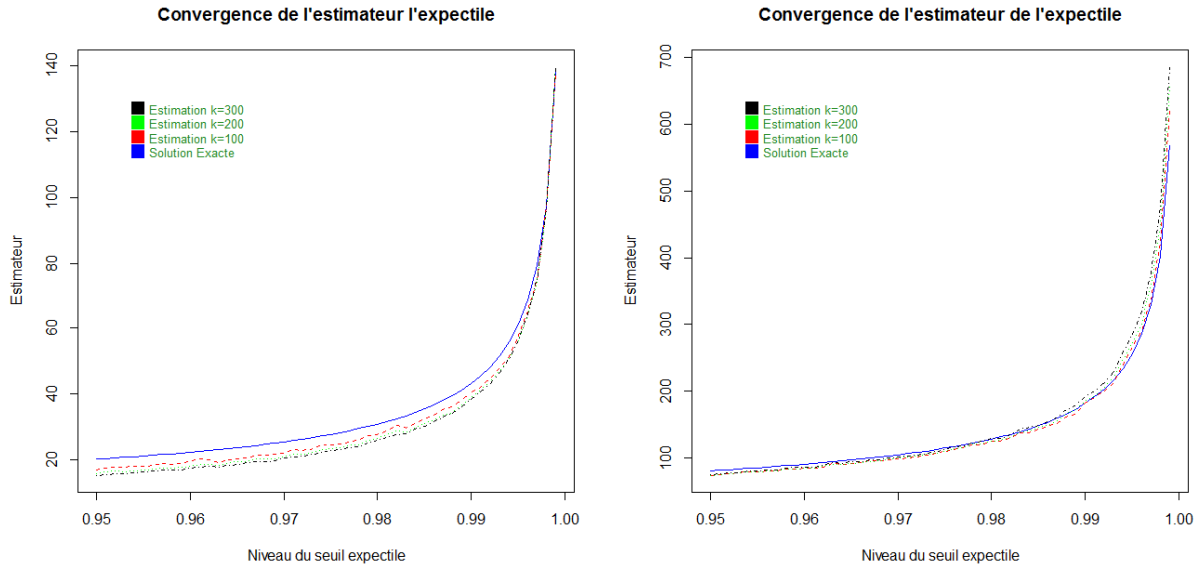


FIGURE 1 – Estimateur expectile multivarié, Cas Fréchet, Indépendance Asymptotique. Cas Pareto ($X_i \sim Par(2, i * 10)$) Convergence de l'estimateur de l'expectile (à gauche X_1 , à droite X_2).

Nous avons aussi étudié les cas des domaines d'attraction de Gumbell et Weibull dans un travail en cours de finalisation et les premiers résultats démontrent que le comportement des expectiles asymptotiques hérite aussi de l'équivalence des queues des distributions marginales. Ils restent comparables à la VaR, ce qui signifie qu'il s'agit bien de mesures du risque extrême, mais le comportement obtenu reste différent pour les limites des rapports des composantes par rapport à celui obtenu dans le cas Fréchet. La nature des queues est finalement ce qui détermine le comportement asymptotique des mesures expectiles multivariées.

Allocation du capital par minimisation du risque

Les normes Solvabilité 2 et le dispositif ORSA *Own Risk and Solvency Assessment* offrent une liberté importante sur le choix des méthodes d'évaluation des risques, et par suite de l'allocation du capital global entre les différentes composantes risquées. Cette liberté engendre la possibilité pour tout assureur de développer sa propre méthode d'allocation du capital économique en fonction de sa vision du risque, et aussi de sa stratégie de développement. C'est dans cette optique que nous allons consacrer la deuxième partie de cette thèse à l'analyse d'une nouvelle méthode d'allocation, basée sur une stratégie de minimisation d'indicateurs multivariés de risque, et dont l'idée était introduite dans Cénac et al. (2014) [23]. Cette approche a pour principal objectif l'intégration de la gestion du risque dans le processus de l'allocation, et la prise en compte de la dépendance entre les différentes activités risquées.

Dans une optique de minimisation de risque, nous définissons l'allocation optimale comme une recherche d'un point dans un simplexe par minimisation d'un indicateur de risque multivarié. Cette méthode ne découle pas du choix d'une mesure de risque univariée, elle est fondée sur une approche multivariée et constitue, contrairement aux méthodes existantes dans la littérature, un outil de gestion de risques et pas seulement une mesure de son poids marginal.

On désigne par le risque X_k , la perte de la $k^{\text{ème}}$ ligne d'affaires durant une période. Il s'agit d'une variable aléatoire positive dans notre contexte. Dans le cas de plusieurs périodes, X_k^n représentera les pertes de la $k^{\text{ème}}$ branche de la compagnie, pendant la $n^{\text{ème}}$ période. On peut bien avoir des dépendances vectorielles entre les $(X_i^n)_i$ (par rapport à k) et/ou temporelles entre les $(X_k^i)_i$ (par rapport à n). On note aussi $Y_k^j = \sum_{l=1}^j X_k^l$ la perte agrégée de la $k^{\text{ème}}$ branche sur j périodes.

Le capital initial d'un groupe d'assurances, ou le capital économique à allouer est noté u . Nous supposons qu'il a été calculé dans une opération d'agrégation de risques.

On désigne par $\mathcal{U}_u^d = \{v = (v_1, \dots, v_d) \in [0, u]^d, \sum_{i=1}^d v_i = u\}$ l'ensemble des allocations possibles du capital u entre les différentes branches. Une allocation peut donc être définie comme une application $\mathbb{R}^+ \rightarrow (\mathbb{R}^+)^d$. Pour tout $i \in \{1, \dots, d\}$ on définit $\alpha_i = \frac{u_i}{u}$ le poids de la participation de la $i^{\text{ème}}$ branche dans le capital u , les α_i vérifient alors $\sum_{i=1}^d \alpha_i = 1$ si $(u_1, \dots, u_d) \in \mathcal{U}_u^d$. De la même façon que \mathcal{U}_u^d , nous définissons $\mathbb{1}_u^d = \{\alpha = (\alpha_1, \dots, \alpha_d) \in [0, 1]^d, \sum_{i=1}^d \alpha_i = 1\}$, l'ensemble des poids d'allocations possibles $\alpha_i = u_i/u$. Une bijection existe donc entre $\mathbb{1}_u^d$ et \mathcal{U}_u^d .

Pour une allocation $(u_1, \dots, u_d) \in \mathcal{U}_u^d$, on définit la réserve de la $k^{\text{ème}}$ ligne d'affaires à l'issue d'une période par $R_k = u_k - X_k$, où u_k désigne la part du capital u allouée à la branche X_k .

Dans le cas de plusieurs périodes, cette réserve sera $R_k^j = u_k - Y_k^j$. La $k^{\text{ème}}$ branche est *insolvable* au temps j si $R_k^j < 0$.

La somme agrégée des différents risques est notée $S = \sum_{i=1}^d X_i$. On définit aussi cette somme privée du risque X_i qui sera notée $S^{-i} = \sum_{j=1; j \neq i}^d X_j$ pour tout $i \in \{1, \dots, d\}$.

Nous avons défini une *allocation optimale* par une recherche d'un minimiseur d'un indicateur de risque multivarié.

Allocation optimale Soit \mathbf{X} un vecteur aléatoire, de composantes positives, dans $(\mathbb{R}^+)^d$, $u \in \mathbb{R}^+$ et $\mathcal{K}_{\mathbf{X}} : \mathcal{U}_u^d \rightarrow \mathbb{R}^+$ un indicateur de risque multivarié associé à \mathbf{X} et u . Une allocation optimale du capital u entre les composantes du vecteur de risque \mathbf{X} est définie comme suit :

$$(u_1, \dots, u_d) \in \underset{(v_1, \dots, v_d) \in \mathcal{U}_u^d}{\arg \inf} \{ \mathcal{K}_{\mathbf{X}}(v_1, \dots, v_d) \}.$$

Nous avons ensuite formulé une axiomatisation de la cohérence d'une méthode d'allocation dans un contexte général. Notre définition de la cohérence est applicable à toute technique d'allocation d'un capital u , indépendamment des choix d'agrégation des risques. Elle généralise ainsi la définition introduite par Denault (2001) [40].

Allocation cohérente Une allocation d'un capital initial $u \in \mathbb{R}^+$, notée $(u_1, \dots, u_d) = A_{X_1, \dots, X_d}(u)$ est considérée cohérente si elle satisfait les propriétés suivantes :

1. *Allocation complète* : L'ensemble du capital $u \in \mathbb{R}^+$ est alloué entre les différentes branches

$$\sum_{i=1}^d u_i = u.$$

2. *Symétrie* : Si la distribution jointe du vecteur aléatoire (X_1, \dots, X_d) est stable par permutation de deux risques X_i et X_j , alors la méthode d'allocation est aussi stable par la même permutation, c.-à-d. les deux branches X_i et X_j ont la même contribution au capital du groupe. Si,

$$\begin{aligned} & (X_1, \dots, X_{i-1}, X_i, X_{i+1}, \dots, X_{j-1}, X_j, X_{j+1}, \dots, X_d) \\ & \quad \underline{\underline{\mathcal{L}}} \\ & (X_1, \dots, X_{i-1}, X_j, X_{i+1}, \dots, X_{j-1}, X_i, X_{j+1}, \dots, X_d), \end{aligned}$$

alors $u_i = u_j$.

3. *Neutralité* : En présence d'un risque déterministe $X = c$, avec c une constante réelle positive $c \in [0, u]$, alors :

$$A_{X, X_1, \dots, X_d}(u) = (c, A_{X_1, \dots, X_d}(u - c)).$$

4. *Sous-additivité* : $\forall M \subseteq \{1, \dots, d\}$, on note $(u^*, u_1^*, \dots, u_r^*) = A_{\sum_{i \in M} X_i, X_{j \in \{1, \dots, d\} \setminus M}}(u)$, avec $r = d - \text{card}(M)$ et $(u_1, \dots, u_d) = A_{X_1, \dots, X_d}(u)$ alors

$$u^* \leq \sum_{i \in M} u_i.$$

5. *Additivité comonotone* : Pour $r \leq d$ risques comonotones,

$$A_{X_{i \in \{1, \dots, d\} \setminus CR}, \sum_{k \in CR} X_k}(u) = (u_{i \in \{1, \dots, d\} \setminus CR}, \sum_{k \in CR} u_k),$$

où $(u_1, \dots, u_d) = A_{X_1, \dots, X_d}(u)$ est l'allocation du capital u entre les d risques (X_1, \dots, X_d) et CR désigne l'ensemble des indices des r risques comonotones.

Chacun des cinq axiomes de cette définition traduit une propriété économiquement désirable pour une allocation du capital dans le sens actuariel. D'autres propriétés désirables ont été définies aussi dans un contexte général, sans être considérées comme des axiomes de cohérence,

– *Invariance par translation* : L'allocation est invariante par translation si elle vérifie pour tout vecteur $(a_1, \dots, a_d) \in \mathbb{R}^d$

$$A_{X_1 - a_1, \dots, X_d - a_d}(u) = A_{X_1, \dots, X_d} \left(u + \sum_{k=1}^d a_k \right) - (a_1, \dots, a_d).$$

– *Homogénéité positive* : L'allocation est positivement homogène si pour toute $\alpha \in \mathbb{R}^+$, elle vérifie

$$A_{\alpha X_1, \dots, \alpha X_d}(\alpha u) = \alpha A_{X_1, \dots, X_d}(u).$$

– *Continuité* : Une allocation de capital peut être considérée comme continue si pour tout $i \in \{1, \dots, d\}$

$$\lim_{\epsilon \rightarrow 0} A_{X_1, \dots, (1+\epsilon)X_i, \dots, X_d}(u) = A_{X_1, \dots, X_i, \dots, X_d}(u).$$

– *Monotonie* : Une méthode d'allocation est monotone si pour $(i, j) \in \{1, \dots, d\}^2$

$$X_i \leq_{st} X_j \Rightarrow u_i \leq u_j.$$

Nous nous focalisons après ces définitions sur l'allocation par minimisation des indicateurs multivariés de risque définis par Cénac et al. (2012) [24]. Deux indicateurs sont proposés. Leurs formes discrètes pour n périodes et d risques sont données par

– Indicateur I :

$$I(u_1, \dots, u_d) = \sum_{k=1}^d \mathbb{E} \left(\sum_{p=1}^n g_k(R_k^p) \mathbb{1}_{\{R_k^p < 0\}} \mathbb{1}_{\{\sum_{j=1}^d R_j^p > 0\}} \right),$$

– Indicateur J :

$$J(u_1, \dots, u_d) = \sum_{k=1}^d \mathbb{E} \left(\sum_{p=1}^n g_k(R_k^p) \mathbb{1}_{\{R_k^p < 0\}} \mathbb{1}_{\{\sum_{j=1}^d R_j^p < 0\}} \right),$$

avec g_k , $k \in \{1, \dots, d\}$ des fonctions de pénalités. $g_k : \mathbb{R}^- \rightarrow \mathbb{R}^+$ sont des fonctions convexes et C^1 , qui vérifient $g_k(0) = 0$, $g_k(x) \geq 0$, $k = 1, \dots, d$, et sont décroissantes sur \mathbb{R}^- . Ces fonctions représentent des coûts que chaque branche doit payer dans le cas de son insolvabilité, quand le groupe est solvable dans le cas de l'indicateur I , et quand il est insolvable dans le cas de l'indicateur J .

Nous démontrons dans le troisième chapitre de cette thèse que l'allocation optimale par minimisation des versions mono-périodiques des indicateurs I et J vérifient un ensemble des propriétés de cohérence. Sous des conditions sur les supports, et pour des fonctions de pénalités égales aux montants de la ruine $g_k(x) = |x|$ pour tout $k \in \{1, \dots, d\}$, l'allocation par minimisation des indicateurs I et J est une allocation totale, neutre, symétrique, invariante par translation, homogène positivement, monotone, continue et vérifie la propriété d'additivité comonotone.

La méthode d'allocation du capital proposée, par minimisation d'indicateurs multivariés de risque, semble plus adéquate avec la nature multivariée de la problématique de l'allocation et son objectif principal de minimisation du risque.

Impact de la dépendance sur l'allocation optimale

Dans le deuxième chapitre de la deuxième partie de cette thèse, nous analysons les propriétés probabilistes de l'allocation par minimisation des indicateurs multivariés de risque. Notre objectif étant d'examiner la composition de l'allocation en fonction de la structure de dépendance et des distributions marginales des risques. Pour cela, nous distinguons quatre cas, l'indépendance, l'indépendance asymptotique, la dépendance asymptotique, et la dépendance parfaite.

Cas de l'indépendance

Dans le cas de l'indépendance, seule la nature des distributions marginales détermine la composition de l'allocation optimale. Nous étudions deux cas de modèles, exponentiel et sous-exponentiel. Le choix de ces deux cas découle de l'importante différence entre leurs queues des distributions, cette différence aura un impact sur le comportement asymptotique de l'allocation optimale.

Modèle exponentiel

Nous considérons des risques représentés par des variables aléatoire X_1, X_2, \dots, X_d indépendantes de distributions exponentielles, de paramètres $0 < \beta_1 < \beta_2 < \dots < \beta_d$ respectivement.

Dans le cas particulier du même paramètre $\beta_1 = \beta_2 = \dots = \beta_d$ la composition de l'allocation optimale est donnée par $u_1 = \dots = u_d = u/d$.

Nous obtenons, dans le cas général, des systèmes d'équations explicites vérifiées par les allocations optimales par minimisation des indicateurs I et J . Ainsi, pour l'allocation optimale par minimisation de la version mono-périodique de l'indicateur multivarié de risque I , dans le cas des fonctions de pénalité $g_k(x) = |x|$ pour tout $k \in \{1, \dots, d\}$, la composition de l'allocation est déterminée par l'unique solution en $(\alpha_1, \dots, \alpha_d)$, dans le simplexe \mathcal{U}_u^d , du système d'équations suivant

$$h(\beta_i \alpha_i) - h(\beta_j \alpha_j) - \sum_{\ell=1}^d A_\ell h(\beta_\ell) [h(\alpha_i \cdot (\beta_i - \beta_\ell)) - h(\alpha_j \cdot (\beta_j - \beta_\ell))] = 0, \forall (i, j) \in \{1, 2, \dots, d\}^2,$$

avec h la fonction définie par $h(x) = \exp(-u \cdot x)$, et $A_\ell = \prod_{j=1, j \neq \ell}^d \frac{\beta_j}{\beta_j - \beta_\ell}$, pour $\ell = 1, \dots, d$.

Quand le capital u tend vers l'infini, l'allocation optimale asymptotique vérifie

$$\lim_{u \rightarrow \infty} \left(\frac{u_1}{u}, \frac{u_2}{u}, \dots, \frac{u_d}{u} \right) = \left(\frac{1}{\beta_i}, \frac{1}{\beta_j}, \dots, \frac{1}{\beta_d} \right)_{i=1,2,\dots,d}.$$

Pour l'allocation optimale par minimisation de la version mono-périodique de l'indicateur multivarié de risque J , dans le cas des fonctions de pénalité $g_k(x) = |x|$ pour tout $k \in \{1, \dots, d\}$, la composition de l'allocation est déterminée par l'unique solution dans le simplexe \mathcal{U}_u^d du système d'équations suivant

$$\forall (i, j) \in \{1, 2, \dots, d\}^2, \quad \sum_{\ell=1}^d A_\ell h(\beta_\ell) [h(\alpha_i \cdot (\beta_i - \beta_\ell)) - h(\alpha_j \cdot (\beta_j - \beta_\ell))] = 0,$$

et l'allocation optimale asymptotique vérifie

$$\lim_{u \rightarrow \infty} \frac{u_1}{u} = 1 \text{ and } \lim_{u \rightarrow \infty} \frac{u_j}{u} = 0 \quad \forall j \in \{2, 3, \dots, d\}.$$

Modèle sous-exponentiel

La famille des distributions sous-exponentielles est constituée de toutes les distributions de support positif dont la fonction de répartition satisfait la relation suivante :

$$\frac{F^{*2}(x)}{F(x)} \xrightarrow{x \rightarrow +\infty} 2,$$

où $\overline{F^{*2}}$ représente le produit de convolution de \overline{F} .

Dans le cas général, on ne peut pas obtenir d'expressions explicites pour la composition de l'allocation optimale. Pour le modèle sous-exponentiel, nous pouvons au moins analyser son comportement asymptotique. Nous avons ainsi démontré, dans le cas de l'allocation optimale par minimisation de l'indicateur I , que pour (X_1, X_2, \dots, X_d) des variables aléatoires positives et continues de support \mathbb{R}^+ qui vérifient pour tout $(i, j) \in \{1, 2, \dots, d\}^2$

1. $\overline{F}_{X_i}(x) \stackrel{x \rightarrow +\infty}{\asymp} O(\overline{F}_{X_j}(x))$,
2. $\overline{F}_{X_i}(s) \stackrel{s \rightarrow +\infty}{\asymp} o(\overline{F}_{X_i}(t))$, si $t = o(s)$,

il existe deux constantes $\kappa_1 > 0$ et $\kappa_2 < 1$ telles que

$$\kappa_1 \leq \frac{u_\ell}{u} \leq \kappa_2 \quad \forall \ell \in \{1, 2, \dots, d\}, \forall u \in \mathbb{R}^+.$$

Si de plus pour tout $i \in \{1, \dots, d\}$, F_{X_i} est une distribution sous-exponentielle telle que

$$\frac{\overline{F}_{X_i}(y)}{\overline{F}_{X_i}(x)} \stackrel{x \rightarrow +\infty}{\asymp} O(1), \quad \text{pour } 0 < \kappa_1 \leq \frac{y}{x} \leq \kappa_2 < 1,$$

alors les contributions u_i et u_j vérifient

$$\overline{F}_{X_i}(u_i) - \overline{F}_{X_i}(u) \stackrel{u \rightarrow +\infty}{\asymp} \overline{F}_{X_j}(u_j) - \overline{F}_{X_j}(u) + o(\overline{F}_{X_i}(u)).$$

Pour l'allocation optimale par minimisation de l'indicateur J , nous avons démontré que s'il existe $i \in \{1, \dots, d\}$ tel que X_i a une distribution sous-exponentielle, alors l'allocation optimale (u_1, \dots, u_d) vérifie pour tout $j \neq i$

$$\lim_{u \rightarrow \infty} \frac{\mathbb{P}(X_j \geq u_j, S \geq u)}{\overline{F}_{X_i}(u)} = 1.$$

L'illustration la plus simple de ce modèle est le cas des distributions de Pareto indépendantes. On considère d variables aléatoires $\{X_1, X_2, \dots, X_d\}$ indépendantes de loi Pareto, de paramètres $(a, b_i)_{\{i=1,2,\dots,d\}}$ respectivement. Nous supposons sans perte de généralité que $b_1 > b_2 > \dots > b_d > 0$. De l'application des résultats obtenus dans le cas sous-exponentiel, nous déduisons directement que la composition de l'allocation asymptotique par minimisation de l'indicateur I vérifie

$$\forall (i, j) \in \{1, 2, \dots, d\}^2, \quad \left(\frac{\lim_{u \rightarrow \infty} \alpha_i}{b_i} \right)^{-a} - \left(\frac{\lim_{u \rightarrow \infty} \alpha_j}{b_j} \right)^{-a} = \left(\frac{1}{b_i} \right)^{-a} - \left(\frac{1}{b_j} \right)^{-a}.$$

Asymptotiquement, la composition de l'allocation par minimisation de l'indicateur J vérifie

$$\lim_{u \rightarrow \infty} \alpha_1 = 1 \quad \text{et} \quad \lim_{u \rightarrow \infty} \alpha_i = 0, \quad \forall i \in \{2, 3, \dots, d\}.$$

L'analyse des deux cas exponentiel et sous-exponentiel nous permet de visualiser l'impact de la nature des distributions marginales sur le comportement asymptotique de l'allocation du capital, pour les deux indicateurs I et J . Nous examinons dans la suite l'impact de la dépendance sur la composition de l'allocation optimale. Pour cela nous considérons des modèles particuliers qui nous permettront de formuler des conclusions.

Cas de l'indépendance asymptotique

Afin d'étudier ce cas, nous allons nous baser sur le modèle FGM. La copule FGM est très étudiée en sciences actuarielles, elle représente une structure de dépendance faible sans dépendance asymptotique. Nous nous limitons sans perte de généralité au cas bivarié, et nous examinons l'impact de la dépendance sur les pourcentages de capital alloués à chacune des deux branches. Nous considérons alors deux risques X_1 et X_2 de distributions marginales exponentielles $X_i \sim \exp(\beta_i)$ tels que $\beta_1 < \beta_2$ et $\beta_2 \neq 2\beta_1$, et une structure de dépendance FGM bivariée de paramètre $-1 \leq \theta \leq 1$. La copule FGM est donnée par son expression analytique

$$C_\theta(u, v) = uv[1 + \theta(1 - u)(1 - v)],$$

et le coefficient de la dépendance supérieure des queues est nul

$$\lambda_U = 0.$$

Nous avons montré que l'allocation optimale d'un capital u par minimisation de l'indicateur I est donnée par $(\alpha_1 u, (1 - \alpha_1)u)$ où $\alpha_1 = u_1/u$ est l'unique solution dans $[0, 1]$ de l'équation

$$\begin{aligned} h(\alpha_1) - h(\gamma(1 - \alpha_1)) &= \frac{1 + \theta}{\gamma - 1} [h(\gamma) + \gamma h(1) - (1 + \gamma)h(\gamma + \alpha_1 - \gamma\alpha_1)] \\ &\quad + \frac{\theta}{\gamma - 1} [h(2\gamma) + \gamma h(2) - (1 + \gamma)h(2\gamma + 2\alpha_1 - 2\gamma\alpha_1)] \\ &\quad - \frac{\theta}{\gamma - 2} [2h(\gamma) + \gamma h(2) - (2 + \gamma)h(\gamma + 2\alpha_1 - \gamma\alpha_1)] \\ &\quad - \frac{\theta}{2\gamma - 1} [h(2\gamma) + 2\gamma h(1) - (1 + 2\gamma)h(2\gamma + \alpha_1 - 2\gamma\alpha_1)], \end{aligned}$$

avec h la fonction définie par $h(x) = \exp(-\beta_1 u x)$, et $\gamma = \beta_2/\beta_1$.

L'allocation optimale du capital u par minimisation de l'indicateur J est déterminée par l'unique solution de l'équation

$$\begin{aligned} \frac{1 + \theta}{\gamma - 1} [h(\gamma) + \gamma h(1) - (1 + \gamma)h(\gamma + \alpha_1 - \gamma\alpha_1)] &+ \frac{\theta}{\gamma - 1} [h(2\gamma) + \gamma h(2) - (1 + \gamma)h(2\gamma + 2\alpha_1 - 2\gamma\alpha_1)] \\ &- \frac{\theta}{\gamma - 2} [2h(\gamma) + \gamma h(2) - (2 + \gamma)h(\gamma + 2\alpha_1 - \gamma\alpha_1)] \\ &- \frac{\theta}{2\gamma - 1} [h(2\gamma) + 2\gamma h(1) - (1 + 2\gamma)h(2\gamma + \alpha_1 - 2\gamma\alpha_1)] = 0. \end{aligned}$$

Ces équations nous ont permis d'examiner de près l'impact de la dépendance sur la composition de l'allocation optimale. Une des particularités de cette méthode d'allocation par rapport aux méthodes classiques est la prise en compte du niveau du capital. Les méthodes classiques produisent le pourcentage alloué à chaque branche en fonction de la mesure de risque utilisée uniquement, alors que l'allocation optimale par minimisation des indicateurs multivariés de risque reste très sensible au niveau du capital de l'allocation. Des illustrations graphiques sont fournies au chapitre 4 de la thèse.

Cas de la dépendance asymptotique

Le cas de la présence d'une dépendance asymptotique regroupe l'ensemble des modèles dont la structure de dépendance permet d'avoir un coefficient de dépendance des queues strictement positif $\lambda_U > 0$. Dans notre contexte, seule la dépendance supérieure des queues nous intéresse. Nous avons choisi deux modèles particuliers, Marshall-Olkin et mélange commun, pour étudier l'impact de la dépendance asymptotique sur le comportement asymptotique de l'allocation. Le modèle Marshall-Olkin présente un cas spécial de construction de la dépendance, car les distributions marginales changent en fonction de l'intensité de la dépendance. En effet, le changement des paramètres de la dépendance engendre au même temps une modification sur les distributions marginales, contrairement au modèle FGM où les lois marginales sont fixées. Le choix du modèle mélange commun est justifié par notre besoin de construction d'un modèle Pareto corrélé afin de visualiser l'impact de la dépendance en le comparant avec le cas indépendant.

Le modèle Marshall-Olkin

Soient $Y_i \sim \exp(\lambda_i)$, avec $i = 0, 1, 2$ trois variables aléatoires indépendantes. On construit deux variables aléatoires avec choc commun : $X_i = \min(Y_i, Y_0)$ pour $i = 1, 2$. Les X_i ont des distributions marginales exponentielles de paramètres $\beta_i = \lambda_i + \lambda_0$. La copule correspondante est donnée par

$$C(u_1, u_2) = \min(u_1^{1-\lambda_0/\beta_1} u_2, u_1 u_2^{1-\lambda_0/\beta_2}).$$

Cette construction permet d'obtenir une corrélation de Pearson de : $\rho_P = \frac{\lambda_0}{\lambda_0 + \lambda_1 + \lambda_2}$, et un rho de Spearman

$$\rho_S = \frac{1}{1 + \frac{2}{3} \frac{\lambda_1 + \lambda_2}{\lambda_0}},$$

ce modèle ne prend donc en compte que la dépendance positive. La dépendance des queues est prise en compte

$$\lambda_U = \min\left(\frac{\lambda_0}{\beta_1}, \frac{\lambda_0}{\beta_2}\right) = \frac{\lambda_0}{\max(\lambda_1, \lambda_2) + \lambda_0},$$

d'où son appartenance aux modèles conduisant à une dépendance asymptotique. On suppose sans perte de généralité que $\lambda_1 < \lambda_2$ et $\lambda_0 \neq \lambda_2 - \lambda_1$. Nous avons démontré que dans le cas de ce modèle, l'allocation optimale d'un capital u par minimisation de l'indicateur I est donnée par $(\alpha_1 u, (1 - \alpha_1)u)$, tel que $\alpha_1 = u_1/u$ est l'unique solution dans $[0, 1]$ de l'équation

$$g(\beta_2(1 - \alpha_1)) - g(\beta_1\alpha_1) + \frac{\beta_1}{\beta_1 - \lambda_2}g((\beta_1 - \lambda_2)\alpha_1 + \lambda_2) + \frac{\lambda_2}{\beta_1 - \lambda_2}g((\lambda_2 - \beta_1)(1 - \alpha_1) + \beta_1) - \frac{\lambda_1}{\lambda_1 - \beta_2}g(\beta_2) = \frac{\lambda_2}{\beta_1 - \lambda_2}g(\beta_1) + g(\lambda_s/2)\left[\frac{\beta_1}{\beta_1 - \lambda_2} - \frac{\lambda_1}{\lambda_1 - \beta_2}\right],$$

avec $\lambda_s = \lambda_0 + \lambda_1 + \lambda_2$, et g la fonction $g(x) = \exp(-ux)$.

L'allocation optimale d'un capital u par minimisation de l'indicateur J est déterminée par $(\alpha_1 u, (1 - \alpha_1)u)$, tel que $\alpha_1 = u_1/u$ est la solution unique dans $[0, 1]$ de l'équation suivante

$$\frac{\beta_1}{\beta_1 - \lambda_2}g((\beta_1 - \lambda_2)\alpha_1 + \lambda_2) + \frac{\lambda_2}{\beta_1 - \lambda_2}g((\lambda_2 - \beta_1)(1 - \alpha_1) + \beta_1) - \frac{\lambda_1}{\lambda_1 - \beta_2}g(\beta_2) = \frac{\lambda_2}{\beta_1 - \lambda_2}g(\beta_1) + g(\lambda_s/2)\left[\frac{\beta_1}{\beta_1 - \lambda_2} - \frac{\lambda_1}{\lambda_1 - \beta_2}\right].$$

Le modèle à mélange commun

L'idée de la construction d'un modèle multivarié par mélange commun est basée sur le choix d'une variable aléatoire dite de mélange Θ de support S_Θ . Elle permet de construire, à partir d'un ensemble de variables aléatoires indépendantes Y_i , un autre ensemble de même taille de variables aléatoires X_i corrélées et indépendantes conditionnellement à Θ tel que

$$\bar{F}_{X_i|\Theta=\theta}(x_i) = (\bar{F}_{Y_i}(x_i))^\theta.$$

À l'aide de ce modèle, nous construisons des variables aléatoires corrélées de distributions marginales Pareto, par un mélange Exponentiel-Gamma. Nous considérons (X_1, \dots, X_d) tel que conditionnellement à θ , les variables aléatoires X_i sont indépendantes et $X_i \sim \mathcal{E}(\beta_i\theta)$ avec $\beta_1 < \beta_2 \dots < \beta_d$. La variable du mélange sera $\theta \sim \Gamma(a, b)$. La fonction de survie d'un risque X_i est donnée dans ce cas par

$$\bar{F}_{X_i}(x) = \int_0^\infty \bar{F}_{X_i|\Theta=\theta} f_\Theta(\theta) d\theta = \int_0^\infty e^{-\beta_i\theta x} f_\Theta(\theta) d\theta = \left(1 + \frac{\beta_i x}{b}\right)^{-a},$$

ainsi X_i suit une loi de Pareto de paramètres $\left(a, \frac{b}{\beta_i}\right)$.

Nous avons démontré que l'allocation optimale par minimisation de l'indicateur I est l'unique solution dans \mathcal{U}_u^d du système d'équations suivant

$$s(\beta_i\alpha_i) - s(\beta_j\alpha_j) - \sum_{\ell=1}^d A_\ell [s(\alpha_i\beta_i + (1 - \alpha_i)\beta_\ell) - s(\alpha_j\beta_j + (1 - \alpha_j)\beta_\ell)] = 0, \quad \forall (i, j) \in \{1, 2, \dots, d\}^2,$$

où s est la fonction définie par $s(x) = (1 + x \frac{u}{b})^{-a}$, $\alpha_i = \frac{u_i}{u}$ pour tout $i \in \{1, \dots, d\}$. et $A_\ell =$

$$\prod_{j=1, j \neq \ell}^d \frac{\beta_j}{\beta_j - \beta_\ell}, \text{ pour } \ell = 1, \dots, d.$$

L'allocation asymptotique optimale obtenue par minimisation de l'indicateur multivarié de risque I est déterminée par l'unique solution dans $\mathbb{1}_u^d$ du système d'équations suivant

$$(\beta_i \alpha_i)^{-a} - (\beta_j \alpha_j)^{-a} - \sum_{\ell=1}^d A_\ell [(\alpha_i \beta_i + (1 - \alpha_i) \beta_\ell)^{-a} - (\alpha_j \beta_j + (1 - \alpha_j) \beta_\ell)^{-a}] = 0, \quad \forall (i, j) \in \{1, 2, \dots, d\}^2.$$

Dans le cas de l'allocation optimale par minimisation de l'indicateur J , nous avons démontré que la composition de l'allocation est déterminée par l'unique solution dans \mathcal{U}_u^d du système d'équations suivant

$$\sum_{\ell=1}^d A_\ell [s(\alpha_i \beta_i + (1 - \alpha_i) \beta_\ell) - s(\alpha_j \beta_j + (1 - \alpha_j) \beta_\ell)] = 0, \quad \forall (i, j) \in \{1, 2, \dots, d\}^2,$$

et que l'allocation optimale asymptotique quand le capital u tend vers $+\infty$ est déterminée par des pourcentages $(\alpha_1, \dots, \alpha_d)$ qui composent la solution unique dans $\mathbb{1}_u^d$ du système d'équations suivant

$$\sum_{\ell=1}^d A_\ell [(\alpha_i \beta_i + (1 - \alpha_i) \beta_\ell)^{-a} - (\alpha_j \beta_j + (1 - \alpha_j) \beta_\ell)^{-a}] = 0, \quad \forall (i, j) \in \{1, 2, \dots, d\}^2.$$

Ces résultats nous ont permis d'étudier l'impact de la dépendance sur la composition de l'allocation optimale et l'allocation optimale asymptotique en les comparant avec ceux obtenus précédemment dans le cas du modèle de Pareto indépendant.

Nous nous sommes ensuite intéressés à l'effet du mélange sur l'allocation optimale, et sa vitesse de convergence vers l'allocation asymptotique. Pour cela, nous avons considéré une version simplifiée du modèle à mélange commun. Nous supposons sans perte de généralité que $b = 1$. On obtient donc $\Theta \sim \Gamma(\gamma, 1)$ et $Y_i \sim \exp(\beta_i)$. Le mélange permet de construire des distributions $Pareto(\gamma, 1/\beta_i)$, et dans le cas bivarié la fonction de survie jointe est donnée par

$$\bar{F}_{X_1, X_2}(x_1, x_2) = \left(\frac{1}{1 + \beta_1 x_1 + \beta_2 x_2} \right)^\gamma = \left(\bar{F}_{X_1}(x_1)^{-1/\gamma} + \bar{F}_{X_1}(x_1)^{-1/\gamma} - 1 \right)^{-\gamma}.$$

Il s'agit d'une construction de la copule de survie de Clayton avec $\theta = 1/\gamma$ comme paramètre de dépendance. Pour cette copule le coefficient de dépendance asymptotique du modèle construit est

$$\lambda_U = 2^{-\gamma},$$

d'où la présence de dépendance asymptotique.

Nous avons démontré que dans cas, l'allocation optimale d'un capital u par minimisation de l'indicateur I entre les deux branches est donnée par $(\alpha_1 u, (1 - \alpha_1)u)$ tel que $\alpha_1 = u_1/u$ est l'unique solution dans $[0, 1]$ de l'équation suivante

$$h(\beta_1 \alpha_1) - h(\beta_2(1 - \alpha_1)) + \frac{\beta_1}{\beta_2 - \beta_1} h(\beta_1) + \frac{\beta_2}{\beta_2 - \beta_1} h(\beta_2) - \frac{\beta_1 + \beta_2}{\beta_2 - \beta_1} h(\beta_1 \alpha_1 + \beta_2(1 - \alpha_1)) = 0,$$

avec h la fonction $h(x) = \frac{1}{(1+xu)^\gamma}$.

Pour les niveaux asymptotiques du capital u , l'allocation optimale $(\alpha_1 u, (1 - \alpha_1)u)$ est déterminée par le pourcentage α_1 , solution unique dans $[0, 1]$ de l'équation suivante

$$g(\beta_1 \alpha_1) - g(\beta_2(1 - \alpha_1)) + \frac{\beta_1}{\beta_2 - \beta_1} g(\beta_1) + \frac{\beta_2}{\beta_2 - \beta_1} g(\beta_2) - \frac{\beta_1 + \beta_2}{\beta_2 - \beta_1} g(\beta_1 \alpha_1 + \beta_2(1 - \alpha_1)) = 0,$$

avec g la fonction $g(x) = \frac{1}{x^\gamma}$.

Nous avons illustré la convergence vers l'allocation optimale pour ce modèle.

Cas de la dépendance parfaite

Nous avons examiné également le cas de la présence d'une dépendance parfaite entre les différents risques. Nous considérons des risques comonotones X_1, \dots, X_d possédant des fonctions de répartition strictement croissantes et dont les supports contiennent l'intervalle $[0, u]$. Nous avons démontré que dans ce cas les allocations optimales par minimisation des indicateurs I , J et I_{loc} défini par

$$I_{loc}(u_1, \dots, u_d) = \sum_{k=1}^d \mathbb{E} \left((X_k - u_k) \mathbb{1}_{\{X_k > u_k\}} \right) = \sum_{k=1}^d \mathbb{E} \left((X_k - u_k)^+ \right),$$

coïncident, et leur composition $(u_1, \dots, u_d) \in \mathcal{U}_u^d$ est déterminée par le système suivant

$$F_{X_i}(u_i) = F_{X_j}(u_j) \quad \forall i, j = 1, \dots, d.$$

De plus, les pourcentages de l'allocation optimale $\alpha_i = \frac{u_i}{u}$ sont indépendants du niveau du capital à allouer u pour tout $i \in \{1, \dots, d\}$ si et seulement si il existe des constantes positives a_{ij} telles que

$$X_i \stackrel{\mathcal{L}}{=} a_{ij} X_j, \quad \forall (i, j) \in \{1, \dots, d\}^2.$$

Ce résultat nous a permis de déterminer la composition explicite des allocations optimales par minimisations des indicateurs I et J . Le Tableau 1 présente un récapitulatif des compositions des allocations obtenues pour les modèles étudiés.

Modèle	Capital alloué à la branche i
I.D : Les risques sont identiquement distribués ($X_i \sim X$)	$u_i = u/d \quad \forall i \in \{1, \dots, d\}$
Exponentiel : $X_i \sim \exp(\beta_i)$	$u_i = \frac{1/\beta_i}{\sum_{j=1}^d 1/\beta_j} u \quad \forall i \in \{1, \dots, d\}$
Weibull : $X_i \sim We(\tau, \beta_i)$	$u_i = \frac{1/\beta_i}{\sum_{j=1}^d 1/\beta_j} u \quad \forall i \in \{1, \dots, d\}$
Normal : $X_i \sim N(\mu_i, \sigma_i)$	$u_i = \mu_i + \frac{\sigma_i}{\sum_{l=1}^d \sigma_l} \left(u - \sum_{l=1}^d \mu_l \right)$
Log-Normal : $X_i \sim LN(\mu_i, \sigma^2)$	$u_i = \frac{\exp(\mu_i)}{\sum_{l=1}^d \exp(\mu_l)} u \quad \forall i \in \{1, \dots, d\}$
Pareto : $X_i \sim Pa(\alpha, \lambda_i)$	$u_i = \frac{\lambda_i}{\sum_{l=1}^d \lambda_l} u \quad \forall i \in \{1, \dots, d\}$

TABLE 1 – Allocations optimales pour des modèles comonotones

Finalement, par ce travail, nous proposons une construction élicitable de mesures de risques multidimensionnelles, utilisables pour des applications actuarielles. Les mesures construites vérifient un ensemble de propriétés de cohérence. Elles sont calculables à l'aide des outils de l'optimisation stochastique. Le risque extrême modélisé par ces mesures possède un comportement asymptotique déterminable en fonction de la nature des lois marginales et de la structure de dépendance extrême. Par ailleurs nous proposons des estimateurs de ces mesures dans le cas des variations régulières.

La méthode d'allocation du capital par minimisation des indicateurs multivariés de risque peut être aussi vue comme une construction de mesure de risque multidimensionnelle. Elle peut être utilisée pour d'autres applications actuarielles, notamment pour répondre à la problématique de la réassurance optimale dans le cas des traités Stop-Loss. Elle sera également utilisable pour évaluer le risque systémique d'une composante d'un système financier.

Des mesures de risque expectiles multivariées

Chapitre 1

Des extensions multivariées des mesures de risque expectiles

Multivariate extensions of expectiles

Un résumé de ce chapitre est présenté dans l'article Maume-Deschamps et al. (2016) [32], soumis pour publication, accessible en prépublication Hal. <https://hal.archives-ouvertes.fr/hal-01367277>

Introduction

Les critères définis par Artzner et al. (1999) [2] pour qualifier une mesure de risque univariée de cohérente ont favorisé l'utilisation pratique de certaines mesures de risque par rapport à d'autres. La *Value-at-Risk* n'étant pas sous additive, est devenue au moins du point de vue académique une mesure de risque incohérente et moins conseillée que la *Tail-Value-at-Risk*, qui vérifie bien tous les axiomes de cohérence. Cette préférence académique entre les deux mesures de risque les plus utilisées dans les pratiques professionnelles n'est toujours pas pleinement prise en compte par les autorités prudentielles. En assurance, la directive Solvabilité 2 offre la possibilité aux assureurs d'utiliser des modèles internes pour le calcul des *capitaux de solvabilités requis* (SCR), et donc de choisir une mesure de risque interne, mais la *Formule Standard* qui représente une base de la nouvelle réglementation, est toujours fondée dans ses calculs sur la VaR. En finance, et en l'absence d'un choix de mesure de risque dans le dispositif prudentiel Bâle 2, la décision est laissée aux gestionnaires de risques, et la "meilleure" mesure dépend du contexte de son utilisation et des procédures internes. Un retour d'expérience basé sur des pratiques professionnelles est présenté par exemple dans Sarykalin et al. (2008) [39]. Dans l'objectif de renforcer la gestion des risques après la crise financière 2008, les propositions Bâle 3 de réglementation bancaire sont en cours de validation et mises en œuvre progressivement. Une réponse issue du monde académique à ces propositions, et qui concerne le choix entre les deux mesures VaR et TVaR, est apparue récemment dans Embrechts et al. (2014) [17]. Elle présente un comparatif détaillé à la fois théorique et pratique qui peut aider à trancher entre les deux, voire à imposer un choix dans les nouvelles normes.

La théorie des mesures de risque est assortie d'une riche littérature en actualisation permanente. La construction mathématique des différents indicateurs et méthodes utilisés en gestion de risque doit être accompagnée par une vérification de leur adéquation avec les pratiques des professionnels. La notion de *cohérence* occupe en ce sens une place importante dans la littérature des sciences actuarielles. La généralisation de la notion de cohérence pour les mesures de risque sous forme d'ensembles de vecteurs est présentée par Jouini et al. (2004) [25]. La cohérence d'une allocation de capital est introduite sous une approche axiomatique reliée aux mesures de risque par Denault (2001) [13], puis étudiée pour la méthode d'Euler dans Buch et Dorfleitner (2008) [7]. Une caractérisation de la cohérence pour les méthodes d'allocation par minimisations d'indicateurs de risque multivariés a été

présentée par Maume-Deschamps et al. (2016) [33]. Les critères définis pour caractériser la cohérence des mesures et des allocations de risques sont justifiés par l'importance économique des propriétés qu'ils représentent en pratique. Sous cet angle, la cohérence ne se limite pas à ces critères, mais touche l'ensemble des traductions mathématiques des propriétés de cohérence souhaitées par les praticiens. Gneiting (2011) [22] a soulevé une question importante concernant la cohérence statistique des mesures de risque usuelles. L'*élicitabilité*, selon lui, est une propriété naturelle qui doit être satisfaite par une mesure afin de garantir la possibilité d'implémenter les procédures de *backtesting* nécessaires d'un point de vue pratique. Ce travail a remis en question la notion de cohérence, notamment pour les mesures qui ne sont pas élicitables même si elles sont cohérentes. Plusieurs travaux ont été présentés récemment sur le sujet, une partie analyse cette propriété dans le cadre des mesures de risque et une autre a été dédiée à la recherche d'une caractérisation des mesures qui sont à la fois cohérentes et élicitables. Les *expectiles* sont alors apparus dans la littérature comme seule famille de mesures parmi les mesures de risque invariantes par loi qui répond à ce besoin, elles sont élicitables par construction et cohérentes pour un intervalle de niveau de seuil.

L'élicitabilité a fait le sujet de l'article de Ziegel (2014) [43], elle est étudiée aussi dans un cadre purement mathématique dans Steinwart et al. (2014) [40]. L'élicitabilité des mesures de risque a été analysée par Bellini et Bigozzi (2015) [3] et dans Wang et Ziegel (2015) [41]. Les expectiles comme mesures de risque ont fait l'objet des travaux de Emmer et al. (2013) [18], de Bellini et al. (2014) [5], et de Bellini et Di Bernardino (2015) [4].

L'importance statistique de la propriété de l'élicitabilité nous invite à revoir la construction mathématique de toute mesure de risque. En effet, la gestion de risque doit être naturellement basée sur la minimisation d'une quantité de risque ou d'une erreur définie en fonction des besoins de l'utilisateur. Dans cette partie, nous présentons la construction d'une nouvelle famille de mesures de risque vectorielles, caractérisée d'abord par son élicitabilité, et puis justifiée par ses interprétations économiques. L'élicitabilité est utilisée dans son sens multivarié pour cette construction. Les mesures de risque construites sont vectorielles (*vector-valued*). La littérature présente un nombre limité de ce type de mesures, nous citons à titre d'exemple, La VaR et CVaR multidimensionnelles introduites dans la thèse de Di Bernardino (2012)[14], et étudiées dans Cousin et Di Bernardino (2013) [9], Cousin et Di Bernardino (2014) [10] et Di Bernardino et al. (2015) [15]. Ce type de mesures est utilisable en pratique pour l'allocation du capital économique et la mesure du risque systémique.

Les mesures proposées dans le cadre de cette thèse sont en général des minimiseurs de fonctions strictement convexes, l'utilisation des algorithmes d'approximation stochastiques pour la détermination du minimum est donc possible. Nous présentons une version de l'algorithme de Robbins-Monro (1951)[36] basée sur la forme multidimensionnelle proposée par Blum (1954) [6], et qui permet d'approximer ces mesures dans le cas général, avec une convergence assez satisfaisante sauf pour le cas asymptotique où le nombre d'observations est limité. L'étude théorique du cas asymptotique semble donc inévitable, nous terminons cette partie par une analyse globale de composition de l'expectile asymptotique afin d'introduire le prochain chapitre constitué d'une analyse plus détaillée du comportement asymptotique des vecteurs expectiles.

Ce chapitre est organisé en plusieurs sections. Dans un premier temps, nous présentons un résumé de la littérature récente sur l'élicitabilité comme propriété de cohérence pour les mesures de risque univariées. Nous présentons dans le même contexte une synthèse sur les mesures expectiles et leurs propriétés. Nous terminons cette partie introductive par l'élicitabilité dans un contexte multivarié et ses différentes définitions récemment introduites dans la théorie du risque.

La section 2 est consacrée à la présentation des constructions d'une famille d'expectiles multivariés. Nous présentons une extension multivariée basée sur les normes et nous discutons ensuite de la possibilité de définir des mesures à l'aide des ensembles d'acceptations. Nous analysons l'interprétation économique pour chaque type d'expectiles et les applications possibles pour chaque construction.

Nous nous focalisons ensuite sur les expectiles matriciels. Nous présentons une méthodologie d'approximation stochastique de l'expectile multivarié par application de l'algorithme Robbins-Monro, et nous illustrons sa convergence à l'aide d'une série d'applications numériques. Nous étudions enfin les différentes propriétés de cohérence vérifiées par ces expectiles, avant d'aborder leur comportement asymptotique dans la dernière partie de ce chapitre.

1.1 Élicitabilité des mesures de risque

Depuis la publication du travail de Gneiting (2011) [22], la notion d'élicitabilité occupe une place importante dans les travaux de recherche sur les propriétés de cohérence des mesures de risque. L'origine de cette notion vient de la théorie de la décision. Nous présentons dans cette section un résumé de la littérature récente sur l'élicitabilité des mesures des risques. Nous abordons dans un premier temps sa définition et les propriétés probabilistes et statistiques qui en découlent. Nous examinons ensuite sa place dans la théorie des mesures de risque, nous passons ainsi par les mesures expectiles, et nous consacrons la dernière partie de cette section à l'élicitabilité multivariée. L'objectif de cette section est de présenter le cadre général qui va constituer la base des constructions de mesures dans le reste de ce document.

Une statistique T est dite *élicitable* si elle peut être écrite pour toute variable aléatoire réelle $X \in \mathbb{R}$ sous la forme d'une solution de minimisation d'une fonction d'erreur notée $s : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$

$$T(X) = \arg \min_x \mathbb{E}_{\mathbb{P}}[s(x, X)], \quad X \sim \mathbb{P}.$$

L'utilité statistique de cette propriété peut être résumée à deux importants avantages :

- La possibilité de comparer à l'aide de la fonction d'erreur les différentes méthodes statistiques et d'estimations utilisées, et donc donner un sens aux procédures de *backtesting*, très importantes en finance et assurance.
- La possibilité de produire des prévisions et des estimations par moyenne de régression dans le cas des mesures élicitables.

Pour plus de détails sur l'importance statistique de la propriété de l'élicitabilité, nous invitons le lecteur à consulter les articles de Gneiting (2011) [22] et Emmer et al. (2013) [18].

1.1.1 La propriété d'élicitabilité

Nous présentons tout d'abord quelques notions statistiques, issues surtout de la théorie de la décision, nécessaires pour formaliser la définition d'élicitabilité. Nous considérons un ensemble d'observations Y , et soit \mathcal{P} un ensemble de mesures de probabilités sur Y .

Définition 1.1.1 (Fonction de score). On considère une statistique $T : \mathcal{P} \rightarrow \mathbb{R}$ et un intervalle de \mathbb{R} noté A , tel que l'image de T est incluse dans A , $\text{Im}(T) \subset A$. Une fonction de score est une fonction de pénalité positive qui permet de quantifier l'erreur d'estimation

$$S : A \times Y \rightarrow \mathbb{R} \\ (x, y) \quad S(x, y)$$

Dans la théorie des mesures de risque, les fonctions de score doivent satisfaire quelques conditions complémentaires pour qu'elles soient d'une utilité pratique. Nous citons à titre d'exemples les trois conditions présentées dans Bellini et Bignozzi (2015) [3]

- (a) $\forall (x, y) \in \mathbb{R}^2 \ S(x, y) \geq 0$ et $S(x, y) = 0$ si et seulement si $x = y$;
- (b) $S(x, y)$ est croissante pour les $x > y$ et décroissante pour $y > x$;
- (c) $\forall y \in Y \ x \rightarrow S(x, y)$ est continue.

Définition 1.1.2 (Sensibilité à l'ordre dans \mathcal{P}). Soit T une statistique et A un intervalle réel tel que $\text{Im}(T) \subset A$. On dit qu'une fonction de score S est sensible à l'ordre de \mathcal{P} pour T si pour tout $\mathbb{P} \in \mathcal{P}$ et tout $(x_1, x_2) \in A^2$ tel que $x_2 < x_1 < T(\mathbb{P})$ ou $T(\mathbb{P}) < x_1 < x_2$

$$\mathbb{E}_{\mathbb{P}}[S(x_1, Y)] < \mathbb{E}_{\mathbb{P}}[S(x_2, Y)].$$

La sensibilité à l'ordre dans \mathcal{P} traduit la prise en compte de la distance par rapport au minimum par la fonction de score. La pénalité moyenne croît avec la distance au minimiseur $T(\mathbb{P})$. On rencontre cette propriété souvent dans la littérature sous le nom *accuracy rewarding*.

Définition 1.1.3 (Consistance et stricte consistance). Une fonction de score S est dite \mathcal{P} -consistante à la statistique T si pour tout $\mathbb{P} \in \mathcal{P}$ et tout $x \in A$

$$\mathbb{E}_{\mathbb{P}}[S(T(\mathbb{P}), Y)] \leq \mathbb{E}_{\mathbb{P}}[S(x, Y)].$$

Elle est dite \mathcal{P} -strictement consistante à T si pour tout $\mathbb{P} \in \mathcal{P}$

$$\mathbb{E}_{\mathbb{P}}[S(x, Y)] = \mathbb{E}_{\mathbb{P}}[S(T(\mathbb{P}), Y)] \Rightarrow x = T(\mathbb{P}).$$

Les propriétés de consistance et de stricte consistance garantissent l'atteinte du minimum de l'espérance de l'erreur en T . Dans le cas général, $T(\mathbb{P})$ est un ensemble. On rencontre dans la littérature trois définitions différentes de la stricte consistance, celle annoncée dans la définition est présentée dans sa forme généralisée par Gneiting (2011) [22]

$$\mathbb{E}_{\mathbb{P}}[S(x, Y)] = \mathbb{E}_{\mathbb{P}}[S(t, Y)] \Rightarrow x \in T(\mathbb{P}), \quad \forall \mathbb{P} \in \mathcal{P}.$$

Selon Osband (1985) [35], une fonction de score consistante à T est strictement consistante si $\mathbb{E}_{\mathbb{P}}[S(x, Y)]$ est unique. Une autre forme de cette définition a été présentée dans Lambert (2013) [38], elle qualifie de stricte consistante toute fonction de score consistante qui vérifie pour tout $t \in A$

$$t \notin T(\mathbb{P}) \Rightarrow t \notin \arg \min_{x \in A} \mathbb{E}_{\mathbb{P}}[S(x, Y)].$$

Ces trois définitions sont finalement équivalentes pour les statistiques à valeur unique.

A partir de maintenant nous pouvons formuler une définition complète de la propriété de l'élicitabilité.

Définition 1.1.4 (Élicitabilité). Une statistique $T : \mathcal{P} \rightarrow \mathbb{R}$ est élicitable s'il existe une fonction de score \mathcal{P} -strictement consistante à T .

Bellini et Bigozzi (2015) [3] définissent directement l'élicitabilité d'une statistique T relativement à la classe \mathcal{P} par l'existence d'une fonction de score $S : \mathbb{R}^2 \rightarrow \mathbb{R}^+$ telle que

$$T(\mathbb{P}) = \arg \min_{x \in \mathbb{R}} \int S(x, y) d\mathbb{P}(y), \quad \forall \mathbb{P} \in \mathcal{P},$$

et dans ce cas S est considérée strictement consistante à T .

Définition 1.1.5 (Identifiabilité). Une statistique T est identifiable s'il existe une fonction $V : A \times Y \rightarrow \mathbb{R}$ telle que, pour tout $x \in \text{Im}T$ et $\mathbb{P} \in \mathcal{P}$

$$\mathbb{E}_{\mathbb{P}}[V(x, Y)] = 0 \iff x = T(\mathbb{P}).$$

V est appelée *la fonction de la \mathcal{P} -identification de T* .

Cette propriété semble naturelle dans le cas des fonctions de score différentiables en x , car elle traduit uniquement la condition d'optimisation de premier ordre qui transforme la recherche d'un minimum d'une fonction en recherche de la racine d'une autre. La notion d'identifiabilité a été étudiée en détail dans Steinwart et al. (2014) [40]. Une méthode de construction de famille des fonctions de score basée sur le choix d'une fonction d'identification a été présentée dans la thèse de Osband (1985) [35].

Dans la suite, nous présentons des propriétés qui sont en relation directe avec l'élicitabilité.

Définition 1.1.6 (Convexité par mélange). Une statistique T est convexe par mélange si pour tout $\mathbb{P} \in \mathcal{P}$, $\mathbb{Q} \in \mathcal{P}$ et tout $\alpha \in [0, 1]$ tel que $\alpha\mathbb{P} + (1 - \alpha)\mathbb{Q} \in \mathcal{P}$,

$$T(\alpha\mathbb{P} + (1 - \alpha)\mathbb{Q}) \leq \alpha T(\mathbb{P}) + (1 - \alpha)T(\mathbb{Q}).$$

Il faut bien distinguer entre les deux propriétés de la convexité d'une mesure de risque et sa convexité par mélange. la première concerne la somme de variables aléatoires alors que la deuxième est définie pour les mélanges de probabilités.

Définition 1.1.7 (Convexité des ensembles de niveaux). Les ensembles de niveaux sont définis comme suit

$$LS_a := \{\mathbb{P} \in \mathcal{P} \mid T(\mathbb{P}) = a\},$$

pour tout $a \in \mathbb{R}$. On dit que T a des ensembles de niveaux convexes, si pour tout $\alpha \in [0, 1]$ et tout $(\mathbb{P}, \mathbb{Q}) \in \mathcal{P}^2$ tel que $\alpha\mathbb{P} + (1 - \alpha)\mathbb{Q} \in \mathcal{P}$,

$$T(\mathbb{P}) = T(\mathbb{Q}) = a \Rightarrow T(\alpha\mathbb{P} + (1 - \alpha)\mathbb{Q}) = a.$$

La convexité des ensembles de niveaux porte dans la littérature le nom de la CxLS propriété. Toute statistique élicitable vérifie la propriété de convexité des ensembles de niveaux, ce résultat a été introduit par Osband (1985) [35]. Selon Bellini et Bigozzi (2015) [3] aussi, et sous les conditions qu'ils ont fixé pour les fonctions de score, si T est élicitable alors elle a des ensembles de niveaux convexes. Delbaen et al. (2014) [12] étudient la famille des mesures de risque univariées qui vérifie cette propriété. Une illustration de l'importance de cette propriété est son utilisation pour démontrer que TVaR n'est pas élicitable, car ses ensembles de niveaux ne sont pas convexes. Un contreexemple que nous rappelons ici est donné dans la démonstration du théorème 11 de Gneiting (2011) [22]. Si on suppose que \mathcal{P} contient les mesures à support borné, on peut choisir des constantes a, b, c, d dans ce support telles que $a < b < c < \frac{1}{2}(b + d)$ pour construire les deux mesures de probabilités suivantes

$$\mathbb{P} = \alpha\delta_a + \frac{1}{2}(1 - \alpha)(\delta_b + \delta_d) \text{ et } \mathbb{Q} = \alpha\delta_c + (1 - \alpha)\delta_{\frac{b+d}{2}},$$

où δ désigne la mesure de Dirac. Dans ce cas, et pour tout $\alpha \geq \frac{1}{3}$,

$$\text{TVaR}_\alpha(\mathbb{P}) = \text{TVaR}_\alpha(\mathbb{Q}) = \frac{b + d}{2} < \text{TVaR}_\alpha\left(\frac{1}{2}(\mathbb{P} + \mathbb{Q})\right) = \frac{b + c + 2d}{4},$$

cela signifie que les ensembles de niveaux ne sont pas convexe, et donc la mesure TVaR n'est pas élicitable. Une autre démonstration de la non élicitabilité de la TVaR est présentée dans Weber (2006) [42].

Il faut distinguer aussi entre les deux propriétés de la convexité par mélange et la convexité des ensembles de niveaux. Bellini et Birgozzi (2015) montrent que si T a des ensembles de niveaux convexes et vérifie les propriétés de monotonie et d'invariance par translation ou la propriété de la convexité par mélange alors les ensembles

$$\{\mathbb{P} \in \mathcal{P} \mid T(\mathbb{P}) \leq a\}; \{\mathbb{P} \in \mathcal{P} \mid T(\mathbb{P}) < a\}; \{\mathbb{P} \in \mathcal{P} \mid T(\mathbb{P}) \geq a\}; \{\mathbb{P} \in \mathcal{P} \mid T(\mathbb{P}) > a\}$$

sont convexes par mélange (Voir Lemme 2.1 Bellini et Bigozzi (2015) [3]).

On présente encore deux propriétés qui résultent de l'élicitabilité des mesures dans les deux prochaines définitions.

Définition 1.1.8 (La continuité par mélange). Une statistique T est continue par mélange si pour tout $\alpha \in [0, 1]$ et tout $(\mathbb{P}, \mathbb{Q}) \in \mathcal{P}^2$ tel que $\alpha\mathbb{P} + (1 - \alpha)\mathbb{Q} \in \mathcal{P}$, la fonction

$$\alpha \longrightarrow T(\alpha\mathbb{P} + (1 - \alpha)\mathbb{Q})$$

est continue en α .

Définition 1.1.9 (La Quasi-linéarité par mélange). Une statistique T est *quasi-convexe par mélange* si pour tout $a \in \mathbb{R}$, son ensemble de niveaux inférieur défini par

$$LLS_a := \{\mathbb{P} \in \mathcal{P} \mid T(\mathbb{P}) \leq a\}$$

est convexe par mélange. De la même façon, T est *quasi-concave par mélange* si et seulement si $-T$ est quasi-convexe par mélange. Finalement, on dit que T est *quasi-linéaire par mélange* si elle est en même temps quasi-convexe et quasi-concave par mélange.

Bellini et Bigozzi (2015) [3] montrent que si T est quasi-linéaire par mélange alors ses ensembles de niveaux sont tous convexes.

Proposition 1.1.10 (Proposition 3.1 Bellini et Bigozzi (2015) [3]). *Si une statistique T est élicitable à l'aide d'une fonction de score qui vérifie les hypothèses (a,b,c) 1.1.1 alors*

1. T possède des ensembles de niveaux convexes ;
2. T est quasi-linéaire par mélange ;
3. T est continue par mélange ;
4. Toute fonction de score qui élicite T vérifie la propriété de la sensibilité à l'ordre dans \mathcal{P} .

Les différents résultats présentés dans la littérature récente de l'élicitabilité valorisent cette propriété et lui affectent une importance majeure dans la cohérence d'une mesure de risque. Cette importance est la base de nouveaux travaux en théorie du risque, qui visent essentiellement à étudier les mesures de risque qui vérifient à la fois la cohérence au sens défini dans Artzner et al. (1999) [2] et l'élicitabilité telle qu'elle est présentée dans Gneiting (2011) [22].

1.1.2 La cohérence et l'élicitabilité

L'élicitabilité des mesures de risque est sans doute une propriété désirable statistiquement. Les mesures de risque usuelles ne vérifient pas toutes cette propriété. La variance et la TVaR sont non élicitables car elles ne vérifient pas la propriété de la convexité des ensembles de niveaux. (Voir Lambert et al. (2008) [29] pour la variance, et Gneiting (2011) [22] pour la TVaR). L'espérance est élicitable car elle peut s'écrire sous la forme

$$\mathbb{E}[X] = \int_{\mathbb{R}} x dF(x) = \arg \min_{x \in \mathbb{R}} \{\mathbb{E}[(X - x)^2]\}.$$

Les quantiles pour tout seuil $\alpha \in [0, 1]$ sont aussi élicitables à l'aide des fonctions de score de Pinball $S_\alpha(x, y) = \alpha(y - x)_+ + (1 - \alpha)(x - y)_+$, ils s'écrivent sous la forme

$$q_\alpha(X) = \min\{x : F_X(x) \geq \alpha\} = \arg \min_{x \in \mathbb{R}} \{\mathbb{E}[\alpha(X - x)_+ + (1 - \alpha)(x - X)_+]\}.$$

L'importance de cette écriture pour les estimations statistiques de l'espérance et des quantiles est étudiée dans Rockafellar et Uryasev (2013) [37]. Ziegel (2014) [43] a étudié l'élicitabilité des mesures de risque spectrales. L'élicitabilité des mesures de distorsion a été étudiée dans un article récent de Wang et Ziegel (2015) [41].

La question naturelle qui se pose concerne la possibilité d'obtenir des mesures de risque à la fois cohérentes et élicitables. La réponse de la récente littérature sur le sujet est unique, parmi les mesures invariantes par lois, seuls les expectiles univariés vérifient les deux, selon Bellini et Bigozzi (2015) [3]. Nous consacrons la suite de cette sous-section aux expectiles univariés.

Les expectiles ont été introduits dans un contexte de modèles de régression statistique dans Newey et Powell (1987) [26].

Définition 1.1.11 (Expectiles (Newey et Powell, 1987)). Pour une variable aléatoire X , l'expectile de seuil α est défini comme suit

$$e_\alpha(X) = \arg \min_{x \in \mathbb{R}} \mathbb{E}[\alpha(X - x)_+^2 + (1 - \alpha)(x - X)_+^2], \quad (1.1)$$

où $(x)_+ = \max(x, 0)$.

Ces mesures sont donc élicitables par construction. Pour $\alpha = 1/2$ l'expectile coïncide avec l'espérance. Cette mesure de risque est cohérente pour $\alpha \geq 1/2$. Pour $\alpha < 1/2$ l'expectile est une mesure sur-additive et donc non cohérente.

L'unicité du minimum est garantie par la stricte convexité de la fonction de score, cela permet de définir aussi les expectiles à l'aide de la condition d'optimalité du premier ordre, comme l'unique solution de l'équation suivante

$$\alpha \mathbb{E}[(X - x)_+] = (1 - \alpha) \mathbb{E}[(x - X)_+]. \quad (1.2)$$

L'interprétation économique des expectiles est plus visible à l'aide de l'écriture suivante de l'équation (1.2)

$$\frac{1 - \alpha}{\alpha} = \frac{\mathbb{E}[(X - x)_+]}{\mathbb{E}[(x - X)_+]},$$

qui aide à voir cette mesure comme seuil permettant d'obtenir un ration gains/pertes fixé à $\frac{1-\alpha}{\alpha}$.

Les propriétés des expectiles comme mesures de risque ont été étudiées dans plusieurs articles récents. Nous rappelons quelques-unes dont les démonstrations ont été présentées dans Emmer et al. (2013) [18] et Bellini et Di Bernardino (2015) [4].

- L'expectile $e_\alpha(X)$ est une mesure strictement croissante en fonction de $\alpha \in [0, 1]$;
- Si de plus la variable aléatoire X a une densité de classe C^1 , alors l'expectile $e_\alpha(X)$ est une fonction C^1 de α est sa dérivée est donnée par l'expression suivante

$$\frac{\partial e_\alpha(X)}{\partial \alpha} = \frac{\mathbb{E}[|X - e_\alpha(X)|]}{(1 - \alpha)F_X(e_\alpha(X)) + \alpha \bar{F}_X(e_\alpha(X))};$$

- Si la distribution de la variable aléatoire X est symétrique par rapport à un point x_0 alors,

$$\frac{e_\alpha(X) + e_{1-\alpha}(X)}{2} = x_0;$$

- L'expectile $e_\alpha(X)$ vérifie la relation

$$e_\alpha(X) = \frac{\alpha \mathbb{E}[X \mathbb{1}_{\{X \geq e_\alpha(X)\}}] + (1 - \alpha) \mathbb{E}[X \mathbb{1}_{\{X < e_\alpha(X)\}}]}{\alpha \mathbb{P}(X \geq e_\alpha(X)) + (1 - \alpha) \mathbb{P}(X < e_\alpha(X))},$$

qui permet de déduire directement qu'il s'agit d'une mesure invariante par loi et positivement homogène pour tout $0 < \alpha < 1$. L'invariance par translation est aussi vérifiée;

- Les expectiles vérifient la relation

$$e_\alpha(-X) = -e_{1-\alpha}(X),$$

pour tout $\alpha \in [0, 1]$;

- Les expectiles $e_\alpha(X)$ sont strictement monotones au sens stochastique par rapport à X , c'est à dire que si $X \leq Y, p.s$ et $\mathbb{P}(X < Y) > 0$, alors

$$e_\alpha(X) < e_\alpha(Y);$$

- Les expectiles sont sous-additifs et par suite cohérents pour $\alpha \in [1/2, 1[$, et ils sont sur-additifs pour $\alpha < 1/2$

$$e_\alpha(X + Y) \leq e_\alpha(X) + e_\alpha(Y), \forall \alpha \in [1/2, 1[,$$

- Les expectiles sont des mesures additives par dépendance linéaire

$$\text{corr}(X, Y) = 1 \Rightarrow e_\alpha(X + Y) = e_\alpha(X) + e_\alpha(Y),$$

mais elles ne sont pas comonotoniquement additives.

La propriété de la robustesse des expectiles dépend de la distance utilisée pour définir la continuité. Un bon choix sera selon la suggestion de Kiesel et al. (2012) [28] la continuité par rapport à la distance de Wasserstein définie par

$$\begin{aligned} d_W(F, G) &:= \inf \mathbb{E}[|X - Y| : X \sim F, Y \sim G] \\ &= \int_{-\infty}^{+\infty} |F(t) - G(t)| dt \\ &= \int_0^1 |F^{-1}(\alpha) - G^{-1}(\alpha)| d\alpha. \end{aligned}$$

Dans ce cas, les expectiles sont Lipschitziens pour cette distance et vérifient selon Emmer et al. (2013) [18], pour tous $X, Y \in L^1$, et $\forall \alpha \in [0, 1]$

$$|e_\alpha(X) - e_\alpha(Y)| \leq \max \left\{ \frac{\alpha}{1 - \alpha}; \frac{1 - \alpha}{\alpha} \right\} d_W(F_X, F_Y),$$

où F_X et F_Y sont respectivement les distributions de X et Y .

Les propriétés asymptotiques de premier ordre des expectiles ont été étudiées dans Bellini et Di Bernardino (2015) [4]. Celles du deuxième ordre ont été présentées dans Mao et Yang (2015) [31]. Une généralisation de ces mesures a été présentée par Bellini et al. (2014) [5], ils introduisent des mesures de risque quantiles généralisées sous la forme d'un minimiseur d'une erreur asymétrique

$$x_\alpha(X) = \arg \min_{x \in \mathbb{R}} \{ \alpha \mathbb{E}[\Phi_+((X - x)_+)] + (1 - \alpha) \mathbb{E}[\Phi_-((X - x)_-)] \}, \quad (1.3)$$

avec Φ_+ et Φ_- des fonctions de score convexes. Le cas $\Phi_-(x) = \Phi_+(x) = x^2$ permet d'obtenir les mesures expectiles. Daouia et al. (2016) [11] ont proposé une estimation des mesures VaR et ES à l'aide des expectiles.

Les expectiles présentent aujourd'hui un remplaçant potentiel à la VaR et la TVaR grâce à leur cohérence et leur élicitabilité. Toutefois, l'interprétation économique de la VaR et TVaR reste plus simple et intuitive que celle des expectiles, ce qui diminue l'attraction pratique pour ces derniers. Cela a orienté de nouveaux travaux vers la proposition de solutions pour remédier au manque d'élicitabilité de la TVaR afin de maintenir sa position forte dans la famille des mesures de risque usuelles.

1.1.3 L'élicitabilité multivariée

Pour pallier la non élicitabilité des mesures de risque usuelles, plusieurs travaux ont utilisé une version multivariée de l'élicitabilité. Lambert (2008) [29] introduit la notion d'élicitabilité indirecte qui permet de considérer par exemple la variance comme élicitable via l'élicitabilité du couple $(\mathbb{E}[X], \mathbb{E}[X^2])$. Le même type de solution a été proposé par exemple dans Fissler et Ziegel (2015) [19] pour l'élicitabilité du couple (VaR, TVaR) afin de permettre une validation des procédures de backtesting de la TVaR.

La nature multidimensionnelle des données statistiques et le contexte multivarié de la gestion des risques font de la généralisation de la propriété de l'élicitabilité des mesures en dimension plus grande que 1, une étape naturelle et d'utilité importante dans la modélisation des risques. L'idée est de pouvoir construire des mesures multidimensionnelles

$$T : \mathcal{P} \rightarrow \mathbb{R}^d,$$

qui s'écrivent sous la forme

$$\arg \inf_{\mathbf{u} \in U} \mathbb{E}[S(\mathbf{X}, \mathbf{u})],$$

où \mathbf{X} est un vecteur aléatoire dans \mathbb{R}^d et \mathbf{u} un vecteur de taille k qui peut être différente de d et $U \subset \mathbb{R}^k$.

Plusieurs auteurs ont publié des contributions importantes sur ce sujet. L'élicitabilité des statistiques vectorielles a été étudiée dans Osband (1985) [35] qui utilise une fonction de score multivariée. Lambert

et al. (2008) [29] ont introduit la notion de *k-élicitabilité*, qui remplace la propriété élicitabilité par une combinaison linéaire de fonctionnelles élicitables. L'élicitabilité multivariée a également une grande importance en apprentissage statistique, nous citons à titre d'exemple Frongillo et Kash (2014) [21] qui étudient l'élicitation de mesures vectorielles en utilisant la notion de *séparabilité* des fonctions de score. Selon ce travail, une statistique vectorielle est élicitable, si elle l'est composante par composante $T_i : \mathcal{P} \rightarrow \mathbb{R}$, et dans ce cas le vecteur est élicitable à l'aide de la fonction somme des fonctions S_i d'élicitation univariée des composantes. La séparabilité des fonctions de score reste une hypothèse forte, Osband (1985) a montré par exemple qu'il n'existe aucune fonction de score séparable qui élicite un quantile bivarié.

Dans la suite de ce travail, nous proposons une définition plus générale de l'élicitabilité dans un contexte multivarié, adaptée aux vecteurs de risques de dimension $k = d$.

Définition 1.1.12 (Elicitabilité des mesures vectorielles). Une mesure de risque vectorielle $T : \mathcal{P} \rightarrow \mathbb{R}^k$ est élicitable s'il existe une fonction de score $s : \mathbb{R}^d \times \mathbb{R}^k \rightarrow \mathbb{R}$, telle que

$$T(\mathbf{X}) \in \arg \inf_{\mathbf{x} \in U \subset \mathbb{R}^k} \mathbb{E}[s(\mathbf{X}, \mathbf{x})].$$

Un article récent de Fissler et Ziegel [19] étudie cette définition dans le cas général. Les auteurs généralisent dans le même sens la notion de consistance des fonctions de score et celle d'identifiabilité en dimension plus grande que 1. En cas de la stricte convexité de la fonction s , la Définition 1.1.12 se simplifie grâce à l'unicité du minimum. C'est ce cas qui va nous intéresser dans le reste de ce travail.

Il faut distinguer entre l'élicitabilité multivariée telle qu'elle est présentée dans la Définition 1.1.12 et la *k-élicitabilité* au sens Lambert et al (2008). Par exemple la variance est dite 2-élicitable car $\text{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2$ et les mesures $\mathbb{E}[X^2], \mathbb{E}[X]$ sont élicitables, ce qui n'implique pas forcément l'élicitabilité multivariée du vecteur $(\mathbb{E}[X^2], \mathbb{E}[X])$ au sens ci-dessus.

Dans le contexte d'allocation de capital, l'allocation peut être considérée une mesure de risque vectorielle avec $k = d$. Dans Maume-Deschamps et al. (2016) [33], une caractérisation axiomatique multivariée de la cohérence d'une allocation a été présentée, et qui concerne notamment les allocations par minimisation des indicateurs multivariés de risque. Cette mesure est alors élicitable dans le sens de la Définition 1.1.12.

La cohérence des mesures de risque multidimensionnelles a été étudiée dans la littérature de la théorie du risque. Jouini et al. (2004) [25] ont introduit des critères de cohérence des mesures de risque multidimensionnelles, leurs axiomes sont applicables aux mesures sous forme d'ensemble et non de vecteurs, il faut donc ajouter un autre critère pour définir une sorte d'optimalité au niveau de ces ensembles. La difficulté de la généralisation de la cohérence des mesures de risque, au sens Artzner et al. (1999) [2], en dimension supérieure à 1, réside dans la perte de l'ordre naturel. Si la propriété de la monotonie est généralisable à l'aide des définitions des ordres stochastiques multivariés, la propriété de la sous-additivité reste liée aux composantes des mesures de risque, ce qui complexifie la formalisation d'une généralisation valide pour les mesures vectorielles. Nous pensons que la notion d'élicitabilité peut aider à pallier ce problème, en introduisant une sorte d'ordre basé sur la comparaison des erreurs obtenues. Cela permettra non seulement de formaliser une généralisation à la sous-additivité pour mesures vectorielles, mais aussi à pouvoir étudier l'impact de la dépendance à l'aide des erreurs obtenues.

1.2 Constructions de mesures expectiles multivariées

L'importance statistique de la propriété de l'élicitabilité des mesures de risque nous invite à la prise en compte de cette propriété de cohérence dès leur construction. C'est le cas par exemple dans la définition des expectiles univariés. La même approche peut être utilisée pour construire des mesures de risque multidimensionnelles élicitables. Dans cette optique, nous présentons ici des extensions possibles des expectiles en dimension plus grande de 1.

1.2.1 Construction à l'aide des normes

L'idée la plus naturelle pour construire des mesures de risque multidimensionnelles élicitables est de partir de la définition univariée. En dimension 1, les expectiles sont définis pour tout seuil $\alpha \in [0, 1]$ par

$$e_\alpha(X) = \arg \min_{x \in \mathbb{R}} \mathbb{E}[\alpha(X - x)_+^2 + (1 - \alpha)(X - x)_-^2].$$

On note L^0 l'espace des vecteurs aléatoires sur un espace de probabilité $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs dans \mathbb{R}^d muni de la tribu borélienne réelle produit $\mathcal{B}(\mathbb{R})^{\otimes d}$. Nous considérons un vecteur aléatoire $\mathbf{X} = (X_1, \dots, X_d)^T \in L^0$.

Soit $\| \cdot \|$ une norme sur \mathbb{R}^d . On désigne par $(\mathbf{X})_+$ le vecteur composé des parties positives des composantes de \mathbf{X} , $(\mathbf{X})_+ = ((X_1)_+, \dots, (X_d)_+)^T$ et par $(\mathbf{X})_-$ le vecteur composé des parties négatives de ses composantes $(\mathbf{X})_- = ((X_1)_-, \dots, (X_d)_-)^T$. On définit la fonction de score

$$s_\alpha(\mathbf{X}, \mathbf{x}) = \alpha \| (\mathbf{X} - \mathbf{x})_+ \|^2 + (1 - \alpha) \| (\mathbf{x} - \mathbf{X})_+ \|^2,$$

pour tout $\mathbf{x} \in \mathbb{R}^d$.

On peut aussi ajouter une constante en définissant $\psi_{\mathbf{X}}^\alpha$ par $\psi_{\mathbf{X}}^\alpha(\mathbf{x}) = s_\alpha(\mathbf{X}, \mathbf{x}) - s_\alpha(\mathbf{X}, \mathbf{0})$, pour garantir $\psi_{\mathbf{X}}^\alpha(\mathbf{0}) = 0$. La fonction de score dans ce cas sera

$$\psi_{\mathbf{X}}^\alpha(\mathbf{x}) = \alpha \left(\| (\mathbf{X} - \mathbf{x})_+ \|^2 - \| (\mathbf{X})_+ \|^2 \right) + (1 - \alpha) \left(\| (\mathbf{x} - \mathbf{X})_+ \|^2 - \| (\mathbf{X})_- \|^2 \right),$$

pour tout $\mathbf{x} \in \mathbb{R}^d$. Le minimum reste le même.

On appelle *expectile multivarié* tout minimiseur

$$\mathbf{x}^* \in \arg \min_{\mathbf{x} \in \mathbb{R}^d} \mathbb{E}[s_\alpha(\mathbf{X}, \mathbf{x})].$$

Considérons l'espace d'Orlicz $M^{\|\cdot\|}$ associé à la norme $\| \cdot \|$ et défini comme suit

$$M^{\|\cdot\|} := \left\{ \mathbf{X} \in L^0 : \mathbb{E}[\| |\mathbf{X}| \|^2] < +\infty \right\},$$

où $|\mathbf{X}|$ désigne le vecteur aléatoire $(|X_1|, \dots, |X_d|)^T$.

On définit l'ensemble des normes

$$\mathcal{N} := \left\{ \| \cdot \| \mid \forall d \in \mathbb{N}^*, \forall (\mathbf{x}, \mathbf{y}) \in \mathbb{R}_+^d, \mathbf{x} \leq \mathbf{y} \Rightarrow \| \mathbf{x} \| \leq \| \mathbf{y} \| \right\},$$

tel que $\forall (\mathbf{x} = (x_1, \dots, x_d)^T, \mathbf{y} = (y_1, \dots, y_d)^T) \in \mathbb{R}_+^d, \mathbf{x} \leq \mathbf{y}$ si et seulement si $x_i \leq y_i$ pour tout $i \in \{1, \dots, d\}$. L'ensemble \mathcal{N} regroupe les normes qui conservent l'ordre naturel sur \mathbb{R}_+^d . Les normes L_p sont toutes dans cet ensemble.

On démontre que pour toute norme $\| \cdot \| \in \mathcal{N}$, pour tout $\mathbf{X} \in M^{\|\cdot\|}$, $\mathbb{E}[s_\alpha(\mathbf{X}, \mathbf{x})]$ est finie, positive et strictement convexe en \mathbf{x} . En effet pour $\| \cdot \| \in \mathcal{N}$

$$\| (\mathbf{X} - \mathbf{x})_+ \| \leq \| |\mathbf{X}| + |\mathbf{x}| \| \leq \| |\mathbf{X}| \| + \| |\mathbf{x}| \|,$$

donc

$$\mathbb{E}[\| (\mathbf{X} - \mathbf{x})_+ \|^2] \leq \mathbb{E}[(\| |\mathbf{X}| \| + \| |\mathbf{x}| \|)^2] \leq \mathbb{E}[2 \| |\mathbf{X}| \|^2] + \mathbb{E}[2 \| |\mathbf{x}| \|^2] < +\infty,$$

car la fonction $x \rightarrow x^2$ est convexe, de la même façon on démontre que $\mathbb{E}[\|(\mathbf{x} - \mathbf{X})_+\|^2] < +\infty$ pour déduire que l'espérance $\mathbb{E}[s_\alpha(\mathbf{X}, \mathbf{x})]$ est finie. Sa positivité est triviale. Pour sa stricte convexité, on considère la fonction $\phi : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$ définie par

$$\phi(\mathbf{t}, \mathbf{x}) = \alpha \|(\mathbf{t} - \mathbf{x})_+\|^2 + (1 - \alpha) \|(\mathbf{x} - \mathbf{t})_+\|^2,$$

et on démontre qu'elle est strictement convexe par rapport à \mathbf{x} . En effet, de la stricte convexité de la fonction $x \rightarrow x^2$ sur \mathbb{R} on a pour tout $\mathbf{t} \in \mathbb{R}^d$, et pour tous $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$ avec $\mathbf{x} \neq \mathbf{y}$, $\forall \lambda \in]0, 1[$

$$(\lambda \|(\mathbf{t} - \mathbf{x})_+\| + (1 - \lambda) \|(\mathbf{t} - \mathbf{y})_+\|)^2 < \lambda \|(\mathbf{t} - \mathbf{x})_+\|^2 + (1 - \lambda) \|(\mathbf{t} - \mathbf{y})_+\|^2,$$

l'inégalité triangulaire vérifiée par les normes permet d'en déduire

$$\|\lambda(\mathbf{t} - \mathbf{x})_+ + (1 - \lambda)(\mathbf{t} - \mathbf{y})_+\|^2 < \lambda \|(\mathbf{t} - \mathbf{x})_+\|^2 + (1 - \lambda) \|(\mathbf{t} - \mathbf{y})_+\|^2,$$

la convexité de la fonction $\psi_{\mathbf{t}}^+(\mathbf{x}) \rightarrow (\mathbf{t} - \mathbf{x})_+$ sur \mathbb{R}^d permet ensuite de conclure que

$$\|(\mathbf{t} - \lambda\mathbf{x} - (1 - \lambda)\mathbf{y})_+\|^2 < \lambda \|(\mathbf{t} - \mathbf{x})_+\|^2 + (1 - \lambda) \|(\mathbf{t} - \mathbf{y})_+\|^2,$$

pour toute norme $\|\cdot\| \in \mathcal{N}$. La convexité de la fonction $\psi_{\mathbf{t}}^-(\mathbf{x}) \rightarrow (\mathbf{x} - \mathbf{t})_+$ sur \mathbb{R}^d permet d'obtenir de la même façon

$$\|(\mathbf{t} - \lambda\mathbf{x} - (1 - \lambda)\mathbf{y})_-\|^2 < \lambda \|(\mathbf{t} - \mathbf{x})_-\|^2 + (1 - \lambda) \|(\mathbf{t} - \mathbf{y})_-\|^2,$$

d'où la stricte convexité de la fonction $\phi(\mathbf{t}, \mathbf{x})$ par rapport à \mathbf{x} et par suite la stricte convexité de $\mathbb{E}[\phi(\mathbf{X}, \mathbf{x})] = \mathbb{E}[s_\alpha(\mathbf{X}, \mathbf{x})]$ en \mathbf{x} . Le minimiseur dans ce cas est unique, et on peut donc définir l'expectile multivarié du vecteur \mathbf{X} de seuil α par

$$\mathbf{e}_\alpha(\mathbf{X}) = \arg \min_{\mathbf{x} \in \mathbb{R}^d} \mathbb{E}[\alpha \|(\mathbf{X} - \mathbf{x})_+\|^2 + (1 - \alpha) \|(\mathbf{x} - \mathbf{X})_+\|^2]. \quad (1.4)$$

Cette mesure de risque est alors élicitable par construction. Le choix d'un seuil α commun à toutes les coordonnées du vecteur aléatoire \mathbf{X} est naturel en gestion des risques. Le niveau du risque accepté doit être le même pour toutes les branches.

Cette construction peut aussi être utilisée pour définir des M-quantiles multivariés si on considère des fonctions de score de la forme

$$s_\alpha(\mathbf{X}, \mathbf{x}) = \alpha \|(\mathbf{X} - \mathbf{x})_+\|^m + (1 - \alpha) \|(\mathbf{x} - \mathbf{X})_+\|^m,$$

pour $m \in \mathbb{N}^*$.

Pour illustrer cette construction, nous présentons deux exemples possibles de familles d'expectiles.

Exemple 1 *Expectiles euclidiens (L^p -expectiles)*

On peut utiliser une norme L^p sur \mathbb{R}^d , $\|\cdot\|_p$, $1 \leq p \leq \infty$, dans ce cas la Définition 1.4 de l'expectile devient

$$\mathbf{e}_\alpha(\mathbf{X}) = \arg \min_{\mathbf{x} \in \mathbb{R}^d} \mathbb{E}[\alpha \|(\mathbf{X} - \mathbf{x})_+\|_p^2 + (1 - \alpha) \|(\mathbf{x} - \mathbf{X})_+\|_p^2].$$

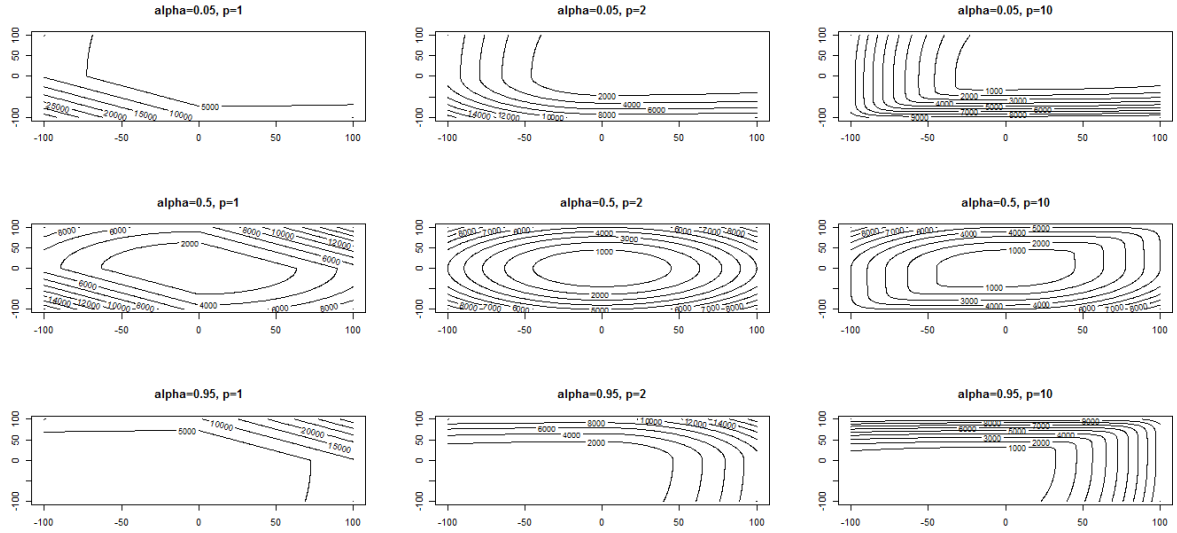
Pour une norme $\|\cdot\|_p$ avec $p < +\infty$,

$$\mathbf{e}_\alpha(\mathbf{X}) = \arg \min_{\mathbf{x} \in \mathbb{R}^d} \mathbb{E} \left[\alpha \left(\sum_{i=1}^d (X_i - x_i)_+^p \right)^{2/p} + (1 - \alpha) \left(\sum_{i=1}^d (x_i - X_i)_+^p \right)^{2/p} \right].$$

Dans la figure 1.1, nous traçons en dimension 2 les courbes de niveaux de la fonction de score

$$\psi(z) = \alpha \|z\|_p^2 + (1 - \alpha) \|z\|_p^2,$$

pour différents niveaux de seuil $\alpha = 0.05, 0.5, 0.95$ et pour les normes L_1, L_2 et L_{10} .

FIGURE 1.1 – Lignes de niveaux de la fonction de score L_p -expectile.

Le cas du L^p -expectile avec $p = 2$ est le seul qui ne prend pas en compte la dépendance. Le L^2 -expectile est le vecteur composé des expectiles univariés des composantes

$$\begin{aligned}
 \mathbf{e}_\alpha(\mathbf{X}) &= \arg \min_{\mathbf{x} \in \mathbb{R}^d} \mathbb{E} \left[\alpha \left(\sum_{i=1}^d (X_i - x_i)_+^2 \right) + (1 - \alpha) \left(\sum_{i=1}^d (x_i - X_i)_+^2 \right) \right] \\
 &= \arg \min_{\mathbf{x} \in \mathbb{R}^d} \mathbb{E} \left[\sum_{i=1}^d \left(\alpha (X_i - x_i)_+^2 + (1 - \alpha) (x_i - X_i)_+^2 \right) \right] \\
 &= (e_\alpha(X_1), \dots, e_\alpha(X_d))^T.
 \end{aligned}$$

De manière générale, pour $p \geq 1$, on peut déterminer la solution optimale à l'aide de la condition du premier ordre d'optimisation :

$$\begin{aligned}
 \nabla(\mathbb{E}[s_\alpha(\mathbf{X}, \mathbf{x})]) &= \mathbb{E} \left[\frac{\partial \left(\alpha \|\mathbf{X} - \mathbf{x}\|_p^2 + (1 - \alpha) \|\mathbf{x} - \mathbf{X}\|_p^2 \right)}{\partial \mathbf{x}} \right] \\
 &= \alpha \mathbb{E} \left[\frac{\partial \left(\|\mathbf{X} - \mathbf{x}\|_p^2 \right)}{\partial \mathbf{x}} \right] + (1 - \alpha) \mathbb{E} \left[\frac{\partial \left(\|\mathbf{x} - \mathbf{X}\|_p^2 \right)}{\partial \mathbf{x}} \right] \\
 &= \alpha \mathbb{E} \left[2 \|\mathbf{X} - \mathbf{x}\|_p \frac{\partial \|\mathbf{X} - \mathbf{x}\|_p}{\partial \mathbf{x}} \right] \\
 &\quad + (1 - \alpha) \mathbb{E} \left[2 \|\mathbf{x} - \mathbf{X}\|_p \frac{\partial \|\mathbf{x} - \mathbf{X}\|_p}{\partial \mathbf{x}} \right] \\
 &= 2\alpha \mathbb{E} \left[\frac{\|\mathbf{X} - \mathbf{x}\|_p}{\|\mathbf{X} - \mathbf{x}\|_p^{p-1}} \frac{\partial \|\mathbf{X} - \mathbf{x}\|_p}{\partial \mathbf{x}} (\mathbf{X} - \mathbf{x})_+^{p-1} \right] \\
 &\quad + 2(1 - \alpha) \mathbb{E} \left[\frac{\|\mathbf{x} - \mathbf{X}\|_p}{\|\mathbf{x} - \mathbf{X}\|_p^{p-1}} \frac{\partial \|\mathbf{x} - \mathbf{X}\|_p}{\partial \mathbf{x}} (\mathbf{x} - \mathbf{X})_+^{p-1} \right],
 \end{aligned}$$

car $\frac{\partial \|\mathbf{x}\|_p}{\partial \mathbf{x}} = \frac{\mathbf{x} \circ |\mathbf{x}|^{p-2}}{\|\mathbf{x}\|_p^{p-2}}$ pour tout $\mathbf{x} \in \mathbb{R}^2$, où \circ présente le produit Hadamard, c'est-à-dire le produit terme à terme des deux matrices. $|\cdot|$ désigne le vecteur composé des valeurs absolues des composantes et $(\cdot)^k$ le vecteur des $k^{\text{ème}}$ puissances des composantes. La matrice $\frac{\partial \|\mathbf{X} - \mathbf{x}\|_p}{\partial \mathbf{x}}$ est la matrice carrée composée des termes $\left(\frac{\partial (X_i - x_i)_+}{\partial x_i} \right)$, elle est donc diagonale et son vecteur diagonal est $(-\mathbb{1}_{\{X_i > x_i\}})_{i=1, \dots, d}$. De même, la matrice $\frac{\partial \|\mathbf{x} - \mathbf{X}\|_p}{\partial \mathbf{x}}$ est la matrice carrée diagonale

dont la diagonale est $(\mathbb{1}_{\{X_i < x_i\}})_{i=1, \dots, d}$.

Pour chaque composante x_k on peut donc écrire

$$\begin{aligned} \nabla_k(\mathbb{E}[s_\alpha(\mathbf{X}, \mathbf{x})]) &= -2\alpha \mathbb{E} \left[\frac{\|(\mathbf{X} - \mathbf{x})_+\|_p}{\|(\mathbf{X} - \mathbf{x})_+\|_p^{p-1}} (X_k - x_k)_+^{p-1} \mathbb{1}_{\{X_k > x_k\}} \right] \\ &\quad + 2(1 - \alpha) \mathbb{E} \left[\frac{\|(\mathbf{x} - \mathbf{X})_+\|_p}{\|(\mathbf{x} - \mathbf{X})_+\|_p^{p-1}} (x_k - X_k)_+^{p-1} \mathbb{1}_{\{X_k < x_k\}} \right]. \end{aligned}$$

La condition de premier ordre s'écrit donc comme suit,

$$\alpha \mathbb{E} \left[\frac{\|(\mathbf{X} - \mathbf{x})_+\|_p}{\|(\mathbf{X} - \mathbf{x})_+\|_p^{p-1}} \frac{\partial(\mathbf{X} - \mathbf{x})_+}{\partial \mathbf{x}} (\mathbf{X} - \mathbf{x})_+^{p-1} \right] = (1 - \alpha) \mathbb{E} \left[\frac{\|(\mathbf{x} - \mathbf{X})_+\|_p}{\|(\mathbf{x} - \mathbf{X})_+\|_p^{p-1}} \frac{\partial(\mathbf{x} - \mathbf{X})_+}{\partial \mathbf{x}} (\mathbf{x} - \mathbf{X})_+^{p-1} \right].$$

L'expectile est l'unique solution du système composé de d équations

$$\alpha \mathbb{E} \left[\frac{\|(\mathbf{X} - \mathbf{x})_+\|_p}{\|(\mathbf{X} - \mathbf{x})_+\|_p^{p-1}} (X_k - x_k)_+^{p-1} \mathbb{1}_{\{X_k > x_k\}} \right] = (1 - \alpha) \mathbb{E} \left[\frac{\|(\mathbf{x} - \mathbf{X})_+\|_p}{\|(\mathbf{x} - \mathbf{X})_+\|_p^{p-1}} (x_k - X_k)_+^{p-1} \mathbb{1}_{\{X_k < x_k\}} \right], \quad (1.5)$$

pour tout $k \in \{1, \dots, d\}$.

Pour $2 < p < +\infty$, le système d'équations (1.5) peut s'écrire sous la forme suivante

$$\alpha \mathbb{E} \left[\frac{(X_k - x_k)_+^{p-1}}{\|(\mathbf{X} - \mathbf{x})_+\|_p^{p-2}} \right] = (1 - \alpha) \mathbb{E} \left[\frac{(X_k - x_k)_-^{p-1}}{\|(\mathbf{X} - \mathbf{x})_-\|_p^{p-2}} \right] \quad \forall k \in \{1, \dots, d\},$$

la dépendance est donc bien prise en compte à l'aide de la norme, et pour visualiser l'interprétation économique de l'expectile multivarié, il suffit d'écrire pour chaque risque X_k

$$\frac{\mathbb{E} \left[(X_k - x_k)_+ \left(\frac{(X_k - x_k)_+}{\|(\mathbf{X} - \mathbf{x})_+\|_p} \right)^{p-2} \right]}{\mathbb{E} \left[(X_k - x_k)_- \left(\frac{(X_k - x_k)_-}{\|(\mathbf{X} - \mathbf{x})_-\|_p} \right)^{p-2} \right]} = \frac{1 - \alpha}{\alpha}.$$

Les composantes de l'expectile multivarié donnent pour chaque X_k un même ratio gains relatives / pertes "relatives". Le terme relatif traduit la présence des poids $\left(\frac{(X_k - x_k)_+}{\|(\mathbf{X} - \mathbf{x})_+\|_p} \right)^{p-2}$ qui représentent une puissance de poids des gains de X_k dans la norme des gains totaux et $\left(\frac{(X_k - x_k)_-}{\|(\mathbf{X} - \mathbf{x})_-\|_p} \right)^{p-2}$ qui représente une puissance de poids des pertes de X_k dans la perte globale. Nous signalons que cette interprétation considère les variables aléatoires X_i représentatives des *Profit and Loss* (P&L) et pas des risques. Cette interprétation est plus compréhensible dans le cas $p=3$.

Le cas $p = 1$ est intéressant, une définition équivalente du L^1 -expectile découle de la condition d'optimalité du premier ordre (1.5). Le L^1 -expectile est l'unique solution dans \mathbb{R}^d du système suivant

$$\alpha \mathbb{E}[\|(\mathbf{X} - \mathbf{x})_+\|_1 \mathbb{1}_{\{X_k > x_k\}}] = (1 - \alpha) \mathbb{E}[\|(\mathbf{X} - \mathbf{x})_-\|_1 \mathbb{1}_{\{X_k < x_k\}}], \quad \forall k \in \{1, \dots, d\}.$$

ou de manière équivalente

$$\frac{\mathbb{E}[\|(\mathbf{X} - \mathbf{x})_+\|_1 \mathbb{1}_{\{X_k > x_k\}}]}{\mathbb{E}[\|(\mathbf{X} - \mathbf{x})_-\|_1 \mathbb{1}_{\{X_k < x_k\}}]} = \frac{1 - \alpha}{\alpha}, \quad \forall k \in \{1, \dots, d\}, \quad (1.6)$$

ce qui s'interprète comme un ratio de la participation d'une activité dans les scénarios positifs sur celui de sa participation dans les scénarios économiquement négatifs. L'expectile multivarié dans ce cas représente le vecteur des niveaux qui permettent d'équilibrer le portefeuille en fixant le même ratio pour toutes les activités. Le L^1 -expectile est le vecteur des seuils de pertes qui permettent de fixer l'impact d'une activité sur un ratio de ruine/solvabilité globale en utilisant la norme L^1 comme outil d'agrégation. L'assureur a intérêt dans ce cas à minimiser ce ratio $\beta =$

$(1 - \alpha)/\alpha$, ce qui revient à augmenter le niveau α . De ce point de vue, l'expectile multivarié peut constituer un bon outil d'allocation du capital. Une dernière utilisation possible en réassurance concerne la mesure de l'impact d'un contrat de type Stop-Loss sur l'ensemble des contrats de même type couverts par le réassureur. L'expectile obtenu permet à un réassureur de fixer une proposition de priorité pour une offre d'un nouveau contrat en maintenant la stabilité de son portefeuille global.

L'Équation 1.6 peut se réécrire à l'aide des espérances bivariées,

$$\alpha \sum_{i=1}^d \mathbb{E}[(X_i - x_i)_+ \mathbb{1}_{\{X_k > x_k\}}] = (1 - \alpha) \sum_{i=1}^d \mathbb{E}[(X_i - x_i)_- \mathbb{1}_{\{X_k < x_k\}}], \forall k \in \{1, \dots, d\}. \quad (1.7)$$

La prise en compte de la dépendance entre les risques se limite au niveau bivarié. Le système (1.7) peut s'écrire à l'aide des fonctions Stop-Loss bivariées transformées, introduites et étudiées avec des applications en actuariat dans Hürlimann (2002) [24].

Pour la norme $\max \|\cdot\|_\infty$

$$\mathbf{e}_\alpha(\mathbf{X}) = \arg \min_{\mathbf{x} \in \mathbb{R}^d} \mathbb{E}[\alpha \max_i \{(X_i - x_i)_+^2\} + (1 - \alpha) \max_i \{(X_i - x_i)_-^2\}].$$

La différentiabilité dans ce cas est perdue. Ce cas ne sera pas traité dans cette thèse.

Exemple 2 Expectiles matriciels

On peut aussi construire des expectiles à l'aide des matrices qui permettent de définir des formes quadratiques sur \mathbb{R}^d . Un exemple simple, basé sur le produit matriciel, sera l'utilisation d'une matrice réelle symétrique Σ de poids. La fonction score dans ce cas peut s'écrire sous la forme suivante

$$s_\alpha(\mathbf{X}, \mathbf{x}) = \alpha(\mathbf{X} - \mathbf{x})_+^T \Sigma (\mathbf{X} - \mathbf{x})_+ + (1 - \alpha)(\mathbf{X} - \mathbf{x})_-^T \Sigma (\mathbf{X} - \mathbf{x})_-,$$

dans le cas de la stricte convexité de $\mathbb{E}[s_\alpha(\mathbf{X}, \mathbf{x})]$ en \mathbf{x} l'expectile multivarié est unique

$$\mathbf{e}_\alpha(\mathbf{X}) = \arg \min_{\mathbf{x} \in \mathbb{R}^d} \mathbb{E}[\alpha(\mathbf{X} - \mathbf{x})_+^T \Sigma (\mathbf{X} - \mathbf{x})_+ + (1 - \alpha)(\mathbf{X} - \mathbf{x})_-^T \Sigma (\mathbf{X} - \mathbf{x})_-].$$

Plusieurs matrices usuelles peuvent être utilisées pour construire des expectiles

- Expectiles de Mahalanobis : Dans le cas où Σ est la matrice inverse de la matrice de covariances, cela revient à utiliser la distance de Mahalanobis, on parlera alors d'expectile de Mahalanobis. Le principal problème de cette construction est que l'inverse de la matrice de covariances n'existe pas toujours. Elle est plus adaptée à des vecteurs aléatoires de lois marginales elliptiques.
- Expectiles corrélés : Cette construction se base sur l'utilisation d'une matrice de corrélations. On peut aussi utiliser d'autres mesures de dépendance entre variables aléatoires comme coefficients de la matrice de construction.

Notons que le L_1 -Expectile défini précédemment est aussi un expectile matriciel : on prend $(\Sigma)_{ij} = 1 \forall i, j$.

1.2.2 Construction à l'aide de l'optimalité et des ensembles d'acceptation

L'idée de ce type de construction est de définir les expectiles multivariés à l'aide d'un système d'équations de solution unique, qui correspond dans le cas de la construction par normes à la condition d'optimalité de premier ordre.

On a vu, par exemple, que le L_1 -expectile est défini par

$$\frac{\mathbb{E}[\|(\mathbf{X} - \mathbf{x})_+\|_1 \mathbb{1}_{\{X_k > x_k\}}]}{\mathbb{E}[\|(\mathbf{X} - \mathbf{x})_-\|_1 \mathbb{1}_{\{X_k < x_k\}}]} = \frac{1 - \alpha}{\alpha}, \forall k \in \{1, \dots, d\}.$$

L'interprétation économique de cette présentation nous donne l'idée de construire des mesures à l'aide d'une interprétation donnée, et en introduisant des fonctions qui permettent à la fois de modéliser la dépendance et garantir l'unicité. L'idée principale de cette méthode est de partir de l'interprétation économique, ou d'un besoin particulier de modélisation pour construire la mesure de risque.

Exemple 1 Dans le cas du choix d'une norme quelconque d'agrégation qui dépend de la nature du portefeuille des risques, on peut chercher les seuils de réserves allouées aux différents risques supportés qui stabilise le ratio de la participation de d'une activité dans les pertes agrégées par rapport à sa participation dans les gains agrégés. Ce besoin se traduit par la recherche des solutions du système suivant

$$\frac{\mathbb{E}[\|(\mathbf{X} - \mathbf{x})_+ \| \mathbb{1}_{\{X_k > x_k\}}]}{\mathbb{E}[\|(\mathbf{X} - \mathbf{x})_- \| \mathbb{1}_{\{X_k < x_k\}}]} = \frac{1 - \alpha}{\alpha}, \forall k \in \{1, \dots, d\},$$

avec $\| \cdot \|$ la norme d'agrégation choisie.

Exemple 2 Dans le cas d'un portefeuille d'assurance de plusieurs activités, on veut stabiliser le ratio des gains en cas de solvabilité de toutes les activités et pertes en cas d'insolvabilité globale, pour toutes les activités. Cela se traduit par une allocation qui vérifie la contrainte suivante

$$\frac{\mathbb{E}[(X_k - x_k)_+ \mathbb{1}_{\{X_1 > x_1, \dots, X_d > x_d\}}]}{\mathbb{E}[(X_k - x_k)_- \mathbb{1}_{\{X_1 < x_1, \dots, X_d < x_d\}}]} = \frac{1 - \alpha}{\alpha}, \forall k \in \{1, \dots, d\}.$$

Exemple 3 Dans le cas de possibilité de transfert des capitaux entre les différentes activités d'un assureur, son état d'insolvabilité est celui du groupe représenté par la somme des risques. Si on veut stabiliser le ratio de la participation de la ruine locale de chaque activité dans la ruine globale par rapport à la participation de sa solvabilité locale dans la solvabilité du groupe, cela se traduit par le système d'équations suivant

$$\frac{\mathbb{E} \left[(X_k - x_k)_+ \mathbb{1}_{\left\{ \sum_{k=1}^d X_k > \sum_{k=1}^d x_k \right\}} \right]}{\mathbb{E} \left[(X_k - x_k)_- \mathbb{1}_{\left\{ \sum_{k=1}^d X_k < \sum_{k=1}^d x_k \right\}} \right]} = \frac{1 - \alpha}{\alpha}, \forall k \in \{1, \dots, d\}.$$

L'utilité donc de cette méthode est de permettre à son utilisateur de construire des expectiles selon une interprétation économique précise, mais son principal inconvénient est que l'on n'est plus assuré d'avoir l'élicitabilité. Pour pallier ce problème, d'autres solutions peuvent être proposées en partant de l'idée initiale des expectiles, et qui consistent, par exemple pour les trois précédentes illustrations, à définir respectivement les expectiles suivants

$$\mathbf{e}_\alpha^{\|\cdot\|}(\mathbf{X}) \in \arg \min_{\mathbf{x} \in \mathbb{R}^d} \mathbb{E} \left[\alpha \sum_{i=1}^d \|(\mathbf{X} - \mathbf{x})_+\|^2 \mathbb{1}_{\{X_i > x_i\}} + (1 - \alpha) \sum_{i=1}^d \|(\mathbf{X} - \mathbf{x})_-\|^2 \mathbb{1}_{\{X_i < x_i\}} \right],$$

$$\mathbf{e}_\alpha^{And}(\mathbf{X}) \in \arg \min_{\mathbf{x} \in \mathbb{R}^d} \mathbb{E} \left[\alpha \sum_{i=1}^d (X_i - x_i)_+^2 \mathbb{1}_{\{X_1 > x_1, \dots, X_d > x_d\}} + (1 - \alpha) \sum_{i=1}^d (x_i - X_i)_+^2 \mathbb{1}_{\{X_1 < x_1, \dots, X_d < x_d\}} \right],$$

$$\mathbf{e}_\alpha^{Sum}(\mathbf{X}) \in \arg \min_{\mathbf{x} \in \mathbb{R}^d} \mathbb{E} \left[\alpha \sum_{i=1}^d (X_i - x_i)_+^2 \mathbb{1}_{\left\{ \sum_{k=1}^d X_k > \sum_{k=1}^d x_k \right\}} + (1 - \alpha) \sum_{i=1}^d (x_i - X_i)_+^2 \mathbb{1}_{\left\{ \sum_{k=1}^d X_k < \sum_{k=1}^d x_k \right\}} \right].$$

Cette méthode se base sur la même idée des constructions univariées par ensemble d'acceptation (*acceptance set*). Pour les expectiles en dimension 1, l'ensemble d'acceptation peut être écrit sous la forme suivante

$$\mathcal{A}_{e_\alpha} = \left\{ X : \frac{\mathbb{E}[(X)_+]}{\mathbb{E}[(X)_-]} \geq \frac{1 - \alpha}{\alpha} \right\}.$$

Cette notion a été utilisée dans un contexte multidimensionnel par Jouini et al. (2004) [25] pour définir des ensembles cohérents de mesures de risque vectorielles, et par Hamel et Heyde (2010) [23] pour donner une première représentation des mesures de risque définies sous forme d'ensembles (*set-valued*

risk measures). Dans notre contexte, et dans le cas de la perte de l'unicité, la mesure obtenue est aussi un ensemble de vecteurs qui peut être vu comme un ensemble d'acceptation multidimensionnel.

On pourrait proposer aussi une construction géométrique analogue à celle des quantiles géométriques. L'approche proposée par Chaudhuri (1996) [8] et celle introduite par Abdous et Theodorescu (1992) [1] peuvent être utilisées pour définir des expectiles géométriques. Cependant l'interprétation de ces mesures ne paraît pas naturelle.

Finalement, Le choix d'une construction parmi l'ensemble des méthodes proposées ne peut dépendre que du besoin de l'utilisation. L'utilisation par exemple d'une matrice de corrélation pour construire un expectile traduit un choix particulier de modélisation de la dépendance entre les risques. Une matrice composée de mesures de dépendance des queues par exemple permettra davantage la prise en compte de la dépendance extrême. C'est pour cela que nous pensons que cette question doit être laissée ouverte à l'utilisateur en fonction de ses besoins d'application. Toutefois, la construction la plus simple reste celle des expectiles matriciels à l'aide de matrices semi-définies positives à coefficients positifs. La condition d'optimalité est simple à obtenir sous forme d'un système d'équations qui relie des transformations des fonctions Stop-Loss, et la stricte convexité est garantie par la positivité des coefficients. Partant de ce constat, nous allons poursuivre cette étude avec ce choix de construction, qui sera fixé dès la prochaine section. Cependant, les propriétés des expectiles dépendent essentiellement de leur construction, l'étude des propriétés des autres méthodes fera le sujet d'un futur travail.

1.3 Propriétés de cohérence

Les exemples présentés précédemment montrent que les expectiles multivariés proposés peuvent être utilisés dans différents contextes de gestion des risques. Le choix de la méthode de construction dépendra de la vision du risque ainsi que l'objectif de couverture derrière cette utilisation. Dans la présente section, nous présentons les propriétés de la construction générale des expectiles à l'aide des matrices. Ces propriétés dépendront directement de la nature de la matrice utilisée.

Nous considérons une construction à l'aide d'une matrice symétrique réelle semi-définie positive. Ce cas présente un intérêt pratique car il permet tout simplement de modéliser la dépendance à l'aide de mesures bivariées de dépendance.

Définition 1.3.1 (Expectile multivarié). Soit $\mathbf{X} = (X_1, \dots, X_d)^T \in \mathbb{R}^d$ un vecteur aléatoire tel que $\mathbb{E}[|X_i X_j|] < +\infty$ pour tout $(i, j) \in \{1, \dots, d\}^2$, et $\Sigma = (\pi_{ij})_{1 \leq i, j \leq d}$ une matrice réelle carrée d'ordre d symétrique et semi-définie positive qui vérifie

1. Pour tout $i \in \{1, \dots, d\}$, $\pi_{ii} = \pi_i > 0$;
2. Pour tout $i, j \in \{1, \dots, d\}$, $\pi_{ii} \geq \pi_{ij}$.

On appelle un Σ -expectile de \mathbf{X} , tout vecteur qui vérifie

$$\mathbf{e}_\alpha^\Sigma(\mathbf{X}) \in \arg \min_{\mathbf{x} \in \mathbb{R}^d} \mathbb{E}[\alpha(\mathbf{X} - \mathbf{x})_+^T \Sigma (\mathbf{X} - \mathbf{x})_+ + (1 - \alpha)(\mathbf{X} - \mathbf{x})_-^T \Sigma (\mathbf{X} - \mathbf{x})_-],$$

dans le cas de l'unicité, on appelle l'expectile multivarié de \mathbf{X} , le vecteur défini par

$$\mathbf{e}_\alpha^\Sigma(\mathbf{X}) = \arg \min_{\mathbf{x} \in \mathbb{R}^d} \mathbb{E}[\alpha(\mathbf{X} - \mathbf{x})_+^T \Sigma (\mathbf{X} - \mathbf{x})_+ + (1 - \alpha)(\mathbf{X} - \mathbf{x})_-^T \Sigma (\mathbf{X} - \mathbf{x})_-]. \quad (1.8)$$

Le caractère semi-définie positive de la matrice est nécessaire pour obtenir une forme quadratique positive ($\mathbf{x}^T \Sigma \mathbf{x}$). La matrice de construction Σ permet la prise en compte de la dépendance bivariée entre les composantes aléatoires du vecteur \mathbf{X} .

L'erreur minimisée a une borne inférieure

$$\mathbb{E}[\alpha(\mathbf{X} - \mathbf{x})_+^T \Sigma (\mathbf{X} - \mathbf{x})_+ + (1 - \alpha)(\mathbf{X} - \mathbf{x})_-^T \Sigma (\mathbf{X} - \mathbf{x})_-] \geq \lambda_{\min} \mathbb{E}[\alpha(\mathbf{X} - \mathbf{x})_+^T (\mathbf{X} - \mathbf{x})_+ + (1 - \alpha)(\mathbf{X} - \mathbf{x})_-^T (\mathbf{X} - \mathbf{x})_-],$$

où λ_{\min} est la plus petite valeur propre de la matrice Σ .

Cette construction permet d'obtenir un cadre général qui englobe plusieurs exemples, notamment les expectiles que nous nommerons *expectiles de corrélation* si on utilise la matrice de corrélations, l'*expectile de Mahalanobis* si la matrice utilisée est l'inverse de la matrice de covariance, et l'*expectile euclidien* L^1 si Σ est la matrice $\Sigma_{ij} = 1, \forall i, j$.

Afin d'assurer l'unicité du vecteur expectile, nous considérons uniquement le cas de matrice à coefficients positifs. En effet, dans le cas où Σ est définie positive, on peut définir la norme $\|\cdot\|$ sur \mathbb{R}^d par

$$\|\mathbf{x}\| = \sqrt{\mathbf{x}^T \Sigma \mathbf{x}}, \quad \forall \mathbf{x} \in \mathbb{R}^d,$$

la positivité des coefficients de la matrice Σ est suffisante pour avoir $\|\cdot\| \in \mathcal{N}$ et donc pour obtenir la stricte convexité de la fonction score. Si la matrice Σ est uniquement semi-définie positive, $\mathbf{x} \rightarrow \sqrt{\mathbf{x}^T \Sigma \mathbf{x}}$ n'est plus une norme, mais la stricte convexité de la fonction score reste vérifiée car

$$\mathbf{x} \leq \mathbf{y} \Rightarrow \mathbf{x}^T \Sigma \mathbf{x} \leq \mathbf{y}^T \Sigma \mathbf{y}, \quad \forall \mathbf{x} \leq \mathbf{y} \in \mathbb{R}_+^d.$$

Sous l'hypothèse de la positivité des coefficients de la matrice de construction, l'expectile multivarié est finalement l'unique solution du système d'équations suivant,

$$\alpha \sum_{i=1}^d \pi_{ki} \mathbb{E}[(X_i - x_i)_+ \mathbb{1}_{\{X_k > x_k\}}] = (1 - \alpha) \sum_{i=1}^d \pi_{ki} \mathbb{E}[(x_i - X_i)_+ \mathbb{1}_{\{x_k > X_k\}}], \quad \forall k \in \{1, \dots, d\}, \quad (1.9)$$

équivalent à

$$\alpha = \frac{\sum_{i=1}^d \pi_{ki} \mathbb{E}[(x_i - X_i)_+ \mathbb{1}_{\{x_k > X_k\}}]}{\sum_{i=1}^d \pi_{ki} \left(\mathbb{E}[(X_i - x_i)_+ \mathbb{1}_{\{X_k > x_k\}}] + \mathbb{E}[(x_i - X_i)_+ \mathbb{1}_{\{x_k > X_k\}}] \right)}, \quad \forall k \in \{1, \dots, d\}, \quad (1.10)$$

et à

$$\frac{\sum_{i=1}^d \pi_{ki} \mathbb{E}[(X_i - x_i)_+ \mathbb{1}_{\{X_k > x_k\}}]}{\sum_{i=1}^d \pi_{ki} \mathbb{E}[(x_i - X_i)_+ \mathbb{1}_{\{x_k > X_k\}}]} = \frac{1 - \alpha}{\alpha}, \quad \forall k \in \{1, \dots, d\},$$

il peut s'écrire sous la forme suivante

$$\sum_{i=1}^d \pi_{ki} x_i \mathbb{E}[Z_{X_i, X_k}^\alpha(x_i, x_k)] = \sum_{i=1}^d \pi_{ki} \mathbb{E}[X_i Z_{X_i, X_k}^\alpha(x_i, x_k)], \quad \forall k \in \{1, \dots, d\}$$

où $Z_{X_i, X_k}^\alpha(x_i, x_k) = \alpha \mathbb{1}_{\{X_i > x_i, X_k > x_k\}} + (1 - \alpha) \mathbb{1}_{\{X_i < x_i, X_k < x_k\}}$.

Le système d'optimalité (1.9) s'écrit aussi

$$\alpha \sum_{i=1}^d \pi_{ki} \int_{x_i}^{+\infty} \mathbb{P}(X_i > t, X_j > x_j) dt = (1 - \alpha) \sum_{i=1}^d \pi_{ki} \int_{-\infty}^{x_i} \mathbb{P}(X_i < t, X_j < x_j) dt, \quad \forall k \in \{1, \dots, d\}. \quad (1.11)$$

On peut aussi l'écrire

$$\alpha \mathbb{E}[\mathbb{1}_{\{\mathbf{X} > \mathbf{x}\}} \Sigma (\mathbf{X} - \mathbf{x})_+] = (1 - \alpha) \mathbb{E}[(I_d - \mathbb{1}_{\{\mathbf{X} > \mathbf{x}\}}) \Sigma (\mathbf{X} - \mathbf{x})_-], \quad (1.12)$$

où $\mathbb{1}_{\{\mathbf{X} > \mathbf{x}\}}$ est la matrice définie par $(\mathbb{1}_{\{\mathbf{X} > \mathbf{x}\}})_{ij} = \delta_{ij} \mathbb{1}_{\{X_i > x_i\}}$, et I_d la matrice identité d'ordre d .

L'expectile multivarié vérifie plusieurs propriétés de cohérence. Nous présentons une partie de ces propriétés dans cette section. À notre connaissance, aucune caractérisation axiomatique de cohérence n'a été introduite pour des mesures de risque vectorielles. La généralisation de la propriété de la sous-additivité reste la tâche la plus délicate en l'absence d'un ordre naturel des vecteurs.

Dans la suite de ce document, on désigne par expectile, les mesures de risque construites uniquement au sens de (1.3.1). Les preuves des trois items de la proposition suivante sont immédiates en utilisant (1.10).

Proposition 1.3.2. *Pour tout vecteur aléatoire $\mathbf{X} = (X_1, \dots, X_d)^T$ d'ordre 2 sur \mathbb{R}^d ,*

(i) *Homogénéité positive : Pour toute constante réelle positive a ,*

$$\mathbf{e}_\alpha(a\mathbf{X}) = a\mathbf{e}_\alpha(\mathbf{X}).$$

(ii) *Invariance par translation : Pour tout vecteur $\mathbf{a} = (a_1, \dots, a_d)^T$ dans \mathbb{R}^d ,*

$$\mathbf{e}_\alpha(\mathbf{X} + \mathbf{a}) = \mathbf{e}_\alpha(\mathbf{X}) + \mathbf{a}.$$

(iii) *Invariance par loi : Pour tout vecteur aléatoire $Y = (Y_1, \dots, Y_d)^T$ d'ordre 2 dans \mathbb{R}^d tel que pour tout $(i, j) \in \{1, \dots, d\}^2$, on a $(X_i, X_j) \sim (Y_i, Y_j)$, alors*

$$\forall \alpha \in [0, 1], \quad \mathbf{e}_\alpha(X) = \mathbf{e}_\alpha(Y).$$

Proposition 1.3.3 (Pseudo-invariance par transformations linéaires). *On désigne par \mathbf{e}_α^Σ l'expectile construit à l'aide de la matrice Σ . Pour tout vecteur aléatoire $\mathbf{X} = (X_1, \dots, X_d)^T$ d'ordre 2 sur \mathbb{R}^d et pour tous vecteurs $\mathbf{a} = (a_1, \dots, a_d)^T \in \mathbb{R}^d$ et $\mathbf{b} = (b_1, \dots, b_d)^T \in (\mathbb{R}^{+*})^d$,*

$$\mathbf{e}_\alpha^\Sigma(V\mathbf{X} + \mathbf{a}) = V\mathbf{e}_\alpha^{V\Sigma V}(\mathbf{X}) + \mathbf{a},$$

où V est la matrice carrée diagonale associée au vecteur \mathbf{b} , $V = \text{Diag}(\mathbf{b}^T)$.

Preuve En utilisant la Proposition 1.3.2-(ii), il suffit de montrer que

$$\mathbf{e}_\alpha^\Sigma(V\mathbf{X}) = V\mathbf{e}_\alpha^{V\Sigma V}(\mathbf{X}).$$

La matrice $V = \text{Diag}(\mathbf{b}^T)$ est définie positive, elle est notamment inversible, donc par définition de l'expectile multivarié

$$\begin{aligned} \mathbf{e}_\alpha^\Sigma(V\mathbf{X}) &= \arg \min_{\mathbf{x} \in \mathbb{R}^d} \mathbb{E}[\alpha(V\mathbf{X} - \mathbf{x})_+^T \Sigma (V\mathbf{X} - \mathbf{x})_+ + (1 - \alpha)(V\mathbf{X} - \mathbf{x})_-^T \Sigma (V\mathbf{X} - \mathbf{x})_-] \\ &= \arg \min_{\mathbf{x} \in \mathbb{R}^d} \mathbb{E}[\alpha(\mathbf{X} - V^{-1}\mathbf{x})_+^T V^T \Sigma V (\mathbf{X} - V^{-1}\mathbf{x})_+ + (1 - \alpha)(\mathbf{X} - V^{-1}\mathbf{x})_-^T V^T \Sigma V (\mathbf{X} - V^{-1}\mathbf{x})_-] \\ &= V\mathbf{e}_\alpha^{V^T \Sigma V}(\mathbf{X}) = V\mathbf{e}_\alpha^{V\Sigma V}(\mathbf{X}), \end{aligned}$$

car V est symétrique. □

La Proposition 1.3.3 n'est pas vérifiée pour une famille plus générale de matrice, car la fonction $(\cdot)_+$ n'est pas additive.

Proposition 1.3.4 (Symétrie par rapport au seuil). *Soit $\mathbf{X} = (X_1, \dots, X_d)^T$ un vecteur aléatoire sur \mathbb{R}^d . Pour tout seuil $\alpha \in [0, 1]$,*

$$\mathbf{e}_\alpha(-\mathbf{X}) = -\mathbf{e}_{1-\alpha}(\mathbf{X}).$$

Preuve On note $\mathbf{x}^* = \mathbf{e}_{1-\alpha}(\mathbf{X})$ l'expectile multivarié de seuil $(1 - \alpha)$ associé au vecteur aléatoire \mathbf{X} et $\mathbf{x} = \mathbf{e}_\alpha(-\mathbf{X})$ l'expectile multivarié de seuil α associé au vecteur aléatoire $-\mathbf{X}$. En utilisant la condition d'optimalité (1.9), \mathbf{x} est l'unique solution du système d'équations suivant

$$\alpha \sum_{i=1}^d \pi_{ki} \mathbb{E}[(-X_i - x_i)_+ \mathbb{1}_{\{-X_k > x_k\}}] = (1 - \alpha) \sum_{i=1}^d \pi_{ki} \mathbb{E}[(x_i + X_i)_+ \mathbb{1}_{\{x_k > -X_k\}}], \quad \forall k \in \{1, \dots, d\},$$

c'est équivalent aussi au système qui suit,

$$\alpha \sum_{i=1}^d \pi_{ki} \mathbb{E}[(-x_i - X_i)_+ \mathbb{1}_{\{X_k < -x_k\}}] = (1 - \alpha) \sum_{i=1}^d \pi_{ki} \mathbb{E}[(X_i - (-x_i))_+ \mathbb{1}_{\{-x_k < X_k\}}], \quad \forall k \in \{1, \dots, d\}.$$

On déduit de ce dernier que $-\mathbf{x}$ est l'unique solution de la condition d'optimalité associée à l'expectile $\mathbf{e}_{1-\alpha}(\mathbf{X})$, c'est à dire $-\mathbf{x} = \mathbf{x}^*$. □

Proposition 1.3.5 (Prise en compte de l'indépendance bivariée). *Pour tout vecteur aléatoire $\mathbf{X} = (X_1, \dots, X_d)^T$ d'ordre 2, tel que il existe $i \in \{1, \dots, d\}$ avec $\pi_{ij} = 0, \forall j \in \{1, \dots, d\} \setminus \{i\}$, on a*

$$\mathbf{e}_\alpha^i(\mathbf{X}) = e_\alpha(X_i),$$

où $\mathbf{e}_\alpha^i(\mathbf{X})$ présente la $i^{\text{ème}}$ coordonnée du vecteur $\mathbf{e}_\alpha(\mathbf{X})$, et $e_\alpha(X_i)$ l'expectile univarié associé à la variable aléatoire X_i . Le vecteur aléatoire $\mathbf{X}^{(-i)} = (X_1, \dots, X_{i-1}, X_{i+1}, \dots, X_d)^T$ vérifie

$$\mathbf{e}_\alpha \left((X_i, \mathbf{X}^{(-i)}) \right) = \left(e_\alpha(X_i), \mathbf{e}_\alpha^{\Sigma_{ii}}(\mathbf{X}^{(-i)}) \right),$$

avec Σ_{ii} la sous-matrice de Σ obtenue par suppression de la $i^{\text{ème}}$ ligne et la $i^{\text{ème}}$ colonne, et $\mathbf{e}_\alpha^{\Sigma_{ii}}(\mathbf{X}^{(-i)})$ est le Σ_{ii} -expectile du vecteur $\mathbf{X}^{(-i)}$.

Cette propriété est importante pour les constructions à l'aide de matrices dont les coefficients sont des mesures de dépendance, par exemple la matrice des corrélations représente le cas parfait pour illustrer cette propriété.

Preuve Comme $\pi_{ij} = 0 \forall j \neq i$, la $i^{\text{ème}}$ équation du système d'optimalité (1.10) devient

$$\alpha \mathbb{E}[(X_i - \mathbf{e}_\alpha^i(\mathbf{X}))_+] = (1 - \alpha) \mathbb{E}[(X_i - \mathbf{e}_\alpha^i(\mathbf{X}))_-],$$

qui est la condition d'optimalité pour la détermination de l'expectile univarié de X_i . Les autres équations constituent le système d'optimalité correspondant à l'expectile $\mathbf{e}_\alpha^{\Sigma_{ii}}(\mathbf{X}^{(-i)})$. \square

Proposition 1.3.6 (Conservation de la symétrie). *Si le vecteur aléatoire $\mathbf{X} = (X_1, \dots, X_d)^T$ est symétrique par rapport à un vecteur $\mathbf{s} = (s_1, \dots, s_d)^T$, l'expectile multivarié est aussi symétrique par seuil par rapport au même vecteur \mathbf{s} ,*

$$\frac{\mathbf{e}_\alpha(\mathbf{X}) + \mathbf{e}_{1-\alpha}(\mathbf{X})}{2} = \mathbf{s}.$$

On remarque en particulier que dans le cas de la symétrie de \mathbf{X} par rapport à \mathbf{s} , l'expectile de seuil $1/2$ est déterminé par $\mathbf{e}_{1/2}(\mathbf{X}) = \mathbf{s} = \mathbb{E}[\mathbf{X}]$.

Preuve Ce résultat découle de la propriété de symétrie de l'expectile par rapport au seuil α . Si le vecteur aléatoire $\mathbf{X} = (X_1, \dots, X_d)^T$ est symétrique par rapport à un vecteur $\mathbf{s} = (s_1, \dots, s_d)^T$, alors $\mathbf{e}_\alpha(\mathbf{X} - \mathbf{s}) = -\mathbf{e}_{1-\alpha}(\mathbf{X} - \mathbf{s})$. L'invariance par translation permet ensuite de déduire le résultat recherché. \square

Proposition 1.3.7 (Stabilité dans le support). *Pour tout vecteur aléatoire $\mathbf{X} = (X_1, \dots, X_d)^T$ d'ordre 2 sur \mathbb{R}^d , alors,*

$$\mathbf{e}_\alpha^i(\mathbf{X}) \in [x_I^i, x_F^i].$$

où $\mathbf{e}_\alpha^i(\mathbf{X})$ présente la $i^{\text{ème}}$ coordonnée du vecteur $\mathbf{e}_\alpha(\mathbf{X})$, et $x_I^i = \text{essInf}(X_i)$ et $x_F^i = \text{essSup}(X_i)$.

Preuve Dans le cas du support marginal infini le résultat est trivial. Supposons qu'il existe $k \in \{1, \dots, d\}$ tel que $x_F^k < +\infty$ et $\mathbf{e}_\alpha^k(\mathbf{X}) > x_F^k$. L'expectile multivarié est défini comme minimiseur

$$\mathbf{e}_\alpha(\mathbf{X}) = \arg \min_{\mathbf{x} \in \mathbb{R}^d} \mathbb{E}[\alpha(\mathbf{X} - \mathbf{x})_+^T \Sigma(\mathbf{X} - \mathbf{x})_+ + (1 - \alpha)(\mathbf{X} - \mathbf{x})_-^T \Sigma(\mathbf{X} - \mathbf{x})_-],$$

et le minimum dans ce cas est

$$\begin{aligned} \min_{\mathbf{x} \in \mathbb{R}^d} \mathbb{E}[s_\alpha(\mathbf{X}, \mathbf{x})] &= \min_{\mathbf{x} \in \mathbb{R}^d} \mathbb{E}[\alpha(\mathbf{X} - \mathbf{x})_+^T \Sigma(\mathbf{X} - \mathbf{x})_+ + (1 - \alpha)(\mathbf{X} - \mathbf{x})_-^T \Sigma(\mathbf{X} - \mathbf{x})_-] \\ &= (1 - \alpha) \left(\pi_k \mathbb{E}[(\mathbf{e}_\alpha^k(\mathbf{X}) - X_k)_+]^2 + \sum_{1 \leq i \leq d, i \neq k} \pi_{ik} \mathbb{E}[(\mathbf{e}_\alpha^i(\mathbf{X}) - X_i)_+(\mathbf{e}_\alpha^k(\mathbf{X}) - X_k)_+] \right) \\ &\quad + (1 - \alpha) \sum_{\substack{1 \leq i \leq d, i \neq k \\ 1 \leq j \leq d, j \neq k}} \pi_{ij} \mathbb{E}[(\mathbf{e}_\alpha^i(\mathbf{X}) - X_i)_+(\mathbf{e}_\alpha^j(\mathbf{X}) - X_j)_+] \\ &\quad + \alpha \sum_{\substack{1 \leq i \leq d, i \neq k \\ 1 \leq j \leq d, j \neq k}} \pi_{ij} \mathbb{E}[(X_i - \mathbf{e}_\alpha^i(\mathbf{X}))_+(X_j - \mathbf{e}_\alpha^j(\mathbf{X}))_+], \end{aligned}$$

car $\mathbb{E}[(X_k - \mathbf{e}_\alpha^k(\mathbf{X}))_+(X_i - \mathbf{e}_\alpha^i(\mathbf{X}))_+] = 0$ pour tout $i \in \{1, \dots, d\}$.

La fonction $x \rightarrow (1 - \alpha) \left(\pi_k \mathbb{E}[(x - X_k)_+]^2 + \sum_{1 \leq i \leq d, i \neq k} \pi_{ik} \mathbb{E}[(\mathbf{e}_\alpha^i(\mathbf{X}) - X_i)_+(x - X_k)_+] \right)$ est croissante en x , considérons x^* avec $\mathbf{e}_\alpha^k(\mathbf{X}) > x^* > x_F^k$, pour $i \in \{1, \dots, d\} \setminus \{k\}$

$$\mathbb{E}[(X_k - x^*)_+] = \mathbb{E}[(X_k - x^*)_+(X_i - \mathbf{e}_\alpha^i(\mathbf{X}))_+] = 0,$$

donc pour $\mathbf{x}^* = (\mathbf{e}_\alpha^1(\mathbf{X}), \dots, \mathbf{e}_\alpha^{k-1}(\mathbf{X}), x^*, \mathbf{e}_\alpha^{k+1}(\mathbf{X}), \dots, \mathbf{e}_\alpha^d(\mathbf{X}))$

$$\begin{aligned} \mathbb{E}[s_\alpha(\mathbf{X}, \mathbf{x}^*)] &= (1 - \alpha) \left(\pi_k \mathbb{E}[(x^* - X_k)_+]^2 + \sum_{1 \leq i \leq d, i \neq k} \pi_{ik} \mathbb{E}[(\mathbf{e}_\alpha^i(\mathbf{X}) - X_i)_+(x^* - X_k)_+] \right) \\ &\quad + (1 - \alpha) \sum_{\substack{1 \leq i \leq d, i \neq k \\ 1 \leq j \leq d, j \neq k}} \pi_{ij} \mathbb{E}[(\mathbf{e}_\alpha^i(\mathbf{X}) - X_i)_+(\mathbf{e}_\alpha^j(\mathbf{X}) - X_j)_+] \\ &\quad + \alpha \sum_{\substack{1 \leq i \leq d, i \neq k \\ 1 \leq j \leq d, j \neq k}} \pi_{ij} \mathbb{E}[(X_i - \mathbf{e}_\alpha^i(\mathbf{X}))_+(X_j - \mathbf{e}_\alpha^j(\mathbf{X}))_+] \\ &\leq \mathbb{E}[s_\alpha(\mathbf{X}, \mathbf{e}_\alpha(\mathbf{X}))] = \min_{\mathbf{x} \in \mathbb{R}^d} \mathbb{E}[s_\alpha(\mathbf{X}, \mathbf{x})], \end{aligned}$$

cela est absurde. On en déduit que

$$\mathbf{e}_\alpha^k(\mathbf{X}) \leq x_F^k, \quad \forall k \in \{1, \dots, d\},$$

et en utilisant la propriété de symétrie par rapport au seuil α que

$$\mathbf{e}_\alpha^k(\mathbf{X}) \geq x_I^k, \quad \forall k \in \{1, \dots, d\}.$$

□

Proposition 1.3.8 (Neutralité). *Soit $\mathbf{X} = (X_1, \dots, X_d)^T$ un vecteur aléatoire à espérance finie. S'il existe $i \in \{1, \dots, d\}$ et une constante réelle c_i tels que $X_i = c_i$, p.s alors l'expectile multivarié construit à l'aide d'une matrice symétrique semi-définie positive $\Sigma = (\pi_{ij})_{1 \leq i, j \leq d}$, à coefficients positifs sur la i ème ligne tel que $\pi_{ii} > 0$, vérifie*

$$\mathbf{e}_\alpha^i(\mathbf{X}) = c_i,$$

et le vecteur aléatoire $\mathbf{X}^{(-i)} = (X_1, \dots, X_{i-1}, X_{i+1}, \dots, X_d)^T$ vérifie

$$\mathbf{e}_\alpha \left((X_i, \mathbf{X}^{(-i)}) \right) = \left(c_i, \mathbf{e}_{\alpha}^{\Sigma_{ii}} \left(\mathbf{X}^{(-i)} \right) \right),$$

avec Σ_{ii} la sous-matrice de Σ obtenue par suppression de la i ème ligne et la i ème colonne, et $\mathbf{e}_{\alpha}^{\Sigma_{ii}} \left(\mathbf{X}^{(-i)} \right)$ est le Σ_{ii} -expectile du vecteur $\mathbf{X}^{(-i)}$.

Preuve Pour tout vecteur $(u_1, \dots, u_d)^T$, on note $\mathbf{u}^{(-i)}$ le vecteur $(u_1, \dots, u_{i-1}, u_{i+1}, \dots, u_d)^T$.

Supposons que $\mathbf{e}_\alpha^i(\mathbf{X}) < c_i$, par définition

$$\mathbf{e}_\alpha(\mathbf{X}) = \arg \min_{\mathbf{x} \in \mathbb{R}^d} \mathbb{E}[\alpha(\mathbf{X} - \mathbf{x})_+^T \Sigma (\mathbf{X} - \mathbf{x})_+ + (1 - \alpha)(\mathbf{X} - \mathbf{x})_-^T \Sigma (\mathbf{X} - \mathbf{x})_-]. \quad (1.13)$$

On note

$$s_\alpha^\Sigma(\mathbf{X}, \mathbf{x}) = \mathbb{E}[\alpha(\mathbf{X} - \mathbf{x})_+^T \Sigma (\mathbf{X} - \mathbf{x})_+ + (1 - \alpha)(\mathbf{X} - \mathbf{x})_-^T \Sigma (\mathbf{X} - \mathbf{x})_-],$$

alors

$$s_\alpha^\Sigma(\mathbf{X}, \mathbf{e}_\alpha(\mathbf{X})) = 2 \sum_{j=1, j \neq i}^d \pi_{ij} (c_i - \mathbf{e}_\alpha^i(\mathbf{X})) \mathbb{E}[(X_j - x_j)_+] + \alpha \pi_{ii} (c_i - \mathbf{e}_\alpha^i(\mathbf{X}))^2 + s_{\alpha}^{\Sigma_{ii}}(\mathbf{X}^{(-i)}, \mathbf{e}_{\alpha}^{(-i)}(\mathbf{X})).$$

Pour tout vecteur $\mathbf{v} = (v_1, \dots, v_{d-1})^T$ tel que $\mathbf{v} \in \arg \min_{\mathbf{x} \in \mathbb{R}^{d-1}} s_{\alpha}^{\Sigma_{ii}}(\mathbf{X}^{(-i)}, \mathbf{x})$, on a alors

$$s_\alpha^\Sigma(\mathbf{X}, \mathbf{e}_\alpha(\mathbf{X})) > s_\alpha^\Sigma(\mathbf{X}, (v_1, \dots, v_{i-1}, c_i, v_i, \dots, v_{d-1})^T),$$

cela est contradictoire avec la Définition 1.13 de $\mathbf{e}_\alpha(\mathbf{X})$. De la même façon, on démontre que le cas $\mathbf{e}_\alpha^i(\mathbf{X}) > c_i$ est aussi impossible, d'où $\mathbf{e}_\alpha^i(\mathbf{X}) = c_i$.

Pour le vecteur aléatoire $\mathbf{X}^{(-i)} = (X_1, \dots, X_{i-1}, X_{i+1}, \dots, X_d)^T$, l'unicité de $\arg \min_{\mathbf{x} \in \mathbb{R}^{d-1}} s_\alpha^{\Sigma_{ii}}(\mathbf{X}^{(-i)}, \mathbf{x})$ est assurée par celle de $\mathbf{e}_\alpha(\mathbf{X})$. On en déduit que $\mathbf{e}_\alpha^{\Sigma_{ii}}(\mathbf{X}^{(-i)}) = \mathbf{e}_\alpha^{(-i)}(\mathbf{X})$ et par suite

$$\mathbf{e}_\alpha \left(\left(X_i, \mathbf{X}^{(-i)} \right) \right) = \left(c_i, \mathbf{e}_\alpha^{\Sigma_{ii}} \left(\mathbf{X}^{(-i)} \right) \right).$$

□

Proposition 1.3.9 (Monotonie interne). *Nous considérons le cas de l'expectile matriciel construit à l'aide d'une matrice $\Sigma = (\pi_{ij})_{1 \leq i, j \leq d}$ symétrique semi-définie-positives à coefficients positifs.*

Pour tout vecteur aléatoire $\mathbf{X} = (X_1, \dots, X_d)^T$ d'ordre 2 dans \mathbb{R}^d , s'il existe un couple $(i, j) \in \{1, \dots, d\}^2$ tel que $X_i \leq X_j$, p.s., $\pi_i \leq \pi_j$, $\pi_{ik} \leq \pi_{jk}$ pour tout $k \in \{1, \dots, d\} \setminus \{i, j\}$, et les couples (X_i, X_k) et (X_j, X_k) ont la même copule pour tout $k \in \{1, \dots, d\} \setminus \{i, j\}$, alors

$$\mathbf{e}_\alpha^i(\mathbf{X}) \leq \mathbf{e}_\alpha^j(\mathbf{X}),$$

pour tout niveau du seuil α .

La dominance presque sûre entre les distributions marginales est donc respectée par les expectiles construits. Ces mesures prennent donc en compte à la fois la dépendance bivariée entre les différents risques et leurs marginales.

Preuve Soit $(X_1, \dots, X_d)^T$ un vecteur aléatoire d'ordre 2 dans \mathbb{R}^d . On suppose qu'il existe i et j tels que $X_i \leq X_j$, p.s., cela implique aussi la dominance stochastique d'ordre 1 $X_i \leq_{st} X_j$. On note les fonctions l_{X_i, X_j}^α pour tout $(i, j) \in \{1, \dots, d\}^2$ comme suit,

$$l_{X_i, X_j}^\alpha(x_i, x_j) = \alpha \mathbb{E}[(X_i - x_i)_+ \mathbb{1}_{\{X_j > x_j\}}] - (1 - \alpha) \mathbb{E}[(X_i - x_i)_- \mathbb{1}_{\{X_j < x_j\}}],$$

pour tout $(x_i, x_j) \in \mathbb{R}^2$. On note également $l_{X_i}^\alpha$ la fonction $l_{X_i}^\alpha(x_i) = l_{X_i, X_i}^\alpha(x_i, x_i)$. L'expectile multivarié est l'unique solution du système d'optimalité (1.9) qui s'écrit à l'aide de des fonctions l^α sous la forme suivante,

$$\sum_{k=1}^d \pi_{ki} l_{X_i, X_k}^\alpha(x_i, x_k) = 0 \quad \forall k \in \{1, \dots, d\}. \quad (1.14)$$

Les fonctions $l_{X_i}^\alpha$ sont décroissantes pour tout $i \in \{1, \dots, d\}$, et les fonctions $l_{X_i, X_j}^\alpha(x_i, x_j)$ sont décroissantes par rapport à x_i pour tout $(i, j) \in \{1, \dots, d\}^2$. En effet, les dérivées partielles sont négatives

$$\frac{\partial l_{X_i}^\alpha(x_i)}{\partial x_i} = -\alpha \mathbb{P}(X_i > x_i) - (1 - \alpha) \mathbb{P}(X_i \leq x_i),$$

$$\frac{\partial l_{X_i, X_j}^\alpha(x_i, x_j)}{\partial x_i} = -\alpha \mathbb{P}(X_i > x_i, X_j > x_j) - (1 - \alpha) \mathbb{P}(X_i \leq x_i, X_j \leq x_j).$$

D'un autre côté, pour tout $(i, j) \in \{1, \dots, d\}^2$

$$l_{X_i, X_j}^\alpha(x_i, x_j) = \mathbb{E}[(\alpha(X_i - x_i)_+ + (1 - \alpha)(X_i - x_i)_-) \mathbb{1}_{\{X_j \geq x_j\}}] - (1 - \alpha) \mathbb{E}[(X_i - x_i)_-],$$

les fonctions l_{X_i, X_j}^α sont donc décroissantes par rapport à x_j aussi.

On peut déduire que si $X_i \leq_{st} X_j$, alors pour tout X_k qui possède la même structure de dépendance bivariée avec X_i et X_j , alors

$$l_{X_k, X_i}^\alpha(x_k, x) \leq l_{X_k, X_j}^\alpha(x_k, x), \quad \forall (x_k, x).$$

Maintenant, on suppose que $x_j = \mathbf{e}_\alpha^j(\mathbf{X}) < \mathbf{e}_\alpha^i(\mathbf{X}) = x_i$.

En utilisant les sens de variation des fonctions l^α par rapport à la première variable et les inégalités issues de la dominance stochastique, cette hypothèse implique

$$l_{X_k, X_i}^\alpha(x_k, x_i) \leq l_{X_k, X_i}^\alpha(x_k, x_j) \leq l_{X_k, X_j}^\alpha(x_k, x_j), \quad \forall k \in \{1, \dots, d\} \setminus \{i, j\}.$$

Les coefficients de la matrice Σ vérifient

$$0 \leq \pi_{ik} \leq \pi_{jk}, \forall k \in \{1, \dots, d\},$$

donc grâce au système d'optimalité (1.14) on peut déduire que

$$\sum_{k=1, k \neq i, k \neq j}^d \pi_{ik} l_{X_k, X_i}^\alpha(x_k, x_i) \leq \sum_{k=1, k \neq i, k \neq j}^d \pi_{jk} l_{X_k, X_j}^\alpha(x_k, x_j). \quad (1.15)$$

La dominance presque sûre $X_i \leq X_j, p.s$ implique que

$$\mathbb{1}_{\{X_i \geq x_i, X_j \leq x_j\}} = 0,$$

car $x_i > x_j$. On peut en déduire que

$$\mathbb{E}[(X_i - x_i)_+] = \mathbb{E}[(X_i - x_i)_+ \mathbb{1}_{\{X_j > x_j\}}],$$

et

$$\mathbb{E}[(X_j - x_j)_-] = \mathbb{E}[(X_j - x_j)_- \mathbb{1}_{\{X_i < x_i\}}],$$

de ces deux égalités on obtient respectivement

$$l_{X_i}^\alpha(x_i) < l_{X_i, X_j}^\alpha(x_i, x_j) \text{ et } l_{X_j, X_i}^\alpha(x_j, x_i) < l_{X_j}^\alpha(x_j),$$

ce qui permet d'arriver à

$$\pi_{ij} l_{X_i}^\alpha(x_i) + \pi_{ij} l_{X_j, X_i}^\alpha(x_j, x_i) < \pi_{jj} l_{X_j}^\alpha(x_j) + \pi_{ij} l_{X_i, X_j}^\alpha(x_i, x_j). \quad (1.16)$$

La dominance stochastique $X_i \leq_{st} X_j$ implique aussi que

$$l_{X_i}^\alpha(x) \leq l_{X_j}^\alpha(x), \quad \forall x,$$

car la fonction $t \rightarrow \alpha(t - x) - (1 - \alpha)(x - t)_+$ est croissante en t . Et par décroissance de la fonction $l_{X_j}^\alpha$ on obtient

$$(\pi_i - \pi_{ij}) l_{X_i}^\alpha(x_i) \leq (\pi_j - \pi_{ij}) l_{X_j}^\alpha(x_j), \quad (1.17)$$

puisque $0 \leq \pi_i - \pi_{ij} \leq \pi_j - \pi_{ij}$.

Finalement, de (1.17) et (1.16) on a

$$\pi_i l_{X_i}^\alpha(x_i) + \pi_{ij} l_{X_j, X_i}^\alpha(x_j, x_i) < \pi_j l_{X_j}^\alpha(x_j) + \pi_{ij} l_{X_i, X_j}^\alpha(x_i, x_j), \quad (1.18)$$

et cette inégalité est contradictoire avec (1.15) et le système d'optimalité (1.14). \square

Les hypothèses de cette proposition sur la matrice de construction de l'expectile sont vérifiées dans le cas du L_1 -expectile.

Lemme 1.3.10 (Impact de la dépendance). *Pour tout vecteur aléatoire $\mathbf{X} = (X_1, \dots, X_d)^T$ d'ordre 2 dans \mathbb{R}^d ,*

1. *s'il existe un couple $(i, j) \in \{1, \dots, d\}^2$ tel que $X_i \stackrel{\mathcal{L}}{=} X_j$, $\pi_i = \pi_j$, $\pi_{ik} = \pi_{jk}$ pour tout $k \in \{1, \dots, d\}$, et les couples (X_i, X_k) et (X_j, X_k) ont la même structure de dépendance pour tout $k \in \{1, \dots, d\} \setminus \{i, j\}$, alors*

$$\mathbf{e}_\alpha^i(\mathbf{X}) = \mathbf{e}_\alpha^j(\mathbf{X}),$$

pour tout niveau du seuil α .

2. *si toutes les marginales sont de la même distribution $X_i \stackrel{\mathcal{L}}{=} X$, $i = 1, \dots, d$, et de même structure de dépendance bivariée et il existe une constante positive π tel que $\pi_{ij} = \pi$ pour tout $i \neq j \in \{1, \dots, d\}$, $\pi_i = \pi$ pour tout $i, j \in \{1, \dots, d\}$, alors*

(a)

$$\mathbf{e}_\alpha(\mathbf{X}) = x^* \mathbb{1}_{\mathbb{R}^d},$$

où x^* est la solution de l'équation

$$\alpha(\mathbb{E}[(X-x)_+] + (d-1)\frac{\pi}{\pi_1}\mathbb{E}[(X_2-x_2)_+ \mathbb{1}_{\{X_1 > x_1\}}]) = (1-\alpha)(\mathbb{E}[(X-x)_-] + (d-1)\frac{\pi}{\pi_1}\mathbb{E}[(X_2-x_2)_- \mathbb{1}_{\{X_1 < x_1\}}]).$$

(b) si le vecteur \mathbf{X} est comonotone alors

$$\mathbf{e}_\alpha(\mathbf{X}) = e_\alpha(X) \mathbb{1}_{\mathbb{R}^d},$$

où $e_\alpha(X)$ est l'expectile univarié de la variable aléatoire X .

(c) si la structure de dépendance bivariée commune est représentée par une copule notée C alors

$$C \preceq C' \Rightarrow \mathbf{e}_\alpha^C(\mathbf{X}) \leq \mathbf{e}_\alpha^{C'}(\mathbf{X}),$$

pour tout seuil $\alpha \geq 1/2$, avec $\mathbf{e}_\alpha^C(\mathbf{X})$ l'expectile multivarié pour une structure bivariée commune de dépendance C .

On rappelle que pour toutes copules bivariées C et C'

$$C \preceq C' \Leftrightarrow C(u, v) \leq C'(u, v) \forall (u, v) \in [0, 1]^2.$$

Preuve Pour le (1) il suffit d'appliquer la propriété de l'invariance par loi de la Proposition 1.3.2, en considérant les v.a $\mathbf{X} = (X_1, \dots, X_d)^T$ et $\mathbf{Y} = (X_1, \dots, X_{i-1}, X_j, X_{i+1}, \dots, X_{j-1}, X_i, X_{j+1}, \dots, X_d)^T$. Pour le (2 - a), on obtient par une application du (1)

$$\mathbf{e}_\alpha^i(\mathbf{X}) = \mathbf{e}_\alpha^j(\mathbf{X}) = x^*,$$

pour tout $i, j \in \{1, \dots, d\}$. L'expectile multivarié est l'unique solution du système d'optimalité (1.9) qui s'écrit sous la forme suivante,

$$\sum_{i=1}^d \pi_{ki} l_{X_i, X_k}^\alpha(x_i, x_k) = 0 \quad \forall k \in \{1, \dots, d\},$$

avec,

$$l_{X_i, X_j}^\alpha(x_i, x_j) = \alpha \mathbb{E}[(X_i - x_i)_+ \mathbb{1}_{\{X_j > x_j\}}] - (1 - \alpha) \mathbb{E}[(X_i - x_i)_- \mathbb{1}_{\{X_j < x_j\}}],$$

pour tout $(x_i, x_j) \in \mathbb{R}^2$. Ces fonctions ne dépendent que de la distribution du couple (X_i, X_j) , alors dans notre cas

$$l_{X_i, X_j}^\alpha(x^*, x^*) = l_{X_1, X_2}^\alpha(x^*, x^*), \quad \forall i \neq j \in \{1, \dots, d\}.$$

On note également $l_{X_i}^\alpha$ la fonction $l_{X_i}^\alpha(x_i) = l_{X_i, X_i}^\alpha(x_i, x_i)$. Cette fonction ne dépend que de la distribution marginale de X_i , on peut donc écrire

$$l_{X_i}^\alpha(x^*) = l_X^\alpha(x^*) \quad \forall i \in \{1, \dots, d\}.$$

En réécrivant le système d'optimalité à l'aide de ces notations, x^* est l'unique solution de l'équation

$$\pi_1 l_X^\alpha(x^*) + (d-1)\pi l_{X_1, X_2}^\alpha(x^*, x^*) = 0.$$

Le (2 - b) est une application directe de (2 - a). En effet, dans le cas comonotone $l_{X_1, X_2}^\alpha(x^*, x^*) = l_X^\alpha(x^*)$, x^* est donc solution de l'équation $l_X^\alpha(x^*) = 0$ qui est exactement l'équation de définition de l'expectile univarié.

Pour le (2 - c), si on note $l_C^\alpha(x) = l_{X_1, X_2}^\alpha(x, x)$, alors

$$l_C^\alpha(x) = \alpha \mathbb{E}[(X_2 - x)_+ \mathbb{1}_{\{X_1 > x\}}] - (1 - \alpha) \mathbb{E}[(X_2 - x)_- \mathbb{1}_{\{X_1 < x\}}],$$

donc pour $\alpha \geq 1/2$

$$\begin{aligned} l_C^\alpha(x) &= (\alpha - 1/2) \left(\mathbb{E}[(X_2 - x)_+ \mathbb{1}_{\{X_1 > x\}}] + \mathbb{E}[(X_2 - x)_- \mathbb{1}_{\{X_1 < x\}}] \right) \\ &\quad + 1/2 \left(\mathbb{E}[(X_2 - x)_+ \mathbb{1}_{\{X_1 > x\}}] - \mathbb{E}[(X_2 - x)_- \mathbb{1}_{\{X_1 < x\}}] \right) \\ &= (\alpha - 1/2) \left(\mathbb{E}[(X_2 - x)_+ \mathbb{1}_{\{X_1 > x\}}] + \mathbb{E}[(X_2 - x)_- \mathbb{1}_{\{X_1 < x\}}] \right) \\ &\quad + 1/2 \left(\mathbb{E}[(X_2 - x)_+ \mathbb{1}_{\{X_1 > x\}}] - \mathbb{E}[(X_2 - x)_-] \right), \end{aligned}$$

on en déduit que pour toutes copules bivariées C et C' si $C \preceq C'$ alors

$$l_C^\alpha(x) \leq l_{C'}^\alpha(x), \quad (1.19)$$

pour tout x .

Considérons deux copules bivariées C et C' telles que $C \preceq C'$, on note $\mathbf{e}_\alpha^C(\mathbf{X})$ l'expectile multivarié correspondant à une copule C , et on suppose que $\mathbf{e}_\alpha^{C'}(\mathbf{X}) < \mathbf{e}_\alpha^C(\mathbf{X})$. La fonction l_C^α est décroissante en x , donc en utilisant (1.19) on obtient

$$l_C^\alpha(\mathbf{e}_\alpha^C(\mathbf{X})) \leq l_{C'}^\alpha(\mathbf{e}_\alpha^C(\mathbf{X})) \leq l_{C'}^\alpha(\mathbf{e}_\alpha^{C'}(\mathbf{X})). \quad (1.20)$$

D'un autre côté, la fonction l_X^α est strictement décroissante pour les distributions continues donc

$$l_X^\alpha(\mathbf{e}_\alpha^C(\mathbf{X})) < l_X^\alpha(\mathbf{e}_\alpha^{C'}(\mathbf{X})),$$

et par suite

$$l_{C'}^\alpha(\mathbf{e}_\alpha^{C'}(\mathbf{X})) < l_C^\alpha(\mathbf{e}_\alpha^C(\mathbf{X})). \quad (1.21)$$

Les deux inégalités (1.20) et (1.21) sont contradictoires, nous avons donc démontré que $\mathbf{e}_\alpha^C(\mathbf{X}) \leq \mathbf{e}_\alpha^{C'}(\mathbf{X})$. \square

Proposition 1.3.11 (Dérivées par rapport au seuil). *Soit $\mathbf{X} = (X_1, \dots, X_d)^T$ un vecteur aléatoire d'ordre 2, de distribution continue dans \mathbb{R}^d . On considère un expectile multivarié construit à l'aide d'une matrice $\Sigma = (\pi_{ij})_{1 \leq i, j \leq d}$ symétrique semi-définie positive, alors le vecteur $X_{\partial_\alpha} = \left(\frac{\partial \mathbf{e}_\alpha^1(\mathbf{X})}{\partial \alpha}, \dots, \frac{\partial \mathbf{e}_\alpha^d(\mathbf{X})}{\partial \alpha} \right)^T$ des dérivées des composantes de l'expectile par rapport au seuil α vérifie le système d'équations suivant*

$$B_k \frac{\partial \mathbf{e}_\alpha^k(\mathbf{X})}{\partial \alpha} + \sum_{i=1}^d \gamma_{ki} \frac{\partial \mathbf{e}_\alpha^i(\mathbf{X})}{\partial \alpha} = A_k, \quad \forall k \in \{1, \dots, d\}, \quad (1.22)$$

avec

$$A_k = \sum_{i=1}^d \pi_{ki} \mathbb{E}[(X_i - \mathbf{e}_\alpha^i(\mathbf{X}))_+ \mathbb{1}_{\{X_k > \mathbf{e}_\alpha^k(\mathbf{X})\}}] + \sum_{i=1}^d \pi_{ki} \mathbb{E}[(\mathbf{e}_\alpha^i(\mathbf{X}) - X_i)_+ \mathbb{1}_{\{X_k < \mathbf{e}_\alpha^k(\mathbf{X})\}}] \geq 0,$$

et

$$B_k = f_{X_k}(\mathbf{e}_\alpha^k(\mathbf{X})) \sum_{i=1, i \neq k}^d \pi_{ki} \left(\alpha \mathbb{E}[(X_i - \mathbf{e}_\alpha^i(\mathbf{X}))_+ | X_k = \mathbf{e}_\alpha^k(\mathbf{X})] + (1 - \alpha) \mathbb{E}[(\mathbf{e}_\alpha^i(\mathbf{X}) - X_i)_+ | X_k = \mathbf{e}_\alpha^k(\mathbf{X})] \right),$$

pour tout $k \in \{1, \dots, d\}$, et

$$\gamma_{ki} = \alpha \pi_{ki} \mathbb{P}(X_i > \mathbf{e}_\alpha^i(\mathbf{X}), X_k > \mathbf{e}_\alpha^k(\mathbf{X})) + (1 - \alpha) \pi_{ki} \mathbb{P}(X_i < \mathbf{e}_\alpha^i(\mathbf{X}), X_k < \mathbf{e}_\alpha^k(\mathbf{X})), \quad \forall (i, k) \in \{1, \dots, d\}^2.$$

Preuve On note $x_i = \mathbf{e}_\alpha^i(\mathbf{X})$. À partir du système d'optimalité (1.10)

$$\alpha \sum_{i=1}^d \pi_{ki} \mathbb{E}[(X_i - x_i)_+ \mathbb{1}_{\{X_k > x_k\}}] = (1 - \alpha) \sum_{i=1}^d \pi_{ki} \mathbb{E}[(x_i - X_i)_+ \mathbb{1}_{\{X_k < x_k\}}], \quad \forall k \in \{1, \dots, d\}$$

on peut obtenir par dérivation par rapport à α

$$\sum_{i=1}^d \pi_{ki} \left((1-\alpha) \frac{\partial}{\partial \alpha} \mathbb{E}[(x_i - X_i)_+ \mathbb{1}_{\{X_k < x_k\}}] - \alpha \frac{\partial}{\partial \alpha} \mathbb{E}[(X_i - x_i)_+ \mathbb{1}_{\{X_k > x_k\}}] \right) = A_k, \quad \forall k \in \{1, \dots, d\},$$

avec $A_k = \sum_{i=1}^d \pi_{ki} \mathbb{E}[(X_i - x_i)_+ \mathbb{1}_{\{X_k > x_k\}}] + \sum_{i=1}^d \pi_{ki} \mathbb{E}[(x_i - X_i)_+ \mathbb{1}_{\{X_k < x_k\}}]$. De cela on déduit que pour tout $k \in \{1, \dots, d\}$

$$\begin{aligned} A_k &= \frac{\partial x_k}{\partial \alpha} \sum_{i=1}^d \pi_{ki} \left((1-\alpha) \frac{\partial}{\partial x_k} \mathbb{E}[(x_i - X_i)_+ \mathbb{1}_{\{X_k < x_k\}}] - \alpha \frac{\partial}{\partial x_k} \mathbb{E}[(X_i - x_i)_+ \mathbb{1}_{\{X_k > x_k\}}] \right) \\ &\quad + \sum_{i=1, i \neq k}^d \pi_{ki} \frac{\partial x_i}{\partial \alpha} \left((1-\alpha) \frac{\partial}{\partial x_i} \mathbb{E}[(x_i - X_i)_+ \mathbb{1}_{\{X_k < x_k\}}] - \alpha \frac{\partial}{\partial x_i} \mathbb{E}[(X_i - x_i)_+ \mathbb{1}_{\{X_k > x_k\}}] \right). \end{aligned} \quad (1.23)$$

Comme le vecteur \mathbf{X} est de distribution continue, alors pour tout $i \neq k \in \{1, \dots, d\}$

$$\begin{aligned} \frac{\partial}{\partial x_k} \mathbb{E}[(X_i - x_i)_+ \mathbb{1}_{\{X_k > x_k\}}] &= \frac{\partial}{\partial x_k} \int_{x_i}^{+\infty} \mathbb{P}(X_i > t, X_k > x_k) dt \\ &= -f_{X_k}(x_k) \mathbb{E}[(X_i - x_i)_+ | X_k = x_k], \end{aligned}$$

et de la même façon

$$\frac{\partial}{\partial x_k} \mathbb{E}[(x_i - X_i)_+ \mathbb{1}_{\{X_k < x_k\}}] = f_{X_k}(x_k) \mathbb{E}[(x_i - X_i)_+ | X_k = x_k].$$

On a aussi

$$\frac{\partial}{\partial x_i} \mathbb{E}[(X_i - x_i)_+ \mathbb{1}_{\{X_k > x_k\}}] = \mathbb{P}(X_i > x_i, X_k > x_k), \quad \forall (i, k) \in \{1, \dots, d\}^2,$$

et

$$\frac{\partial}{\partial x_i} \mathbb{E}[(x_i - X_i)_+ \mathbb{1}_{\{X_k < x_k\}}] = \mathbb{P}(X_i < x_i, X_k < x_k), \quad \forall (i, k) \in \{1, \dots, d\}^2.$$

De (1.23), on peut déduire que le vecteur $X_{\partial \alpha} = \left(\frac{\partial x_1}{\partial \alpha}, \dots, \frac{\partial x_d}{\partial \alpha} \right)^T$ des dérivées des composantes de l'expectile par rapport au seuil α vérifie le système d'équations (1.22) annoncé dans la proposition. \square

Corollaire 1.3.12 (Variations par rapport au seuil). *Soit $\mathbf{X} = (X_1, \dots, X_d)^T$ un vecteur aléatoire d'ordre 2 et de distribution continue dans \mathbb{R}^d . On considère un expectile multivarié construit à l'aide d'une matrice $\Sigma = (\pi_{ij})_{1 \leq i, j \leq d}$ symétrique semi-définie positive de coefficients positifs. On note \otimes le produit matriciel Hadamard (produit terme à terme), et $X_{\partial \alpha} = \left(\frac{\partial \mathbf{e}_\alpha^1(\mathbf{X})}{\partial \alpha}, \dots, \frac{\partial \mathbf{e}_\alpha^d(\mathbf{X})}{\partial \alpha} \right)^T$ le vecteur des dérivées des composantes de l'expectile par rapport au seuil α . Alors,*

$$X_{\partial \alpha} = [B + \Sigma \otimes \Gamma]^{-1} A,$$

avec Γ la matrice définie par

$$(\Gamma)_{ij} = \alpha \mathbb{P}(X_i > \mathbf{e}_\alpha^i(\mathbf{X}), X_k > \mathbf{e}_\alpha^k(\mathbf{X})) + (1-\alpha) \mathbb{P}(X_i < \mathbf{e}_\alpha^i(\mathbf{X}), X_k < \mathbf{e}_\alpha^k(\mathbf{X})), \quad \forall i, j \in \{1, \dots, d\},$$

et $A = (A_1, \dots, A_d)^T$ et $B = (B_{ij})_{1 \leq i, j \leq d} = \text{Diag}(B_1, \dots, B_d)^T$ où les A_k et les B_k sont définis dans la Proposition 1.3.11.

Preuve La matrice Γ est aussi symétrique semi-définie positive.

De la Proposition 1.3.11, le vecteur $X_{\partial \alpha} = \left(\frac{\partial \mathbf{e}_\alpha^1(\mathbf{X})}{\partial \alpha}, \dots, \frac{\partial \mathbf{e}_\alpha^d(\mathbf{X})}{\partial \alpha} \right)^T$ vérifie le système matriciel suivant

$$[B + \Sigma \otimes \Gamma] X_{\partial \alpha} = A.$$

Pour tout $k \in \{1, \dots, d\}$, si $\forall i \neq k \in \{1, \dots, d\}$, $\pi_{ik} = 0$, alors $B_k = 0$ et

$$\frac{\partial \mathbf{e}_\alpha^k(\mathbf{X})}{\partial \alpha} = \frac{\mathbb{E}[|X_k - \mathbf{e}_\alpha^k(\mathbf{X})|]}{\alpha \mathbb{P}(X_k > \mathbf{e}_\alpha^k(\mathbf{X})) + (1 - \alpha) \mathbb{P}(X_k < \mathbf{e}_\alpha^k(\mathbf{X}))} \geq 0.$$

En utilisant la Proposition 1.3.5, on peut donc considérer que $\forall k \in \{1, \dots, d\}$, il existe $i \neq k \in \{1, \dots, d\}$ tel que $\pi_{ik} \neq 0$. Dans ce cas, par la Proposition 1.3.7 et la continuité des distributions jointes bivariées

$$B_k > 0, \quad \forall k \in \{1, \dots, d\}.$$

La matrice diagonale B est donc définie positive, et puisque la matrice $\Sigma \otimes \Gamma$ est semi-définie positive, alors $B + \Sigma \otimes \Gamma$ est une matrice inversible. D'où

$$X_{\partial_\alpha} = [B + \Sigma \otimes \Gamma]^{-1} A.$$

□

Nous avons donc montré que l'expectile multivarié que nous avons défini vérifie des propriétés naturelles attendues pour une mesure de risque multivariée, ce qui confirme une utilité potentielle. La section suivante donne une méthode efficace d'approximation.

1.4 Approximation stochastique

Nous présentons ici une méthode d'approximation stochastique des expectiles multivariés définis précédemment. Nous utilisons la famille des algorithmes stochastique d'optimisation et de recherche de zéros présentes dans la littérature. En général, l'expectile multivarié ne peut pas être calculé directement, mais il peut être estimé en utilisant des observations bruitées.

Dans notre contexte, la stricte convexité garantit l'existence et l'unicité du minimum, et la condition d'optimalité du premier ordre peut s'écrire explicitement sous forme de système d'équations composées d'espérances. Nous pouvons donc utiliser le célèbre algorithme stochastique de Robbins-Monro (1951) [36], ou plus précisément sa version multidimensionnelle introduite dans la littérature par Blum (1954) [6]. En effet, pour un expectile multivarié construit à l'aide d'une matrice semi-définie positive à coefficients positifs $\Sigma = (\pi_{ij})_{1 \leq i \leq j \leq d}$, le vecteur expectile est l'unique solution du système d'équations (1.9),

$$\alpha \sum_{i=1}^d \pi_{ki} \mathbb{E}[(X_i - x_i)_+ \mathbb{1}_{\{X_k > x_k\}}] = (1 - \alpha) \sum_{i=1}^d \pi_{ki} \mathbb{E}[(x_i - X_i)_+ \mathbb{1}_{\{x_k > X_k\}}], \quad \forall k \in \{1, \dots, d\},$$

qui peut s'écrire sous la forme,

$$\phi(\mathbf{x}) = \mathbb{E}[\Phi(\mathbf{x}, \mathbf{X})] = 0,$$

où $\Phi(\cdot, \mathbf{X})$ est une fonction de \mathbb{R}^d dans \mathbb{R}^d , avec

$$\Phi_k(\mathbf{x}, \mathbf{X}) = \sum_{i=1}^d \pi_{ki} \left(\alpha (X_i - x_i)_+ \mathbb{1}_{\{X_k > x_k\}} - (1 - \alpha) (x_i - X_i)_+ \mathbb{1}_{\{x_k > X_k\}} \right), \quad k \in \{1, \dots, d\}.$$

Les méthodes de l'approximation stochastique sont des algorithmes itératifs de la forme

$$\mathbf{x}_{n+1} = \mathbf{x}_n + \gamma_n \Phi(\mathbf{x}_n, \mathbf{X}_{n+1}),$$

où (γ_n) est une suite de pas déterministe qui vérifie certaines conditions que nous allons préciser, et (\mathbf{X}_n) est une suite de vecteurs aléatoires indépendants et identiquement distribués, de même loi qu'un v.a. générique \mathbf{X} . Ces algorithmes permettent de résoudre des problèmes de recherche de zéros, mais aussi de recherche de minimum qui se résout à l'aide de l'algorithme de Kiefer-Wolfowitz (1952) [27] en cas de non observation du gradient. Ces algorithmes peuvent être considérés comme des versions perturbées de l'algorithme de Newton d'optimisation déterministe. Les conditions de régularité et de convergence dans le cas déterministe sont présentées en détail dans le livre de Dufflo (1997) [16].

1.4.1 L'algorithme de Robbins-Monro

Nous appliquons l'algorithme de Robbins-Monro pour obtenir l'expectile multivarié. On désigne par \mathbf{x} l'expectile multivarié recherché. L'algorithme consiste à définir une suite (\mathbf{x}_n)

$$\mathbf{x}_{n+1} = \mathbf{x}_n + \gamma_n \mathbf{Z}_{n+1},$$

où $\mathbf{x}_0 \in \mathbb{R}^d$ est le point initial, et \mathbf{Z}_{n+1} le vecteur aléatoire qui permet d'observer $\phi(\mathbf{x}_n)$ pour l'itération n , elle vérifie $\phi(\mathbf{x}_n) = \mathbb{E}[\mathbf{Z}_{n+1} \mid \mathcal{F}_n]$ avec \mathcal{F}_n la σ -algèbre des informations présentes à l'instant n

$$\mathcal{F}_n = \sigma(\mathbf{x}_0, \mathbf{Z}_1, \mathbf{x}_1, \dots, \mathbf{x}_{n-1}, \mathbf{Z}_n).$$

Nous utilisons la version du théorème de Robbins-Monro énoncée dans la thèse de Fraysse (2013) [20] (Théorème 1.3.1).

Théorème 1.4.1 (Robbins-Monro). *Sous les hypothèses suivantes :*

1. ϕ est une fonction continue ;

2. Pour tout $\mathbf{x} \neq \mathbf{x}^*$,

$$(\mathbf{x} - \mathbf{x}^*)^T \phi(\mathbf{x}) < 0;$$

3. pour tout $n \geq 0$,

$$\mathbb{E}[\mathbf{Z}_{n+1} \mid \mathcal{F}_n] = \phi(\mathbf{x}_n), \text{ p.s.};$$

4. Il existe $K > 0$ telle que

$$\mathbb{E}[\|\mathbf{Z}_{n+1}\|^2 \mid \mathcal{F}_n] \leq K \left(1 + \|\mathbf{x}_n - \mathbf{x}^*\|^2\right) \text{ p.s.};$$

5. La suite (γ_n) est décroissante vers 0 et vérifie

$$\sum_{n=0}^{+\infty} \gamma_n = +\infty \text{ et } \sum_{n=0}^{+\infty} \gamma_n^2 < +\infty,$$

la suite (\mathbf{x}_n) définie par

$$\mathbf{x}_{n+1} = \mathbf{x}_n + \gamma_n \mathbf{Z}_{n+1},$$

converge presque sûrement vers \mathbf{x}^* solution de l'équation $\phi(\mathbf{x}^*) = \mathbf{0}$.

Les deux premières hypothèses du théorème concernent la régularité de la fonction ϕ , elles garantissent l'unicité de la solution. Les hypothèses 3 et 4 concernent la suite des observations et garantissent respectivement sa proche distance de la valeur exacte de la fonction et le contrôle de sa variance. La dernière hypothèse sur la suite de pas est nécessaire pour obtenir la convergence. D'autres versions modifiées de l'algorithme sont proposées dans la littérature en cas de non vérification de ces hypothèses, notamment celles qui concernent la suite des observations.

1.4.2 Application aux expectiles multivariés

Pour un vecteur aléatoire \mathbf{X} de distributions marginales continues, la fonction ϕ définie par $\phi(\mathbf{x}) = \mathbb{E}[\Phi(\mathbf{x}, \mathbf{X})]$ est clairement continue.

D'un autre côté, l'expectile multivarié \mathbf{x}^* est le minimum d'une fonction strictement convexe sur \mathbb{R}^d de gradient $-2\phi(\mathbf{x})$, elle vérifie donc pour tout $\mathbf{x} \in \mathbb{R}^d \setminus \{\mathbf{x}^*\}$

$$\langle -2\phi(\mathbf{x}) + 2\phi(\mathbf{x}^*), \mathbf{x} - \mathbf{x}^* \rangle = \langle -2\phi(\mathbf{x}), \mathbf{x} - \mathbf{x}^* \rangle > 0,$$

on en déduit que pour tout $\mathbf{x} \neq \mathbf{x}^*$,

$$(\mathbf{x} - \mathbf{x}^*)^T \phi(\mathbf{x}) < 0.$$

Les hypothèses du théorème de Robbins-Monro (1.4.1) sur la fonction ϕ sont donc vérifiées.

On considère la suite des vecteurs aléatoires indépendants et identiquement distribués $(\mathbf{X}_n) \stackrel{\mathcal{L}}{=} \mathbf{X}$. On définit la suite des observations (Z_n) pour chaque itération $n \geq 1$ de l'algorithme par

$$\mathbf{Z}_n^{(k)} = \sum_{i=1}^d \pi_{ki} \left(\alpha (\mathbf{X}_n^{(i)} - \mathbf{x}_{n-1}^{(i)})_+ \mathbb{1}_{\{\mathbf{X}_n^{(k)} > \mathbf{x}_{n-1}^{(k)}\}} - (1 - \alpha) (\mathbf{x}_{n-1}^{(i)} - \mathbf{X}_n^{(i)})_+ \mathbb{1}_{\{\mathbf{X}_n^{(k)} < \mathbf{x}_{n-1}^{(k)}\}} \right)$$

pour tout $k \in \{1, \dots, d\}$, où $\mathbf{x}^{(i)}$ désigne la $i^{\text{ième}}$ coordonnée du vecteur \mathbf{x} . Cette suite vérifie pour tout $k \in \{1, \dots, d\}$ et tout $n \geq 0$

$$\mathbb{E}[\mathbf{Z}_{n+1}^{(k)} \mid \mathcal{F}_n] = \phi_k(\mathbf{x}_n), \text{ p.s.,}$$

d'où

$$\forall n \geq 0, \quad \mathbb{E}[\mathbf{Z}_{n+1} \mid \mathcal{F}_n] = \phi(\mathbf{x}_n), \text{ p.s..}$$

On note $\mathbf{Z}_{i,n}^{(k)}$

$$\begin{aligned} \mathbf{Z}_{i,n}^{(k)} &= \alpha \left(\mathbf{X}_n^{(i)} - \mathbf{x}_{n-1}^{(i)} \right)_+ \mathbb{1}_{\{\mathbf{X}_n^{(k)} > \mathbf{x}_{n-1}^{(k)}\}} - (1 - \alpha) \left(\mathbf{x}_{n-1}^{(i)} - \mathbf{X}_n^{(i)} \right)_+ \mathbb{1}_{\{\mathbf{X}_n^{(k)} < \mathbf{x}_{n-1}^{(k)}\}} \\ &= \left(\mathbf{X}_n^{(i)} - \mathbf{x}_{n-1}^{(i)} \right) \left[\alpha \mathbb{1}_{\{\mathbf{X}_n^{(k)} > \mathbf{x}_{n-1}^{(k)}, \mathbf{X}_n^{(i)} > \mathbf{x}_{n-1}^{(i)}\}} + (1 - \alpha) \mathbb{1}_{\{\mathbf{X}_n^{(k)} < \mathbf{x}_{n-1}^{(k)}, \mathbf{X}_n^{(i)} < \mathbf{x}_{n-1}^{(i)}\}} \right], \end{aligned}$$

alors $\|\mathbf{Z}_{n+1}\|^2 = \sum_{k=1}^d |\mathbf{Z}_{n+1}^{(k)}|^2 = \sum_{k=1}^d \left| \sum_{i=1}^d \pi_{ik} \mathbf{Z}_{i,n+1}^{(k)} \right|^2$. Pour simplifier les notations, on définit les variables aléatoires

$$T_{X_i, X_k}^\alpha(x_i, x_k) = \alpha \mathbb{1}_{\{X_k > x_k, X_i > x_i\}} + (1 - \alpha) \mathbb{1}_{\{X_k < x_k, X_i < x_i\}},$$

on peut donc écrire

$$\begin{aligned} |\mathbf{Z}_{n+1}^{(k)}|^2 &= \left| \sum_{i=1}^d \pi_{ik} \mathbf{Z}_{i,n+1}^{(k)} \right|^2 = \left| \sum_{i=1}^d \pi_{ik} \left(\mathbf{X}_{n+1}^{(i)} - \mathbf{x}_n^{(i)} \right) T_{\mathbf{X}_{n+1}^{(i)}, \mathbf{X}_{n+1}^{(k)}}^\alpha(\mathbf{x}_n^{(i)}, \mathbf{x}_n^{(k)}) \right|^2 \\ &= \left| \sum_{i=1}^d \pi_{ik} \left[\left(\mathbf{X}_{n+1}^{(i)} - \mathbf{x}_i^* \right) T_{\mathbf{X}_{n+1}^{(i)}, \mathbf{X}_{n+1}^{(k)}}^\alpha(\mathbf{x}_n^{(i)}, \mathbf{x}_n^{(k)}) + \left(\mathbf{x}_i^* - \mathbf{x}_n^{(i)} \right) T_{\mathbf{X}_{n+1}^{(i)}, \mathbf{X}_{n+1}^{(k)}}^\alpha(\mathbf{x}_n^{(i)}, \mathbf{x}_n^{(k)}) \right] \right|^2, \end{aligned}$$

où \mathbf{x}_i^* est la $i^{\text{ième}}$ composante de l'expectile multivarié \mathbf{x}^* unique zéro de la fonction objectif. Nous appliquons l'inégalité triviale $(a + b)^2 \leq 2(a^2 + b^2)$ pour déduire que

$$|\mathbf{Z}_{n+1}^{(k)}|^2 \leq 2 \left| \sum_{i=1}^d \pi_{ik} \left(\mathbf{X}_{n+1}^{(i)} - \mathbf{x}_i^* \right) T_{\mathbf{X}_{n+1}^{(i)}, \mathbf{X}_{n+1}^{(k)}}^\alpha(\mathbf{x}_n^{(i)}, \mathbf{x}_n^{(k)}) \right|^2 + 2 \left| \sum_{i=1}^d \pi_{ik} \left(\mathbf{x}_i^* - \mathbf{x}_n^{(i)} \right) T_{\mathbf{X}_{n+1}^{(i)}, \mathbf{X}_{n+1}^{(k)}}^\alpha(\mathbf{x}_n^{(i)}, \mathbf{x}_n^{(k)}) \right|^2,$$

et à l'aide de l'inégalité triangulaire pour la valeur absolue, puis de l'inégalité de Hölder ($p = q = 2$) que

$$\begin{aligned} |\mathbf{Z}_{n+1}^{(k)}|^2 &\leq 2 \left[\sum_{i=1}^d \left| \pi_{ik} \left(\mathbf{X}_{n+1}^{(i)} - \mathbf{x}_i^* \right) T_{\mathbf{X}_{n+1}^{(i)}, \mathbf{X}_{n+1}^{(k)}}^\alpha(\mathbf{x}_n^{(i)}, \mathbf{x}_n^{(k)}) \right| \right]^2 + 2 \left[\sum_{i=1}^d \left| \pi_{ik} \left(\mathbf{x}_i^* - \mathbf{x}_n^{(i)} \right) T_{\mathbf{X}_{n+1}^{(i)}, \mathbf{X}_{n+1}^{(k)}}^\alpha(\mathbf{x}_n^{(i)}, \mathbf{x}_n^{(k)}) \right| \right]^2 \\ &\leq 2 \left(\sum_{i=1}^d \left| \pi_{ik} T_{\mathbf{X}_{n+1}^{(i)}, \mathbf{X}_{n+1}^{(k)}}^\alpha(\mathbf{x}_n^{(i)}, \mathbf{x}_n^{(k)}) \right|^2 \right) \left[\sum_{i=1}^d \left(\mathbf{X}_{n+1}^{(i)} - \mathbf{x}_i^* \right)^2 + \sum_{i=1}^d \left(\mathbf{x}_i^* - \mathbf{x}_n^{(i)} \right)^2 \right] \\ &= 2 \left(\sum_{i=1}^d \left| \pi_{ik} T_{\mathbf{X}_{n+1}^{(i)}, \mathbf{X}_{n+1}^{(k)}}^\alpha(\mathbf{x}_n^{(i)}, \mathbf{x}_n^{(k)}) \right|^2 \right) \left[\|\mathbf{X}_{n+1} - \mathbf{x}^*\|^2 + \|\mathbf{x}^* - \mathbf{x}_n\|^2 \right]. \end{aligned}$$

On remarque que pour tout $n \in \mathbb{N}$ et tous $(i, k) \in \{1, \dots, d\}^2$

$$0 \leq T_{\mathbf{X}_{n+1}^{(i)}, \mathbf{X}_{n+1}^{(k)}}^\alpha(\mathbf{x}_n^{(i)}, \mathbf{x}_n^{(k)}) \leq \max(\alpha, 1 - \alpha),$$

et on en déduit que

$$|\mathbf{Z}_{n+1}^{(k)}|^2 \leq 2 \max(\alpha, 1 - \alpha)^2 \left(\sum_{i=1}^d |\pi_{ik}|^2 \right) \left[\|\mathbf{X}_{n+1} - \mathbf{x}^*\|^2 + \|\mathbf{x}^* - \mathbf{x}_n\|^2 \right],$$

pour tout $k \in \{1, \dots, d\}$, et par suite

$$\|\mathbf{Z}_{n+1}\|^2 = \sum_{k=1}^d |\mathbf{Z}_{n+1}^{(k)}|^2 \leq 2 \max(\alpha, 1 - \alpha)^2 \left(\sum_{k=1}^d \sum_{i=1}^d |\pi_{ik}|^2 \right) \left[\|\mathbf{X}_{n+1} - \mathbf{x}^*\|^2 + \|\mathbf{x}^* - \mathbf{x}_n\|^2 \right].$$

Cette inégalité implique que

$$\mathbb{E}[\|\mathbf{Z}_{n+1}\|^2 | \mathcal{F}_n] \leq \mathbb{E}[2 \max(\alpha, 1 - \alpha)^2 \left(\sum_{k=1}^d \sum_{i=1}^d |\pi_{ik}|^2 \right) \left[\|\mathbf{X}_{n+1} - \mathbf{x}^*\|^2 + \|\mathbf{x}^* - \mathbf{x}_n\|^2 \right] | \mathcal{F}_n] \text{ p.s.,}$$

et puisque $\|\mathbf{X}_{n+1} - \mathbf{x}^*\|^2 = \|\mathbf{X}_n - \mathbf{x}^*\|^2 = \|\mathbf{X} - \mathbf{x}^*\|^2$ et $\|\mathbf{x}^* - \mathbf{x}_n\|^2$ sont \mathcal{F}_n -mesurables alors

$$\mathbb{E}[\|\mathbf{Z}_{n+1}\|^2 | \mathcal{F}_n] \leq 2 \max(\alpha, 1 - \alpha)^2 \left(\sum_{k=1}^d \sum_{i=1}^d |\pi_{ik}|^2 \right) \left[\|\mathbf{X} - \mathbf{x}^*\|^2 + \|\mathbf{x}^* - \mathbf{x}_n\|^2 \right] \text{ p.s..}$$

Finalement, en considérant $K = 2 \max(\alpha, 1 - \alpha)^2 \left(\sum_{k=1}^d \sum_{i=1}^d |\pi_{ik}|^2 \right) \max(\|\mathbf{X} - \mathbf{x}^*\|^2, 1)$ on obtient

$$\mathbb{E}[\|\mathbf{Z}_{n+1}\|^2 | \mathcal{F}_n] \leq K \left(1 + \|\mathbf{x}_n - \mathbf{x}^*\|^2 \right) \text{ p.s..}$$

Les hypothèses du théorème de Robbins-Monro 1.4.1 concernant la suite des observations sont vérifiées. Le théorème est donc applicable pour toute suite de pas (γ_n) décroissante vers 0 choisie telle que

$$\sum_{n=0}^{+\infty} \gamma_n = +\infty \quad \text{et} \quad \sum_{n=0}^{+\infty} \gamma_n^2 < +\infty.$$

Le choix naturel pour la suite des pas sera $\gamma_n = 1/n$ ou $1/n^\kappa$ avec $1/2 < \kappa < 1$. De manière générale, le choix de cette suite se fait par ajustement des constantes κ, a, b sachant que $\gamma_n = a/(b+n)^\kappa$. Nous allons dans les illustrations numériques examiner l'impact de ce choix sur la vitesse de la convergence de l'algorithme.

1.4.3 Illustrations numérique

Nous considérons un modèle exponentiel simple bivarié. $X_1 \sim \exp(\beta_1)$ et $X_2 \sim \exp(\beta_2)$. Nous examinons tout d'abord le cas du L_1 -expectile $(\pi_{ij} = 1, \forall (i, j) \in \{1, 2\}^2)$ sous l'hypothèse de l'indépendance. Dans ce cas, le système d'optimalité est explicite, et composé des deux équations suivantes :

$$(2\alpha - 1) \frac{1}{\beta_i} e^{-\beta_i x_i} - (1 - \alpha) \left(x_i - \frac{1}{\beta_i} \right) = (1 - \alpha) \left(x_j - \frac{1}{\beta_j} (1 - e^{-\beta_j x_j}) \right) (1 - e^{-\beta_i x_i}) - \alpha \frac{1}{\beta_j} e^{-\beta_j x_j} e^{-\beta_i x_i},$$

pour $(i, j) \in \{(1, 2), (2, 1)\}$. La solution exacte peut être obtenue à l'aide de méthodes d'optimisation numériques. Pour notre illustration nous utilisons l'algorithme de Newton-Raphson multidimensionnel. La figure 1.2 présente le résultat obtenu pour $\beta_1 = 0.05$ (représenté en bleu) et $\beta_2 = 0.25$ (représenté en rouge). La valeur exacte de l'expectile (20.02, 3.22) est représentée par les droites en traits pointillés de même couleur respectivement. Pour garantir la robustesse du résultat, nous lançons l'algorithme Robbins-Monro 100 fois, et nous calculons la moyenne.

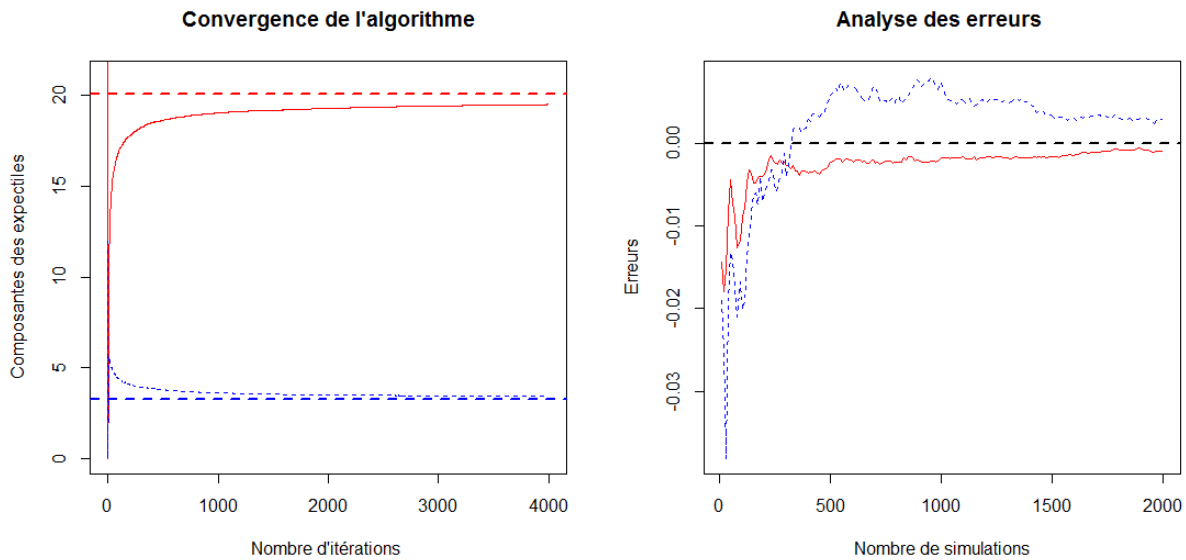


FIGURE 1.2 – Convergence de l'algorithme : L_1 -expectile, modèle exponentiel indépendant (rouge $\beta_1 = 0.05$, bleu $\beta_2 = 0.25$).

La convergence est assez satisfaisante pour un nombre d'itération inférieur à 1000. L'analyse graphique de la normalité des erreurs est présentée dans la figure 1.3. L'impact du choix de la suite des pas sur la vitesse de convergence est illustré dans la Figure 1.4. La convergence n'est pas très sensible au choix du point initial.

La convergence de l'algorithme Robbins-Monro a été étudiée dans la thèse de Lelong (2007) [30]. Il présente deux TCL, un pour le choix d'une suite de pas $1/n$, et le deuxième pour une suite sous la forme $\gamma_n = \gamma/n^\alpha$, avec γ une constante et $1/2 < \alpha < 1$.

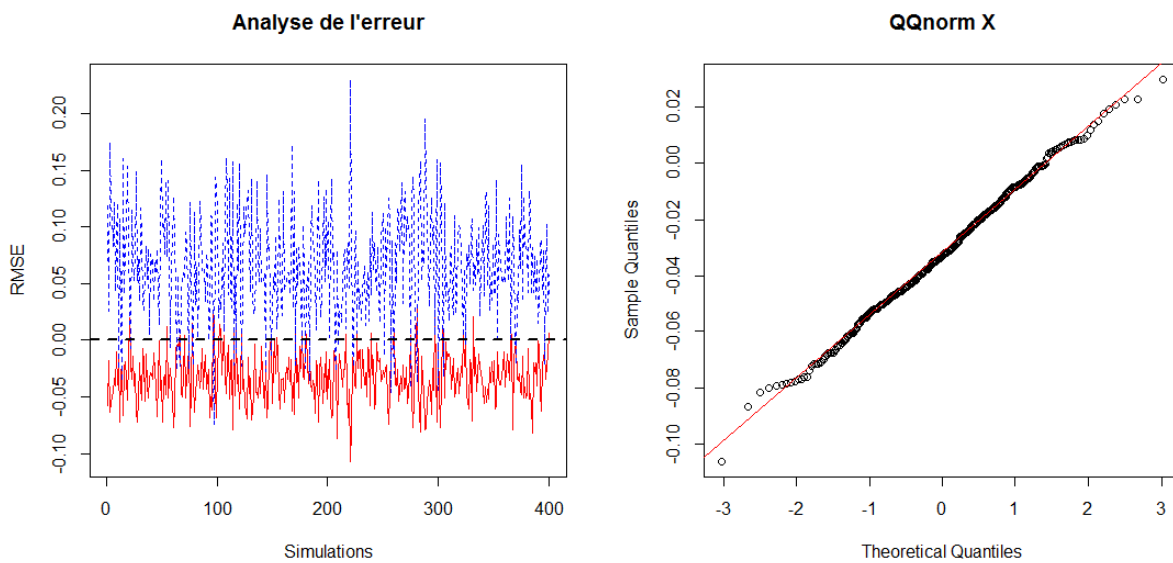


FIGURE 1.3 – Analyse de l'erreur, L_1 -expectile, modèle EI.

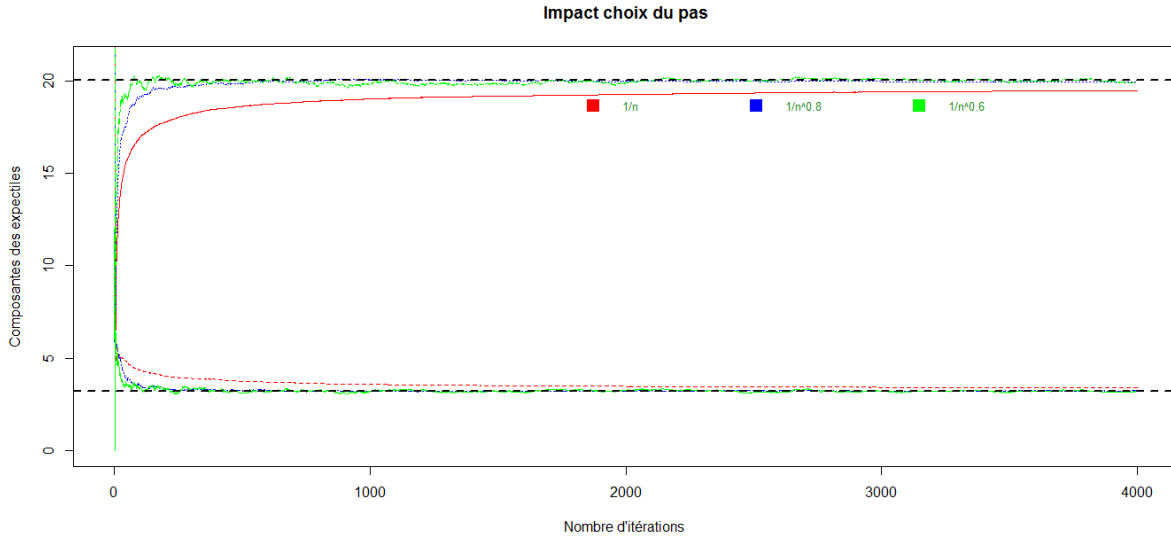


FIGURE 1.4 – Impact du choix de la suite de pas sur la convergence.

Nous considérons aussi le cas d'un modèle bivarié Pareto indépendant. Les deux variables aléatoires sont de lois Pareto, $X_i \sim Pa(a, b_i)$, $i \in \{1, 2\}$, telles que $a > 1$ et $b_i > 0$ pour tout $i \in \{1, 2\}$. Le système d'optimalité du L_1 -expectile est dans ce cas explicite. L'expectile $(x_1, x_2)^T$ est la solution unique du système suivant

$$l_{X_i}^\alpha(x_i) = -l_{X_j, X_i}^\alpha(x_j, x_i), (i, j) \in \{(1, 2), (2, 1)\},$$

avec,

$$l_{X_i}^\alpha(x_i) = (2\alpha - 1) \frac{b_i}{a - 1} \left(\frac{b_i}{b_i + x_i} \right)^{a-1} - (1 - \alpha) \left(x_i - \frac{b_i}{a - 1} \right),$$

et

$$l_{X_j, X_i}^\alpha(x_j, x_i) = \frac{b_j}{a - 1} \left(\left(\frac{b_i}{b_i + x_i} \right)^a - (1 - \alpha) \right) \left(\frac{b_j}{b_j + x_j} \right)^{a-1} - (1 - \alpha) \left(1 - \left(\frac{b_i}{b_i + x_i} \right)^a \right) \left(x_j - \frac{b_j}{a - 1} \right).$$

La méthode de Newton-Raphson permet d'obtenir la solution exacte.

Nous pouvons donc tester la convergence de l'algorithme stochastique vers la solution exacte. La Figure 1.5 illustre cette convergence pour les deux niveaux d'expectile $\alpha = 0.7$ et $\alpha = 0.99$.

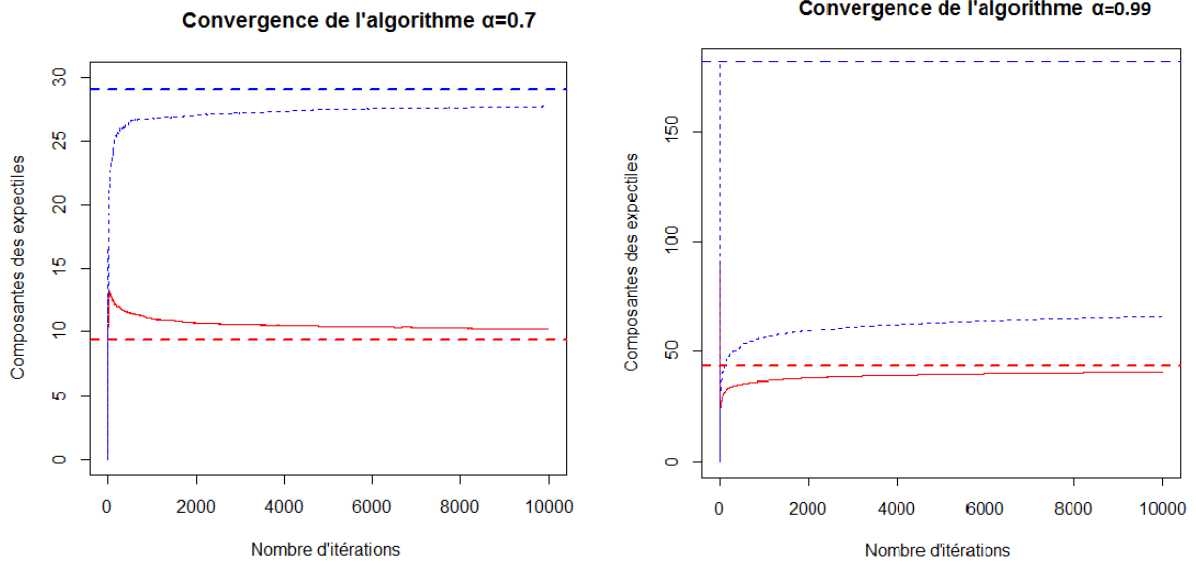


FIGURE 1.5 – Analyse de la convergence, L_1 –expectile, modèle Pareto indépendant ($a = 2$, rouge $b_1 = 10$, bleu $b_2 = 20$).

La convergence n'est pas bonne pour des valeurs de α proches de 1. L'algorithme n'est donc pas efficace pour estimer les expectiles asymptotiques. L'étude asymptotique semble incontournable pour pouvoir comprendre le comportement de l'expectile, notamment dans les cas où l'on ne dispose pas de solution analytique. La prochaine section sera dédiée à cet objectif.

Pour illustrer le cas de la dépendance, nous considérons un vecteur aléatoire $\mathbf{X} = (X_1, X_2)^T$ de lois marginales exponentielles $X_i \sim \mathcal{E}(\beta_i)$, $i = 1, 2$ et de copule FGM bivariée de paramètre $\theta \in [-1, 1]$. Pour tout $(i \neq j) \in \{1, 2\}$, l'expression de la copule FGM bivariée est

$$C_{\theta}^{\text{FGM}}(u, v) = uv[1 + \theta(1 - u)(1 - v)], \quad \forall (u, v) \in [0, 1]^2.$$

La copule FGM est présentée dans Nelsen (2007) [34] (Exemple 3.12., section 3.2.5). Nous rappelons qu'il s'agit d'une structure de dépendance faible qui ne peut pas prendre en compte les dépendances des queues des distributions. La fonction de répartition jointe est donnée par

$$\begin{aligned} F_{X_1, X_2}(x_1, x_2) &= C_{\theta}^{\text{FGM}}(F_{X_1}(x_1), F_{X_2}(x_2)) \\ &= F_{X_1}(x_1)F_{X_2}(x_2)[1 + \theta\bar{F}_{X_1}(x_1)\bar{F}_{X_2}(x_2)] \\ &= (1 - e^{-\beta_1 x_1})(1 - e^{-\beta_2 x_2}) + \theta(1 - e^{-\beta_1 x_1})(1 - e^{-\beta_2 x_2})e^{-\beta_1 x_1}e^{-\beta_2 x_2}. \end{aligned}$$

Le L_1 -expectile est l'unique solution du système d'équations suivant

$$\begin{aligned} (2\alpha - 1)\frac{1}{\beta_j}e^{-\beta_j x_j} - (1 - \alpha)\left(x_j - \frac{1}{\beta_j}\right) &= (1 - \alpha)(1 - e^{-\beta_j x_j})(1 - \theta e^{-\beta_j x_j})\left(x_i - \frac{1}{\beta_i}(1 - e^{-\beta_i x_i})\right) \\ &\quad + (1 - \alpha)\frac{\theta}{2}(1 - e^{-\beta_j x_j})e^{-\beta_j x_j}\left(2x_i - \frac{1}{\beta_i}(1 - e^{-2\beta_i x_i})\right) \\ &\quad - \alpha\frac{1}{\beta_i}e^{-\beta_i x_i}e^{-\beta_j x_j}\left(1 + \theta\left(1 - \frac{e^{-\beta_i x_i}}{2}\right)(1 - e^{-\beta_j x_j})\right), \end{aligned}$$

pour $(i, j) \in \{(1, 2), (2, 1)\}$.

La Figure 1.6 présente le résultat obtenu pour $\beta_1 = 0.05$ (rouge) et $\beta_2 = 0.25$ (bleu), dans les cas de dépendances négative et positive.

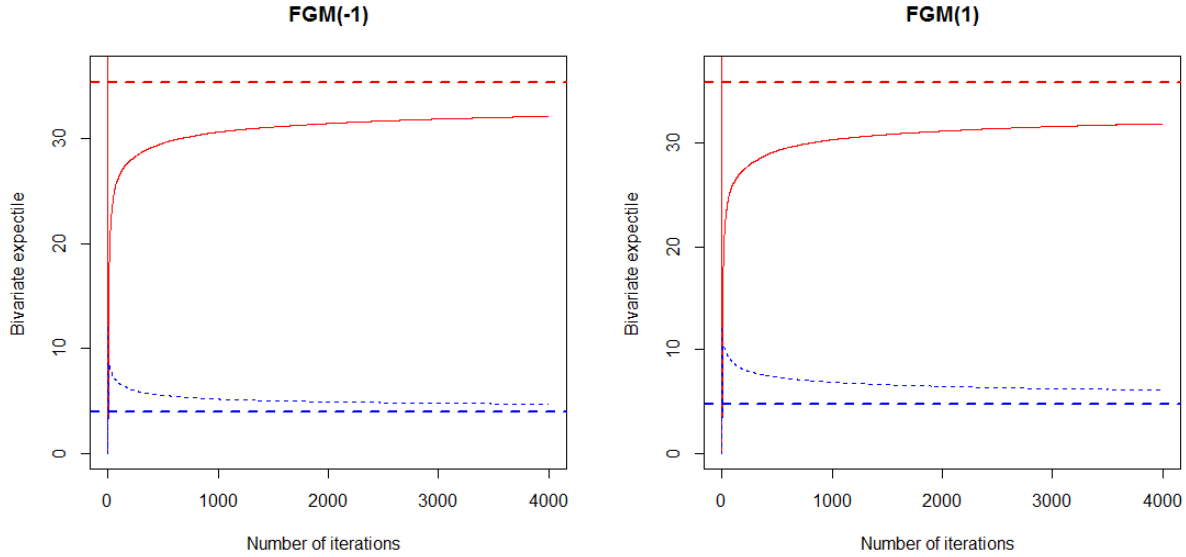


FIGURE 1.6 – Convergence de l’algorithme : L_1 -expectile $\alpha = 0.85$, modèle FGM (à gauche $\theta = -1$, à droite $\theta = 1$, rouge $\beta_1 = 0.05$, bleu $\beta_2 = 0.25$).

D’autres illustrations numériques sont proposées en annexe.

1.5 Comportement asymptotique

Soit $(X_1, \dots, X_d)^T$ un vecteur aléatoire dans \mathbb{R}^d . Le support de chaque variable aléatoire X_i est noté $[x_I^i, x_F^i]$, avec $x_I^i \in \mathbb{R} \cup \{-\infty\}$ et $x_F^i \in \mathbb{R} \cup \{+\infty\}$. On note \mathbf{X}_F le vecteur $(x_F^1, \dots, x_F^d)^T$, et \mathbf{X}_I le vecteur $(x_I^1, \dots, x_I^d)^T$.

Rappelons que nous avons défini les fonctions l_{X_i, X_j}^α pour tout $(i, j) \in \{1, \dots, d\}^2$ comme suit,

$$l_{X_i, X_j}^\alpha(x_i, x_j) = \alpha \mathbb{E}[(X_i - x_i)_+ \mathbb{1}_{\{X_j > x_j\}}] - (1 - \alpha) \mathbb{E}[(X_i - x_i)_- \mathbb{1}_{\{X_j < x_j\}}], \quad (1.24)$$

pour tout $(x_i, x_j) \in \mathbb{R}^2$. On note aussi $l_{X_i}^\alpha$ la fonction $l_{X_i}^\alpha(x_i) = l_{X_i, X_i}^\alpha(x_i, x_i)$.

Le système d’optimalité (1.9) s’écrit à l’aide de ces fonctions sous la forme suivante,

$$\sum_{i=1}^d \pi_{ki} l_{X_i, X_k}^\alpha(x_i, x_k) = 0 \quad \forall k \in \{1, \dots, d\}. \quad (1.25)$$

La question qui se pose ici concerne la nature du comportement asymptotique d’un expectile multivarié. La matrice de construction Σ est supposée à coefficients positifs ($\pi_{ij} \geq 0$, $\forall (i, j) \in \{1, \dots, d\}^2$).

Proposition 1.5.1 (Expectiles asymptotiques). *Pour tout vecteur aléatoire $\mathbf{X} = (X_1, \dots, X_d)^T$ dans \mathbb{R}^d , avec $\mathbb{E}[|X_i|] < +\infty$ pour tout $i \in \{1, \dots, d\}$, on a :*

$$\lim_{\alpha \rightarrow 1} \mathbf{e}_\alpha(\mathbf{X}) = \mathbf{X}_F, \quad \text{et} \quad \lim_{\alpha \rightarrow 0} \mathbf{e}_\alpha(\mathbf{X}) = \mathbf{X}_I.$$

Si de plus tous les supports sont infinis, on a pour tout $k \in \{1, \dots, d\}$

$$\lim_{x_k \rightarrow +\infty} \alpha(x_1, \dots, x_d) = 1, \quad \text{et} \quad \lim_{x_k \rightarrow -\infty} \alpha(x_1, \dots, x_d) = 0,$$

avec

$$\alpha(x_1, \dots, x_d) = \frac{\sum_{i=1}^d \pi_{ki} \mathbb{E}[(x_i - X_i)_+ \mathbb{1}_{\{x_k > X_k\}}]}{\sum_{i=1}^d \pi_{ki} \left(\mathbb{E}[(X_i - x_i)_+ \mathbb{1}_{\{X_k > x_k\}}] + \mathbb{E}[(x_i - X_i)_+ \mathbb{1}_{\{x_k > X_k\}}] \right)}. \quad (1.26)$$

Preuve Nous démontrons que

$$\lim_{\alpha \rightarrow 1} \mathbf{e}_\alpha(\mathbf{X}) = \mathbf{X}_F,$$

et nous déduisons la limite pour $\alpha \rightarrow 0$ à l'aide de la propriété de symétrie par rapport à α $\mathbf{e}_\alpha(-\mathbf{X}) = -\mathbf{e}_{1-\alpha}(\mathbf{X})$.

Nous considérons sans perte de généralité que $\pi_{ij} = 1 \forall (i, j) \in \{1, \dots, d\}^2$.

Quitte à prendre une sous suite, on considère que la limite $\lim_{\alpha \rightarrow 1} \mathbf{e}_\alpha(\mathbf{X})$ existe.

On définit l'ensemble J_∞ des indices dans $\{1, \dots, d\}$ pour lesquels la composante dans l'expectile multivarié tend vers le point terminal du support

$$J_\infty := \{i \in \{1, \dots, d\} \mid \lim_{\alpha \rightarrow 1} \mathbf{e}_\alpha^i(\mathbf{X}) = x_F^i\}.$$

Son ensemble complémentaire sera noté \bar{J}_∞

$$\bar{J}_\infty := \{i \in \{1, \dots, d\} \mid i \notin J_\infty\} = \{i \in \{1, \dots, d\} \mid \lim_{\alpha \rightarrow 1} \mathbf{e}_\alpha^i(\mathbf{X}) < x_F^i\}.$$

Supposons tout d'abord que $J_\infty = \emptyset$. Alors pour tout $i, j \in \bar{J}_\infty = \{1, \dots, d\}$, on peut calculer les limites des fonctions l^α ,

$$\lim_{\alpha \rightarrow 1} l_{X_i}^\alpha(x_i) = \mathbb{E}[(X_i - \lim_{\alpha \rightarrow 1} \mathbf{e}_\alpha^i(\mathbf{X}))_+]$$

et

$$\lim_{\alpha \rightarrow 1} l_{X_i, X_j}^\alpha(x_i, x_j) = \mathbb{E}[(X_i - \lim_{\alpha \rightarrow 1} \mathbf{e}_\alpha^i(\mathbf{X}))_+ \mathbb{1}_{\{X_j > \lim_{\alpha \rightarrow 1} \mathbf{e}_\alpha^j(\mathbf{X})\}}],$$

qui sont toutes des limites strictement positives. Cela est absurde, car contradictoire avec le système d'optimalité (1.25). On en déduit que $J_\infty \neq \emptyset$.

Il existe alors au moins un $k \in \{1, \dots, d\}$ tel que $\lim_{\alpha \rightarrow 1} \mathbf{e}_\alpha^k(\mathbf{X}) = x_F^k$. On a donc

$$\lim_{\alpha \rightarrow 1} l_{X_i, X_k}^\alpha(\mathbf{e}_\alpha^i(\mathbf{X}), \mathbf{e}_\alpha^k(\mathbf{X})) = 0, \quad \forall i \in \bar{J}_\infty,$$

et

$$\lim_{\alpha \rightarrow 1} \sum_{i \in J_\infty} l_{X_i, X_k}^\alpha(\mathbf{e}_\alpha^i(\mathbf{X}), \mathbf{e}_\alpha^k(\mathbf{X})) = - \lim_{\alpha \rightarrow 1} \sum_{i \in J_\infty} \left((1 - \alpha) \mathbb{E}[(X_i - \mathbf{e}_\alpha^i(\mathbf{X}))_- \mathbb{1}_{\{X_k < \mathbf{e}_\alpha^k(\mathbf{X})\}}] \right) = 0,$$

par (1.25).

On déduit pour $i = k$

$$\lim_{\alpha \rightarrow 1} l_{X_i}^\alpha(\mathbf{e}_\alpha^i(\mathbf{X})) = - \lim_{\alpha \rightarrow 1} \left((1 - \alpha) \mathbb{E}[(X_k - \mathbf{e}_\alpha^i(\mathbf{X}))_-] \right) = 0, \quad \forall i \in J_\infty.$$

Par passage à la limite ($\alpha \rightarrow 1$) dans toutes les $i^{\text{ème}}$, $i \in J_\infty$ équations du système d'optimisation (1.25), on déduit que

$$\lim_{\alpha \rightarrow 1} l_{X_k}^\alpha(\mathbf{e}_\alpha^k(\mathbf{X})) = - \lim_{\alpha \rightarrow 1} \left((1 - \alpha) \mathbb{E}[(X_k - \mathbf{e}_\alpha^k(\mathbf{X}))_-] \right) = 0, \quad \forall k \in J_\infty. \quad (1.27)$$

Maintenant, on suppose que $\bar{J}_\infty \neq \emptyset$, alors il existe $\ell \in \{1, \dots, d\}$ tel que $\lim_{\alpha \rightarrow 1} \mathbf{e}_\alpha^\ell(\mathbf{X}) < x_F^\ell$. Dans ce cas,

$$\lim_{\alpha \rightarrow 1} l_{X_i, X_\ell}^\alpha(\mathbf{e}_\alpha^i(\mathbf{X}), \mathbf{e}_\alpha^\ell(\mathbf{X})) = \mathbb{E}[(X_i - \lim_{\alpha \rightarrow 1} \mathbf{e}_\alpha^i(\mathbf{X}))_+ \mathbb{1}_{\{X_\ell > \lim_{\alpha \rightarrow 1} \mathbf{e}_\alpha^\ell(\mathbf{X})\}}] \in \mathbb{R}^+ \setminus \{+\infty\}, \quad \forall i \in \bar{J}_\infty,$$

et en utilisant (1.27)

$$\lim_{\alpha \rightarrow 1} l_{X_i, X_\ell}^\alpha(\mathbf{e}_\alpha^i(\mathbf{X}), \mathbf{e}_\alpha^\ell(\mathbf{X})) = - \lim_{\alpha \rightarrow 1} \left((1 - \alpha) \mathbb{E}[(X_i - \mathbf{e}_\alpha^i(\mathbf{X}))_- \mathbb{1}_{\{X_\ell < \lim_{\alpha \rightarrow 1} \mathbf{e}_\alpha^\ell(\mathbf{X})\}}] \right) = 0, \quad \forall i \in J_\infty,$$

car

$$\mathbb{E}[(X_i - \mathbf{e}_\alpha^i(\mathbf{X}))_- \mathbb{1}_{\{X_\ell < \lim_{\alpha \rightarrow 1} \mathbf{e}_\alpha^\ell(\mathbf{X})\}}] \leq \mathbb{E}[(X_i - \mathbf{e}_\alpha^i(\mathbf{X}))_-],$$

et

$$\lim_{\alpha \rightarrow 1} \left((1 - \alpha) \mathbb{E}[(X_i - \mathbf{e}_\alpha^i(\mathbf{X}))_-] \right) = 0,$$

pour tout $i \in J_\infty$.

Finalement, l'utilisation de la $\ell^{\text{ème}}$ équation du système d'optimalité (1.25) permet de conclure que

$$\mathbb{E}[(X_i - \lim_{\alpha \rightarrow 1} \mathbf{e}_\alpha^i(\mathbf{X}))_+ \mathbb{1}_{\{X_\ell > \lim_{\alpha \rightarrow 1} \mathbf{e}_\alpha^\ell(\mathbf{X})\}}] = 0, \quad \forall i \in \bar{J}_\infty,$$

et en particulier que

$$\mathbb{E}[(X_\ell - \lim_{\alpha \rightarrow 1} \mathbf{e}_\alpha^\ell(\mathbf{X}))_+] = 0,$$

cela est contradictoire avec l'hypothèse $\lim_{\alpha \rightarrow 1} \mathbf{e}_\alpha^\ell(\mathbf{X}) < x_F^\ell$. Nous déduisons alors que $\bar{J}_\infty = \emptyset$.

La deuxième partie de la preuve est évidente dans le cas des supports infinis. \square

L'expectile multivarié tend alors vers le vecteur des points terminaux de ses distributions marginales dans le cas asymptotique où $\alpha \rightarrow 1$. Nous pouvons donc étudier à partir de maintenant le comportement extrême de cette mesure de risque. Le cas asymptotique modélise la situation d'un risque extrême, d'où son importance pratique, notamment en assurance. Le chapitre suivant est dédié à l'étude de l'expectile asymptotique. Il s'appuie sur la théorie des valeurs extrêmes et sur le concept des variations régulières.

Conclusion

Dans ce premier chapitre, nous avons présenté différentes approches de construction de mesures de risque multivariées. Le point de départ de ces méthodes était la propriété de l'élicitabilité. Nous avons choisi ensuite une construction matricielle pour étudier ses propriétés de cohérence. Les expectiles multivariés construits à l'aide de matrices semi-définies positives à coefficients positifs sont capables de modéliser la dépendance et de prendre en compte la nature des distributions marginales. Une série d'illustrations numériques de différents expectiles matriciels pour des modèles de dépendance est présentée dans l'annexe 1.6 de chapitre. Nous avons également proposé une méthode d'approximation stochastique à cette famille de mesures issue de l'algorithme de Robbins-Monro. Pour les niveaux asymptotiques du seuil, l'approximation ne fournit pas une information pertinente sur le comportement du vecteur expectile. Une analyse théorique du comportement asymptotique semble plus utile pour construire une idée sur l'impact de la dépendance asymptotique sur ce comportement. C'est l'objet du prochain chapitre.

Bibliographie

- [1] B. Abdous and R. Theodorescu. Note on the spatial quantile of a random vector. *Statistics & Probability Letters*, 13(4) :333–336, 1992. 54
- [2] P. Artzner, F. Delbaen, J-M. Eber, and D. Heath. Coherent measures of risk. *Mathematical Finance*, 9(3) :203–228, 1999. 4, 39, 44, 47, 130, 131, 177
- [3] F. Bellini and V. Bignozzi. On elicitable risk measures. *Quantitative Finance*, 15(5) :725–733, 2015. 6, 40, 41, 42, 43, 44, 128
- [4] F. Bellini and E. Di Bernardino. Risk management with expectiles. *The European Journal of Finance*, pages 1–20, 2015. 40, 45, 46, 87, 104, 110
- [5] F. Bellini, B. Klar, A. Müller, and E.R. Gianin. Generalized quantiles as risk measures. *Insurance : Mathematics and Economics*, 54 :41–48, 2014. 6, 40, 46
- [6] J.R. Blum. Multidimensional stochastic approximation methods. *The Annals of Mathematical Statistics*, pages 737–744, 1954. 40, 64
- [7] A. Buch and G. Dornflaetter. Coherent risk measures, coherent capital allocations and the gradient allocation principle. *Insurance : Mathematics and Economics*, 42(1) :235–242, 2008. 39
- [8] P. Chaudhuri. On a geometric notion of quantiles for multivariate data. *Journal of the American Statistical Association*, 91(434) :862–872, 1996. 54
- [9] A. Cousin and E. Di Bernardino. On multivariate extensions of value-at-risk. *Journal of Multivariate Analysis*, 119 :32–46, 2013. 40
- [10] A. Cousin and E. Di Bernardino. On multivariate extensions of conditional-tail-expectation. *Insurance : Mathematics and Economics*, 55 :272–282, 2014. 40
- [11] A. Daouia, S. Girard, and G. Stupfler. Estimation of tail risk based on extreme expectiles. In *Workshop Extremes-Copulas-Actuarial science*, 2016. 46
- [12] F. Delbaen, F. Bellini, V. Bignozzi, and J.F. Ziegel. Risk measures with the cxls property. *arXiv preprint arXiv :1411.0426*, 2014. 43
- [13] M. Denuit. Coherent allocation of risk capital. *Journal of risk*, 4 :1–34, 2001. 28, 39
- [14] E. Di Bernardino. *Modélisation de la dépendance et mesures de risque multidimensionnelles*. PhD thesis, Citeseer, 2013. 40
- [15] E. Di Bernardino, JM. Fernández-Ponce, F. Palacios-Rodríguez, and MR. Rodríguez-Griñolo. On multivariate extensions of the conditional value-at-risk measure. *Insurance : Mathematics and Economics*, 61 :1–16, 2015. 40
- [16] M. Dufflo. Random iterative models, volume 34 of applications of mathematics, 1997. 64
- [17] P. Embrechts, G. Puccetti, L. Rüschendorf, R. Wang, and A. Beleraaj. An academic response to basel 3.5. *Risks*, 2(1) :25–48, 2014. 39
- [18] E. Emmer, M. Kratz, and D. Tasche. What is the best risk measure in practice ? A comparison of standard measures. December 2013. 40, 41, 45, 46
- [19] T. Fissler and J.F. Ziegel. Higher order elicibility and osband’s principle. *arXiv preprint arXiv :1503.08123*, 2015. 46, 47
- [20] P. Fraysse. *Estimation récursive dans certains modeles de déformation*. PhD thesis, Université Sciences et Technologies-Bordeaux I, 2013. 65

- [21] R. Frongillo and I. Kash. Vector-valued property elicitation. *Preprint*, 2014. 47
- [22] T. Gneiting. Making and Evaluating Point Forecasts . *Journal of the American Statistical Association*, 106(494) :746–762, June 2011. 6, 40, 41, 42, 43, 44, 128
- [23] A. H. Hamel and F. Heyde. Duality for set-valued measures of risk. *SIAM Journal on Financial Mathematics*, 1(1) :66–95, 2010. 53, 128
- [24] W. Hürlimann. On higher-degree bivariate stop-loss transforms, with applications. *Blätter der DGVMF*, 25(3) :485–502, 2002. 52
- [25] E. Jouini, M. Meddeb, and N. Touzi. Vector-valued coherent risk measures. *Finance and Stochastics*, 8(4) :531–552, 2004. 4, 39, 47, 53, 128, 144, 177
- [26] W K. Newey and J. L. Powell. Asymmetric Least Squares Estimation and Testing. *Econometrica*, 55(4) :819–847, July 1987. 6, 44
- [27] J. Kiefer, J. Wolfowitz, et al. Stochastic estimation of the maximum of a regression function. *The Annals of Mathematical Statistics*, 23(3) :462–466, 1952. 64
- [28] R. Kiesel, R. Rühlcke, G. Stahl, and J. Zheng. Conceptualizing robustness in risk management. *Available at SSRN 2065723*, 2012. 46
- [29] N.S. Lambert, D.M. Pennock, and Y. Shoham. Eliciting properties of probability distributions. In *Proceedings of the 9th ACM Conference on Electronic Commerce*, pages 129–138. ACM, 2008. 44, 46, 47
- [30] J. Lelong. *Etude asymptotique des algorithmes stochastiques et calcul du prix des options Parisiennes*. PhD thesis, Ecole des Ponts ParisTech, 2007. 68
- [31] T. Mao and F. Yang. Risk concentration based on expectiles for extreme risks under fgm copula. *Insurance : Mathematics and Economics*, 64 :429–439, 2015. 46, 87
- [32] V. Maume-Deschamps, D. Rullière, and K. Said. Multivariate extensions of expectiles risk measures. working paper or preprint, September 2016. 39
- [33] V. Maume-Deschamps, D. Rulliere, and K. Said. On a capital allocation by minimization of some risk indicators. *European Actuarial Journal*, pages 1–20, 2016. 40, 47, 125, 143, 144, 145, 146, 156
- [34] R.B. Nelsen. *An introduction to copulas*. Springer, 2007. 70, 77, 92, 111, 157, 161
- [35] K.H. Osband. *Providing Incentives for Better Cost Forecasting*. University of California, Berkeley, 1985. 42, 43, 46
- [36] H. Robbins and S. Monro. A stochastic approximation method. *The annals of mathematical statistics*, pages 400–407, 1951. 40, 64
- [37] R.T. Rockafellar and S. Uryasev. The fundamental risk quadrangle in risk management, optimization and statistical estimation . *Surveys in Operations Research and Management Science*, (18) :33–53, March 2013. 44
- [38] N. S. Lambert. Elicitation and Evaluation of Statistical Forecasts . July 2013. 42
- [39] S. Sarykalin, G. Serraino, and S. Uryasev. Value-at-risk vs. conditional value-at-risk in risk management and optimization. *Tutorials in Operations Research. INFORMS, Hanover, MD*, pages 270–294, 2008. 39
- [40] I. Steinwart, C. Pasin, R. C. Williamson, and S. Zhang. Elicitation and Identification of Properties. *Workshop and Conference Proceedings*, 34 :1–45, 2014. 40, 42, 128
- [41] R. Wang and J.F. Ziegel. Elicitable distortion risk measures : A concise proof. *Statistics & Probability Letters*, 100 :172–175, 2015. 40, 44
- [42] S. Weber. Distribution-invariant risk measures, information, and dynamic consistency. *Mathematical Finance*, 16(2) :419–441, 2006. 43
- [43] J. F. Ziegel. Coherence and elicibility. *Mathematical Finance*, 2014. 40, 44, 128

1.6 Annexe : Illustrations numériques

1.6.1 Modèle FGM Bivarié

Considérons un vecteur aléatoire $\mathbf{X} = (X_1, \dots, X_d)^T$ de distributions marginales exponentielles $X_i \sim \mathcal{E}(\beta_i)$, $i = 1, \dots, d$ et de structures de dépendance bivariées FGM de paramètres θ_{ij} , $i \neq j = 1, \dots, d$. Pour tout $i \in \{1, \dots, d\}$, la VaR de seuil α est donnée explicitement par la formule

$$\text{VaR}_\alpha(X_i) = -\frac{1}{\beta_i} \ln(1 - \alpha). \quad (1.28)$$

L'espérance tronquée à droite est donnée par

$$\mathbb{E}[X_i \mathbb{1}_{\{X_i \leq x_i\}}] = \frac{1}{\beta_i} (1 - e^{-\beta_i x_i}) - x_i e^{-\beta_i x_i},$$

l'espérance tronquée à gauche par

$$\mathbb{E}[X_i \mathbb{1}_{\{x_i \leq X_i\}}] = \frac{1}{\beta_i} e^{-\beta_i x_i} + x_i e^{-\beta_i x_i},$$

la fonction Stop-Loss est par suite

$$\mathbb{E}[(X_i - x_i)_+] = \frac{1}{\beta_i} e^{-\beta_i x_i},$$

et

$$\mathbb{E}[(X_i - x_i)_-] = x_i - \frac{1}{\beta_i} (1 - e^{-\beta_i x_i}),$$

nous en déduisons l'expression de la fonction $l_{X_i}^\alpha$

$$l_{X_i}^\alpha(x_i) = (2\alpha - 1) \frac{1}{\beta_i} e^{-\beta_i x_i} + (1 - \alpha) \left(\frac{1}{\beta_i} - x_i \right).$$

L'expectile univarié $e_\alpha(X_i)$ est donc solution unique de l'équation

$$(1 - \alpha) \left(x_i - \frac{1}{\beta_i} \right) = (2\alpha - 1) \frac{1}{\beta_i} e^{-\beta_i x_i}. \quad (1.29)$$

Pour tout $(i \neq j) \in \{1, \dots, 2\}$, l'expression de la copule FGM bivariée est donnée par

$$C_{\theta_{ij}}^{\text{FGM}}(u, v) = uv[1 + \theta_{ij}(1 - u)(1 - v)], \quad \forall (u, v) \in [0, 1]^2.$$

Cette copule est présentée en détail dans Nelsen (2007) [34] (Exemple 3.12., section 3.2.5). Nous rappelons qu'il s'agit uniquement d'une structure de dépendance faible qui ne prend pas en compte les dépendances extrêmes. Le paramètre de la copule $\theta \in [-1, 1]$. On peut donc écrire la fonction de répartition bivariée sous la forme

$$\begin{aligned} F_{X_i, X_j}(x_i, x_j) &= C_{\theta_{ij}}^{\text{FGM}}(F_{X_i}(x_i), F_{X_j}(x_j)) \\ &= F_{X_i}(x_i) F_{X_j}(x_j) [1 + \theta_{ij} \bar{F}_{X_i}(x_i) \bar{F}_{X_j}(x_j)] \\ &= (1 - e^{-\beta_i x_i})(1 - e^{-\beta_j x_j}) + \theta_{ij} (1 - e^{-\beta_i x_i})(1 - e^{-\beta_j x_j}) e^{-\beta_i x_i} e^{-\beta_j x_j}, \end{aligned}$$

pour tout $(x_i, x_j) \in \mathbb{R}_+^2$.

Les fonctions Stop-Loss transformée sont données pour tout $(x_i, x_j) \in \mathbb{R}_+^2$ par

$$\begin{aligned} \mathbb{E}[(X_i - x_i)_+ \mathbb{1}_{\{X_j \geq x_j\}}] &= \int_{x_i}^{+\infty} \bar{F}_{X_i, X_j}(t, x_j) dt \\ &= \int_{x_i}^{+\infty} \left(\bar{F}_{X_i}(t) + F_{X_i, X_j}(t, x_j) - F_{X_i}(x_j) \right) dt \\ &= \left(\bar{F}_{X_j}(x_j) + \theta_{ij} \bar{F}_{X_j}(x_j) F_{X_j}(x_j) \right) \int_{x_i}^{+\infty} F_{X_i}(t) dt - \theta_{ij} \bar{F}_{X_j}(x_j) F_{X_j}(x_j) \int_{x_i}^{+\infty} F_{X_i}^2(t) dt \\ &= \left(\bar{F}_{X_j}(x_j) + \theta_{ij} \bar{F}_{X_j}(x_j) F_{X_j}(x_j) \right) \mathbb{E}[(X_i - x_i)_+] - \frac{\theta_{ij}}{2} \bar{F}_{X_j}(x_j) F_{X_j}(x_j) \mathbb{E}[(X_i - 2x_i)_+] \\ &= \frac{1}{\beta_i} e^{-\beta_i x_i} e^{-\beta_j x_j} \left(1 + \theta_{ij} \left(1 - \frac{e^{-\beta_i x_i}}{2} \right) (1 - e^{-\beta_j x_j}) \right), \end{aligned}$$

et leurs fonctions duales par

$$\begin{aligned}
\mathbb{E}[(X_i - x_i)_- \mathbb{1}_{\{X_j \leq x_j\}}] &= \int_0^{x_i} F_{X_i, X_j}(t, x_j) dt \\
&= (F_{X_j}(x_j) \int_0^{x_i} F_{X_i}(t) dt + \theta_{ij} \bar{F}_{X_j}(x_j) F_{X_j}(x_j) \int_0^{x_i} (F_{X_i}(t) - F_{X_i}(t)^2) dt \\
&= ((F_{X_j}(x_j) - \theta_{ij} \bar{F}_{X_j}(x_j) F_{X_j}(x_j)) \int_0^{x_i} F_{X_i}(t) dt + \theta_{ij} \bar{F}_{X_j}(x_j) F_{X_j}(x_j) \int_0^{x_i} F_{X_i}(2t) dt \\
&= ((F_{X_j}(x_j) - \theta_{ij} \bar{F}_{X_j}(x_j) F_{X_j}(x_j)) \mathbb{E}[(X_i - x_i)_-] + \frac{\theta_{ij}}{2} \bar{F}_{X_j}(x_j) F_{X_j}(x_j) \mathbb{E}[(X_i - 2x_i)_-] \\
&= (1 - e^{-\beta_j x_j})(1 - \theta_{ij} e^{-\beta_j x_j}) \left(x_i - \frac{1}{\beta_i} (1 - e^{-\beta_i x_i}) \right) \\
&\quad + \frac{\theta_{ij}}{2} (1 - e^{-\beta_j x_j}) e^{-\beta_j x_j} \left(2x_i - \frac{1}{\beta_i} (1 - e^{-2\beta_i x_i}) \right).
\end{aligned}$$

On a donc obtenu une expression explicite des fonctions

$$l_{X_i, X_j}^\alpha(x_i, x_j) = \alpha \mathbb{E}[(X_i - x_i)_+ \mathbb{1}_{\{X_j \geq x_j\}}] - (1 - \alpha) \mathbb{E}[(X_i - x_i)_- \mathbb{1}_{\{X_j \leq x_j\}}],$$

l'expectile multivarié construit à l'aide d'une matrice $\Sigma = (\pi_{ij})_{1 \leq i, j \leq d}$ vérifiant les conditions de la stricte convexité est l'unique solution du système d'équations

$$l_{X_k}^\alpha(x_k) = - \sum_{i=1, i \neq k}^d \frac{\pi_{ik}}{\pi_k} l_{X_i, X_k}^\alpha(x_i, x_k), \quad \forall k \in \{1, \dots, d\}.$$

L_1 -expectile

Dans le cas du L_1 -expectile, tous les coefficients de la matrice Σ sont égaux à 1. L'expectile multivarié est l'unique solution du système d'équations

$$l_{X_k}^\alpha(x_k) = - \sum_{i=1, i \neq k}^d l_{X_i, X_k}^\alpha(x_i, x_k), \quad \forall k \in \{1, \dots, d\}.$$

La Figure 1.7 montre l'impact de l'approche multivariée par comparaison avec l'expectile univarié et le quantile du même seuil. L'impact est plus important au niveau de la branche la moins risquée.

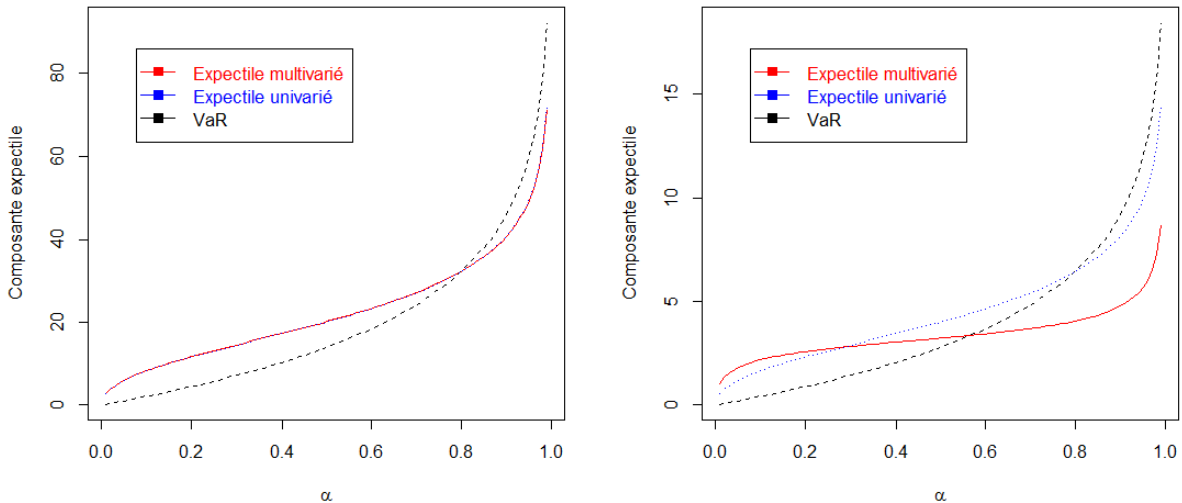


FIGURE 1.7 – Impact de l'approche multivariée (à gauche $X_1 \sim \mathcal{E}(0.05)$, à droite $X_2 \sim \mathcal{E}(0.25)$).

La Figure 1.8 illustre l'impact de la dépendance sur l'expectile de seuil $\alpha = 0.95$. Le L_1 -expectile semble converger vers l'expectile univarié dans le cas comonotone. Cette convergence est plus rapide pour la branche la plus risquée X_1 .

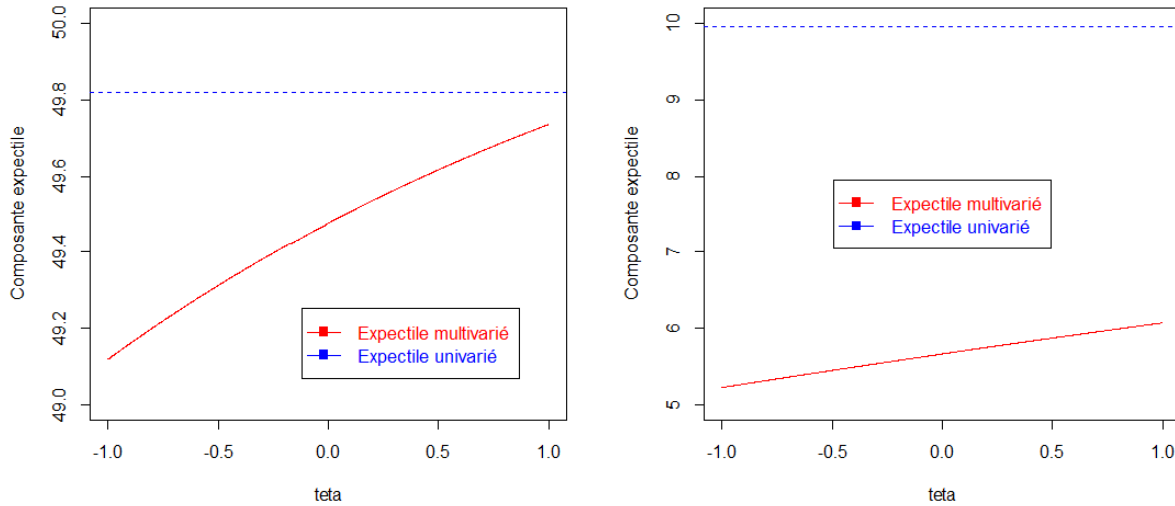


FIGURE 1.8 – Impact de la dépendance $\alpha = 0.95$ (à gauche $X_1 \sim \mathcal{E}(0.05)$, à droite $X_2 \sim \mathcal{E}(0.25)$).

Le même comportement est généralisable à tous les expectiles construits à l'aide de matrices statique symétrique semi-définies positives à coefficients positifs et non nuls sur la diagonale.

Expectiles de corrélation

Nous considérons uniquement le cas des dépendances positives. L'objectif de cette sous-section est de comparer l'impact du choix de la mesure de corrélation sur le comportement de l'expectile multivarié obtenu. Ainsi que l'impact du choix de la matrice. Pour cela, nous étudions trois vecteurs expectiles :

Expectile de Pearson La matrice de la construction sera dans ce cas composée des coefficients de corrélation de Pearson

$$\pi_{ij} = (\rho_P)_{ij} = \frac{\theta_{ij}}{4}.$$

Expectile de Spearman La matrice de la construction sera dans ce cas composée les rho de Pearson bivariés

$$\pi_{ij} = (\rho_S)_{ij} = \frac{\theta_{ij}}{3}.$$

Expectile de Kendall Dans cette troisième illustration, la matrice sera composée des τ de Kendall bivariée

$$\pi_{ij} = \tau_{ij} = \frac{2\theta_{ij}}{9}.$$

Dans les trois cas, les diagonales des matrices sont composées d'éléments égaux à 1. Nous rappelons que le vecteur expectile est l'unique solution du système d'équations

$$l_{X_k}^\alpha(x_k) = - \sum_{i=1, i \neq k}^d \frac{\pi_{ik}}{\pi_k} l_{X_i, X_k}^\alpha(x_i, x_k), \quad \forall k \in \{1, \dots, d\}.$$

La Figure 1.9 illustre les expectiles obtenus.

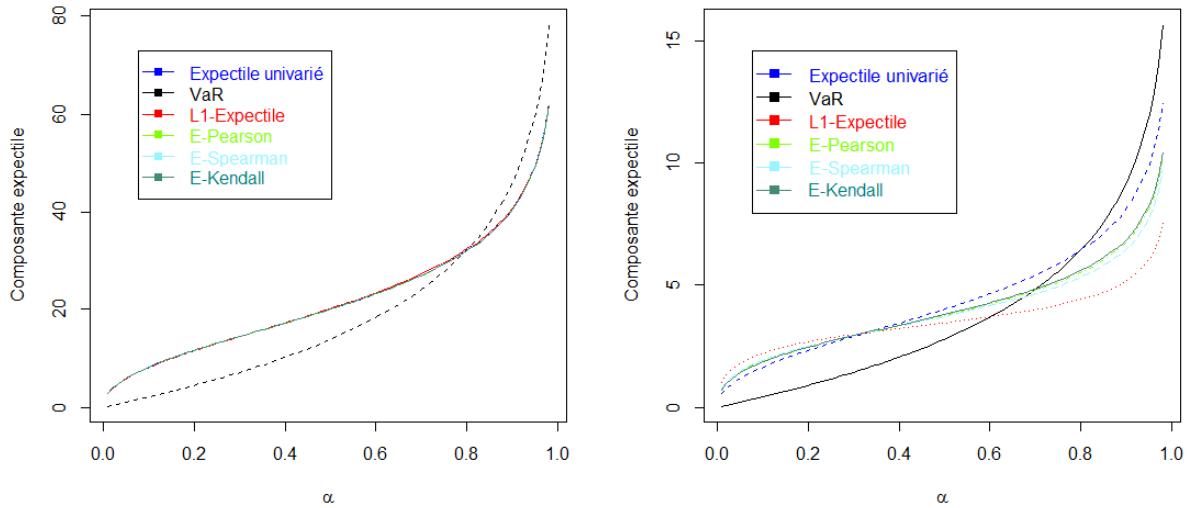


FIGURE 1.9 – Impact du choix de la construction $\theta = 1$ (à gauche $X_1 \sim \mathcal{E}(0.05)$, à droite $X_2 \sim \mathcal{E}(0.25)$).

L'ordre des courbes suit l'ordre des coefficients de la dépendance. Plus la dépendance est surévaluée, plus on s'éloigne de la courbe du cas univarié.

L'impact du choix de la matrice Σ est visible dans le comportement de l'expectile obtenu face à la dépendance. La Figure 1.10 montre que l'impact de la dépendance diffère entre les deux choix de construction. Si le L_1 expectile semble croissant en fonction du degré de dépendance et converge dans le cas comonomotes vers l'expectile univarié, les expectiles de corrélation sont décroissants en fonction de la dépendance.

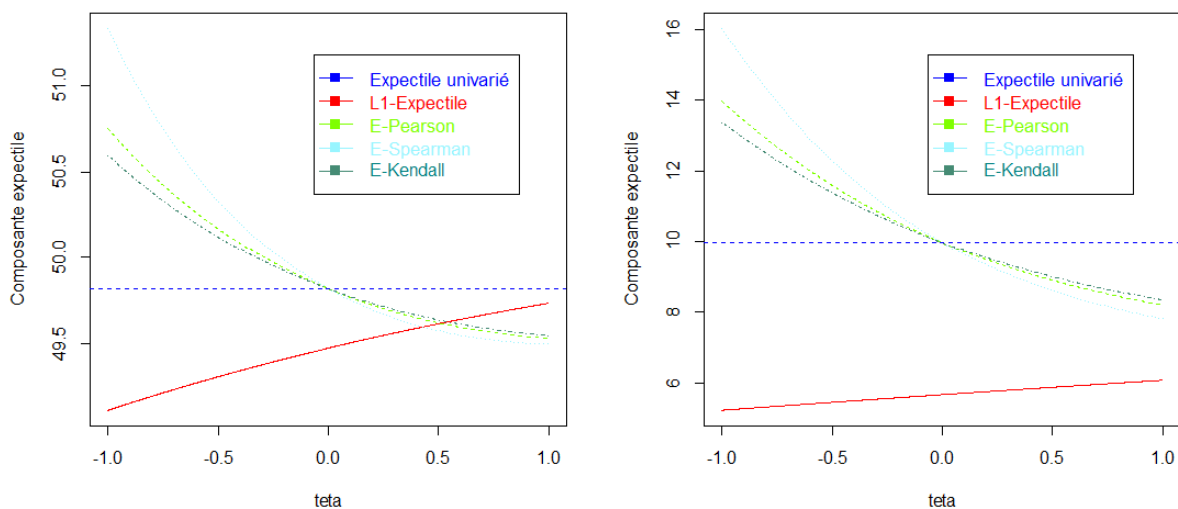


FIGURE 1.10 – Impact de la dépendance, $\alpha = 0.95$ (à gauche $X_1 \sim \mathcal{E}(0.05)$, à droite $X_2 \sim \mathcal{E}(0.25)$).

1.6.2 Modèle Pareto

Cas d'indépendance

Nous considérons deux variables aléatoires de loi Pareto, $X_i \sim Pa(a, b_i)$, $i \in \{1, \dots, d\}$, telles que $a > 1$ et $b_i > 0$ pour tout $i \in \{1, \dots, d\}$.

Pour $i \in \{1, \dots, d\}$, la fonction de survie est donnée par

$$\bar{F}_{X_i}(x) = \left(\frac{b_i}{b_i + x} \right)^a,$$

l'espérance est

$$\mathbb{E}[X_i] = \frac{b_i}{a-1},$$

et la fonction Stop-Loss est

$$\mathbb{E}[(X_i - x)_+] = \frac{b_i}{a-1} \left(\frac{b_i}{b_i + x} \right)^{a-1}.$$

On déduit de ces expressions celle de la fonction $l_{X_i}^\alpha$

$$\begin{aligned} l_{X_i}^\alpha(x_i) &= (2\alpha - 1)\pi_{X_i}(x_i) - (1 - \alpha)(x_i - \mathbb{E}[X_i]) \\ &= (2\alpha - 1)\frac{b_i}{a-1} \left(\frac{b_i}{b_i + x_i} \right)^{a-1} - (1 - \alpha) \left(x_i - \frac{b_i}{a-1} \right), \end{aligned}$$

l'expectile univarié $e_\alpha(X_i)$ est donc l'unique solution de l'équation

$$x = \frac{2\alpha - 1}{1 - \alpha} \frac{b_i}{a-1} \left(\frac{b_i}{b_i + x} \right)^{a-1} + \frac{b_i}{a-1}.$$

La VaR de seuil α est donnée par l'expression

$$\text{VaR}_\alpha(X_i) = b_i \left((1 - \alpha)^{-\frac{1}{a}} - 1 \right).$$

Dans le cas de l'indépendance, nous obtenons directement l'expression de la fonction l_{X_j, X_i}^α pour tout $j \in \{1, \dots, d\} \setminus \{i\}$,

$$\begin{aligned} l_{X_j, X_i}^\alpha(x_j, x_i) &= \alpha \pi_{X_j}(x_j) \bar{F}_{X_i}(x_i) - (1 - \alpha)(\pi_{X_j}(x_j) + x_j - \mathbb{E}[X_j])(1 - \bar{F}_{X_i}(x_i)) \\ &= (\bar{F}_{X_i}(x_i) - (1 - \alpha))\pi_{X_j}(x_j) - (1 - \alpha)(x_j - \mathbb{E}[X_j])(1 - \bar{F}_{X_i}(x_i)) \\ &= \frac{b_j}{a-1} \left(\left(\frac{b_i}{b_i + x_i} \right)^a - (1 - \alpha) \right) \left(\frac{b_j}{b_j + x_j} \right)^{a-1} - (1 - \alpha) \left(1 - \left(\frac{b_i}{b_i + x_i} \right)^a \right) \left(x_j - \frac{b_j}{a-1} \right). \end{aligned}$$

Ces expressions sont suffisantes pour obtenir des systèmes d'optimalité explicites. Dans le cas bivarié, le L_1 -expectile bivarié est l'unique solution du système

$$l_{X_i}^\alpha(x_i) = -l_{X_j, X_i}^\alpha(x_j, x_i), (i, j) \in \{(1, 2), (2, 1)\}.$$

La méthode de Newton-Raphson permet d'obtenir la solution exacte.

Les deux distributions possèdent le même paramètres d'échelle a , elles sont donc à queues équivalentes, et le paramètre de leur équivalence c_2 est déterminé explicitement

$$c_2 = \lim_{x \rightarrow +\infty} \frac{\bar{F}_{X_2}(x)}{\bar{F}_{X_1}(x)} = \lim_{x \rightarrow +\infty} \frac{\left(\frac{b_2}{b_2 + x} \right)^a}{\left(\frac{b_1}{b_1 + x} \right)^a} = \left(\frac{b_2}{b_1} \right)^a.$$

L'étude asymptotique présentée dans le chapitre 2 de la thèse est donc applicable pour ce modèle de distributions marginales.

Construction par mélange

On considère un vecteur aléatoire (X_1, \dots, X_d) construit par un mélange de distributions exponentielles par une distribution Gamma. Pour $i \in \{1, 2, \dots, d\}$, $X_i \sim \mathcal{E}(\beta_i \theta)$, avec $(\beta_1 < \beta_2 \dots < \beta_d)$, et $\theta \sim \Gamma(a, b)$. Pour tout $i \in \{1, \dots, d\}$, la fonction de survie d'une variable aléatoire X_i sera donc de la forme suivante

$$\bar{F}_{X_i}(x) = \int_0^\infty \bar{F}_{X_i|\Theta=\theta} f_\Theta(\theta) d\theta = \int_0^\infty e^{-\beta_i \theta x} f_\Theta(\theta) d\theta = \left(1 + \frac{\beta_i x}{b}\right)^{-a},$$

et par conséquent, X_i suit une distribution de Pareto de paramètres $(a, \frac{b}{\beta_i})$. Ces variables aléatoires sont donc indépendantes conditionnellement.

Il s'agit également d'une construction de la copule de survie de Clayton.

Pour tout $(i, j) \in \{1, \dots, d\}^2$, la fonction de répartition bivariée est donnée par

$$\bar{F}_{X_i, X_j}(x_i, x_j) = \left(1 + \frac{\beta_i}{b} x_i + \frac{\beta_j}{b} x_j\right)^{-a},$$

la densité jointe du couple (X_i, X_j) est

$$f_{X_i, X_j}(x_i, x_j) = \frac{a(a+1)}{b^2} \beta_1 \beta_2 \left(\frac{1}{1 + \frac{\beta_i}{b} x_i + \frac{\beta_j}{b} x_j}\right)^{a+2}.$$

Nous considérons pour simplifier les calculs le cas $b = 1$. Nous pouvons déterminer explicitement toutes les quantités probabilistes nécessaires pour obtenir un système d'optimalité explicite pour une construction matricielle d'un expectile multivarié, ainsi pour tout $(X_i, X_j) \in \{X_1, \dots, X_d\}^2$

$$\begin{aligned} \mathbb{E}[X_i \mathbb{1}_{\{X_i \leq x_i\}}] &= \frac{1}{(a-1)\beta_i} \left(1 - \frac{1}{(1 + \beta_i x_i)^{a-1}}\right) - x_i \frac{1}{(1 + \beta_i x_i)^a}; \\ \mathbb{E}[X_i \mathbb{1}_{\{X_i \geq x_i\}}] &= x_i \frac{1}{(1 + \beta_i x_i)^a} + \frac{1}{(a-1)\beta_i} \frac{1}{(1 + \beta_i x_i)^{a-1}}; \\ \mathbb{E}[(X_i - x_i)_+] &= \frac{1}{(a-1)\beta_i} \frac{1}{(1 + \beta_i x_i)^{a-1}}; \\ \mathbb{E}[(X_i - x_i)_-] &= \frac{1}{(a-1)\beta_i} \frac{1}{(1 + \beta_i x_i)^{a-1}} + \left(x_i - \frac{1}{(a-1)\beta_i}\right); \\ \mathbb{E}[(X_i - x_i)_+ \mathbb{1}_{\{X_j \geq x_j\}}] &= \frac{1}{(a-1)\beta_i} \frac{1}{(1 + \beta_i x_i + \beta_j x_j)^{a-1}}; \\ \mathbb{E}[(X_i - x_i)_- + \mathbb{1}_{\{X_j \leq x_j\}}] &= x_i - \frac{1}{(a-1)\beta_i} - \frac{x_i}{(1 + \beta_j x_j)^a} \\ &\quad + \frac{1}{(a-1)\beta_i} \left[\frac{1}{(1 + \beta_i x_i)^{a-1}} + \frac{1}{(1 + \beta_j x_j)^{a-1}} - \frac{1}{(1 + \beta_i x_i + \beta_j x_j)^{a-1}}\right]. \end{aligned}$$

Dans ce modèle les marginales changent au même temps que la copule.

L_1 -Expectile

Pour le L_1 -Expectile, la Figure 1.11 illustre la différence observée entre les composantes de l'expectile multivarié, les expectiles univariés et les quantiles.

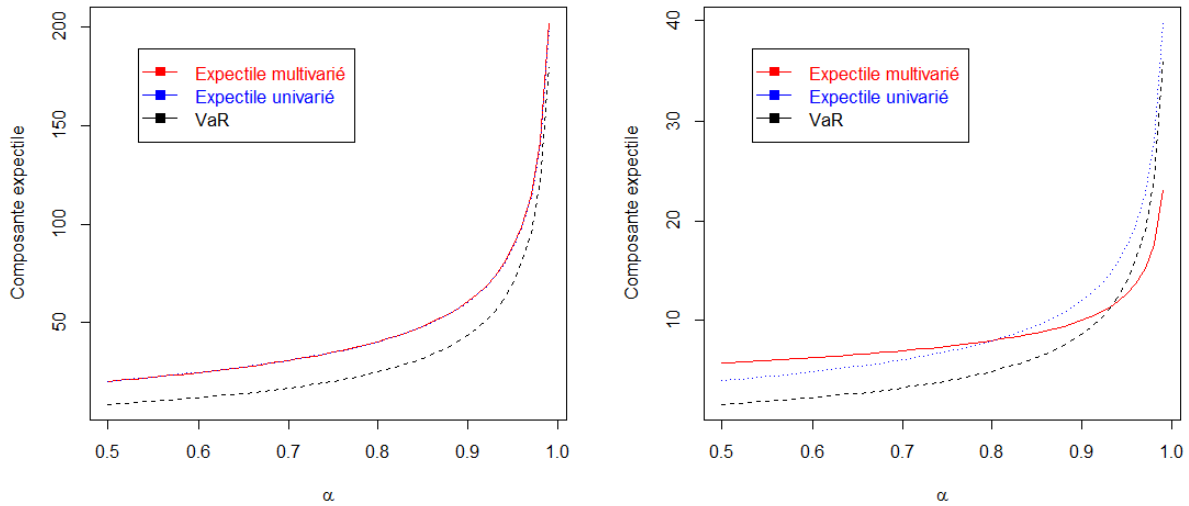


FIGURE 1.11 – Impact de l’approche multivariée $\beta_1 = 0.05, \beta_2 = 0.25, \theta = 1/a = 0.9$ (à gauche X_1 , à droite X_2).

L’impact de l’approche multivariée est plus visible sur la branche la moins risquée. La position par rapport au quantile n’est plus la même entre l’expectile univarié et la composante de l’expectile bivarié. Afin de mieux comprendre ce comportement, nous examinons dans la Figure 1.12 cet impact pour d’autres niveaux de la dépendance $\theta = 0.1$ et $\theta = 0.5$.

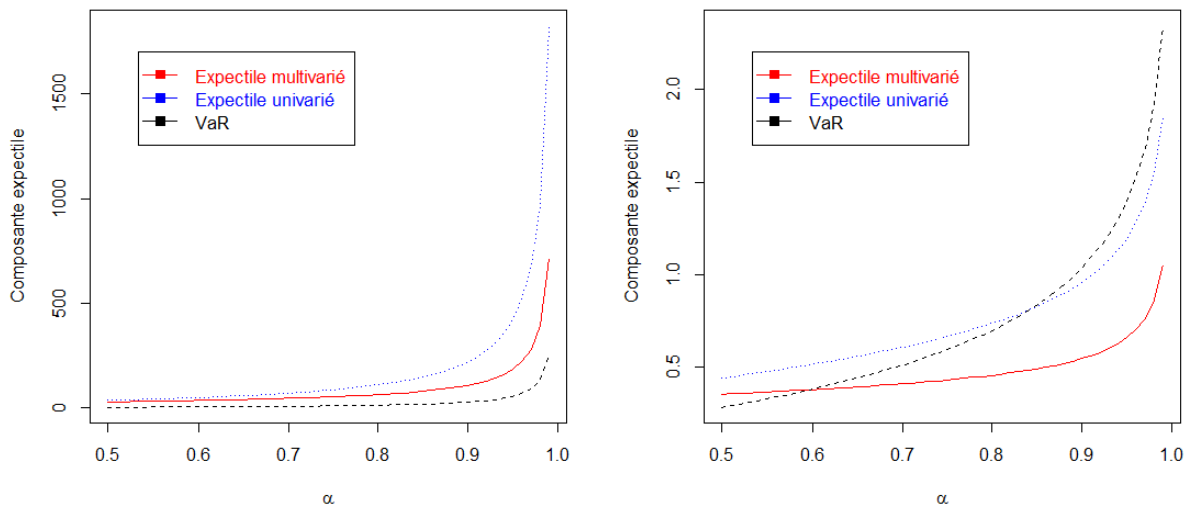


FIGURE 1.12 – Impact de l’approche multivariée X_2 . $\beta_1 = 0.05, \beta_2 = 0.25$ (à gauche $\theta = 0.1$, à droite $\theta = 0.5$).

La position du quantile par rapport à l’expectile change en fonction de la dépendance. La dépendance joue un rôle essentiel dans la détermination du comportement de l’expectile multivarié pour les grands niveaux du seuil.

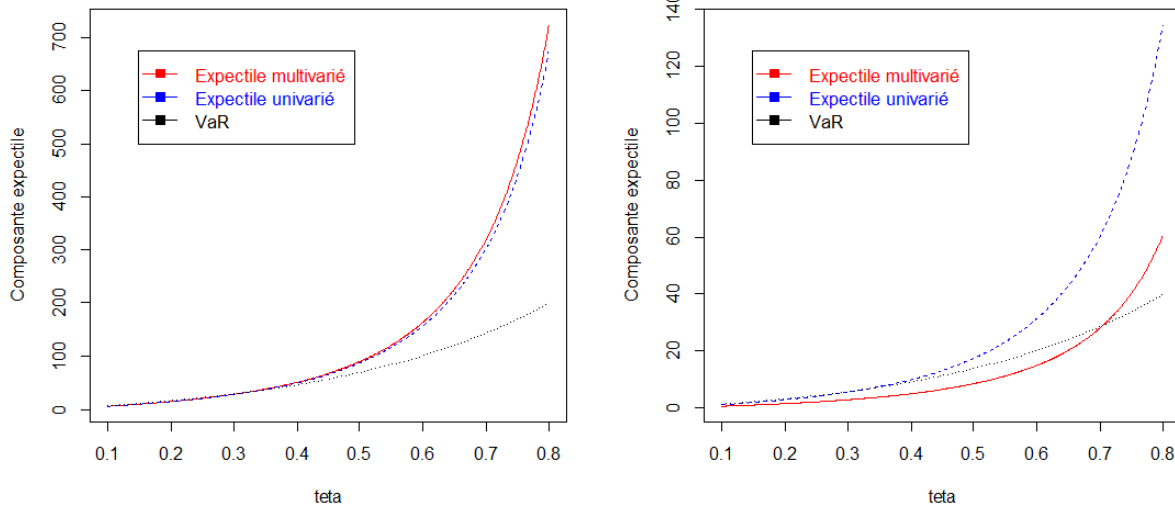


FIGURE 1.13 – L_1 -expectile. Impact de la dépendance. $\beta_1 = 0.05, \beta_2 = 0.25, \alpha = 0.95$. (à gauche X_1 , à droite X_2).

La Figure 1.13 illustre l’impact du poids de la dépendance sur la différence entre les deux approches univariée et bivariée. Cet impact est toujours plus visible au niveau de la branche la moins risquée et négligeable pour la plus risquée.

Expectile de Kendall

La Figure 1.14 illustre la différence entre le comportement du L_1 -expectile construit à l’aide d’une matrice statique, et celui de l’expectile de Kendall construit à l’aide de la matrice composée des τ de Kendall bivariés

$$\pi_{ij} = \tau_{ij} = \frac{\theta}{\theta + 2} = \frac{\frac{1}{a}}{\frac{1}{a} + 2}.$$

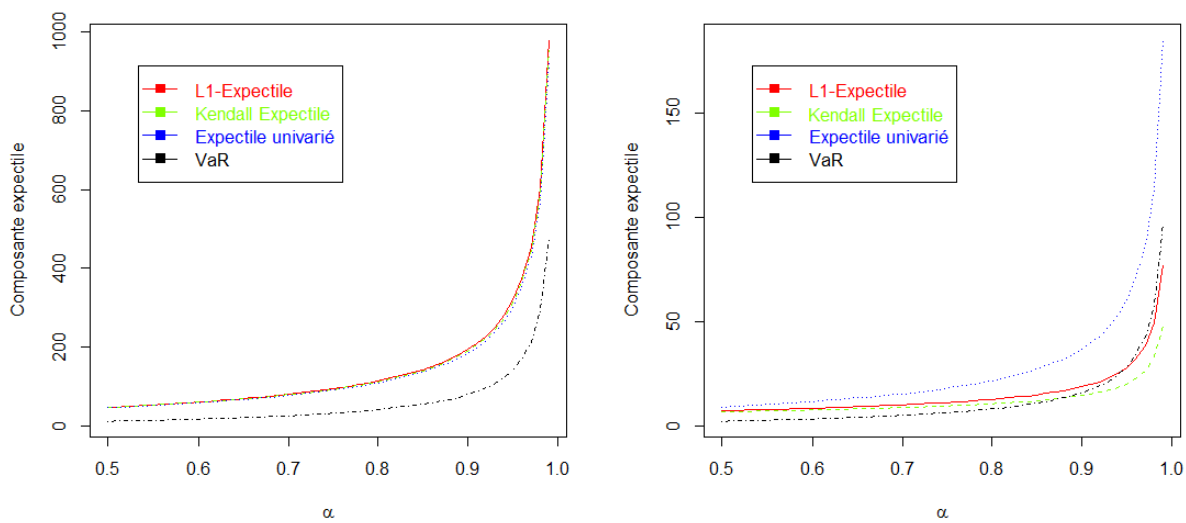


FIGURE 1.14 – Impact du choix de la construction $\beta_1 = 0.05, \beta_2 = 0.25, \theta = 1/a = 0.7$ (à gauche X_1 , à droite X_2).

La Figure 1.15 permet de comparer l’impact de la dépendance sur les deux constructions.

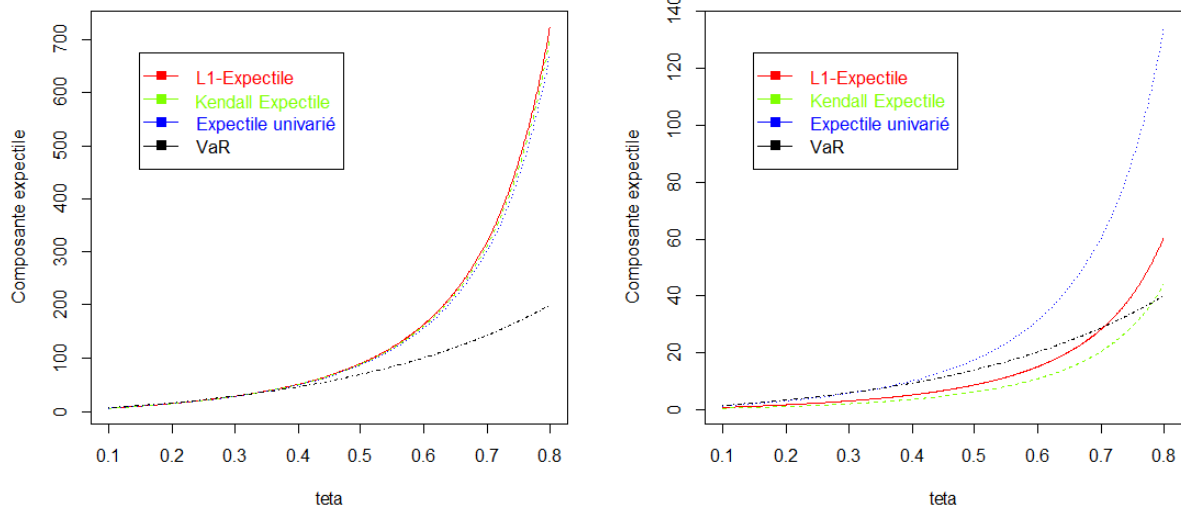


FIGURE 1.15 – Impact de la dépendance en fonction du choix de la construction de l'expectile. $\beta_1 = 0.05$, $\beta_2 = 0.25$, $\alpha = 0.95$ (à gauche X_1 , à droite X_2).

Finalement, le comportement de l'expectile multivarié dépend d'abord de la construction choisie. Mais aussi de la structure de la dépendance et la nature des marginales.

Chapitre 2

Comportement asymptotique des expectiles multivariés



Asymptotic behavior of multivariate expectiles

Introduction

Le niveau asymptotique, lorsque le niveau α de l'expectile tend vers un, représente économiquement le cas d'un risque extrême. Nous avons constaté précédemment que les algorithmes d'approximation stochastique ne sont pas efficaces au voisinage de ce niveau. Nous avons aussi démontré que l'expectile extrême tend vers le vecteur composé des points terminaux des distributions de ses composantes.

Nous consacrons ce chapitre à une étude plus approfondie du comportement asymptotique des vecteurs expectiles. Nous considérons le cas des distributions à variations régulières. Ce concept est très important dans l'étude des événements extrêmes et applicable en assurance et en gestion des risques en général. Nous abordons un ensemble de cas, en fonction de la nature de la dépendance asymptotique, pour des distributions marginales du domaine d'attractions des valeurs extrêmes de Fréchet. Plusieurs cas particuliers sont étudiés, notamment la présence d'une indépendance asymptotique et la comonotonie des risques. Une section de ce chapitre est dédiée au modèle multivarié par mélange commun. Nous nous basons dans la suite sur les résultats obtenus pour proposer des estimateurs du vecteur expectile asymptotique dans les cas de l'indépendance asymptotique et de la dépendance parfaite. Un ensemble d'illustrations théoriques et numériques accompagne les résultats présentés dans ce chapitre.

2.1 Variations régulières multivariées et dépendance extrême

Les distributions à variations régulières constituent une classe particulière utilisable en actuariat pour modéliser les charges des sinistres. Plusieurs travaux se sont intéressés à l'analyse asymptotique de différentes mesures de risque usuelles pour cette classe de distributions. L'étude de la concentration du risque au niveau extrême a produit plusieurs résultats basés sur les propriétés de la queue de la distribution de la somme des risques à variations régulières. Les domaines d'attractions dans la théorie des valeurs extrêmes présentent chacun une équivalence avec une famille des distributions à variations régulières, l'étude du comportement extrêmes des mesures de risque peut donc se réduire, en grande partie, à une étude de cette famille des distributions (Voir Embrechts et al. (1997) [7], De Haan et Ferreira (2007) [5] et Bingham et al. (1989) [3]). Pour les expectiles univariés, Bellini et Di Bernardino (2015) [2] ont présenté récemment des équivalents asymptotiques en fonction du quantile de même niveau pour des distributions des trois domaines d'attractions extrêmes. Mao et Yang (2015) [13] ont

présenté des asymptotiques du premier et second ordre pour l'expectile de la somme sous une structure de dépendance type FGM.

Dans cette partie nous revenons sur la notion des variations régulières. Nous présentons par la suite la généralisation classique des variations régulières au contexte multivarié, afin de l'utiliser pour étudier le comportement asymptotique des expectiles multivariés définis précédemment.

2.1.1 Les variations régulières

La théorie des variations régulières présente un intérêt particulier dans la modélisation des sinistres en science actuarielle. Les distributions qui possèdent une fonction de survie à variations régulières d'indice positif sont à queues lourdes (*heavy-tailed distribution*). Cette catégorie englobe donc les trois grandes familles des distributions, à queue épaisse, à longue queue et les distributions sous-exponentielles qui sont très utilisées pour la modélisation des charges de sinistres, notamment dans la théorie de la ruine.

Définition 2.1.1 (Fonctions à variations lentes). Une fonction réelle L mesurable positive est dite à variations lentes en $a \in \{0, +\infty\}$, si pour tout réel $t > 0$

$$\lim_{x \rightarrow a} \frac{L(tx)}{L(x)} = 1,$$

et on note $L \in \text{RV}_0(a)$.

Théorème 2.1.2 (Représentation de Karamata, Resnick (2013) [18]). *Pour toute fonction L à variations lentes en $+\infty$, il existe une fonction mesurable strictement positive $c(\cdot)$ telle que $\lim_{x \rightarrow +\infty} c(x) = c \in]0, +\infty[$, et une fonction mesurable $\varepsilon(\cdot)$ qui vérifie $\lim_{x \rightarrow +\infty} \varepsilon(x) = 0$, telles que*

$$L(x) = c(x) \exp\left(\int_1^x \frac{\varepsilon(t)}{t} dt\right).$$

La fonction L est dite normalisée si la fonction $c(\cdot)$ est une constante, dans ce cas

$$L'(x) = cL(x) \frac{\varepsilon(x)}{x},$$

et par suite

$$\lim_{x \rightarrow +\infty} \frac{xL'(x)}{L(x)} = 0.$$

Cette dernière limite est souvent utilisée dans les démonstrations de la normalité asymptotique des estimateurs extrêmes.

Définition 2.1.3 (Fonctions à variations régulières). Une fonction mesurable positive f est dite à variations régulières en $a \in \{0, +\infty\}$ d'indice ρ , si pour tout $t > 0$,

$$\lim_{x \rightarrow a} \frac{f(tx)}{f(x)} = t^\rho,$$

et on note dans ce cas $f \in \text{RV}_\rho(a)$.

Une fonction à variations lentes est donc à variations régulières d'indice $\rho = 0$. On note aussi qu'une fonction $f \in \text{RV}_\rho(+\infty)$ si et seulement si il existe une fonction L à variations lentes $f \in \text{RV}_0(+\infty)$ telle que

$$f(x) = x^\rho L(x).$$

La représentation de Karamata est généralisée pour les fonctions à variations régulières. Une fonction $f \in \text{RV}_\rho(+\infty)$ si et seulement si elle s'écrit sous la forme

$$f(x) = c(x) \int_1^x \frac{\rho(t)}{t} dt,$$

avec $\rho(t) = \rho$ et $c(t) = c \in]0, +\infty[$.

On dit que f est à variations rapides si l'indice $\rho \in \{-\infty, +\infty\}$, et on note $f \in \text{RV}_{+\infty}(a)$, pour $a \in \{0, +\infty\}$. Dans ce cas

$$\lim_{x \rightarrow a} \frac{f(tx)}{f(x)} = \begin{cases} \infty & \text{si } t > 1 \\ 0 & \text{si } 0 < t < 1 \end{cases}.$$

On a d'une façon équivalente $1/f \in \text{RV}_{-\infty}(a)$.

Lemme 2.1.4 (Dérivée d'une fonction RV, Bingham et al. (1989) [3]). *Si f' est la dérivée d'une fonction $f \in \text{RV}_{\rho}(a)$, avec $a \in \{0, +\infty\}$ est monotone, alors*

$$f' \in \text{RV}_{\rho-1}(a).$$

Lemme 2.1.5 (Inverse d'une fonction RV, Bingham et al. (1989) [3]). *L'inverse f^{\leftarrow} d'une fonction à variations régulières $f \in \text{RV}_{\rho}(a)$ est une fonction à variations régulières $f^{\leftarrow} \in \text{RV}_{1/\rho}(f(a))$ si $\rho > 0$ et $f^{\leftarrow} \in \text{RV}_{-1/\rho}(-f(a))$ si $\rho < 0$.*

Une démonstration de ce résultat en équivalence, dans le cas d'une fonction $f(\cdot)$ positive mesurable, définie sur \mathbb{R}^+ , telle que $\lim_{x \rightarrow +\infty} f(x) = +\infty$, est présentée dans Resnick (2007) [17]

$$f \in \text{RV}_{\rho}(+\infty) \text{ si et seulement si } f^{\leftarrow} \in \text{RV}_{\frac{1}{\rho}}(+\infty).$$

Lemme 2.1.6 (Intégrale d'une fonction RV (Théorème de Karamata), Mikosch (2003) [15]). *Soit $f(\cdot)$ une fonction positive mesurable à variations régulières en $+\infty$ d'indice ρ , localement bornée sur $[x_0, +\infty)$ avec $x_0 \geq 0$ alors*

- Si $\rho > -1$

$$\lim_{x \rightarrow +\infty} \frac{\int_{x_0}^x f(t) dt}{x f(x)} = \frac{1}{\rho + 1},$$

- Si $\rho < -1$

$$\lim_{x \rightarrow +\infty} \frac{\int_x^{+\infty} f(t) dt}{x f(x)} = -\frac{1}{\rho + 1}.$$

Lemme 2.1.7 (Bornes de Potter, Bingham et al. (1989) [3]). *Soit $f \in \text{RV}_{\rho}(a)$, avec $a \in \{0, \infty\}$ et $\rho \in \mathbb{R}$. Pour tout $0 < \epsilon < 1$ et tout x et y au voisinage de a*

$$(1 - \epsilon) \min \left(\left(\frac{x}{y} \right)^{\rho - \epsilon}, \left(\frac{x}{y} \right)^{\rho + \epsilon} \right) \leq \frac{f(x)}{f(y)} \leq (1 + \epsilon) \max \left(\left(\frac{x}{y} \right)^{\rho - \epsilon}, \left(\frac{x}{y} \right)^{\rho + \epsilon} \right).$$

Les bornes de Potter sont très utiles pour l'utilisation du théorème de la convergence dominée dans le cas d'études des intégrations des fonctions à variations régulières.

Plusieurs autres résultats sur les fonctions à variations régulières sont présentés dans le livre de Bingham et al. (1989) [3].

Définition 2.1.8 (Fonctions à variations régulières prolongées). Une fonction mesurable $f : \mathbb{R}^+ \rightarrow \mathbb{R}$ est dite à variations régulières prolongées d'indice réel γ au voisinage de $+\infty$, s'il existe une fonction auxiliaire $a : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ telle que

$$\forall t > 0 \quad \lim_{x \rightarrow +\infty} \frac{f(tx) - f(x)}{a(x)} = \frac{t^{\gamma} - 1}{\gamma},$$

on note $f \in \text{ERV}_{\gamma}(+\infty)$. Dans le cas $\gamma = 0$ cette limite est interprétée comme la fonction log.

La fonction queue (*tail quantile function*) est définie par $U : \left(\frac{1}{F}\right)$, qui peut s'écrire pour tout $t > 0$ directement sous la forme

$$U(t) := \inf \left\{ x : F_X(x) \geq 1 - \frac{1}{t} \right\}.$$

Cette fonction est très utilisée dans la théorie des variations régulières, elle relie la loi des valeurs extrêmes aux fonctions à variations régulières prolongées. Cette relation présentée dans le théorème 1.1.6 du livre de De Haan et Ferreira (2007) [5] est donnée par l'équivalence suivante

$$F \in \text{MDA}(H_\gamma) \Leftrightarrow U \in \text{ERV}_\gamma.$$

Les fonctions ERV sont aussi liées aux fonctions RV via ce théorème. Cette relation est présentée dans le théorème B.2.2 de de Haan et Ferreira (2007) [5]. Pour toute fonction $f \in \text{ERV}_\gamma$, si $\gamma > 0$ alors

$$f \in \text{RV}_\gamma \text{ et } \lim_{x \rightarrow +\infty} \frac{f(x)}{a(x)} = \frac{1}{\gamma},$$

et si $\gamma < 0$ ce qui correspond au domaine d'attraction de Weibull pour les valeurs extrêmes, alors

$$\lim_{x \rightarrow +\infty} f(x) = \ell \text{ existe, et } \lim_{x \rightarrow +\infty} \frac{\ell - f(x)}{a(x)} = \frac{1}{\gamma}, \text{ et } \ell - f \in \text{RV}_\gamma.$$

2.1.2 Les variations régulières multivariées

Les distributions à queues lourdes représentent un intérêt important en applications actuarielles et financières, et notamment dans la gestion des risques. Il est connu par exemple que les rendements financiers sont souvent à queues lourdes. En assurance, le besoin d'analyser les risques extrêmes fait de l'étude de cette famille de distributions une nécessité incontestable (voir Resnick (2007) [17] et De Haan et Ferreira (2007) [5]). L'extension multivariée de la notion des variations régulières a été introduite dans la littérature par De Haan et Resnick (1977) [6], depuis, d'autres définitions équivalentes ont été présentées.

Nous utiliserons la notion de la vague convergence telle qu'expliquée en détail dans Kallenberg (1983) [10] et dans le second chapitre de la thèse de Basrak (2000) [1]. On note $\mu_n \xrightarrow{v} \mu$ la convergence vague des mesures de Radon.

On peut se limiter à des distributions à supports positifs car il est possible d'en déduire les résultats pour le cas général à l'aide de la décomposition $X = X_+ - X_-$.

Définition 2.1.9 (Les variations régulières multivariées). La distribution d'un vecteur aléatoire \mathbf{X} dans $[0, \infty]^d$ est dite à variations régulières s'il existe une mesure de Radon non nulle $\mu_{\mathbf{X}}$ sur la tribu borélienne \mathcal{B}_d σ -algèbre sur $[0, \infty]^d \setminus \mathbf{0}$, et une fonction de normalisation $b : \mathbb{R} \rightarrow \mathbb{R}$ qui vérifie $\lim_{x \rightarrow +\infty} b(x) = +\infty$ telles que

$$u\mathbb{P} \left(\frac{\mathbf{X}}{b(u)} \in \cdot \right) \xrightarrow{v} \mu_{\mathbf{X}}(\cdot) \text{ quand } u \rightarrow +\infty, \quad (2.1)$$

sur l'espace des mesures de Radon positives sur $[0, \infty]^d$.

Il existe différentes caractérisations des distributions multivariées à queue à variations régulières. Nous présentons ici quelques-unes qui seront utilisées en fonction du besoin.

Définition 2.1.10 (MRV Définitions équivalentes). Soit \mathbf{X} un vecteur aléatoire dans \mathbb{R}^d , les définitions suivantes sont équivalentes :

- Le vecteur aléatoire \mathbf{X} est à queue à variations régulières d'exposant θ .
- Il existe une mesure finie μ sur la sphère unité \mathbb{S}^{d-1} , et une fonction de normalisation $b : (0, \infty) \rightarrow (0, \infty)$ telles que

$$\lim_{t \rightarrow +\infty} \mathbb{P} \left(\|\mathbf{X}\| > xb(t), \frac{\mathbf{X}}{\|\mathbf{X}\|} \in \cdot \right) = x^{-\theta} \mu(\cdot), \quad (2.2)$$

pour tout $x > 0$. La mesure μ dépend de la norme utilisée et de \mathbf{X} , on l'appelle la *mesure spectrale* de \mathbf{X} .

– Il existe une mesure finie μ sur la sphère unité \mathbb{S}^{d-1} , une fonction à variations lentes L , et un réel strictement positif $\theta > 0$ tels que

$$\lim_{x \rightarrow +\infty} \frac{x^\theta}{L(x)} \mathbb{P} \left(\|\mathbf{X}\| > x, \frac{\mathbf{X}}{\|\mathbf{X}\|} \in B \right) = \mu(B), \quad (2.3)$$

pour tout $B \in \mathcal{B}(\mathbb{S}^{d-1})$ avec $\mu(\partial B) = 0$.

On désigne à partir de maintenant par MRV (*Multivariate Regular Variation*) les distributions à variations régulières multivariées, et par $\text{MRV}(\theta, \mu)$ les vecteurs aléatoires à queue à variation régulières d'exposant θ et de mesure spectrale μ . Il faut noter que la mesure spectrale dépend de la norme choisie. Pour mieux comprendre l'interprétation du concept des MRV, on peut examiner sans perte de généralité le cas d'une mesure μ normalisée, c.-à-d. qui vérifie $\mu(\mathbb{S}^{d-1}) = 1$. En appliquant la définition MRV (2.3) à $B = \mathbb{S}^{d-1}$, On déduit que

$$\mathbb{P}(\|\mathbf{X}\| > x) \stackrel{x \uparrow +\infty}{\sim} x^{-\theta} L(x),$$

la variable aléatoire positive $\|\mathbf{X}\|$ est donc à queue à variations régulières d'indice $-\theta$. D'un autre côté,

$$\begin{aligned} \lim_{x \rightarrow +\infty} \mathbb{P} \left(\frac{\mathbf{X}}{\|\mathbf{X}\|} \in B \mid \|\mathbf{X}\| > x \right) &= \lim_{x \rightarrow +\infty} \mathbb{P} \left(\|\mathbf{X}\| > x, \frac{\mathbf{X}}{\|\mathbf{X}\|} \in B \right) \mathbb{P}(\|\mathbf{X}\| > x)^{-1} \\ &= \lim_{x \rightarrow +\infty} \frac{x^\theta}{L(x)} \mathbb{P} \left(\|\mathbf{X}\| > x, \frac{\mathbf{X}}{\|\mathbf{X}\|} \in B \right) x^{-\theta} L(x) \mathbb{P}(\|\mathbf{X}\| > x)^{-1} \\ &= \mu(B), \end{aligned}$$

pour tout $B \in \mathcal{B}(\mathbb{S}^{d-1})$ avec $\mu(\partial B) = 0$. Cela signifie que la loi conditionnelle de $\frac{\mathbf{X}}{\|\mathbf{X}\|}$ sachant $\{\|\mathbf{X}\| > x\}$ converge faiblement vers μ .

Il existe plusieurs caractérisations du concept des variations régulières multivariées dans la littérature. Les différentes définitions des MRV sont présentées dans le livre de Mikosch (2003) [15] et dans la thèse de Basrak (2000) [1].

2.1.3 Caractérisation à l'aide des fonctions de dépendance extrême.

Nous considérons le cas d'un vecteur aléatoire $\mathbf{X} = (X_1, \dots, X_d)$ à queues équivalentes et à variations régulières. Ce modèle se traduit par deux hypothèses :

H1 X_1 est à distributions à variations régulières

$$\bar{F}_{X_1} \in \text{RV}_{-\theta},$$

avec $\theta > 0$.

H2 Les queues des $X_i, i = 1, \dots, d$ sont équivalentes. C.-à-d., pour tout $i \in \{2, \dots, d\}$, il existe une constante strictement positive c_i telle que

$$\lim_{x \rightarrow +\infty} \frac{\bar{F}_{X_i}(x)}{\bar{F}_{X_1}(x)} = c_i.$$

Les deux hypothèses impliquent que les queues sont toutes à variations régulières de même indice θ . Pour la suite nous utilisons la définition de la fonction de la dépendance des queues droites. Nous considérons la définition introduite dans Klüppelberg et al. (2008) [11] de la fonction de la dépendance extrême supérieure.

Définition 2.1.11 (La fonction de dépendance extrême supérieure). Soit \mathbf{X} un vecteur aléatoire dans \mathbb{R}^d de distributions marginales continues. La fonction de dépendance extrême supérieure est définie comme suit

$$\lambda_U^{\mathbf{X}}(x_1, \dots, x_d) = \lim_{t \rightarrow 0} t^{-1} \mathbb{P}(1 - F_{X_1}(X_1) \leq tx_1, \dots, 1 - F_{X_d}(X_d) \leq tx_d), \quad (2.4)$$

lorsque la limite existe.

Cette définition peut s'écrire à l'aide de la copule de dépendance comme suit

$$\lambda_U^k(u_1, \dots, u_k) = \lim_{t \rightarrow 0^+} \frac{\bar{C}_k(tu_1, \dots, tu_k)}{t}, \quad (2.5)$$

si cette limite existe. \bar{C} désigne la copule de survie associé à la copule C (voir Nelsen (2007) [16]). La notation λ^k désigne la fonction de dépendance extrême associé à un sous vecteur de \mathbf{X} de taille k . On peut aussi définir d'une façon équivalente la fonction de dépendance extrême à gauche sous la forme

$$\lambda_L^k(u_1, \dots, u_k) = \lim_{t \rightarrow 0^+} \frac{C_k(tu_1, \dots, tu_k)}{t},$$

lorsque la limite existe.

Nous nous limitons dans le cas de cette étude à la dépendance extrême à droite, car nous sommes dans un contexte de gestion des risques.

L'objectif de cette section est de visualiser l'équivalence entre l'existence des fonctions de la dépendance extrême et le caractère MRV des vecteurs aléatoires, sous les hypothèses H1 et H2. Les Théorèmes 2.1.12 et 2.1.13 énoncent cette équivalence et relient les fonctions de la dépendance extrême à la mesure spectrale MRV. Les variations régulières multivariées sont utilisées dans ces théorèmes dans la forme de la Définition 2.1.9.

Théorème 2.1.12 (Théorème 2.3, Li et Sun (2009) [12]). *Soit $\mathbf{X} = (X_1, \dots, X_d)$ un vecteur aléatoire dans \mathbb{R}^d , dont les lois marginales $F_{X_i}, i = 1, \dots, d$ sont continues et satisfont les hypothèses H1 et H2. Si \mathbf{X} a une distribution multivariée à variations régulières au sens de la Définition 2.1.9, alors la fonction de la dépendance de la queue droite existe, et elle est donnée par*

$$\lambda_U^k(u_1, \dots, u_k) = \lim_{x \rightarrow +\infty} x \mathbb{P} \left(X_1 > b(x) \left(\frac{u_1}{c_1} \right)^{-1/\theta}, \dots, X_k > b(x) \left(\frac{u_d}{c_d} \right)^{-1/\theta} \right),$$

pour tout $k \in \{1, \dots, d\}$. L'existence de cette limite découle naturellement de la définition des distributions MRV.

La réciproque de ce résultat est présentée dans le théorème suivant. L'utilité d'avoir ces deux équivalences est de pouvoir caractériser les distributions MRV qui vérifient les hypothèses H1 et H2 à l'aide de leurs copules de survie.

Théorème 2.1.13 (Théorème 3.2, Weng et Zhang (2012) [20]). *Soit $\mathbf{X} = (X_1, \dots, X_d)$ un vecteur aléatoire dans \mathbb{R}^d dont les lois marginales $F_{X_i}, i = 1, \dots, d$ sont continues et satisfont les hypothèses H1 et H2. Si la fonction de dépendance de la queue droite λ_U^k existe pour tout $k \in \{1, \dots, d\}$, alors \mathbf{X} est à distribution MRV dont la fonction de normalisation est donnée par $b(u) = \left(\frac{1}{F_{X_1}} \right)^{\leftarrow} (u)$ et la mesure spectrale est donnée par*

$$\mu([\mathbf{0}, \mathbf{x}]^c) = \sum_{i=1}^d c_i x_i^{-\theta} - \sum_{1 \leq i < j \leq d} \lambda_U^2(c_i x_i^{-\theta}, c_j x_j^{-\theta}) + \dots + (-1)^{d+1} \lambda_U^d(c_1 x_1^{-\theta}, \dots, c_d x_d^{-\theta}).$$

La démonstration de ce théorème est une application directe du principe d'inclusion-exclusion.

Les deux théorèmes permettent de relier l'existence des fonctions λ_U sous H1 et H2 à la caractérisation MRV.

Dans le cas des expectiles multivariés, nous avons besoin uniquement des dépendances bivariées. Pour tout $(i, j) \in \{1, \dots, d\}^2$, Le couple (X_i, X_j) est trivialement à queues à variations régulières si le vecteur \mathbf{X} l'est. Pour le cas bivarié, on note λ_U^{ik} la fonction $\lambda_U^{(X_i, X_k)}$, pour tout $(i, k) \in \{1, \dots, d\}^2$. Dans un cadre MRV, et sous les hypothèse H1 et H2, l'existence des fonctions λ^{ik} est donc assurée. Nous supposons par la suite que ces fonctions sont continues.

Finalement, le modèle à queues équivalentes est d'une utilité pratique importante, il est souvent utilisé pour la modélisation des charges de sinistres en assurance, dans l'étude des événements extrêmes en présence de dépendance et dans la théorie de la ruine. Il englobe notamment les modèles à distributions identiques, ou à distributions pseudo-identiques obtenues par exemple par un changement d'échelle. Dans la suite de ce chapitre, nous allons examiner le comportement asymptotique des expectiles multivariés dans le cas d'un modèle MRV à queues équivalentes. Nous nous focalisons sur le cas du domaine d'attraction des valeurs extrêmes de Fréchet pour les distributions marginales.

Nous considérons dans ce chapitre le cas des expectiles multivariés construits à l'aide d'une matrice symétrique $\Sigma = (\pi_{ij})_{1 \leq i \leq j \leq d}$ semi-définie positive à coefficients positifs. Pour l'ensemble des résultats présentés, nous supposons sans perte de généralité que $\pi_{ij} = 1$ pour tout $i, j \in \{1, \dots, d\}$, ce cas caractérise les L_1 -expectiles, et nous donnons ensuite les formules générales pour des coefficients positifs quelconques.

2.2 Modèle Fréchet à queues équivalentes

Dans cette partie, nous considérons les deux hypothèses H1 et H2 sur les distributions marginales des risques X_i , avec $\theta > 1$. L'hypothèse H1 traduit le cas de l'appartenance de la distribution de X_1 au domaine d'attraction extrême de Fréchet $\text{MDA}(\Phi_\theta)$. Pour les distributions de ce domaine, $x_F = \sup\{x : F(x) < 1\} = +\infty$, donc quand $\alpha \rightarrow 1$ on a $\mathbf{e}_\alpha^i(\mathbf{X}) \rightarrow +\infty \forall i$.

Nous déduisons également par application du théorème de Karamata (2.1.6) que

$$0 < \lim_{x \rightarrow +\infty} \frac{\mathbb{E}[(X_i - x)_+]}{x \bar{F}_{X_i}(x)} = \frac{1}{\theta - 1} < +\infty, \quad (2.6)$$

pour tout $i \in \{1, \dots, d\}$.

Proposition 2.2.1. *Sous les hypothèses H1 et H2, et pour des variables aléatoires $X_i, i = 1, \dots, d$, les composantes de l'expectile multivarié asymptotique vérifient*

$$0 < \lim_{\alpha \rightarrow 1} \frac{\mathbf{e}_\alpha^i(\mathbf{X})}{\mathbf{e}_\alpha^1(\mathbf{X})} \leq \overline{\lim}_{\alpha \rightarrow 1} \frac{\mathbf{e}_\alpha^i(\mathbf{X})}{\mathbf{e}_\alpha^1(\mathbf{X})} < +\infty, \forall i \in \{2, \dots, d\}.$$

Les distributions à queues équivalentes ont donc des composantes d'expectile multivarié comparables. Pour un vecteur aléatoire $\mathbf{X} = (X_1, \dots, X_d)^T$ vérifiant les hypothèses H1 et H2, on note $x_i = \mathbf{e}_\alpha^i(\mathbf{X})$ pour tout $i \in \{1, \dots, d\}$. On utilise dans les preuves de cette partie les fonctions l_{X_i, X_j}^α précédemment définies pour tout $(i, j) \in \{1, \dots, d\}^2$ par

$$l_{X_i, X_j}^\alpha(x_i, x_j) = \alpha \mathbb{E}[(X_i - x_i)_+ \mathbb{1}_{\{X_j > x_j\}}] - (1 - \alpha) \mathbb{E}[(X_i - x_i)_- \mathbb{1}_{\{X_j < x_j\}}],$$

pour tout $(x_i, x_j) \in \mathbb{R}^2$, avec $l_{X_i}^\alpha(x_i) = l_{X_i, X_i}^\alpha(x_i, x_i)$.

Le système d'optimalité (1.9) s'écrit à l'aide de ces fonctions sous la forme suivante,

$$l_{X_k}^\alpha(x_k) = - \sum_{i=1, i \neq k}^d \frac{\pi_{ki}}{\pi_{kk}} l_{X_i, X_k}^\alpha(x_i, x_k) \quad \forall k \in \{1, \dots, d\}. \quad (2.7)$$

Pour faciliter la rédaction des preuves, nous définissons les trois ensembles suivants :

$$J_0^i = \{j \in \{1, \dots, d\} \setminus \{i\} \mid \lim_{\alpha \rightarrow 1} \frac{x_j}{x_i} = 0\},$$

$$J_C^i = \{j \in \{1, \dots, d\} \setminus \{i\} \mid \lim_{\alpha \rightarrow 1} \frac{x_j}{x_i} \text{ avec } 0 < \underline{\lim}_{\alpha \rightarrow 1} \frac{x_j}{x_i} < \overline{\lim}_{\alpha \rightarrow 1} \frac{x_j}{x_i} < +\infty\},$$

$$\text{et } J_\infty^i = \{j \in \{1, \dots, d\} \setminus \{i\} \mid \overline{\lim}_{\alpha \rightarrow 1} \frac{x_j}{x_i} = +\infty\}.$$

Nous considérons sans perte de généralité que $\pi_{ij} = 1$, pour tout $(i, j) \in \{1, \dots, d\}^2$.

Pour la démonstration de la Proposition 2.2.1, nous aurons besoin du Lemme 2.2.2 et des deux Propositions 2.2.3 et 2.2.4 ci-après.

Lemme 2.2.2. *Sous les hypothèses H1 et H2, si $t = o(s)$ alors*

$$\lim_{t \uparrow +\infty} \frac{s \bar{F}_{X_i}(s)}{t \bar{F}_{X_j}(t)} = 0,$$

pour tout $(i, j) \in \{1, \dots, d\}^2$.

Preuve du Lemme 2.2.2 Sous les hypothèses H1 et H2, pour tout $i \in \{1, \dots, d\}$

$$\bar{F}_{X_i} \in \text{RV}_{-\theta}(+\infty),$$

il existe pour tout i une fonction mesurable positive $L_i \in \text{RV}_0(+\infty)$ telle que

$$\bar{F}_{X_i}(x) = x^{-\theta} L_i(x), \forall x > 0,$$

alors pour tout $(i, j) \in \{1, \dots, d\}^2$ et tout $t, s > 0$

$$\frac{s \bar{F}_{X_i}(s)}{t \bar{F}_{X_j}(t)} = \left(\frac{s}{t}\right)^{-\theta+1} \frac{L_i(s)}{L_j(t)} = \left(\frac{s}{t}\right)^{-\theta+1} \frac{L_i(s)}{L_i(t)} \frac{L_i(t)}{L_j(t)}, \quad (2.8)$$

et de l'hypothèse H2

$$\lim_{x \uparrow +\infty} \frac{L_i(x)}{L_j(x)} = \frac{c_i}{c_j}. \quad (2.9)$$

En utilisant la représentation de Karamata pour les fonctions à variations lentes (2.1.2), pour $i \in \{1, \dots, d\}$, il existe une constante réelle $c > 0$ et une fonction mesurable strictement positive $c(\cdot)$ telle que $\lim_{x \uparrow +\infty} c(x) = c > 0$, et une fonction mesurable $\varepsilon(\cdot)$ qui vérifie $\lim_{x \uparrow +\infty} \varepsilon(x) = 0$ telles que

$$L_i(x) = c(x) \exp\left(\int_1^x \frac{\varepsilon(y)}{y} dy\right),$$

alors pour tout $t, s > 0$

$$\frac{L_i(s)}{L_i(t)} = \frac{c(s)}{c(t)} \exp\left(\int_t^s \frac{\varepsilon(y)}{y} dy\right).$$

Nous considérons le cas $t = o(s)$, alors $s > t$ pour t assez grand.

La fonction $\varepsilon(\cdot)$ vérifie $\lim_{x \uparrow +\infty} \varepsilon(x) = 0$, alors pour tout $\epsilon > 0$ petit, $\exists t_0$ tel que $\forall t > t_0$

$$\int_t^s \frac{\varepsilon(y)}{y} dy \leq \epsilon \int_t^s \frac{dy}{y},$$

on a donc

$$\left(\frac{s}{t}\right)^{-\theta+1} \frac{L_i(s)}{L_i(t)} \leq \left(\frac{s}{t}\right)^{-\theta+1+\epsilon} \frac{c(s)}{c(t)}.$$

Ainsi, en prenant $0 < \epsilon < \theta - 1$, on obtient

$$\lim_{t \uparrow +\infty} \frac{s \bar{F}_{X_i}(s)}{t \bar{F}_{X_j}(t)} = 0, \quad \forall (i, j) \in \{1, \dots, d\}^2.$$

□

Proposition 2.2.3. *Pour tout vecteur aléatoire $\mathbf{X} = (X_1, \dots, X_d)$, et sous les hypothèses H1 et H2, les composantes de l'expectile multivarié asymptotique vérifient*

$$\overline{\lim}_{\alpha \rightarrow 1} \frac{1 - \alpha}{\bar{F}_{X_i}(\mathbf{e}_\alpha^i(\mathbf{X}))} < +\infty, \quad \forall i \in \{2, \dots, d\}.$$

Preuve de la Proposition 2.2.3 On cherche à démontrer que

$$\overline{\lim}_{\alpha \rightarrow 1} \frac{1 - \alpha}{\bar{F}_{X_1}(\mathbf{e}_\alpha^1(\mathbf{X}))} < +\infty,$$

cela est suffisant pour déduire le résultat de la proposition puisque les variables aléatoires sont échangeables.

On suppose que $\overline{\lim}_{\alpha \rightarrow 1} \frac{1 - \alpha}{\bar{F}_{X_1}(\mathbf{e}_\alpha^1(\mathbf{X}))} = +\infty$. Quitte à prendre une sous suite, on considère que

$$\lim_{\alpha \rightarrow 1} \frac{1 - \alpha}{\bar{F}_{X_1}(\mathbf{e}_\alpha^1(\mathbf{X}))} = +\infty.$$

On a

$$\begin{aligned} \frac{l_{X_1}^\alpha(x_1)}{(1 - \alpha)x_1} &= \left(\frac{\alpha \mathbb{E}[(X_1 - x_1)_+] - (1 - \alpha) \mathbb{E}[(X_1 - x_1)_-]}{(1 - \alpha)x_1} \right) \\ &= \left((2\alpha - 1) \frac{\mathbb{E}[(X_1 - x_1)_+]}{(1 - \alpha)x_1} - \frac{(1 - \alpha)(x_1 - \mathbb{E}[X_1])}{(1 - \alpha)x_1} \right) \\ &= \left((2\alpha - 1) \frac{\mathbb{E}[(X_1 - x_1)_+]}{x_1 \bar{F}_{X_1}(x_1)} \frac{\bar{F}_{X_1}(x_1)}{1 - \alpha} - 1 + \frac{\mathbb{E}[X_1]}{x_1} \right) \xrightarrow{\alpha \uparrow 1} -1. \end{aligned}$$

De plus, pour tout $i \in \{2, \dots, d\}$

$$\begin{aligned} \lim_{\alpha \rightarrow 1} \frac{l_{X_i, X_1}^\alpha(x_i, x_1)}{(1 - \alpha)x_1} &= \lim_{\alpha \rightarrow 1} \left(\frac{\alpha \mathbb{E}[(X_i - x_i)_+ \mathbb{1}_{\{X_1 > x_1\}}] - (1 - \alpha) \mathbb{E}[(X_i - x_i)_- \mathbb{1}_{\{X_1 < x_1\}}]}{(1 - \alpha)x_1} \right) \\ &= \lim_{\alpha \rightarrow 1} \left(\frac{\mathbb{E}[(X_i - x_i)_+ \mathbb{1}_{\{X_1 > x_1\}}]}{(1 - \alpha)x_1} - \frac{x_i}{x_1} \right), \end{aligned}$$

$$\text{car } \frac{\mathbb{E}[(X_i - x_i)_- \mathbb{1}_{\{X_1 < x_1\}}]}{x_1} = \frac{\mathbb{E}[(X_i - x_i)_+ \mathbb{1}_{\{X_1 < x_1\}}] + x_i \mathbb{P}(X_1 < x_1) - \mathbb{E}[X_i \mathbb{1}_{\{X_1 < x_1\}}]}{x_1} \underset{\alpha \uparrow 1}{\sim} \frac{x_i}{x_1}.$$

D'un autre côté,

$$\frac{\mathbb{E}[(X_i - x_i)_+ \mathbb{1}_{\{X_1 > x_1\}}]}{(1 - \alpha)x_1} \leq \frac{\mathbb{E}[(X_i - x_i)_+]}{(1 - \alpha)x_1} = \frac{\bar{F}_{X_1}(x_1)}{1 - \alpha} \frac{\mathbb{E}[(X_i - x_i)_+]}{x_i \bar{F}_{X_i}(x_i)} \frac{x_i \bar{F}_{X_i}(x_i)}{x_1 \bar{F}_{X_1}(x_1)}, \forall i \in \{2, \dots, d\}.$$

On en déduit que

$$\lim_{\alpha \rightarrow 1} \frac{\mathbb{E}[(X_k - x_k)_+ \mathbb{1}_{\{X_1 > x_1\}}]}{(1 - \alpha)x_1} = 0, \forall k \in J_C^1, \quad (2.10)$$

car

$$\frac{\bar{F}_{X_1}(x_1)}{1 - \alpha} \frac{\mathbb{E}[(X_k - x_k)_+]}{x_k \bar{F}_{X_k}(x_k)} \frac{x_k \bar{F}_{X_k}(x_k)}{x_1 \bar{F}_{X_1}(x_1)} \underset{\alpha \rightarrow 1}{\sim} \frac{\bar{F}_{X_1}(x_1)}{1 - \alpha} \frac{c_k}{\theta - 1} \left(\frac{x_k}{x_1} \right)^{-\theta + 1}.$$

Par application du Lemme 2.2.2, nous déduisons aussi

$$\lim_{\alpha \rightarrow 1} \frac{\mathbb{E}[(X_k - x_k)_+ \mathbb{1}_{\{X_1 > x_1\}}]}{(1 - \alpha)x_1} = 0, \forall k \in J_\infty^1, \quad (2.11)$$

car

$$\lim_{\alpha \rightarrow 1} \frac{x_k \bar{F}_{X_k}(x_k)}{x_1 \bar{F}_{X_1}(x_1)} = 0.$$

Pour tout $k \in J_0^1$, nous pouvons écrire

$$\frac{\mathbb{E}[(X_k - x_k)_+ \mathbb{1}_{\{X_1 > x_1\}}]}{(1 - \alpha)x_1} = \frac{\int_{x_k}^{x_1} \mathbb{P}(X_k > t, X_1 > x_1) dt}{(1 - \alpha)x_1} + \frac{\int_{x_1}^{+\infty} \mathbb{P}(X_k > t, X_1 > x_1) dt}{(1 - \alpha)x_1}. \quad (2.12)$$

On a

$$\frac{\int_{x_k}^{x_1} \mathbb{P}(X_k > t, X_1 > x_1) dt}{(1 - \alpha)x_1} \leq \frac{\int_{x_k}^{x_1} \mathbb{P}(X_1 > x_1) dt}{(1 - \alpha)x_1} = \frac{\bar{F}_{X_1}(x_1)}{1 - \alpha} \left(1 - \frac{x_k}{x_1} \right).$$

Ainsi,

$$\lim_{\alpha \rightarrow 1} \frac{\int_{x_k}^{x_1} \mathbb{P}(X_k > t, X_1 > x_1) dt}{(1 - \alpha)x_1} = 0. \quad (2.13)$$

Pour le deuxième terme,

$$\frac{\int_{x_1}^{+\infty} \mathbb{P}(X_k > t, X_1 > x_1) dt}{(1 - \alpha)x_1} \leq \frac{\int_{x_1}^{+\infty} \mathbb{P}(X_k > t) dt}{(1 - \alpha)x_1},$$

et par application du Théorème de Karamata (2.1.6)

$$\frac{\int_{x_1}^{+\infty} \mathbb{P}(X_k > t) dt}{(1 - \alpha)x_1} \underset{\alpha \uparrow 1}{\sim} \frac{1}{\theta - 1} \frac{\bar{F}_{X_k}(x_1)}{1 - \alpha},$$

donc

$$\lim_{\alpha \rightarrow 1} \frac{\int_{x_1}^{+\infty} \mathbb{P}(X_k > t, X_1 > x_1) dt}{(1 - \alpha)x_1} = 0. \quad (2.14)$$

Par passage à la limite dans (2.12), et utilisation de (2.13) et (2.14) nous déduisons que

$$\lim_{\alpha \rightarrow 1} \frac{\mathbb{E}[(X_k - x_k)_+ \mathbb{1}_{\{X_1 > x_1\}}]}{(1 - \alpha)x_1} = 0, \forall k \in J_0^1. \quad (2.15)$$

En combinant (2.10), (2.11) et (2.15), nous avons alors montré que

$$\lim_{\alpha \rightarrow 1} \frac{\mathbb{E}[(X_k - x_k)_+ \mathbb{1}_{\{X_1 > x_1\}}]}{(1 - \alpha)x_1} = 0, \forall k \in \{2, \dots, d\},$$

donc la première équation du système d'optimalité (2.7) implique que

$$-\lim_{\alpha \rightarrow 1} \left(\sum_{k \in J_0^1} \frac{l_{X_k, X_1}^\alpha(x_k, x_1)}{(1 - \alpha)x_1} + \sum_{k \in J_C^1} \frac{l_{X_k, X_1}^\alpha(x_k, x_1)}{(1 - \alpha)x_1} + \sum_{k \in J_\infty^1} \frac{l_{X_k, X_1}^\alpha(x_k, x_1)}{(1 - \alpha)x_1} \right) = \lim_{\alpha \rightarrow 1} \sum_{k=2}^d \frac{x_k}{x_1} = -1,$$

cela est absurde, et par suite

$$\overline{\lim}_{\alpha \rightarrow 1} \frac{1 - \alpha}{\bar{F}_{X_1}(x_1)} < +\infty.$$

□

Proposition 2.2.4. *Pour tout vecteur aléatoire $\mathbf{X} = (X_1, \dots, X_d)$, et sous les hypothèses H1 et H2, les composantes de l'expectile multivarié asymptotique vérifient*

$$0 < \lim_{\alpha \rightarrow 1} \frac{1 - \alpha}{\bar{F}_{X_i}(\mathbf{e}_\alpha^i(\mathbf{X}))}, \forall i \in \{2, \dots, d\}.$$

Preuve de la Proposition 2.2.4 Il suffit de démontrer que

$$0 < \lim_{\alpha \rightarrow 1} \frac{1 - \alpha}{\bar{F}_{X_i}(\mathbf{e}_\alpha^i(\mathbf{X}))}.$$

Pour cela, on suppose que $\lim_{\alpha \rightarrow 1} \frac{1 - \alpha}{\bar{F}_{X_i}(\mathbf{e}_\alpha^i(\mathbf{X}))} = 0$. Quitte à prendre une sous suite, on considère que

$\lim_{\alpha \rightarrow 1} \frac{1 - \alpha}{\bar{F}_{X_1}(\mathbf{e}_\alpha^1(\mathbf{X}))} = 0$. Dans ce cas,

$$\frac{l_{X_1}^\alpha(x_1)}{x_1 \bar{F}_{X_1}(x_1)} = \left((2\alpha - 1) \frac{\mathbb{E}[(X_1 - x_1)_+]}{x_1 \bar{F}_{X_1}(x_1)} - \frac{1 - \alpha}{\bar{F}_1(x_1)} \left(1 - \frac{\mathbb{E}[X_1]}{x_1} \right) \right) \underset{\alpha \uparrow 1}{\rightarrow} \frac{1}{\theta - 1} > 0.$$

D'un autre côté, pour tout $k \in J_\infty^1$, on a forcément

$$\lim_{\alpha \rightarrow 1} \frac{1 - \alpha}{\bar{F}_1(x_1)} \frac{x_k}{x_1} = 0,$$

car si ce n'est pas le cas, alors par application du Lemme 2.2.2

$$\lim_{\alpha \rightarrow 1} \frac{1 - \alpha}{\bar{F}_k(x_k)} = \lim_{\alpha \rightarrow 1} \frac{1 - \alpha}{\bar{F}_1(x_1)} \frac{x_k}{x_1} \left(\frac{x_1 \bar{F}_k(x_1)}{x_k \bar{F}_k(x_k)} \right) = +\infty,$$

et cela est absurde car contradictoire avec le résultat de la Proposition 2.2.3 appliqué à k

$$\lim_{\alpha \rightarrow 1} \frac{1 - \alpha}{\bar{F}_{X_k}(x_k)} < +\infty.$$

Nous démontrons aussi à l'aide de (2.6) et du Lemme 2.2.2 que

$$\lim_{\alpha \rightarrow 1} \frac{\mathbb{E}[(X_k - x_k)_+ \mathbb{1}_{\{X_1 > x_1\}}]}{x_1 \bar{F}_{X_1}(x_1)} = 0, \quad \forall k \in J_\infty^1,$$

puisque

$$\frac{\mathbb{E}[(X_i - x_i)_+ \mathbb{1}_{\{X_1 > x_1\}}]}{x_1 \bar{F}_{X_1}(x_1)} \leq \frac{\mathbb{E}[(X_i - x_i)_+]}{x_1 \bar{F}_{X_1}(x_1)} = \frac{\mathbb{E}[(X_i - x_i)_+]}{x_i \bar{F}_{X_i}(x_i)} \frac{x_i \bar{F}_{X_i}(x_i)}{x_1 \bar{F}_{X_1}(x_1)},$$

pour en déduire que

$$\begin{aligned} \lim_{\alpha \rightarrow 1} \frac{l_{X_k, X_1}^\alpha(x_k, x_1)}{x_1 \bar{F}_{X_1}(x_1)} &= \lim_{\alpha \rightarrow 1} \left(\frac{\alpha \mathbb{E}[(X_k - x_k)_+ \mathbb{1}_{\{X_1 > x_1\}}] - (1 - \alpha) \mathbb{E}[(X_k - x_k)_- \mathbb{1}_{\{X_1 < x_1\}}]}{x_1 \bar{F}_{X_1}(x_1)} \right) \\ &= \lim_{\alpha \rightarrow 1} \left(\frac{\mathbb{E}[(X_k - x_k)_+ \mathbb{1}_{\{X_1 > x_1\}}]}{x_1 \bar{F}_{X_1}(x_1)} - \frac{1 - \alpha}{\bar{F}_{X_1}(x_1)} \frac{x_k}{x_1} \right) = 0, \quad \forall k \in J_\infty^1. \end{aligned}$$

Par passage à la limite ($\alpha \rightarrow 1$) dans la première équation du système d'optimalité 2.7 divisée par $x_1 \bar{F}_{X_1}(x_1)$, nous obtenons alors

$$\lim_{\alpha \rightarrow 1} \sum_{k \in J_0^1 \cup J_C^1} \frac{l_{X_k, X_1}^\alpha(x_k, x_1)}{x_1 \bar{F}_{X_1}(x_1)} = -\frac{1}{\theta - 1},$$

cette conclusion est absurde car

$$\begin{aligned} \lim_{\alpha \rightarrow 1} \sum_{k \in J_0^1 \cup J_C^1} \frac{l_{X_k, X_1}^\alpha(x_k, x_1)}{x_1 \bar{F}_{X_1}(x_1)} &= \lim_{\alpha \rightarrow 1} \sum_{k \in J_0^1 \cup J_C^1} \left(\frac{\alpha \mathbb{E}[(X_k - x_k)_+ \mathbb{1}_{\{X_1 > x_1\}}] - (1 - \alpha) \mathbb{E}[(X_k - x_k)_- \mathbb{1}_{\{X_1 < x_1\}}]}{x_1 \bar{F}_{X_1}(x_1)} \right) \\ &= \lim_{\alpha \rightarrow 1} \sum_{k \in J_0^1 \cup J_C^1} \left(\frac{\mathbb{E}[(X_k - x_k)_+ \mathbb{1}_{\{X_1 > x_1\}}]}{x_1 \bar{F}_{X_1}(x_1)} - \frac{1 - \alpha}{\bar{F}_{X_1}(x_1)} \frac{x_k}{x_1} \right) \\ &= \lim_{\alpha \rightarrow 1} \sum_{k \in J_0^1 \cup J_C^1} \left(\frac{\mathbb{E}[(X_k - x_k)_+ \mathbb{1}_{\{X_1 > x_1\}}]}{x_1 \bar{F}_{X_1}(x_1)} \right) \geq 0. \end{aligned}$$

On peut finalement en déduire que

$$\lim_{\alpha \rightarrow 1} \frac{1 - \alpha}{\bar{F}_{X_1}(x_1)} > 0,$$

et par le même raisonnement que

$$\lim_{\alpha \rightarrow 1} \frac{1 - \alpha}{\bar{F}_{X_i}(x_i)} > 0, \quad \forall i \in \{1, \dots, d\}.$$

□

La combinaison des Propositions 2.2.3 et 2.2.4 donne

$$0 < \lim_{\alpha \rightarrow 1} \frac{1 - \alpha}{\bar{F}_{X_i}(\mathbf{e}_\alpha^i(\mathbf{X}))} \leq \overline{\lim}_{\alpha \rightarrow 1} \frac{1 - \alpha}{\bar{F}_{X_i}(\mathbf{e}_\alpha^i(\mathbf{X}))} < +\infty, \forall i \in \{2, \dots, d\}.$$

Cette conclusion fait de la comparaison entre les composantes de l'expectile multivarié asymptotique et les quantiles une question naturelle. Nous pouvons maintenant démontrer la Proposition 2.2.1.

Preuve de la Proposition 2.2.1 Il suffit de montrer que $J_\infty^1 = \emptyset$, cela implique que pour tout $i \in \{1, \dots, d\}$, $J_0^i = J_\infty^i = \emptyset$.

Supposons que $J_\infty^1 \neq \emptyset$, alors $\exists i \in \{2, \dots, d\}$ et une suite $(\alpha_n)_{n \in \mathbb{N}}$ telle que $x_i/x_1 \rightarrow +\infty$ quand $\alpha_n \rightarrow 1$.

On a, des Propositions 2.2.3 et 2.2.4

$$0 < \lim_{\alpha \rightarrow 1} \frac{1 - \alpha}{\bar{F}_{X_i}(\mathbf{e}_\alpha^i(\mathbf{X}))} \leq \overline{\lim}_{\alpha \rightarrow 1} \frac{1 - \alpha}{\bar{F}_{X_i}(\mathbf{e}_\alpha^i(\mathbf{X}))} < +\infty, \forall i \in \{2, \dots, d\},$$

donc quitte à prendre une sous suite, on considère qu'il existe $\ell \in \mathbb{R}^* \setminus \{+\infty\}$ tel que

$$\lim_{\alpha \rightarrow 1} \frac{1 - \alpha}{\bar{F}_{X_1}(x_1)} = \ell.$$

Dans ce cas,

$$\lim_{\alpha \rightarrow 1} \frac{l_{X_1}^\alpha(x_1)}{x_1 \bar{F}_{X_1}(x_1)} = \lim_{\alpha \rightarrow 1} \left((2\alpha - 1) \frac{\mathbb{E}[(X_1 - x_1)_+]}{x_1 \bar{F}_{X_1}(x_1)} - \frac{1 - \alpha}{\bar{F}_1(x_1)} \left(1 - \frac{\mathbb{E}[X_1]}{x_1}\right) \right) = \frac{1}{\theta - 1} - \ell < +\infty.$$

L'inégalité

$$\frac{\mathbb{E}[(X_i - x_i)_+ \mathbb{1}_{\{X_1 > x_1\}}]}{x_1 \bar{F}_{X_1}(x_1)} \leq \frac{\mathbb{E}[(X_i - x_i)_+]}{x_1 \bar{F}_{X_1}(x_1)} = \frac{\mathbb{E}[(X_i - x_i)_+]}{x_i \bar{F}_{X_i}(x_i)} \frac{x_i \bar{F}_{X_i}(x_i)}{x_1 \bar{F}_{X_1}(x_1)},$$

valide $\forall i \in \{2, \dots, d\}$, nous permet de déduire par application du Lemme 2.2.2 que

$$\lim_{\alpha \rightarrow 1} \frac{\mathbb{E}[(X_k - x_k)_+ \mathbb{1}_{\{X_1 > x_1\}}]}{x_1 \bar{F}_{X_1}(x_1)} = 0, \forall k \in J_\infty^1,$$

puis que

$$\begin{aligned} \lim_{\alpha \rightarrow 1} \frac{l_{X_k, X_1}^\alpha(x_k, x_1)}{x_1 \bar{F}_{X_1}(x_1)} &= \lim_{\alpha \rightarrow 1} \left(\frac{\alpha \mathbb{E}[(X_k - x_k)_+ \mathbb{1}_{\{X_1 > x_1\}}] - (1 - \alpha) \mathbb{E}[(X_k - x_k)_- \mathbb{1}_{\{X_1 < x_1\}}]}{x_1 \bar{F}_{X_1}(x_1)} \right) \\ &= \lim_{\alpha \rightarrow 1} \left(\frac{\mathbb{E}[(X_k - x_k)_+ \mathbb{1}_{\{X_1 > x_1\}}]}{x_1 \bar{F}_{X_1}(x_1)} - \frac{1 - \alpha}{\bar{F}_{X_1}(x_1)} \frac{x_k}{x_1} \right) = -\infty, \forall k \in J_\infty^1. \end{aligned}$$

Par passage à la limite ($\alpha \rightarrow 1$) dans la première équation du système d'optimalité (2.7) divisée par $x_1 \bar{F}_{X_1}(x_1)$, nous déduisons que

$$\lim_{\alpha \rightarrow 1} \sum_{k \in J_0^1 \cup J_C^1} \frac{l_{X_k, X_1}^\alpha(x_k, x_1)}{x_1 \bar{F}_{X_1}(x_1)} = -\infty. \quad (2.16)$$

Pour tout $k \in J_0^1$

$$\begin{aligned} \frac{\mathbb{E}[(X_k - x_k)_+ \mathbb{1}_{\{X_1 > x_1\}}]}{x_1 \bar{F}_{X_1}(x_1)} &= \frac{\int_{x_k}^{x_1} \mathbb{P}(X_k > t, X_1 > x_1) dt}{x_1 \bar{F}_{X_1}(x_1)} + \frac{\int_{x_1}^{+\infty} \mathbb{P}(X_k > t, X_1 > x_1) dt}{x_1 \bar{F}_{X_1}(x_1)} \\ &\leq \frac{\int_{x_k}^{x_1} \mathbb{P}(X_1 > x_1) dt}{x_1 \bar{F}_{X_1}(x_1)} + \frac{\int_{x_1}^{+\infty} \mathbb{P}(X_k > t) dt}{x_1 \bar{F}_{X_1}(x_1)}, \end{aligned}$$

avec, par application du théorème de Karamata (2.1.6)

$$\lim_{\alpha \rightarrow 1} \frac{\int_{x_k}^{x_1} \mathbb{P}(X_1 > x_1) dt}{x_1 \bar{F}_{X_1}(x_1)} + \frac{\int_{x_1}^{+\infty} \mathbb{P}(X_k > t) dt}{x_1 \bar{F}_{X_1}(x_1)} = 1 + \frac{c_k}{\theta - 1}, \forall k \in J_0^1.$$

Pour tout $k \in J_C^1$

$$\frac{\mathbb{E}[(X_k - x_k)_+ \mathbb{1}_{\{X_1 > x_1\}}]}{x_1 \bar{F}_{X_1}(x_1)} \leq \frac{\mathbb{E}[(X_k - x_k)_+]}{x_1 \bar{F}_{X_1}(x_1)} = \frac{\mathbb{E}[(X_k - x_k)_+]}{x_k \bar{F}_{X_k}(x_k)} \frac{x_k \bar{F}_{X_k}(x_k)}{x_1 \bar{F}_{X_1}(x_1)},$$

avec

$$\frac{\mathbb{E}[(X_k - x_k)_+]}{x_k \bar{F}_{X_k}(x_k)} \frac{x_k \bar{F}_{X_k}(x_k)}{x_1 \bar{F}_{X_1}(x_1)} \underset{\alpha \rightarrow 1}{\sim} \frac{c_k}{\theta - 1} \left(\frac{x_k}{x_1} \right)^{-\theta+1}.$$

Nous déduisons finalement que

$$\begin{aligned} - \sum_{k \in J_0^1 \cup J_C^1} \overline{\lim}_{\alpha \rightarrow 1} \frac{x_k}{x_1} &\leq \overline{\lim}_{\alpha \rightarrow 1} \sum_{k \in J_0^1 \cup J_C^1} \frac{l_{X_k, X_1}^\alpha(x_k, x_1)}{x_1 \bar{F}_{X_1}(x_1)} \\ &\leq \overline{\lim}_{\alpha \rightarrow 1} \sum_{k \in J_0^1 \cup J_C^1} \frac{l_{X_k, X_1}^\alpha(x_k, x_1)}{x_1 \bar{F}_{X_1}(x_1)} \\ &\leq \sum_{k \in \cup J_C^1} \left(\frac{c_k}{\theta - 1} \left(\overline{\lim}_{\alpha \rightarrow 1} \frac{x_k}{x_1} \right)^{-\theta+1} - \ell \overline{\lim}_{\alpha \rightarrow 1} \frac{x_k}{x_1} \right) + \sum_{k \in \cup J_0^1} \left(1 + \frac{c_k}{\theta - 1} \right). \end{aligned}$$

Cela est contradictoire avec la limite (2.16), et par suite J_∞^1 est forcément un ensemble vide. \square

Proposition 2.2.5 (Expectile multivarié asymptotique). *Soit $\mathbf{X} = (X_1, \dots, X_d)$ un vecteur aléatoire dans \mathbb{R}^d de lois marginales $F_{X_i}, i = 1, \dots, d$ continues, et satisfont les hypothèses H1, H2. Si \mathbf{X} a une distribution multivariée à variations régulières au sens de la Définition 2.1.9, alors tout vecteur d'adhérence $(\eta, \beta_2, \dots, \beta_d)$ du vecteur $\left(\frac{1-\alpha}{\bar{F}_{X_1}(\mathbf{e}_\alpha^1(\mathbf{X}))}, \frac{\mathbf{e}_\alpha^2(\mathbf{X})}{\mathbf{e}_\alpha^1(\mathbf{X})}, \dots, \frac{\mathbf{e}_\alpha^d(\mathbf{X})}{\mathbf{e}_\alpha^1(\mathbf{X})} \right)$ vérifie le système d'équations suivant*

$$\frac{1}{\theta - 1} - \eta \frac{(\beta_k)^\theta}{c_k} = - \sum_{i=1, i \neq k}^d \left(\int_{\frac{\beta_i}{\beta_k}}^{+\infty} \lambda_U^{ik} \left(\frac{c_i}{c_k} t^{-\theta}, 1 \right) dt - \eta \frac{\beta_k^{\theta-1}}{c_k} \beta_i \right), \forall k \in \{1, \dots, d\}. \quad (2.17)$$

Le système obtenu est composé de d variables $(\eta, \beta_2, \dots, \beta_d)$, sa résolution permet d'obtenir à la fois l'expectile multivarié asymptotique, et des équivalences qui relient ses composantes aux quantiles univariés des risques.

Preuve Le système d'optimalité (2.7) s'écrit aussi sous la forme suivante

$$\begin{aligned} (2\alpha - 1) \frac{\mathbb{E}[(X_k - x_k)_+]}{x_k \bar{F}_{X_k}(x_k)} - \frac{1 - \alpha}{\bar{F}_{X_k}(x_k)} \left(1 - \frac{\mathbb{E}[X_k]}{x_k} \right) &= \sum_{i=1, i \neq k}^d \left((1 - \alpha) \frac{\mathbb{E}[(X_i - x_i)_- \mathbb{1}_{\{X_k < x_k\}}]}{x_k \bar{F}_{X_k}(x_k)} \right) \\ &\quad - \sum_{i=1, i \neq k}^d \alpha \frac{\mathbb{E}[(X_i - x_i)_+ \mathbb{1}_{\{X_k > x_k\}}]}{x_k \bar{F}_{X_k}(x_k)}, \quad \forall k \in \{1, \dots, d\}. \end{aligned}$$

Pour tout $k \in \{1, \dots, d\}$, on a

$$\lim_{\alpha \rightarrow 1} (2\alpha - 1) \frac{\mathbb{E}[(X_k - x_k)_+]}{x_k \bar{F}_{X_k}(x_k)} - \frac{1 - \alpha}{\bar{F}_{X_k}(x_k)} \left(1 - \frac{\mathbb{E}[X_k]}{x_k} \right) = \frac{1}{\theta - 1} - \eta \frac{(\beta_k)^\theta}{c_k},$$

et pour tout $i \in \{1, \dots, d\} \setminus \{k\}$

$$\begin{aligned} \lim_{\alpha \rightarrow 1} (1 - \alpha) \frac{\mathbb{E}[(X_i - x_i)_- \mathbb{1}_{\{X_k < x_k\}}]}{x_k \bar{F}_{X_k}(x_k)} &= \lim_{\alpha \rightarrow 1} \frac{1 - \alpha}{\bar{F}_{X_k}(x_k)} \left(\frac{x_i}{x_k} \mathbb{P}(X_i < x_i, X_k < x_k) - \frac{\mathbb{E}[X_i \mathbb{1}_{\{X_i < x_i, X_k < x_k\}}]}{x_k} \right) \\ &= \eta \frac{\beta_k^\theta}{c_k} \frac{\beta_i}{\beta_k} = \eta \frac{\beta_k^{\theta-1}}{c_k} \beta_i. \end{aligned}$$

Par ailleurs, on peut aussi écrire

$$\begin{aligned} \frac{\mathbb{E}[(X_i - x_i)_+ \mathbb{1}_{\{X_k > x_k\}}]}{x_k \bar{F}_{X_k}(x_k)} &= \frac{1}{x_k \bar{F}_{X_k}(x_k)} \int_{x_i}^{+\infty} \mathbb{P}(X_i > t, X_k > x_k) dt \\ &= \int_{\frac{x_i}{x_k}}^{+\infty} \frac{\mathbb{P}(X_i > tx_k, X_k > x_k)}{\bar{F}_{X_k}(x_k)} dt \\ &= \int_{\frac{x_i}{x_k}}^{+\infty} \frac{\mathbb{P}\left(\bar{F}_{X_i}(X_i) < \bar{F}_{X_i}(tx_k), \bar{F}_{X_k}(X_k) < \bar{F}_{X_k}(x_k)\right)}{\bar{F}_{X_k}(x_k)} dt. \end{aligned}$$

Tout d'abord, on remarque que

$$\left| \int_{\frac{\beta_i}{\beta_k}}^{\frac{x_i}{x_k}} \frac{\mathbb{P}\left(\bar{F}_{X_i}(X_i) < \bar{F}_{X_i}(tx_k), \bar{F}_{X_k}(X_k) < \bar{F}_{X_k}(x_k)\right)}{\bar{F}_{X_k}(x_k)} dt \right| \leq \left| \frac{x_i}{x_k} - \frac{\beta_i}{\beta_k} \right|.$$

La fonction λ_U étant supposée continue, on a

$$\lim_{\alpha \rightarrow 1} \frac{\mathbb{P}\left(\bar{F}_{X_i}(X_i) < \bar{F}_{X_i}(tx_k), \bar{F}_{X_k}(X_k) < \bar{F}_{X_k}(x_k)\right)}{\bar{F}_{X_k}(x_k)} = \lambda_U^{ik} \left(\frac{c_i}{c_k} t^{-\theta}, 1 \right). \quad (2.18)$$

Pour montrer que

$$\lim_{\alpha \rightarrow 1} \alpha \frac{\mathbb{E}[(X_i - x_i)_+ \mathbb{1}_{\{X_k > x_k\}}]}{x_k \bar{F}_{X_k}(x_k)} = \int_{\frac{\beta_i}{\beta_k}}^{+\infty} \lambda_U^{ik} \left(\frac{c_i}{c_k} t^{-\theta}, 1 \right) dt,$$

on utilise le théorème de convergence dominée à l'aide des bornes Potter (1942) (2.1.7) pour les fonctions à variations régulières.

Tout d'abord,

$$\frac{\mathbb{P}\left(\bar{F}_{X_i}(X_i) < \bar{F}_{X_i}(tx_k), \bar{F}_{X_k}(X_k) < \bar{F}_{X_k}(x_k)\right)}{\bar{F}_{X_k}(x_k)} \leq \min \left\{ 1, \frac{\bar{F}_{X_i}(tx_k)}{\bar{F}_{X_k}(x_k)} \right\},$$

on a $\frac{\bar{F}_{X_i}(tx_k)}{\bar{F}_{X_k}(x_k)} = \frac{\bar{F}_{X_i}(tx_k)}{\bar{F}_{X_k}(tx_k)} \frac{\bar{F}_{X_k}(tx_k)}{\bar{F}_{X_k}(x_k)}$ et $\lim_{\alpha \rightarrow 1} \frac{\bar{F}_{X_i}(tx_k)}{\bar{F}_{X_k}(tx_k)} = \frac{c_i}{c_k}$, en appliquant les bornes Potter (2.1.7), pour tout $\varepsilon_1 > 0$ et $0 < \varepsilon_2 < \theta - 1$, il existe $x_k^0(\varepsilon_2, \varepsilon_1)$ tel que pour $\min\{x_k, tx_k\} \geq x_k^0(\varepsilon_2, \varepsilon_1)$

$$\frac{\bar{F}_{X_i}(tx_k)}{\bar{F}_{X_k}(x_k)} \leq \left(\frac{c_i}{c_k} + 2\varepsilon_1 \right) t^{-\theta} \max(t^{\varepsilon_2}, t^{-\varepsilon_2}).$$

La fonction $t \rightarrow \min \left\{ 1, \left(\frac{c_i}{c_k} + 2\varepsilon_1 \right) t^{-\theta} \max(t^{\varepsilon_2}, t^{-\varepsilon_2}) \right\}$ est intégrable sur \mathbb{R}^+ , on peut donc appliquer le théorème de la convergence dominée et obtenir par la suite

$$\lim_{\alpha \rightarrow 1} \int_{\frac{x_i}{x_k}}^{+\infty} \frac{\mathbb{P}\left(\bar{F}_{X_i}(X_i) < \bar{F}_{X_i}(tx_k), \bar{F}_{X_k}(X_k) < \bar{F}_{X_k}(x_k)\right)}{\bar{F}_{X_k}(x_k)} dt = \int_{\frac{\beta_i}{\beta_k}}^{+\infty} \lambda_U^{ik} \left(\frac{c_i}{c_k} t^{-\theta}, 1 \right) dt,$$

et par la suite pour tout $(i \neq k) \in \{1, \dots, d\}^2$

$$\lim_{\alpha \rightarrow 1} \frac{\mathbb{E}[(X_i - x_i)_+ \mathbb{1}_{\{X_k > x_k\}}]}{x_k \bar{F}_{X_k}(x_k)} = \int_{\frac{\beta_i}{\beta_k}}^{+\infty} \lambda_U^{ik} \left(\frac{c_i}{c_k} t^{-\theta}, 1 \right) dt,$$

d'où le système annoncé dans la proposition. \square

Dans le cas d'une matrice à coefficients positifs le système (2.17) devient

$$\frac{1}{\theta - 1} - \eta \frac{(\beta_k)^\theta}{c_k} = - \sum_{i=1, i \neq k}^d \frac{\pi_{ik}}{\pi_k} \left(\int_{\frac{\beta_i}{\beta_k}}^{+\infty} \lambda_U^{ik} \left(\frac{c_i}{c_k} t^{-\theta}, 1 \right) dt - \eta \frac{\beta_k^{\theta-1}}{c_k} \beta_i \right), \forall k \in \{1, \dots, d\}.$$

La résolution du système (2.17) obtenu dans la Proposition 2.2.5 est équivalente à celle du système d'équations suivant

$$\sum_{i=1}^d \int_{\frac{\beta_i}{\beta_k}}^{+\infty} \lambda_U^{ik} \left(c_i t^{-\theta}, c_k \beta_k^{-\theta} \right) dt = \sum_{i=1}^d \int_{\beta_i}^{+\infty} \lambda_U^{i1} \left(c_i t^{-\theta}, 1 \right) dt, \forall k \in \{2, \dots, d\}. \quad (2.19)$$

Les limites β_i sont donc complètement déterminées par la nature des dépendances asymptotiques bivariées entre les composantes du vecteur \mathbf{X} .

Proposition 2.2.6. *Soit $\mathbf{X} = (X_1, \dots, X_d)$ un vecteur aléatoire dans \mathbb{R}^d , dont les lois marginales $F_{X_i}, i = 1, \dots, d$ sont continues, et satisfont les hypothèses H1 et H2. Si \mathbf{X} a une distribution multivariée à variations régulières au sens de la définition (2.1.9), alors tout vecteur d'adhérence $(\eta, \beta_2, \dots, \beta_d)$ du vecteur $\left(\frac{1-\alpha}{F_{X_1}(\mathbf{e}_\alpha^1(\mathbf{X}))}, \frac{\mathbf{e}_\alpha^2(\mathbf{X})}{\mathbf{e}_\alpha^1(\mathbf{X})}, \dots, \frac{\mathbf{e}_\alpha^d(\mathbf{X})}{\mathbf{e}_\alpha^1(\mathbf{X})} \right)$ vérifie le système d'équations suivant, $\forall k \in \{1, \dots, d\}$*

$$\frac{1}{\theta - 1} - \eta \frac{(\beta_k)^\theta}{c_k} = - \sum_{i=1, i \neq k}^d \left(\frac{c_i}{c_k} \left(\frac{\beta_i}{\beta_k} \right)^{-\theta+1} \int_1^{+\infty} \lambda_U^{ik} \left(t^{-\theta}, \frac{c_k}{c_i} \left(\frac{\beta_k}{\beta_i} \right)^{-\theta} \right) dt - \eta \frac{\beta_k^{\theta-1}}{c_k} \beta_i \right). \quad (2.20)$$

Preuve Il suffit d'appliquer un changement de variable au niveau des intégrales du système (2.17) $u = \frac{\beta_k}{\beta_i} t$, puis d'utiliser la propriété de la positive homogénéité des fonctions λ_U^{ik} (Proposition 2.2, Joe et al. (2010) [9]) pour obtenir (2.20). \square

L'utilité principale de l'écriture du système d'optimalité asymptotique sous la forme (2.20) est de pouvoir donner facilement une forme explicite aux intégrales dans plusieurs cas de structure de dépendance.

Exemple Dans le cas des copules de survie archimédiennes qu'on peut présenter sous la forme

$$\bar{C}(x_1, \dots, x_d) = \psi(\psi^{\leftarrow}(x_1) + \dots + \psi^{\leftarrow}(x_d)),$$

avec $\psi^{\leftarrow}(x) = \inf\{t \geq 0 | \psi(t) \leq x\}$. Le générateur $\psi : [0, +\infty) \rightarrow [0, 1]$ est une fonction continue décroissante convexe et vérifie $\psi(0) = 1$ et $\lim_{x \rightarrow +\infty} \psi(x) = 0$.

Cette construction est étudiée dans McNeil et Nešlehová (2009) [14]. Une condition nécessaire et suffisante pour que \bar{C} soit bien une copule est que le générateur ψ est d-monotone sur \mathbb{R}^+ . Si de plus ψ est à variation régulière d'indice strictement négatif $\psi \in \text{RV}_{-\theta_\psi}$, avec $\theta_\psi > 0$, alors selon Charpentier et Segers (2009) [4], les fonctions de dépendance extrême à droite existent, et on peut obtenir leurs formes explicites

$$\lambda_U^k(x_1, \dots, x_k) = \left(\sum_{i=1}^k x_i^{-\frac{1}{\theta_\psi}} \right)^{-\theta_\psi}.$$

Dans notre cas, pour tout $(i, k) \in \{1, \dots, d\}^2$ avec $i \neq k$

$$\lambda_U^{ik} \left(t^{-\theta}, \frac{c_k}{c_i} \left(\frac{\beta_k}{\beta_i} \right)^{-\theta} \right) = \left(t^{\frac{\theta}{\theta_\psi}} + \left(\frac{c_i}{c_k} \right)^{\frac{1}{\theta_\psi}} \left(\frac{\beta_k}{\beta_i} \right)^{\frac{\theta}{\theta_\psi}} \right)^{-\theta_\psi},$$

en particulier si $\theta = \theta_\psi$ alors

$$\int_1^{+\infty} \lambda_U^{ik} \left(t^{-\theta}, \frac{c_k}{c_i} \left(\frac{\beta_k}{\beta_i} \right)^{-\theta} \right) dt = \frac{1}{\theta - 1} \left(1 + \left(\frac{c_i}{c_k} \right)^{\frac{1}{\theta}} \frac{\beta_k}{\beta_i} \right)^{-\theta+1}.$$

Le système d'équations (2.20) se simplifie et devient

$$\frac{1}{\theta-1} - \eta \frac{(\beta_k)^\theta}{c_k} = - \sum_{i=1, i \neq k}^d \left(\frac{1}{\theta-1} \frac{c_i}{c_k} \left(\frac{\beta_i}{\beta_k} + \left(\frac{c_i}{c_k} \right)^{\frac{1}{\theta}} \right)^{-\theta+1} - \eta \frac{\beta_k^{\theta-1}}{c_k} \beta_i \right).$$

□

Lemme 2.2.7 (Cas Fréchet comonotone). *Si le vecteur $\mathbf{X} = (X_1, \dots, X_d)$ est comonotone, sous les hypothèses H1 et H2, et si la limite*

$$(\eta, \beta_2, \dots, \beta_d) = \lim_{\alpha \rightarrow 1} \left(\frac{1-\alpha}{\bar{F}_{X_1}(\mathbf{e}_\alpha^1(\mathbf{X}))}, \frac{\mathbf{e}_\alpha^2(\mathbf{X})}{\mathbf{e}_\alpha^1(\mathbf{X})}, \dots, \frac{\mathbf{e}_\alpha^d(\mathbf{X})}{\mathbf{e}_\alpha^1(\mathbf{X})} \right),$$

existe, alors

$$\lim_{\alpha \rightarrow 1} \frac{1-\alpha}{\bar{F}_{X_k}(\mathbf{e}_\alpha^k(\mathbf{X}))} = \frac{1}{\theta-1} \quad \text{et} \quad \beta_k = c_k^{1/\theta}, \quad \forall k \in \{1, \dots, d\}.$$

Ce résultat est valide pour toute matrice de construction semi-définie positive à coefficients positifs.

Preuve Le vecteur \mathbf{X} est comonotone, sa copule de survie est

$$\hat{C}_{\mathbf{X}}(u_1, \dots, u_d) = \min(u_1, \dots, u_d), \quad \forall (u_1, \dots, u_d) \in [0, 1]^d.$$

On en déduit l'expression de la fonction λ_U^{ij}

$$\lambda_U^{ij}(x_i, x_j) = \min(x_i, x_j), \quad \forall (x_i, x_j) \in \mathbb{R}_+^2.$$

On peut donc écrire

$$\begin{aligned} \int_1^{+\infty} \lambda_U^{ik} \left(t^{-\theta}, \frac{c_k}{c_i} \left(\frac{\beta_k}{\beta_i} \right)^{-\theta} \right) dt &= \int_1^{+\infty} \min \left(t^{-\theta}, \frac{c_k}{c_i} \left(\frac{\beta_k}{\beta_i} \right)^{-\theta} \right) dt \\ &= \left(\frac{\beta_k}{\beta_i} \left(\frac{c_k}{c_i} \right)^{-\frac{1}{\theta}} - 1 \right)_+ \frac{c_k}{c_i} \left(\frac{\beta_k}{\beta_i} \right)^{-\theta} \\ &\quad + \frac{1}{\theta-1} \left(1 + \left(\frac{\beta_k}{\beta_i} \left(\frac{c_k}{c_i} \right)^{-\frac{1}{\theta}} - 1 \right)_+ \right)^{-\theta+1} \end{aligned}$$

Sous les hypothèses H1 et H2, et par application de la Proposition 2.20, le vecteur $(\eta, \beta_2, \dots, \beta_d)$ est alors l'unique solution du système d'équations suivant

$$\begin{aligned} \eta \sum_{i=1}^d \beta_i - \frac{1}{\theta-1} \sum_{i=1}^d c_i \beta_i^{-\theta+1} &= \sum_{i=1, i \neq k}^d c_k \beta_k^{-\theta} \beta_i \left(\frac{\beta_k}{\beta_i} \left(\frac{c_k}{c_i} \right)^{-\frac{1}{\theta}} - 1 \right)_+ \\ &\quad + \frac{1}{\theta-1} \sum_{i=1, i \neq k}^d c_i \beta_i^{-\theta+1} \left[\left(1 + \left(\frac{\beta_k}{\beta_i} \left(\frac{c_k}{c_i} \right)^{-\frac{1}{\theta}} - 1 \right)_+ \right)^{-\theta+1} - 1 \right], \end{aligned}$$

$\forall k \in \{1, \dots, d\}$.

Finalement, il suffit que remarquer que l'annulation des termes en $(\cdot)_+$ permet de visualiser la solution du système. □

Proposition 2.2.8 (Cas d'indépendance asymptotique). *Dans le cas de l'indépendance asymptotique entre les variables aléatoires $X_i, i = 1, \dots, d$, $\exists!$ vecteur d'adhérence $(\eta, \beta_2, \dots, \beta_d)$ du vecteur $\left(\frac{1-\alpha}{\bar{F}_{X_1}(\mathbf{e}_\alpha^1(\mathbf{X}))}, \frac{\mathbf{e}_\alpha^2(\mathbf{X})}{\mathbf{e}_\alpha^1(\mathbf{X})}, \dots, \frac{\mathbf{e}_\alpha^d(\mathbf{X})}{\mathbf{e}_\alpha^1(\mathbf{X})} \right)$ qui vérifie*

$$\eta = \frac{1}{(\theta-1) \left(1 + \sum_{j=2}^d c_j^{\frac{1}{\theta-1}} \right)} \quad \text{et} \quad \beta_k = c_k^{\frac{1}{\theta-1}},$$

pour tout $k \in \{1, \dots, d\}$.

Preuve Le cas d'indépendance asymptotique se traduit par

$$\lim_{\alpha \rightarrow 1} \frac{\mathbb{P}(X_i > x_i, X_j > x_j)}{\mathbb{P}(X_j > x_j)} = \lim_{\alpha \rightarrow 1} \frac{\mathbb{P}(X_i > tx_j, X_j > x_j)}{\mathbb{P}(X_j > x_j)} = 0,$$

pour tout $(i, j) \in \{1, \dots, d\}^2$ et tout réel $t > 0$, donc

$$\begin{aligned} \lim_{\alpha \rightarrow 1} \frac{\mathbb{E}[(X_i - x_i)_+ \mathbb{1}_{\{X_j > x_j\}}]}{x_j \bar{F}_{X_j}(x_j)} &= \lim_{\alpha \rightarrow 1} \int_{\frac{x_i}{x_j}}^{+\infty} \frac{\mathbb{P}(X_i > tx_j, X_j > x_j)}{\mathbb{P}(X_j > x_j)} dt \\ &= \lim_{\alpha \rightarrow 1} \int_{\frac{\beta_i}{\beta_j}}^{+\infty} \frac{\mathbb{P}(X_i > tx_j, X_j > x_j)}{\mathbb{P}(X_j > x_j)} dt \\ &= \int_{\frac{\beta_i}{\beta_j}}^{+\infty} \lim_{\alpha \rightarrow 1} \frac{\mathbb{P}(X_i > tx_j, X_j > x_j)}{\mathbb{P}(X_j > x_j)} dt = 0, \end{aligned}$$

car la même méthode de la preuve de la Proposition 2.2.5 s'applique pour utiliser le théorème de convergence dominée.

L'expectile multivarié vérifie donc asymptotiquement le système d'équations suivant

$$\frac{1}{\theta - 1} - \frac{\eta}{c_k} \beta_k^\theta = + \sum_{i=1, i \neq k}^d \frac{\eta}{c_k} \beta_k^{\theta-1} \beta_i, \quad \forall k \in \{1, \dots, d\},$$

qui peut s'écrire encore

$$\frac{c_k}{\eta(\theta - 1)\beta_k^{\theta-1}} = \sum_{i=1}^d \beta_i, \quad \forall k \in \{1, \dots, d\}, \quad (2.21)$$

d'où $\beta_k = c_k^{\frac{1}{\theta-1}}$ pour tout $k \in \{1, \dots, d\}$, et en utilisant l'équation pour $k = 1$ dans le système (2.21) on obtient aussi la valeur de η en fonction des constantes c_i

$$\eta = \frac{1}{(\theta - 1) \left(1 + \sum_{j=2}^d c_j^{\frac{1}{\theta-1}} \right)},$$

on en déduit que

$$\lim_{\alpha \rightarrow 1} \frac{1 - \alpha}{\bar{F}_{X_i}(x_i)} = \eta \frac{1}{c_i} \beta_i^\theta = \frac{c_i^{\frac{1}{\theta-1}}}{(\theta - 1) \left(1 + \sum_{j=2}^d c_j^{\frac{1}{\theta-1}} \right)},$$

pour tout $i \in \{2, \dots, d\}$. □

Dans le cas plus général d'une matrice à coefficients π_{ij} , $i, j \in \{1, \dots, d\}$ positifs, les limites β_i , $i = 2, \dots, d$ restent les mêmes, mais la limite η change, l'expectile multivarié asymptotique vérifie dans ce cas

$$\lim_{\alpha \rightarrow 1} \frac{\mathbf{e}_\alpha^k(\mathbf{X})}{\mathbf{e}_\alpha^1(\mathbf{X})} = c_k^{\frac{1}{\theta-1}} \text{ et } \lim_{\alpha \rightarrow 1} \frac{1 - \alpha}{\bar{F}_{X_k}(\mathbf{e}_\alpha^k(\mathbf{X}))} = \frac{c_k^{\frac{1}{\theta-1}}}{(\theta - 1) \left(1 + \sum_{j=2}^d \frac{\pi_{jk}}{\pi_k} c_j^{\frac{1}{\theta-1}} \right)},$$

pour tout $k \in \{1, \dots, d\}$.

Il faut signaler que plusieurs définitions de l'indépendance asymptotique sont proposées dans la littérature. Nous utilisons ici la définition la plus simple qui se résume dans une fonction de dépendance de queues nulle.

On remarque que dans ce cas,

$$\lim_{\alpha \rightarrow 1} \frac{1 - \alpha}{\bar{F}_{X_i}(x_i)} \leq \frac{c_i^{\frac{1}{\theta-1}}}{\theta - 1},$$

ce qui permet d'avoir une comparaison entre le niveau du quantile et celui de la composante de l'expectile multivarié, et en utilisant la régularité des variations de la fonction $F_{X_k}^{-1}(1 - \cdot)$ en 0 pour tout $k \in \{1, \dots, d\}$ et avec un indice de variation $-\frac{1}{\theta}$ (Lemme 2.1.5) on obtient

$$x_k \underset{\alpha \rightarrow 1}{\sim} \text{VaR}_\alpha(X_k) (\theta - 1)^{-\frac{1}{\theta}} \left(\frac{1 + \sum_{i=2}^d c_i^{\frac{1}{\theta-1}}}{c_i^{\frac{1}{\theta-1}}} \right)^{-\frac{1}{\theta}}.$$

Ces conclusions coïncident avec le résultat obtenu en dimension 1 pour les distributions du domaine d'attraction extrême de Fréchet dans Bellini et Di Bernardino (2015) [2]. Les valeurs des constantes c_i jouent un rôle important dans la position du quantile par rapport à la composante de l'expectile pour chaque risque.

2.3 Modèle Fréchet à queue dominante

Un autre cas que l'on peut étudier est celui de la dominance d'un risque traduite par des $c_i = 0$ pour tout $i \in \{2, \dots, d\}$ dans l'hypothèse H1. Un risque domine asymptotiquement les autres. On démontre que sa composante dans l'expectile multivarié domine elle aussi celles des autres ($\beta_i = 0, \forall i \in \{2, \dots, d\}$).

Proposition 2.3.1 (Cas de dominance asymptotique). *Sous l'hypothèse H1, si*

$$\lim_{x \uparrow +\infty} \frac{\bar{F}_{X_i}(x)}{\bar{F}_{X_1}(x)} = 0, \quad \forall i \in \{2, \dots, d\},$$

alors

$$\beta_i = \lim_{\alpha \uparrow 1} \frac{\mathbf{e}_\alpha^i(\mathbf{X})}{\mathbf{e}_\alpha^1(\mathbf{X})} = 0, \quad \lim_{\alpha \uparrow 1} \frac{1 - \alpha}{\bar{F}_{X_k}(\mathbf{e}_\alpha^i(\mathbf{X}))} = 0, \quad \forall i \in \{2, \dots, d\},$$

et

$$\lim_{\alpha \uparrow 1} \frac{1 - \alpha}{\bar{F}_{X_1}(\mathbf{e}_\alpha^1(\mathbf{X}))} = \frac{1}{\theta - 1}.$$

Pour la démonstration de la Proposition 2.3.1, nous aurons besoin des Lemmes 2.3.2, 2.3.3 et 2.3.4 suivants.

Lemme 2.3.2. *Sous l'hypothèse H1, si*

$$\lim_{x \uparrow +\infty} \frac{\bar{F}_{X_i}(x)}{\bar{F}_{X_1}(x)} = 0, \quad \forall i \in \{2, \dots, d\},$$

alors

$$\lim_{\alpha \uparrow 1} \frac{\mathbb{E}[(X_i - x_i)_+ \mathbb{1}_{\{X_1 > x_1\}}]}{x_1 \bar{F}_{X_1}(x_1)} = 0, \quad \forall i \in \{2, \dots, d\}.$$

Preuve du Lemme 2.3.2 Quitte à prendre une suite $(\alpha_n)_{n \in \mathbb{N}}$ $\alpha_n \rightarrow 1$, on considère que les limites

$$\lim_{\alpha \rightarrow 1} \frac{x_i}{x_1} = \beta_i \text{ existent.}$$

On note $J_C = \{i \in \{2, \dots, d\} \mid 0 < \beta_i < +\infty\}$, alors pour tout $i \in J_C$

$$\lim_{\alpha \uparrow 1} \frac{\mathbb{E}[(X_i - x_i)_+ \mathbb{1}_{\{X_1 > x_1\}}]}{x_1 \bar{F}_{X_1}(x_1)} = \lim_{\alpha \uparrow 1} \int_{\beta_i}^{+\infty} \frac{\mathbb{P}(X_i > tx_1, X_1 > x_1)}{\mathbb{P}(X_1 > x_1)} dt,$$

car

$$\frac{\mathbb{P}(X_i > tx_1, X_1 > x_1)}{\mathbb{P}(X_1 > x_1)} = \mathbb{P}(X_i > tx_1 \mid X_1 > x_1) \leq 1.$$

D'un autre côté,

$$\frac{\mathbb{P}(X_i > tx_1, X_1 > x_1)}{\mathbb{P}(X_1 > x_1)} \leq \min\left\{1, \frac{\mathbb{P}(X_i > tx_1)}{\mathbb{P}(X_1 > x_1)}\right\},$$

et

$$\frac{\mathbb{P}(X_i > tx_1)}{\mathbb{P}(X_1 > x_1)} = \frac{\bar{F}_{X_i}(tx_1) \bar{F}_{X_1}(tx_1)}{\bar{F}_{X_i}(tx_1) \bar{F}_{X_1}(x_1)},$$

alors, en utilisant que $\lim_{\alpha \uparrow 1} \frac{\bar{F}_{X_i}(tx_1)}{\bar{F}_{X_1}(tx_1)} = 0$ et les bornes de Potter (2.1.7) associées pour \bar{F}_{X_1} , on déduit que pour tout $\epsilon_1 > 0$ et $0 < \epsilon_2 < 1$, il existe $x_1^0(\epsilon_1, \epsilon_2)$ tel que pour $x_1 \geq \frac{x_1^0(\epsilon_1, \epsilon_2)}{\min\{1, \beta_i\}}$

$$\frac{\mathbb{P}(X_i > tx_1)}{\mathbb{P}(X_1 > x_1)} \leq \epsilon_1(1 + \epsilon_2) \max(t^{-\theta + \epsilon_2}, t^{-\theta - \epsilon_2}).$$

On applique le théorème de la convergence dominée pour déduire que

$$\lim_{\alpha \uparrow 1} \frac{\mathbb{E}[(X_i - x_i)_+ \mathbb{1}_{\{X_1 > x_1\}}]}{x_1 \bar{F}_{X_1}(x_1)} = \int_{\beta_i}^{+\infty} \lim_{\alpha \uparrow 1} \frac{\mathbb{P}(X_i > tx_1, X_1 > x_1)}{\mathbb{P}(X_1 > x_1)} dt = 0, \quad \forall i \in J_C.$$

On note $J_\infty = \{i \in \{2, \dots, d\} | \beta_i = +\infty\}$. Pour tout $i \in J_\infty$, on a $x_1 = o(x_i)$ et

$$\begin{aligned} \frac{\mathbb{E}[(X_i - x_i)_+ \mathbb{1}_{\{X_1 > x_1\}}]}{x_1 \bar{F}_{X_1}(x_1)} &= \int_{x_i}^{+\infty} \frac{\mathbb{P}(X_i > t, X_1 > x_1)}{x_1 \mathbb{P}(X_1 > x_1)} dt \\ &\leq \int_{x_1}^{+\infty} \frac{\mathbb{P}(X_i > t, X_1 > x_1)}{x_1 \mathbb{P}(X_1 > x_1)} dt = \int_1^{+\infty} \frac{\mathbb{P}(X_i > tx_1, X_1 > x_1)}{\mathbb{P}(X_1 > x_1)} dt. \end{aligned}$$

De la même méthode que le cas précédent, et en utilisant les bornes de Potter, on démontre que

$$\lim_{\alpha \uparrow 1} \int_1^{+\infty} \frac{\mathbb{P}(X_i > tx_1, X_1 > x_1)}{\mathbb{P}(X_1 > x_1)} dt = \int_1^{+\infty} \lim_{\alpha \uparrow 1} \frac{\mathbb{P}(X_i > tx_1, X_1 > x_1)}{\mathbb{P}(X_1 > x_1)} dt = 0,$$

et on en déduit que

$$\lim_{\alpha \uparrow 1} \frac{\mathbb{E}[(X_i - x_i)_+ \mathbb{1}_{\{X_1 > x_1\}}]}{x_1 \bar{F}_{X_1}(x_1)} = 0, \quad \forall i \in J_\infty.$$

On note $J_0 = \{i \in \{2, \dots, d\} | \beta_i = 0\}$. Pour tout $i \in J_0$ on a $x_i = o(x_1)$ donc

$$\lim_{\alpha \uparrow 1} \frac{\mathbb{E}[(X_i - x_i)_+ \mathbb{1}_{\{X_1 > x_1\}}]}{x_1 \bar{F}_{X_1}(x_1)} = \lim_{\alpha \uparrow 1} \int_{x_i}^{+\infty} \frac{\mathbb{P}(X_i > t, X_1 > x_1)}{x_1 \mathbb{P}(X_1 > x_1)} dt = \lim_{\alpha \uparrow 1} \int_0^{+\infty} \frac{\mathbb{P}(X_i > tx_1, X_1 > x_1)}{\mathbb{P}(X_1 > x_1)} dt,$$

car $\lim_{\alpha \uparrow 1} \frac{x_i}{x_1} = 0$ et $\frac{\mathbb{P}(X_i > tx_1, X_1 > x_1)}{\mathbb{P}(X_1 > x_1)} \leq 1$. Et pour tout $\epsilon > 0$ on a

$$\lim_{\alpha \uparrow 1} \int_\epsilon^{+\infty} \frac{\mathbb{P}(X_i > tx_1, X_1 > x_1)}{\mathbb{P}(X_1 > x_1)} dt = 0,$$

car le théorème de la convergence dominée est applicable à l'aide des bornes de Potter, et $\lim_{\alpha \uparrow 1} \frac{\mathbb{P}(X_i > tx_1, X_1 > x_1)}{\mathbb{P}(X_1 > x_1)} = 0$ pour tout $t > 0$ parce que $c_i = 0$.

Soit $\kappa > 0$, $\forall \epsilon > 0 \exists \alpha_0$ tel que $\forall \alpha > \alpha_0$

$$\int_\epsilon^{+\infty} \frac{\mathbb{P}(X_i > tx_1, X_1 > x_1)}{\mathbb{P}(X_1 > x_1)} dt < \kappa,$$

donc

$$\int_0^{+\infty} \frac{\mathbb{P}(X_i > tx_1, X_1 > x_1)}{\mathbb{P}(X_1 > x_1)} dt = \int_0^\epsilon \frac{\mathbb{P}(X_i > tx_1, X_1 > x_1)}{\mathbb{P}(X_1 > x_1)} dt + \int_\epsilon^{+\infty} \frac{\mathbb{P}(X_i > tx_1, X_1 > x_1)}{\mathbb{P}(X_1 > x_1)} dt < \epsilon + \kappa,$$

on en déduit que

$$\lim_{\alpha \uparrow 1} \int_0^{+\infty} \frac{\mathbb{P}(X_i > tx_1, X_1 > x_1)}{\mathbb{P}(X_1 > x_1)} dt = 0,$$

donc

$$\lim_{\alpha \uparrow 1} \frac{\mathbb{E}[(X_i - x_i)_+ \mathbb{1}_{\{X_1 > x_1\}}]}{x_1 \bar{F}_{X_1}(x_1)} = 0, \quad \forall i \in J_0.$$

Nous avons donc démontré que

$$\lim_{\alpha \uparrow 1} \frac{\mathbb{E}[(X_i - x_i)_+ \mathbb{1}_{\{X_1 > x_1\}}]}{x_1 \bar{F}_{X_1}(x_1)} = 0, \quad \forall i \in \{2, \dots, d\}.$$

□

Lemme 2.3.3. *Sous l'hypothèse H1, si*

$$\lim_{x \uparrow +\infty} \frac{\bar{F}_{X_i}(x)}{\bar{F}_{X_1}(x)} = 0, \quad \forall i \in \{2, \dots, d\},$$

alors

$$\overline{\lim}_{\alpha \uparrow 1} \frac{1 - \alpha}{\bar{F}_{X_1}(\mathbf{e}_\alpha^1(\mathbf{X}))} < +\infty.$$

Preuve du Lemme 2.3.3 Supposons que $\overline{\lim}_{\alpha \uparrow 1} \frac{1 - \alpha}{\bar{F}_{X_1}(\mathbf{e}_\alpha^1(\mathbf{X}))} = +\infty$. Quitte à prendre une suite $(\alpha_n)_{n \in \mathbb{N}}$

$\alpha_n \rightarrow 1$, on considère que les limites $\lim_{\alpha \rightarrow 1} \frac{x_i}{x_1} = \beta_i$ existent et que $\lim_{\alpha \uparrow 1} \frac{1 - \alpha}{\bar{F}_{X_1}(x_1)} = +\infty$.

On note $J_C = \{i \in \{2, \dots, d\} \mid 0 < \beta_i < +\infty\}$, $J_0 = \{i \in \{2, \dots, d\} \mid \beta_i = 0\}$, et $J_\infty = \{i \in \{2, \dots, d\} \mid \beta_i = +\infty\}$.

La première équation du système d'optimalité (1.9) divisée par $x_1 \bar{F}_{X_1}(x_1)$ peut s'écrire sous la forme suivante

$$\begin{aligned} \frac{(2\alpha - 1)\mathbb{E}[(X_1 - x_1)_+]}{x_1 \bar{F}_{X_1}(x_1)} + \sum_{i=2}^d \frac{\alpha \mathbb{E}[(X_i - x_i)_+ \mathbb{1}_{\{X_1 > x_1\}}]}{x_1 \bar{F}_{X_1}(x_1)} &= \frac{1 - \alpha}{\bar{F}_{X_1}(x_1)} \left(\frac{x_1 - \mathbb{E}[X_1]}{x_1} \right) \\ &+ \frac{1 - \alpha}{\bar{F}_{X_1}(x_1)} \left(\sum_{i=2}^d \frac{\mathbb{E}[(X_i - x_i)_- \mathbb{1}_{\{X_1 < x_1\}}]}{x_1} \right), \end{aligned}$$

et on a par (2.6)

$$\lim_{\alpha \uparrow 1} \frac{(2\alpha - 1)\mathbb{E}[(X_1 - x_1)_+]}{x_1 \bar{F}_{X_1}(x_1)} = \frac{1}{\theta - 1},$$

et par le Lemme 2.3.2

$$\lim_{\alpha \uparrow 1} \sum_{i=2}^d \frac{\alpha \mathbb{E}[(X_i - x_i)_+ \mathbb{1}_{\{X_1 > x_1\}}]}{x_1 \bar{F}_{X_1}(x_1)} = 0,$$

donc par passage à la limite ($\alpha \rightarrow 1$) dans l'équation précédente

$$\lim_{\alpha \uparrow 1} \frac{1 - \alpha}{\bar{F}_{X_1}(x_1)} \left(\frac{x_1 - \mathbb{E}[X_1]}{x_1} + \sum_{i=2}^d \frac{\mathbb{E}[(X_i - x_i)_- \mathbb{1}_{\{X_1 < x_1\}}]}{x_1} \right) = \frac{1}{\theta - 1},$$

or,

$$\lim_{\alpha \uparrow 1} \frac{1 - \alpha}{\bar{F}_{X_1}(x_1)} \left(\frac{x_1 - \mathbb{E}[X_1]}{x_1} + \sum_{i=2}^d \frac{\mathbb{E}[(X_i - x_i)_- \mathbb{1}_{\{X_1 < x_1\}}]}{x_1} \right) = \lim_{\alpha \uparrow 1} \frac{1 - \alpha}{\bar{F}_{X_1}(x_1)} \left(1 + \sum_{i=2}^d \frac{x_i}{x_1} \right) = +\infty.$$

Cette contraction nous permet de conclure que le cas $\lim_{\alpha \uparrow 1} \frac{1 - \alpha}{\bar{F}_{X_1}(x_1)} = +\infty$ est impossible. □

Lemme 2.3.4. *Sous l'hypothèse H1, si*

$$\lim_{x \uparrow +\infty} \frac{\bar{F}_{X_i}(x)}{\bar{F}_{X_1}(x)} = 0, \quad \forall i \in \{2, \dots, d\},$$

alors

$$0 < \lim_{\alpha \uparrow 1} \frac{1 - \alpha}{\bar{F}_{X_1}(e_\alpha^1(\mathbf{X}))}.$$

Preuve du Lemme 2.3.4 Supposons que $\lim_{\alpha \uparrow 1} \frac{1 - \alpha}{\bar{F}_{X_1}(e_\alpha^1(\mathbf{X}))} = 0$. Quitte à prendre une suite $(\alpha_n)_{n \in \mathbb{N}}$

$\alpha_n \rightarrow 1$, on considère que les limites $\lim_{\alpha \rightarrow 1} \frac{x_i}{x_1} = \beta_i$ existent et que $\lim_{\alpha \uparrow 1} \frac{1 - \alpha}{\bar{F}_{X_1}(x_1)} = 0$.

On note $J_C = \{i \in \{2, \dots, d\} \mid 0 < \beta_i < +\infty\}$, $J_0 = \{i \in \{2, \dots, d\} \mid \beta_i = 0\}$, et $J_\infty = \{i \in \{2, \dots, d\} \mid \beta_i = +\infty\}$.

Par passage à la limite ($\alpha \rightarrow 1$) dans la première équation du système d'optimalité (1.9) divisée par $x_1 \bar{F}_{X_1}(x_1)$, et à l'aide du Lemme 2.3.2 et de l'Équation 2.6, on obtient

$$\lim_{\alpha \uparrow 1} \left(\frac{1 - \alpha}{\bar{F}_{X_1}(x_1)} \sum_{i \in J_\infty} \frac{x_i}{x_1} \right) = \frac{1}{\theta - 1}. \quad (2.22)$$

Si $J_\infty \neq \emptyset$, alors il existe $i \in \{2, \dots, d\}$ tel que $i \in J_\infty$. Dans ce cas,

$$\lim_{\alpha \uparrow 1} \frac{\mathbb{E}[(X_i - x_i)_+]}{x_1 \bar{F}_{X_1}(x_1)} = \lim_{\alpha \uparrow 1} \frac{\bar{F}_{X_i}(x_i)}{\bar{F}_{X_1}(x_i)} \frac{x_i \bar{F}_{X_1}(x_i)}{x_1 \bar{F}_{X_1}(x_1)} \frac{\mathbb{E}[(X_i - x_i)_+]}{x_i \bar{F}_{X_i}(x_i)} = 0,$$

car $\lim_{\alpha \uparrow 1} \frac{\mathbb{E}[(X_i - x_i)_+]}{x_i \bar{F}_{X_i}(x_i)} = \frac{1}{\theta - 1}$, $\lim_{\alpha \uparrow 1} \frac{\bar{F}_{X_i}(x_i)}{\bar{F}_{X_1}(x_i)} = 0$, et par application du Lemme 2.2.2 (cas $X_i = X_j = X_1$)

$$\lim_{\alpha \uparrow 1} \frac{x_i \bar{F}_{X_1}(x_i)}{x_1 \bar{F}_{X_1}(x_1)} = 0.$$

D'un autre côté, pour tout $j \in \{1, \dots, d\} \setminus \{i\}$,

$$\frac{\mathbb{E}[(X_j - x_j)_+ \mathbb{1}_{\{X_i > x_i\}}]}{x_1 \bar{F}_{X_1}(x_1)} = \int_{x_j}^{+\infty} \frac{\mathbb{P}(X_j > t, X_i > x_i)}{x_1 \mathbb{P}(X_1 > x_1)} dt,$$

donc si $j \in J_C$, alors

$$\lim_{\alpha \uparrow 1} \frac{\mathbb{E}[(X_j - x_j)_+ \mathbb{1}_{\{X_i > x_i\}}]}{x_1 \bar{F}_{X_1}(x_1)} = \lim_{\alpha \uparrow 1} \int_{\frac{x_j}{x_1}}^{+\infty} \frac{\mathbb{P}(X_j > tx_1, X_i > x_i)}{\mathbb{P}(X_1 > x_1)} dt = \lim_{\alpha \uparrow 1} \int_{\beta_j}^{+\infty} \frac{\mathbb{P}(X_j > tx_1, X_i > x_i)}{\mathbb{P}(X_1 > x_1)} dt,$$

car $\frac{\mathbb{P}(X_j > tx_1, X_i > x_i)}{\mathbb{P}(X_1 > x_1)} \leq \frac{\mathbb{P}(X_j > tx_1)}{\mathbb{P}(X_1 > x_1)}$ et $\lim_{\alpha \uparrow 1} \frac{\mathbb{P}(X_j > tx_1)}{\mathbb{P}(X_1 > x_1)} = 0$ pour tout $t > 0$. Par utilisation des bornes

de Potter pour la fonction \bar{F}_{X_1} pour appliquer le théorème de la convergence dominée, on a

$$\lim_{\alpha \uparrow 1} \frac{\mathbb{E}[(X_j - x_j)_+ \mathbb{1}_{\{X_i > x_i\}}]}{x_1 \bar{F}_{X_1}(x_1)} = \int_{\beta_j}^{+\infty} \lim_{\alpha \uparrow 1} \frac{\mathbb{P}(X_j > tx_1, X_i > x_i)}{\mathbb{P}(X_1 > x_1)} dt,$$

et puisque

$$\frac{\mathbb{P}(X_j > tx_1, X_i > x_i)}{\mathbb{P}(X_1 > x_1)} \leq \frac{\mathbb{P}(X_i > x_i)}{\mathbb{P}(X_1 > x_1)} = \frac{x_1 \bar{F}_{X_1}(x_i) \bar{F}_{X_i}(x_i)}{x_i \bar{F}_{X_1}(x_1) \bar{F}_{X_1}(x_i)},$$

alors par application du Lemme 2.2.2,

$$\lim_{\alpha \uparrow 1} \frac{\mathbb{P}(X_j > tx_1, X_i > x_i)}{\mathbb{P}(X_1 > x_1)} = \lim_{\alpha \uparrow 1} \frac{\mathbb{P}(X_i > x_i)}{\mathbb{P}(X_1 > x_1)} = 0,$$

et on en déduit finalement que

$$\lim_{\alpha \uparrow 1} \frac{\mathbb{E}[(X_j - x_j)_+ \mathbb{1}_{\{X_i > x_i\}}]}{x_1 \bar{F}_{X_1}(x_1)} = 0, \quad \forall j \in J_C.$$

Si $j \in J_\infty \setminus \{i\}$, alors

$$\begin{aligned} \frac{\mathbb{E}[(X_j - x_j)_+ \mathbb{1}_{\{X_i > x_i\}}]}{x_1 \bar{F}_{X_1}(x_1)} &= \int_{x_j}^{+\infty} \frac{\mathbb{P}(X_j > t, X_i > x_i)}{x_1 \mathbb{P}(X_1 > x_1)} dt \\ &\leq \int_{x_1}^{+\infty} \frac{\mathbb{P}(X_j > t, X_i > x_i)}{x_1 \mathbb{P}(X_1 > x_1)} dt = \int_1^{+\infty} \frac{\mathbb{P}(X_j > tx_1, X_i > x_i)}{\mathbb{P}(X_1 > x_1)} dt, \end{aligned}$$

de la même façon que le cas précédent on démontre que

$$\lim_{\alpha \uparrow 1} \int_1^{+\infty} \frac{\mathbb{P}(X_j > tx_1, X_i > x_i)}{\mathbb{P}(X_1 > x_1)} dt = \int_1^{+\infty} \lim_{\alpha \uparrow 1} \frac{\mathbb{P}(X_j > tx_1, X_i > x_i)}{\mathbb{P}(X_1 > x_1)} dt = 0,$$

donc

$$\lim_{\alpha \uparrow 1} \frac{\mathbb{E}[(X_j - x_j)_+ \mathbb{1}_{\{X_i > x_i\}}]}{x_1 \bar{F}_{X_1}(x_1)} = 0, \quad \forall j \in J_\infty \setminus \{i\}.$$

Si $j \in J_0$, alors

$$\begin{aligned} \frac{\mathbb{E}[(X_j - x_j)_+ \mathbb{1}_{\{X_i > x_i\}}]}{x_1 \bar{F}_{X_1}(x_1)} &= \int_{x_j}^{x_1} \frac{\mathbb{P}(X_j > tx_1, X_i > x_i)}{x_1 \mathbb{P}(X_1 > x_1)} dt + \int_{x_1}^{+\infty} \frac{\mathbb{P}(X_j > tx_1, X_i > x_i)}{x_1 \mathbb{P}(X_1 > x_1)} dt \\ &\leq \frac{x_1 - x_j}{x_1} \frac{\bar{F}_{X_i}(x_i)}{\bar{F}_{X_1}(x_1)} + \int_1^{+\infty} \frac{\mathbb{P}(X_j > tx_1, X_i > x_i)}{\mathbb{P}(X_1 > x_1)} dt, \end{aligned}$$

on sait que $\lim_{\alpha \uparrow 1} \int_1^{+\infty} \frac{\mathbb{P}(X_j > tx_1, X_i > x_i)}{\mathbb{P}(X_1 > x_1)} dt = 0$, et par application du Lemme 2.2.2, on démontre que

$$\lim_{\alpha \uparrow 1} \frac{x_1 - x_j}{x_1} \frac{\bar{F}_{X_i}(x_i)}{\bar{F}_{X_1}(x_1)} = \lim_{\alpha \uparrow 1} \frac{x_1 - x_j}{x_1} \frac{\bar{F}_{X_i}(x_i)}{\bar{F}_{X_1}(x_1)} \frac{x_i \bar{F}_{X_1}(x_i)}{x_1 \bar{F}_{X_1}(x_1)} \frac{x_1}{x_i} = 0,$$

alors

$$\lim_{\alpha \uparrow 1} \frac{\mathbb{E}[(X_j - x_j)_+ \mathbb{1}_{\{X_i > x_i\}}]}{x_1 \bar{F}_{X_1}(x_1)} = 0, \quad \forall j \in J_0,$$

et par suite

$$\lim_{\alpha \uparrow 1} \frac{\mathbb{E}[(X_j - x_j)_+ \mathbb{1}_{\{X_i > x_i\}}]}{x_1 \bar{F}_{X_1}(x_1)} = 0, \quad \forall j \in \{1, \dots, d\} \setminus \{i\}.$$

La $i^{\text{ème}}$ équation du système d'optimalité (1.9) divisée par $x_1 \bar{F}_{X_1}(x_1)$ s'écrit

$$\begin{aligned} \frac{(2\alpha - 1)\mathbb{E}[(X_i - x_i)_+]}{x_1 \bar{F}_{X_1}(x_1)} - \frac{1 - \alpha}{\bar{F}_{X_1}(x_1)} \frac{x_i - \mathbb{E}[X_i]}{x_1} &= \sum_{\substack{j=1 \\ j \neq i}}^d \frac{(1 - \alpha)\mathbb{E}[(X_j - x_j)_- \mathbb{1}_{\{X_i < x_i\}}]}{x_1 \bar{F}_{X_1}(x_1)} \\ &\quad - \sum_{\substack{j=1 \\ j \neq i}}^d \frac{\alpha \mathbb{E}[(X_j - x_j)_+ \mathbb{1}_{\{X_i > x_i\}}]}{x_1 \bar{F}_{X_1}(x_1)}, \end{aligned}$$

par passage à la limite ($\alpha \rightarrow 1$) on obtient

$$-\lim_{\alpha \uparrow 1} \frac{1 - \alpha}{\bar{F}_{X_1}(x_1)} \frac{x_i}{x_1} = \lim_{\alpha \uparrow 1} \frac{1 - \alpha}{\bar{F}_{X_1}(x_1)} \sum_{\substack{j=1 \\ j \neq i}}^d \frac{x_j}{x_1},$$

cela n'est possible que si $\lim_{\alpha \uparrow 1} \frac{1 - \alpha}{\bar{F}_{X_1}(x_1)} \frac{x_i}{x_1} = 0$, ce qui est contradictoire avec l'équation (2.22). \square

Maintenant, on peut démontrer la Proposition 2.3.1.

Preuve de la Proposition 2.3.1 En combinant les Lemmes 2.3.3 et 2.3.4, on a

$$0 < \liminf_{\alpha \uparrow 1} \frac{1 - \alpha}{\bar{F}_{X_1}(x_1)} \leq \overline{\lim}_{\alpha \uparrow 1} \frac{1 - \alpha}{\bar{F}_{X_1}(x_1)} < +\infty.$$

Quitte à prendre une suite $(\alpha_n)_{n \in \mathbb{N}}$ $\alpha_n \rightarrow 1$, on suppose que $\frac{1 - \alpha}{\bar{F}_{X_1}(x_1)}$ converge vers une limite notée $0 < \eta < +\infty$ et que les limites $\lim_{\alpha \rightarrow 1} \frac{x_i}{x_1} = \beta_i$ existent.

Par passage à la limite ($\alpha \rightarrow 1$) dans la 1^{ère} équation du système d'optimalité (1.9) divisée par $x_1 \bar{F}_{X_1}(x_1)$, et à l'aide du Lemme 2.3.2, on obtient

$$\lim_{\alpha \uparrow 1} \left(\frac{1 - \alpha}{\bar{F}_{X_1}(x_1)} \sum_{i \in J_C \cup J_\infty} \frac{x_i}{x_1} \right) = \lim_{\alpha \uparrow 1} \left(\eta \sum_{i \in J_C \cup J_\infty} \frac{x_i}{x_1} \right) = \frac{1}{\theta - 1} - \eta, \quad (2.23)$$

on en déduit que $J_\infty = \emptyset$.

On suppose que $J_C \neq \emptyset$, donc il existe au moins $i \in \{2, \dots, d\}$ tel que $i \in J_C$. Pour tout $j \in \{1, \dots, d\} \setminus \{i\}$, on démontre que

$$\lim_{\alpha \uparrow 1} \frac{\mathbb{E}[(X_j - x_j)_+ \mathbb{1}_{\{X_i > x_i\}}]}{x_1 \bar{F}_{X_1}(x_1)} = 0.$$

En effet, si $j \in J_C \setminus \{i\}$

$$\lim_{\alpha \uparrow 1} \frac{\mathbb{E}[(X_j - x_j)_+ \mathbb{1}_{\{X_i > x_i\}}]}{x_1 \bar{F}_{X_1}(x_1)} = \lim_{\alpha \uparrow 1} \int_{\beta_j}^{+\infty} \frac{\mathbb{P}(X_j > tx_1, X_i > x_i)}{\mathbb{P}(X_1 > x_1)} dt,$$

car

$$\frac{\mathbb{P}(X_j > tx_j, X_i > x_i)}{\mathbb{P}(X_1 > x_1)} = \mathbb{P}(X_j > tx_j | X_i > x_i) \frac{\mathbb{P}(X_i > x_i)}{\mathbb{P}(X_1 > x_1)},$$

et $\lim_{\alpha \uparrow 1} \frac{\mathbb{P}(X_i > x_i)}{\mathbb{P}(X_1 > x_1)} = \lim_{\alpha \uparrow 1} \frac{\mathbb{P}(X_i > x_i)}{\mathbb{P}(X_1 > x_i)} \frac{\mathbb{P}(X_1 > x_i)}{\mathbb{P}(X_1 > x_1)} = \lim_{\alpha \uparrow 1} \frac{\mathbb{P}(X_i > x_i)}{\mathbb{P}(X_1 > x_i)} \left(\frac{x_i}{x_1} \right)^{-\theta} = 0$. Et puisque

$$\frac{\mathbb{P}(X_j > tx_1, X_i > x_i)}{\mathbb{P}(X_1 > x_1)} \leq \min \left(\frac{\mathbb{P}(X_j > tx_1)}{\mathbb{P}(X_1 > x_1)}, \frac{\mathbb{P}(X_i > x_i)}{\mathbb{P}(X_1 > x_1)} \right),$$

et par utilisation des bornes de Potter pour la fonction à variation régulière \bar{F}_{X_1} pour appliquer le théorème de la convergence dominée, on a

$$\lim_{\alpha \uparrow 1} \frac{\mathbb{E}[(X_j - x_j)_+ \mathbb{1}_{\{X_i > x_i\}}]}{x_1 \bar{F}_{X_1}(x_1)} = \int_{\beta_j}^{+\infty} \lim_{\alpha \uparrow 1} \frac{\mathbb{P}(X_j > tx_1, X_i > x_i)}{\mathbb{P}(X_1 > x_1)} dt = 0, \quad \forall j \in J_C \setminus \{i\}.$$

Et si $j \in J_0$ alors

$$\begin{aligned} \frac{\mathbb{E}[(X_j - x_j)_+ \mathbb{1}_{\{X_i > x_i\}}]}{x_1 \bar{F}_{X_1}(x_1)} &= \int_{x_j}^{x_1} \frac{\mathbb{P}(X_j > t, X_i > x_i)}{x_1 \mathbb{P}(X_1 > x_1)} dt + \int_1^{+\infty} \frac{\mathbb{P}(X_j > tx_1, X_i > x_i)}{\mathbb{P}(X_1 > x_1)} dt \\ &\leq \left(1 - \frac{x_j}{x_1} \right) \frac{\mathbb{P}(X_i > x_i)}{\mathbb{P}(X_1 > x_1)} + \int_1^{+\infty} \frac{\mathbb{P}(X_j > tx_1, X_i > x_i)}{\mathbb{P}(X_1 > x_1)} dt, \end{aligned}$$

donc

$$\lim_{\alpha \uparrow 1} \frac{\mathbb{E}[(X_j - x_j)_+ \mathbb{1}_{\{X_i > x_i\}}]}{x_1 \bar{F}_{X_1}(x_1)} = 0, \quad \forall j \in J_0,$$

car $\lim_{\alpha \uparrow 1} \frac{\mathbb{P}(X_i > x_i)}{\mathbb{P}(X_1 > x_1)} = 0$ et $\lim_{\alpha \uparrow 1} \int_1^{+\infty} \frac{\mathbb{P}(X_j > tx_1, X_i > x_i)}{\mathbb{P}(X_1 > x_1)} dt = 0$.

Par passage à la limite ($\alpha \rightarrow 1$) dans la i ^{ème} équation du système d'optimalité (1.9) divisée par $x_1 \bar{F}_{X_1}(x_1)$, on obtient

$$-\eta \beta_i = \eta \left(1 + \sum_{j \in J_C \setminus \{i\}} \beta_j \right),$$

cela est absurde, et par suite $J_C = \emptyset$.

On a donc démontré que $J_0 = \{2, \dots, d\}$ ce que signifie

$$\lim_{\alpha \uparrow 1} \frac{\mathbf{e}_\alpha^i(\mathbf{X})}{\mathbf{e}_\alpha^1(\mathbf{X})} = \beta_i = 0, \quad \forall i \in \{2, \dots, d\}.$$

L'Équation 2.23 permet ensuite de déduire que

$$\eta = \lim_{\alpha \uparrow 1} \frac{1 - \alpha}{\bar{F}_{X_1}(\mathbf{e}_\alpha^1(\mathbf{X}))} = \frac{1}{\theta - 1},$$

et par application du Lemme 2.2.2 d'en déduire que

$$\lim_{\alpha \uparrow 1} \frac{1 - \alpha}{\bar{F}_{X_k}(\mathbf{e}_\alpha^i(\mathbf{X}))} = 0, \quad \forall i \in \{2, \dots, d\}.$$

□

Remarquons que la dominance asymptotique implique l'indépendance asymptotique entre le risque dominant et les autres composantes du vecteur \mathbf{X} . Finalement, le risque dominant se comporte asymptotiquement comme dans le cas univarié, et sa composante dans l'expectile asymptotique vérifie

$$\mathbf{e}_\alpha^1(\mathbf{X}) \stackrel{\alpha \uparrow 1}{\sim} (\theta - 1)^{-\frac{1}{\theta}} \text{VaR}_\alpha(X_1) \stackrel{\alpha \uparrow 1}{\sim} \mathbf{e}_\alpha(X_1),$$

l'équivalence à droite est démontrée dans le cas univarié par Bellini et Di Bernardino (2015) (Proposition 2.3, [2]).

Exemple On considère des distributions de Pareto, $X_i \sim Pa(a_i, b)$, $i = 1, \dots, d$, telles que $a_i > a_1$ pour tout $i \in \{1, \dots, d\}$. La dominance des queues par celle de X_1 est donc vérifiée,

$$\lim_{x \uparrow +\infty} \frac{\bar{F}_{X_i}(x)}{\bar{F}_{X_1}(x)} = \lim_{x \uparrow +\infty} \left(\frac{b}{b+x} \right)^{(a_i - a_1)} = 0, \quad \forall i \in \{2, \dots, d\},$$

et la proposition (2.3.1) est donc applicable.

2.4 Modèle par mélange d'échelle

Dans cette partie, nous considérons un vecteur aléatoire \mathbf{X} construit par mélange commun

$$\mathbf{X} = (X_1, \dots, X_d)^T = (RT_1, \dots, RT_d)^T,$$

où le vecteur aléatoire $(T_1, \dots, T_d)^T$ est indépendant de R , et les distributions marginales T_i sont à supports positifs et vérifient certaines conditions sur l'existence des moments qu'on va préciser à chaque fois en fonction du besoin. La variable aléatoire de mélange R est supposée positive et possède une queue droite à variations régulières $\bar{F}_R \in \text{RV}_{-\theta}$. Si les variables T_i sont indépendantes entre elles, cette construction génère donc une dépendance positive.

Comme $\bar{F}_R \in \text{RV}_{-\theta}$ avec $\theta > 1$, on a $\int_x^{+\infty} \bar{F}_R(t) dt \in \text{RV}_{-\theta+1}$ en utilisant le théorème de Karamata

$$\int_x^{+\infty} \bar{F}_R(r) dr \stackrel{+\infty}{\sim} \frac{x}{\theta - 1} \bar{F}_R(x). \quad (2.24)$$

On obtient alors, pour toute constante $c > 0$

$$\int_{cx}^{+\infty} \bar{F}_R(r) dr \stackrel{+\infty}{\sim} \frac{xc^{-\theta+1}}{\theta - 1} \bar{F}_R(x). \quad (2.25)$$

Cette équivalence est très utile pour les calculs des fonctions Stop-Loss.

Proposition 2.4.1. *Si il existe $0 < \delta < 1$ telle que les variables aléatoires T_i , $i = 1, \dots, d$ admettent toutes des moments d'ordre $\theta + \delta$, strictement positifs, $0 < \mathbb{E}[T_i^{\theta+\delta}] < +\infty$, alors le vecteur aléatoire $\mathbf{X} = (X_1, \dots, X_d)^T = (RT_1, \dots, RT_d)^T$ satisfait les hypothèses H1 et H2.*

Preuve Pour tout $i \in \{1, \dots, d\}$

$$\frac{\bar{F}_{X_i}(x)}{\bar{F}_R(x)} = \int_{\mathbb{R}_+} \frac{\mathbb{P}(Rt \geq x)}{\bar{F}_R(x)} f_{T_i}(t) dt.$$

Pour tout $t > 0$ on a par utilisation des Bornes de Potter (2.1.7) pour la fonction à variations régulières \bar{F}_R , il existe $x^0(\delta)$ tel que pour tout $x \geq x^0(\delta)$

$$\frac{\mathbb{P}(Rt \geq x)}{\bar{F}_R(x)} \leq (1 + \delta) \max(t^{\theta-\delta}, t^{\theta+\delta}),$$

et puisque la fonction $t \rightarrow t^{\theta+\delta} f_{T_i}(t)$ est intégrale, on peut appliquer le théorème de la convergence dominée pour obtenir

$$\begin{aligned} \lim_{x \rightarrow +\infty} \frac{\bar{F}_{X_i}(x)}{\bar{F}_R(x)} &= \int_{\mathbb{R}_+} \lim_{x \rightarrow +\infty} \frac{\mathbb{P}(Rt \geq x)}{\bar{F}_R(x)} f_{T_i}(t) dt \\ &= \int_{\mathbb{R}_+} t^\theta f_{T_i}(t) dt = \mathbb{E}[T_i^\theta], \end{aligned}$$

et on en déduit que pour tout $i \in \{1, \dots, d\}$, $\bar{F}_{X_i} \in \text{RV}_{-\theta}$, l'hypothèse H1 est donc vérifiée. L'hypothèse H2 est également vérifiée car

$$\lim_{x \rightarrow +\infty} \frac{\bar{F}_{X_i}(x)}{\bar{F}_{X_1}(x)} = \frac{\mathbb{E}[T_i^\theta]}{\mathbb{E}[T_1^\theta]} > 0.$$

□

Un exemple de l'application de cette proposition est le modèle mélange commun exponentiel-Gamma. En effet, on sait que cette construction permet d'obtenir la copule de survie de Clayton (voir Nelsen (2007) [16]), qui possède une fonction de dépendance de queue gauche. On peut donc appliquer le Théorème 2.1.13 pour en déduire que ce modèle génère une distribution multivariée à queue à variations régulières, puis pour déterminer la composition de l'expectile multivarié asymptotique à l'aide de la Proposition 2.2.5)

Proposition 2.4.2 (Expectile multivarié asymptotique). *Soit $\mathbf{X} = (X_1, \dots, X_d) = (RT_1, \dots, RT_d)$ un vecteur aléatoire construit par modèle de mélange d'échelle commun, tel que il existe $0 < \delta < 1$ tel que les variables aléatoires $T_i, i = 1, \dots, d$ admettent toutes des moments d'ordre $\theta + \delta$, strictement positifs, $0 < \mathbb{E}[T_i^{\theta+\delta}] < +\infty$. Tout vecteur d'adhérence $(\eta, \beta_2, \dots, \beta_d)$ du vecteur $\left(\frac{1-\alpha}{\bar{F}_{X_1}(\mathbf{e}_\alpha^1(\mathbf{X}))}, \frac{\mathbf{e}_\alpha^2(\mathbf{X})}{\mathbf{e}_\alpha^1(\mathbf{X})}, \dots, \frac{\mathbf{e}_\alpha^d(\mathbf{X})}{\mathbf{e}_\alpha^1(\mathbf{X})} \right)$ vérifie le système d'équations suivant*

$$\eta \beta_k^{\theta-1} \sum_{i=1}^d \beta_i = \frac{[T_k^\theta]}{\theta-1} + \sum_{i=1, i \neq k}^d A_{ik}, \quad \forall k \in \{1, \dots, d\}, \quad (2.26)$$

avec

$$A_{ik} = \mathbb{E} \left[\left(\frac{T_i}{T_k} - \frac{\beta_i}{\beta_k} \right)_+ T_k^\theta \right] + \frac{\left(\frac{\beta_i}{\beta_k} \right)^{-\theta+1} \mathbb{E} \left[T_i^\theta \bar{F}_{T_k} \left(\frac{\beta_k}{\beta_i} T_i \right) \right] + \mathbb{E} \left[T_k^{\theta-1} \left(T_i - T_k \frac{\beta_i}{\beta_k} \right)_+ + \frac{\beta_i}{\beta_k} T_k^\theta \bar{F}_{T_i} \left(\frac{\beta_i}{\beta_k} T_k \right) \right]}{\theta-1}.$$

Preuve Par application de la Proposition 2.2.5, tout vecteur d'adhérence $(\eta, \beta_2, \dots, \beta_d)$ du vecteur $\left(\frac{1-\alpha}{\bar{F}_{X_1}(\mathbf{e}_\alpha^1(\mathbf{X}))}, \frac{\mathbf{e}_\alpha^2(\mathbf{X})}{\mathbf{e}_\alpha^1(\mathbf{X})}, \dots, \frac{\mathbf{e}_\alpha^d(\mathbf{X})}{\mathbf{e}_\alpha^1(\mathbf{X})} \right)$ vérifie le système d'équations suivant

$$\frac{1}{\theta-1} - \eta \frac{(\beta_k)^\theta}{c_k} = - \sum_{i=1, i \neq k}^d \left(\int_{\frac{\beta_i}{\beta_k}}^{+\infty} \lambda_U^{ik} \left(\frac{c_i}{c_k} t^{-\theta}, 1 \right) dt - \eta \frac{\beta_k^{\theta-1}}{c_k} \beta_i \right), \quad \forall k \in \{1, \dots, d\},$$

équivalent à

$$\eta\beta_k^{\theta-1} \sum_{i=1}^d \beta_i = \frac{c_k}{\theta-1} + c_k \sum_{i=1, i \neq k}^d \int_{\frac{\beta_i}{\beta_k}}^{+\infty} \lambda_U^{ik} \left(\frac{c_i}{c_k} t^{-\theta}, 1 \right) dt, \forall k \in \{1, \dots, d\}.$$

Et puisque $c_k = \mathbb{E}[T_k^\theta] \forall k \in \{1, \dots, d\}$, alors il suffit donc de calculer explicitement les termes

$$\int_{\frac{\beta_i}{\beta_j}}^{+\infty} \lambda_U^{ij} \left(\frac{c_i}{c_j} t^{-\theta}, 1 \right) dt,$$

pour tout $(i, j) \in \{1, \dots, d\}^2$ dans le cas du modèle par mélange d'échelle pour obtenir (2.26). Dans notre cas, il est plus simple de calculer ces termes directement comme limite

$$\int_{\frac{\beta_i}{\beta_j}}^{+\infty} \lambda_U^{ij} \left(\frac{c_i}{c_j} t^{-\theta}, 1 \right) dt = \lim_{\alpha \uparrow 1} \frac{\mathbb{E}[(X_i - x_i)_+ \mathbb{1}_{\{X_j > x_j\}}]}{x_j \bar{F}_{X_j}(x_j)}, \quad \forall (i, j) \in \{1, \dots, d\}^2.$$

En effet,

$$\begin{aligned} \lim_{\alpha \uparrow 1} \frac{\mathbb{E}[(X_i - x_i)_+ \mathbb{1}_{\{X_j > x_j\}}]}{x_j \bar{F}_{X_j}(x_j)} &= \lim_{\alpha \rightarrow 1} \int_{\frac{x_i}{x_j}}^{+\infty} \frac{\mathbb{P}(X_i > tx_j, X_j > x_j)}{\mathbb{P}(X_j > x_j)} dt \\ &= \lim_{\alpha \uparrow 1} \frac{1}{\mathbb{P}(X_j > x_j)} \int_{\mathbb{R}_+^2} \int_{\frac{x_i}{x_j}}^{+\infty} \mathbb{P}\left(R > \frac{tx_j}{t_1}, R > \frac{x_j}{t_2}\right) dt dF_{T_i}(t_1) dF_{T_j}(t_2) \\ &= \lim_{\alpha \uparrow 1} \frac{1}{\mathbb{P}(X_j > x_j)} \int_{\mathbb{R}_+^2} \int_{\frac{x_i}{x_j}}^{+\infty} \mathbb{P}\left(R > \max\left\{\frac{tx_j}{t_1}, \frac{x_j}{t_2}\right\}\right) dF_{T_i}(t_1) dF_{T_j}(t_2). \end{aligned}$$

On peut écrire

$$\int_{\frac{x_i}{x_j}}^{+\infty} \mathbb{P}\left(R > \max\left\{\frac{tx_j}{t_1}, \frac{x_j}{t_2}\right\}\right) dt = \int_{\max\left\{\frac{x_i}{x_j}, \frac{t_1}{t_2}\right\}}^{+\infty} \mathbb{P}\left(R > \frac{tx_j}{t_1}\right) dt + \int_{\frac{x_i}{x_j}}^{\max\left\{\frac{x_i}{x_j}, \frac{t_1}{t_2}\right\}} \mathbb{P}\left(R > \frac{x_j}{t_2}\right) dt.$$

On commence par la deuxième terme de cette somme

$$\begin{aligned} A_1 &= \lim_{\alpha \uparrow 1} \frac{1}{\mathbb{P}(X_j > x_j)} \int_{\mathbb{R}_+^2} \int_{\frac{x_i}{x_j}}^{\max\left\{\frac{x_i}{x_j}, \frac{t_1}{t_2}\right\}} \mathbb{P}\left(R > \frac{x_j}{t_2}\right) dt dF_{T_i}(t_1) dF_{T_j}(t_2) \\ &= \lim_{\alpha \uparrow 1} \frac{1}{\mathbb{E}[T_j^\theta] \mathbb{P}(R > x_j)} \int_{\mathbb{R}_+^2} \left(\frac{t_1}{t_2} - \frac{x_i}{x_j}\right)_+ \mathbb{P}\left(R > \frac{x_j}{t_2}\right) dF_{T_i}(t_1) dF_{T_j}(t_2) \\ &= \lim_{\alpha \uparrow 1} \frac{1}{\mathbb{E}[T_j^\theta]} \int_{\mathbb{R}_+^2} \left(\frac{t_1}{t_2} - \frac{x_i}{x_j}\right)_+ \frac{\mathbb{P}\left(R > \frac{x_j}{t_2}\right)}{\mathbb{P}(R > x_j)} dF_{T_i}(t_1) dF_{T_j}(t_2) \\ &= \frac{1}{\mathbb{E}[T_j^\theta]} \int_{\mathbb{R}_+^2} \left(\frac{t_1}{t_2} - \frac{x_i}{x_j}\right)_+ t_2^\theta dF_{T_i}(t_1) dF_{T_j}(t_2) \\ &= \frac{\mathbb{E}\left[\left(\frac{T_i}{T_j} - \frac{\beta_i}{\beta_j}\right)_+ T_j^\theta\right]}{\mathbb{E}[T_j^\theta]}. \end{aligned}$$

L'interversion de la limite et l'intégrale dans l'avant dernière ligne est faite par application du théorème de la convergence dominée, applicable par utilisation des bornes de Potter pour la

fonction à variations régulières \bar{F}_R .

Le deuxième terme peut s'écrire

$$\begin{aligned} A_2 &= \lim_{\alpha \uparrow 1} \frac{1}{\mathbb{P}(X_j > x_j)} \int_{\mathbb{R}_+^2} \int_{\max\left\{\frac{x_i}{x_j}, \frac{t_1}{t_2}\right\}}^{+\infty} \mathbb{P}\left(R > \frac{tx_j}{t_1}\right) dt dF_{T_i}(t_1) dF_{T_j}(t_2) \\ &= \lim_{\alpha \uparrow 1} \int_{\mathbb{R}_+^2} \int_{\max\left\{\frac{x_i}{x_j}, \frac{t_1}{t_2}\right\}}^{+\infty} \frac{\mathbb{P}\left(R > \frac{tx_j}{t_1}\right)}{\mathbb{P}(X_j > x_j)} dt dF_{T_i}(t_1) dF_{T_j}(t_2). \end{aligned}$$

L'intégration est construite à la base sous forme d'intégrale d'une probabilité, donc Lipschitzienne en fonction des bornes, le passage à la limite au niveau des bornes ne pose alors aucun problème. D'un côté, le théorème de convergence dominée s'applique par utilisation des bornes de Potter des fonctions à variations régulières.

Par application du théorème de Karamata 2.24 à la fonction de survie de R on obtient

$$\begin{aligned} \int_{\max\left\{\frac{x_i}{x_j}, \frac{t_1}{t_2}\right\}}^{+\infty} \mathbb{P}\left(R > \frac{tx_j}{t_1}\right) dt &= \frac{t_1}{x_j} \int_{\max\left\{\frac{x_i}{t_1}, \frac{x_j}{t_2}\right\}}^{+\infty} \mathbb{P}(R > t) dt \\ &\underset{\alpha \uparrow 1}{\sim} \frac{t_1}{x_j} \frac{1}{\theta - 1} \max\left\{\frac{x_i}{t_1}, \frac{x_j}{t_2}\right\} \mathbb{P}\left(R > \max\left\{\frac{x_i}{t_1}, \frac{x_j}{t_2}\right\}\right), \end{aligned}$$

car $\lim_{\alpha \uparrow 1} \frac{1}{\max\left\{\frac{x_i}{x_j}, \frac{t_1}{t_2}\right\}} = +\infty$. Donc,

$$\begin{aligned} \int_{\max\left\{\frac{x_i}{x_j}, \frac{t_1}{t_2}\right\}}^{+\infty} \frac{\mathbb{P}\left(R > \frac{tx_j}{t_1}\right)}{\mathbb{P}(X_j > x_j)} dt &\underset{\alpha \uparrow 1}{\sim} \frac{t_1}{x_j} \frac{1}{\theta - 1} \max\left\{\frac{x_i}{t_1}, \frac{x_j}{t_2}\right\} \frac{\mathbb{P}\left(R > \max\left\{\frac{x_i}{t_1}, \frac{x_j}{t_2}\right\}\right)}{\mathbb{P}(X_j > x_j)} \\ &\underset{\alpha \uparrow 1}{\sim} \frac{t_1}{x_j} \frac{1}{\theta - 1} \max\left\{\frac{x_i}{t_1}, \frac{x_j}{t_2}\right\} \frac{1}{\mathbb{E}[T_j^\theta]} \frac{\mathbb{P}\left(R > \max\left\{\frac{x_i}{t_1}, \frac{x_j}{t_2}\right\}\right)}{\mathbb{P}(R > x_j)} \\ &= \frac{t_1}{x_j} \frac{1}{\theta - 1} \max\left\{\frac{x_i}{t_1}, \frac{x_j}{t_2}\right\} \frac{1}{\mathbb{E}[T_j^\theta]} \min\left\{\frac{\mathbb{P}\left(R > \frac{x_i}{t_1}\right)}{\mathbb{P}(R > x_j)}, \frac{\mathbb{P}\left(R > \frac{x_j}{t_2}\right)}{\mathbb{P}(R > x_j)}\right\} \\ &\underset{\alpha \uparrow 1}{\sim} \frac{t_1}{x_j} \frac{1}{\theta - 1} \max\left\{\frac{x_i}{t_1}, \frac{x_j}{t_2}\right\} \frac{1}{\mathbb{E}[T_j^\theta]} \min\left\{t_1^\theta \left(\frac{x_i}{x_j}\right)^{-\theta}, t_2^\theta\right\}. \end{aligned}$$

On en déduit que

$$\begin{aligned} A_2 &= \int_{\mathbb{R}_+} \int_0^{t_2 \frac{\beta_i}{\beta_j}} \frac{1}{\theta - 1} \frac{1}{\mathbb{E}[T_j^\theta]} \left(\frac{\beta_i}{\beta_j}\right)^{-\theta+1} t_1^\theta dF_{T_i}(t_1) dF_{T_j}(t_2) + \int_{\mathbb{R}_+} \int_{t_2 \frac{\beta_i}{\beta_j}}^{+\infty} \frac{1}{\theta - 1} \frac{1}{\mathbb{E}[T_j^\theta]} t_1 t_2^{\theta-1} dF_{T_i}(t_1) dF_{T_j}(t_2) \\ &= \frac{1}{\theta - 1} \frac{1}{\mathbb{E}[T_j^\theta]} \left(\frac{\beta_i}{\beta_j}\right)^{-\theta+1} \int_{\mathbb{R}_+} \int_{t_1 \frac{\beta_i}{\beta_j}}^{+\infty} t_1^\theta dF_{T_i}(t_1) dF_{T_j}(t_2) + \frac{1}{\theta - 1} \frac{1}{\mathbb{E}[T_j^\theta]} \int_{\mathbb{R}_+} \int_{t_2 \frac{\beta_i}{\beta_j}}^{+\infty} t_1 t_2^{\theta-1} dF_{T_i}(t_1) dF_{T_j}(t_2) \\ &= \frac{1}{\theta - 1} \frac{1}{\mathbb{E}[T_j^\theta]} \left(\left(\frac{\beta_i}{\beta_j}\right)^{-\theta+1} \mathbb{E}\left[T_i^\theta \bar{F}_{T_j}\left(\frac{\beta_j}{\beta_i} T_i\right)\right] + \mathbb{E}\left[T_j^{\theta-1} \left(T_i - T_j \frac{\beta_i}{\beta_j}\right)_+ + \frac{\beta_i}{\beta_j} T_j^\theta \bar{F}_{T_i}\left(\frac{\beta_i}{\beta_j} T_j\right)\right]. \end{aligned}$$

La somme des deux termes A_1 et A_2 permet d'obtenir le résultat annoncé dans la proposition.

□

2.5 Estimation de l'expectile multivarié asymptotique

On construit dans cette partie des estimateurs de l'expectile multivarié asymptotique sous les hypothèses H1 et H2, à l'aide des quantiles univariés. Les estimateurs proposés concernent uniquement les cas d'indépendance asymptotique et de dépendance parfaite. Nous présentons tout d'abord l'idée générale de la méthodologie, puis nous nous appuyons sur les statistiques des extrêmes pour construire ces estimateurs.

Proposition 2.5.1 (Idée de l'estimation). *Sous les hypothèses H1 et H2, ainsi que celle de l'existence des limites, et en utilisant les notations des parties précédentes,*

$$\mathbf{e}_\alpha(\mathbf{X}) \underset{\alpha \rightarrow 1}{\sim} \text{VaR}_\alpha(X_1) \eta^{\frac{1}{\theta}} (1, \beta_2, \dots, \beta_d)^T.$$

Preuve Une fois les limites $(\eta, \beta_2, \dots, \beta_d)$ déterminées en fonction du modèle utilisé, nous savons que

$$\mathbf{e}_\alpha(\mathbf{X})^T \underset{\alpha \rightarrow 1}{\sim} x_1 (1, \beta_2, \dots, \beta_d)^T.$$

Et puisque $\lim_{\alpha \rightarrow 1} \frac{1-\alpha}{F_{X_1}(x_1)} = \eta$, alors

$$x_1 = F_{X_1}^{-1} \left(1 - \frac{1-\alpha + o(1)}{\eta} \right).$$

La fonction $F_{X_1}^{-1}(1 - \cdot)$ est à variation régulière en 0, car par application du théorème 1.5.12 de Bingham et al. (1989) [3], on a $F_{X_1}^{-1}(1 - \cdot) \in \text{RV}_{-\frac{1}{\theta}}(0)$. Cela permet de déduire que

$$x_1 \underset{\alpha \rightarrow 1}{\sim} F_{X_1}^{-1}(\alpha) \left(\frac{1}{\eta} \right)^{-\frac{1}{\theta}},$$

ce qui permet d'obtenir l'équivalence de la proposition. \square

Nous nous basons dans la suite sur la Proposition 2.5.1 pour proposer des estimations de l'expectile multivarié extrême pour des cas précédemment étudiés. Pour un vecteur aléatoire $\mathbf{X} = (X_1, \dots, X_d)^T$, on considère un échantillon indépendant de taille n noté $\{\mathbf{X}_1, \dots, \mathbf{X}_n\}$ avec $\mathbf{X}_i = (X_{1,i}, \dots, X_{d,i})^T$ pour tout $i \in \{1, \dots, n\}$. Pour chaque risque X_i , on note $X_{i,1,n} \leq X_{i,2,n} \leq \dots \leq X_{i,n,n}$ l'échantillon ordonné. La dépendance entre les composantes du vecteur \mathbf{X} est prise en compte dans l'échantillon. Nous avons obtenus précédemment des résultats asymptotiques qui permettent de relier les composantes de l'expectile extrême aux quantiles extrêmes, on peut donc obtenir des estimateurs de l'expectile en fonction des estimateurs des quantiles.

Dans l'introduction de la thèse, nous avons présenté le théorème des valeurs extrêmes introduit par Gnedenko (1943) [8] et qui permet de caractériser la distribution limite du maximum $X_{i,n,n}$. Nous avons aussi présenté les outils qui nous permettront d'estimer l'indices des variations régulières des distributions marginales dans l'environnement MRV étudié dans les sections précédentes. Dans cette section, nous utilisons les statistiques des extrêmes pour construire des estimateurs des expectiles multivariés dans les cas de l'indépendance asymptotique et la comonotonie du vecteur des risques.

2.5.1 Estimateur des expectiles dans le cas Fréchet

Le cas de Fréchet correspond au modèle étudié dans la Section 2.2.

Pour estimer l'expectile multivarié dans le cas de l'indépendance asymptotique, on peut se baser sur les Propositions 2.2.8 et 2.5.1. Nous avons démontré dans la Proposition 2.2.8 que pour tout $i \in \{1, \dots, d\}$

$$\beta_i = c_i^{\frac{1}{\theta-1}},$$

et

$$\lim_{\alpha \rightarrow 1} \frac{1-\alpha}{F_{X_i}(\mathbf{e}_\alpha^i(\mathbf{X}))} = \frac{c_i^{\frac{1}{\theta-1}}}{(\theta-1) \left(1 + \sum_{j=2}^d c_j^{\frac{1}{\theta-1}} \right)}.$$

En appliquant la Proposition 2.5.1, on déduit que

$$\mathbf{e}_\alpha(\mathbf{X}) \underset{\alpha \uparrow 1}{\sim} \text{VaR}_\alpha(X_1) (\theta-1)^{-\frac{1}{\theta}} \left(1 + \sum_{i=2}^d c_i^{\frac{1}{\theta-1}} \right)^{-\frac{1}{\theta}} (1, c_2^{\frac{1}{\theta-1}}, \dots, c_d^{\frac{1}{\theta-1}})^T.$$

L'expectile asymptotique peut donc être estimé à l'aide de l'estimation du quantile univarié de X_1 et les paramètres de l'équivalence des queues.

De même, et dans le cas de la comonotonie des risques, on peut utiliser le résultat de la Proposition 2.2.7

$$\lim_{\alpha \rightarrow 1} \frac{1 - \alpha}{\bar{F}_{X_i}(\mathbf{e}_\alpha^i(\mathbf{X}))} = \frac{1}{\theta - 1} \quad \text{et} \quad \beta_i = c_i^{1/\theta}, \quad \forall i \in \{1, \dots, d\},$$

et par application de la Proposition 2.5.1 on obtient

$$\mathbf{e}_\alpha(\mathbf{X}) \stackrel{\alpha \uparrow 1}{\sim} \text{VaR}_\alpha(X_1) (\theta - 1)^{-\frac{1}{\theta}} (1, c_2^{\frac{1}{\theta}}, \dots, c_d^{\frac{1}{\theta}})^T.$$

Nous pouvons donc utiliser les statistiques des extrêmes pour proposer des estimateurs à l'expectile multivarié extrême dans les deux cas précédant. Pour cela, nous utilisons, par exemple, l'estimateur de Hill, valable dans le cas Fréchet, pour estimer l'indice des variations régulières, commun à toutes les distributions

$$\hat{\theta} = \frac{1}{\hat{\gamma}_{Hill}} = \frac{1}{\frac{1}{k(n)} \sum_{\ell=1}^{k(n)} [\log(X_{1,n-\ell+1,n}) - \log(X_{1,n-\ell,n})]}.$$

Pour estimer les indices d'équivalences des queues c_i , on peut utiliser l'estimation des paramètres de la loi des valeurs extrêmes. Puisque pour tout $i \in \{1, \dots, d\}$, $\mathbb{P}(X_{i,n,n} \leq x) = F_{X_i}^n(x)$, on peut approcher asymptotiquement $F_{X_i}(x)$ par $H_\gamma^{1/n} \left(\frac{x - a_{i,n}}{b_{i,n}} \right)$ pour avoir

$$\log(1 - \bar{F}_{X_i}(x)) \simeq \frac{1}{n} \log H_\gamma \left(\frac{x - a_{i,n}}{b_{i,n}} \right),$$

et en appliquant un développement limité d'ordre 1 à la fonction $\log(1 + x)$ au voisinage de 0 on obtient une approximation de la fonction de survie sous la forme suivante

$$\bar{F}_{X_i}(x) \simeq -\frac{1}{n} \log H_\gamma \left(\frac{x - a_{i,n}}{b_{i,n}} \right) = \frac{1}{n} \left(1 + \frac{1}{\theta} \left(\frac{x - a_n}{b_n} \right) \right)^{-\theta}.$$

Pour des échantillons de même taille, on peut donc estimer les c_i par

$$\hat{c}_i = \left(\frac{\hat{b}_i}{\hat{b}_1} \right)^{\hat{\theta}} = \left(\frac{\hat{\mu}_{i,0} - 2\hat{\mu}_{i,1}}{\hat{\mu}_{1,0} - 2\hat{\mu}_{1,1}} \right)^{\hat{\theta}}, \quad (2.27)$$

où $\hat{\mu}_{i,m}$ est l'estimateur des moments pondérés d'ordre m correspondant à l'échantillon des extrema de X_i , de taille $k(n)$ commune à toutes les variables et tous les estimateurs. Il s'agit d'un estimateur simple construit par substitution (plug-in) à l'aide de fonctions simples et régulières, la normalité asymptotique est donc héritée par le nouvel estimateur et la vitesse de convergence est obtenue à l'aide de la méthode Delta.

L'estimateur (2.27) est directement lié à l'estimation des suites de normalisation dont la convergence n'est pas toujours assurée. La structure de dépendance entre les variables aléatoires X_i , $i = 1, \dots, d$ crée de la dépendance entre ces suites même en cas d'indépendance asymptotique, l'estimateur sera donc très sensible à cette dépendance.

Pour pallier ce problème, on peut utiliser la méthode GPD pour construire des estimateurs non paramétriques des c_i . L'idée sera d'utiliser l'approximation (1) utilisée dans la construction de l'estimateur de Hill pour les distributions du domaine de Fréchet au-delà d'un seuil u

$$\bar{F}(x) \simeq \bar{F}(u) \left(\frac{x}{u} \right)^{-\frac{1}{\gamma}},$$

en fixant la même probabilité de dépassement des seuils pour les X_i

$$\bar{F}_{X_i}(u_i) = \bar{F}_{X_1}(u_1) = \frac{k}{n}, \quad \forall i \in \{1, \dots, d\},$$

le seuil de dépassement pour chaque X_i sera donc estimé par $X_{i,n-k+1,n}$. Sous les hypothèses $k \rightarrow \infty$ et $k/n \rightarrow 0$ quand $n \rightarrow \infty$, nous pouvons avec cette approche approximer c_i pour tout $i \in \{2, \dots, d\}$ et $x > \max_{i \in \{1, \dots, d\}} (u_i)$ comme suit

$$c_i \simeq \frac{\bar{F}_{X_i}(x)}{\bar{F}_{X_1}(x)} \simeq \frac{\bar{F}_{X_i}(u_i) \left(\frac{x}{u_i}\right)^{-\frac{1}{\gamma}}}{\bar{F}_{X_1}(u_1) \left(\frac{x}{u_1}\right)^{-\frac{1}{\gamma}}} = \left(\frac{u_i}{u_1}\right)^{\frac{1}{\gamma}},$$

on peut donc proposer d'estimer c_i par

$$\hat{c}_i = \left(\frac{X_{i,n-k+1,n}}{X_{1,n-k+1,n}} \right)^{\frac{1}{\hat{\gamma}_H(k)}}, \quad (2.28)$$

où $\hat{\gamma}_H(k)$ est l'estimateur de Hill de l'indice des valeurs extrêmes construit à l'aide de k plus grandes observations.

Nous décrivons ci-dessous la preuve de la cohérence de l'estimateur (2.28) sous les mêmes hypothèses utilisées pour la convergence de l'estimateur de Hill. Les équivalences et les limites peuvent nécessiter une attention plus technique, surtout pour prouver que

$$\frac{F_1(X_{1,n-k,n})}{F_i(X_{i,n-k,n})} \rightarrow 1 \text{ p.s.}$$

et

$$F_i(x) \sim F_i(u_i)(x/u_i)^{1/\gamma}.$$

Le choix de la même probabilité de dépassement de seuil k/n implique

$$\frac{\bar{F}_{X_i}(X_{i,n-k+1,n})}{\bar{F}_{X_1}(X_{1,n-k+1,n})} = \frac{\bar{F}_{X_i}(X_{i,n-k+1,n})}{\bar{F}_{X_1}(X_{i,n-k+1,n})} \frac{\bar{F}_{X_1}(X_{i,n-k+1,n})}{\bar{F}_{X_1}(X_{1,n-k+1,n})} \underset{p.s.}{\overset{n \rightarrow \infty}{\sim}} c_i \left(\frac{X_{i,n-k+1,n}}{X_{1,n-k+1,n}} \right)^{-\frac{1}{\gamma}},$$

car $\bar{F}_{X_1} \in \text{RV}_{-\frac{1}{\gamma}}$, et puisque $0 < \gamma < 1$ sous l'hypothèse H1, on peut déduire que

$$\lim_{n \rightarrow \infty} \left(\frac{X_{i,n-k+1,n}}{X_{1,n-k+1,n}} \right)^{\frac{1}{\gamma}} = c_i \text{ p.s.},$$

pour tout $i \in \{2, \dots, d\}$. L'estimateur de Hill est convergent, on peut donc conclure (composition exponentielle à gauche ou *continuous map theorem*) que

$$\lim_{n \rightarrow \infty} \hat{c}_i(k) = \lim_{n \rightarrow \infty} \left(\frac{X_{i,n-k+1,n}}{X_{1,n-k+1,n}} \right)^{\frac{1}{\hat{\gamma}_H(k)}} = \lim_{n \rightarrow \infty} \left(\frac{X_{i,n-k+1,n}}{X_{1,n-k+1,n}} \right)^{\frac{1}{\gamma}} = c_i \text{ p.s.},$$

pour tout $i \in \{2, \dots, d\}$.

Pour montrer la normalité asymptotique, nous pourrions utiliser la version multidimensionnelle de la méthode Delta. La normalité asymptotique de $X_{n-k+1,n}$ comme estimateur du quantile extrême d'un niveau α avec $k = \lfloor \alpha n \rfloor$ est donnée par le Théorème 2.5.2 suivant.

Théorème 2.5.2 (Théorème 2.2.1, De Haan et Ferreira (2007) [5]). *Sous les hypothèses $k \rightarrow \infty$ et $k/n \rightarrow 0$ quand $n \rightarrow \infty$, si*

$$\lim_{x \uparrow x_F^-} \frac{\bar{F}(x)F''(x)}{F'(x)^2} = -\gamma - 1,$$

alors

$$\sqrt{k} \left(\frac{X_{n-k+1,n} - U(\frac{n}{k})}{\frac{n}{k} U'(\frac{n}{k})} \right) \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1).$$

Il faut donc calculer la matrice de corrélation du vecteur $(X_{1,n-k+1,n}, X_{i,n-k+1,n}, \hat{\gamma}_H(k))$ pour l'application de la méthode Delta.

Pour estimer le quantile extrême, on peut également utiliser l'estimateur de Weissman (1978) [19], pour tout $i \in \{1, \dots, d\}$

$$\widehat{\text{VaR}}_\alpha(X_i) = \hat{q}_{\text{Weissman}}^i(1 - \alpha) = X_{i,n-k(n)+1,n} \left(\frac{k(n)}{(1 - \alpha)n} \right)^{\hat{\gamma}_{\text{Hill}}}.$$

Les propriétés de l'estimateur de Weissman sont présentées dans l'ouvrage de Embrechts et al. (1997) [7]. Sa normalité asymptotique découle directement de celle de l'estimateur de Hill et du théorème (2.5.2).

On peut maintenant déduire des estimateurs de l'expectile multivarié extrême, dans les cas de l'indépendance asymptotique et la comonotonie, à l'aide des précédents estimateurs.

Définition 2.5.3 (Estimateur expectile multivarié, Cas Fréchet, Indépendance Asymptotique). Sous les hypothèses du modèle Fréchet à variations régulières multivariées, en cas de l'indépendance asymptotique deux à deux entre les composantes du vecteur aléatoire $\mathbf{X} = (X_1, \dots, X_d)^T$, on définit l'estimateur du L_1 -expectile par la manière suivante

$$\hat{\mathbf{e}}_\alpha^\perp(\mathbf{X}) = X_{1,n-k(n)+1,n} \left(\frac{k(n)}{(1 - \alpha)n} \right)^{\hat{\gamma}} \left(\frac{\hat{\gamma}}{1 - \hat{\gamma}} \right)^{\hat{\gamma}} \left(\frac{1}{1 + \sum_{\ell=2}^d \hat{c}_\ell^{\frac{\hat{\gamma}}{1 - \hat{\gamma}}}} \right)^{\hat{\gamma}} (1, \hat{c}_2^{\frac{\hat{\gamma}}{1 - \hat{\gamma}}}, \dots, \hat{c}_d^{\frac{\hat{\gamma}}{1 - \hat{\gamma}}})^T.$$

L'estimateur de la Définition 2.5.3 est obtenu par application de la Proposition 2.2.8, et utilisation des estimateurs précédemment présentés.

Définition 2.5.4 (Estimateur expectile multivarié, Cas Fréchet comotone). Sous les hypothèses du modèle Fréchet à variations régulières multivariées, si le vecteur aléatoire $\mathbf{X} = (X_1, \dots, X_d)^T$ est comotone, on définit l'estimateur du L_1 -expectile par le vecteur suivant

$$\hat{\mathbf{e}}_\alpha^+(\mathbf{X}) = X_{1,n-k(n)+1,n} \left(\frac{k(n)}{(1 - \alpha)n} \right)^{\hat{\gamma}} \left(\frac{\hat{\gamma}}{1 - \hat{\gamma}} \right)^{\hat{\gamma}} (1, \hat{c}_2^{\hat{\gamma}}, \dots, \hat{c}_d^{\hat{\gamma}})^T.$$

L'estimateur de la Définition 2.5.4 est obtenu par application de la Proposition 2.2.7, et utilisation des estimateurs précédemment présentés.

L'estimateur de l'indice des valeurs extrêmes $\hat{\gamma}$ peut être tout autre estimateur que Hill. Le nombre d'observation $k(n)$ est le même pour l'ensemble des estimateurs utilisés dans l'opération de *plug-in*, la convergence des estimateurs construits découle de celles des estimateurs utilisés, et la normalité asymptotique peut être démontrée à l'aide de la méthode Delta. Une prise en compte des dépendances entre les différents estimateurs est nécessaire, nous avons décidé de consacrer un travail futur à l'étude statistique complète des mesures expectiles multivariées (normalité asymptotique, efficacité,...).

2.5.2 Illustration numérique

Le domaine d'attraction de Fréchet pour les valeurs extrêmes contient les lois usuelles Pareto, Student et Cauchy. Nous pouvons illustrer la convergence des estimateurs proposés à l'aide des exemples issus de ces lois.

Nous considérons deux variables aléatoires de loi Pareto, $X_i \sim Pa(a, b_i)$, $i \in \{1, 2\}$. Les deux distributions possèdent le même paramètre d'échelle a , elles sont donc à queues équivalentes, et le paramètre de leur équivalence c_2 est déterminé explicitement

$$c_2 = \lim_{x \rightarrow +\infty} \frac{\bar{F}_{X_2}(x)}{\bar{F}_{X_1}(x)} = \lim_{x \rightarrow +\infty} \frac{\left(\frac{b_2}{b_2 + x} \right)^a}{\left(\frac{b_1}{b_1 + x} \right)^a} = \left(\frac{b_2}{b_1} \right)^a.$$

La figure 2.1 est une illustration de la convergence de l'estimateur \hat{c}_2 quand $k \rightarrow +\infty$. L'illustration numérique est faite pour $X_i \sim Pa(2, 10 \cdot i)_{i=1,2}$, la valeur exacte de c_2 est égale à 4 dans ce cas.

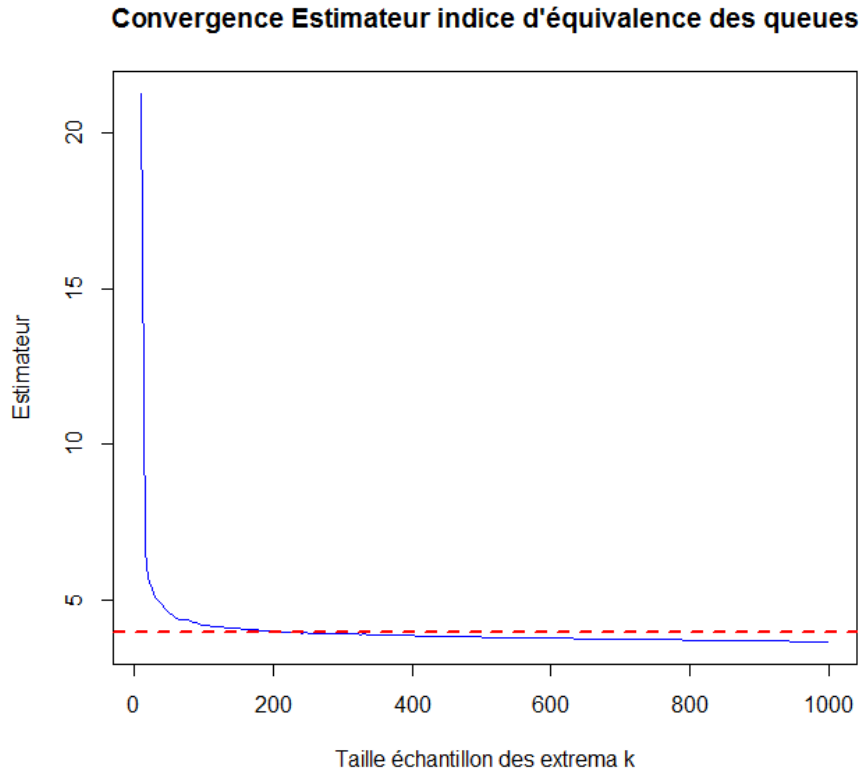


FIGURE 2.1 – Convergence de l'estimateur \hat{c}_2 .

Dans le cas de l'indépendance qui implique l'indépendance asymptotique, on peut déterminer la solution explicite, et par la suite vérifier la convergence des estimateurs. Dans le figure 2.2 nous présentons les résultats obtenus pour différents choix de la suite $k(n)$.

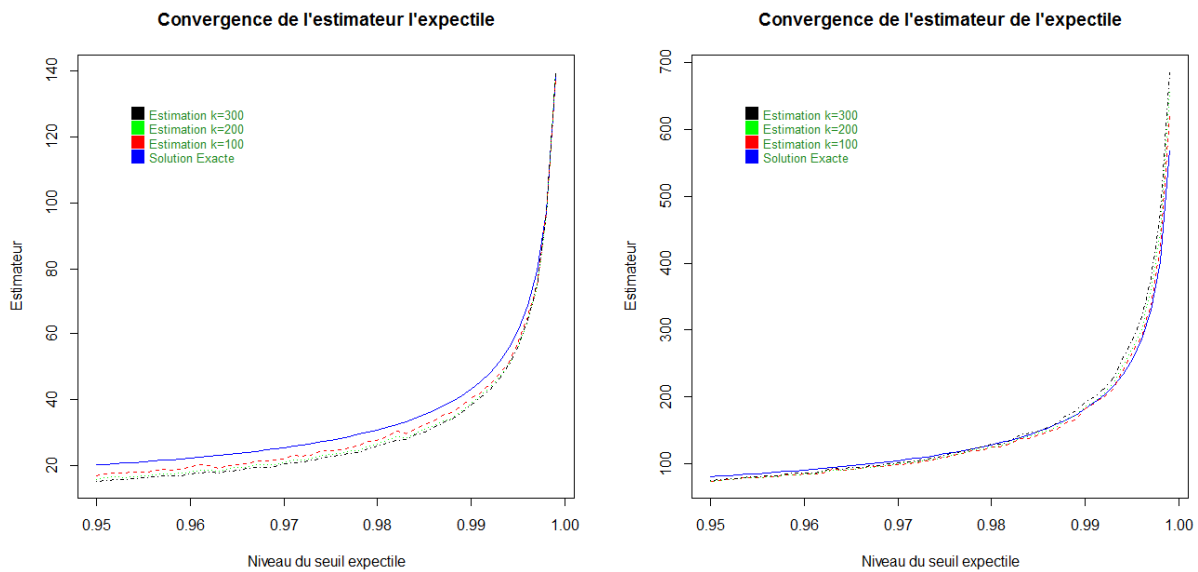


FIGURE 2.2 – Convergence de l'estimateur de l'expectile (gauche X_1 , droite X_2).

Conclusion

En dimension 1 on peut voir l'expectile comme un quantile d'une transformation de la distribution. L'étude asymptotique que nous avons présentée dans un environnement à variations régulières montre que les vecteurs expectiles restent liés asymptotiquement au comportement des queues marginales et leur composition dépend de la nature de la dépendance asymptotique. Nous avons constaté que l'équivalence des queues peut engendrer une équivalence entre les composantes de l'expectile asymptotique.

L'estimation statistique des intégrales des fonctions de dépendance des queues permettra de construire des estimateurs du vecteurs expectile asymptotique. Nous nous sommes limités au cas de l'indépendance asymptotique et de la comonotonie qui ne nécessitent pas l'estimation de la fonction de dépendance des queues.

L'étude du comportement asymptotique reste incomplète sans l'analyse des cas d'appartenance aux autres domaines d'attractions des valeurs extrêmes Gumbel et Weibull. Bien qu'actuellement en cours d'investigation, cette question n'est pas abordée dans cette thèse.

Bibliographie

- [1] B. Basrak. *The sample autocorrelation function of non-linear time series*. PhD thesis, University of Groningen, 2000. 90, 91
- [2] F. Bellini and E. Di Bernardino. Risk management with expectiles. *The European Journal of Finance*, pages 1–20, 2015. 40, 45, 46, 87, 104, 110
- [3] N. Bingham, C. M Goldie, and J. Teugels. *Regular variation*, volume 27. Cambridge university press, 1989. 87, 89, 114
- [4] A. Charpentier and J. Segers. Tails of multivariate archimedean copulas. *Journal of Multivariate Analysis*, 100(7) :1521–1537, 2009. 101
- [5] L. De Haan and A. Ferreira. *Extreme value theory : an introduction*. Springer Science & Business Media, 2007. 87, 90, 116
- [6] L. De Haan and S. Resnick. Limit theory for multivariate sample extremes. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 40(4) :317–337, 1977. 90
- [7] P. Embrechts, C. Klüppelberg, and T. Mikosch. *Modelling extremal events*, volume 33. Springer Science & Business Media, 1997. 14, 87, 117
- [8] B. Gnedenko. Sur la distribution limite du terme maximum d’une serie aleatoire. *Annals of mathematics*, pages 423–453, 1943. 9, 114
- [9] H. Joe, H. Li, and A.K. Nikoloulopoulos. Tail dependence functions and vine copulas. *Journal of Multivariate Analysis*, 101(1) :252–270, 2010. 101
- [10] O. Kallenberg. Random measures, 3rd. *Edition, Akademie Verlag, Berlin, Academic Press, New York*, 1983. 90
- [11] C. Klüppelberg, G.l Kuhn, and L. Peng. Semi-parametric models for the multivariate tail dependence function—the asymptotically dependent case. *Scandinavian Journal of Statistics*, 35(4) :701–718, 2008. 8, 91
- [12] H. Li and Y. Sun. Tail dependence for heavy-tailed scale mixtures of multivariate distributions. *Journal of Applied Probability*, pages 925–937, 2009. 92
- [13] T. Mao and F. Yang. Risk concentration based on expectiles for extreme risks under fgm copula. *Insurance : Mathematics and Economics*, 64 :429–439, 2015. 46, 87
- [14] A.J. McNeil and J. Nešlehová. Multivariate archimedean copulas, d-monotone functions and ℓ_1 -norm symmetric distributions. *The Annals of Statistics*, pages 3059–3097, 2009. 101
- [15] T. Mikosch. Modeling dependence and tails of financial time series. *Extreme Values in Finance, Telecommunications, and the Environment*, pages 185–286, 2003. 89, 91
- [16] R.B. Nelsen. *An introduction to copulas*. Springer, 2007. 70, 77, 92, 111, 157, 161
- [17] S. Resnick. *Heavy-tail phenomena : probabilistic and statistical modeling*. Springer Science & Business Media, 2007. 89, 90
- [18] S.I. Resnick. *Extreme values, regular variation and point processes*. Springer, 2013. 88
- [19] I. Weissman. Estimation of parameters and large quantiles based on the k largest observations. *Journal of the American Statistical Association*, 73(364) :812–815, 1978. 117
- [20] C. Weng and Y. Zhang. Characterization of multivariate heavy-tailed distribution families via copula. *Journal of Multivariate Analysis*, 106 :178–186, 2012. 92

Allocation du capital par minimisation du risque

Chapitre 3

Allocation du capital par minimisation d'indicateurs multivariés de risque

On a capital allocation by minimization of some risk indicators

Ce chapitre est composé de l'article Maume-Deschamps et al. (2016) [22] qui porte le même titre,
publié dans *European Actuarial Journal*.

<http://link.springer.com/article/10.1007/s13385-016-0123-1>

abstract

European insurance sector will soon be faced with the application of the Solvency 2 regulation norms. It will create a real change in the risk management of insurance practices. The ORSA (Own Risk and Solvency Assessment) approach of the second pillar makes the capital allocation an important exercise for all insurers, especially when it comes to groups.

Considering multi-branches firms, a capital allocation has to be based on multivariate risk modeling. Several allocation methods are present in the actuarial literature and insurance practices. In this paper, we focus on a risk allocation method. By minimizing some of the multivariate risk indicators, we study the coherence of the risk allocation using an axiomatic approach. Furthermore, we discuss what can be the best allocation choice for an insurance group.

Keywords : *Multivariate risk indicators, Solvency 2, Solvency Capital Requirement SCR, Own Risk and Solvency Assessment ORSA, dependence modeling, coherence properties, risk theory, optimal capital allocation..*

Introduction

Solvency 2 standards will involve radical changes in the risk management practices of the actuarial sector. They target the strengthening of the control of risk and the minimization of the ruin probabilities. Under this prudential mechanism, the calculation of the economic regulatory capital will be faced with a methodological revolution. Choices of both dependence structures between different risks and aggregation methods are required. In the prescribed Standard Formula, the risk aggregation is derived via correlation matrices that connect families and subfamilies of risks. Once the Solvency Capital Requirement (SCR) is determined, its allocation between the different risk activity branches becomes the new operational challenge.

Capital allocation is an internal exercise. Although not controlled by the first pillar of Solvency 2, it plays a crucial role in determining performance of the insurer overall activity. The case of insurance groups requires special treatment in the context of the Own Risk and Solvency Assessment (ORSA). In this context, a multivariate analysis of risk becomes relevant.

The challenge of capital allocation in a multivariate context arises from the dependence between various risky activities which may generate a diversification effect.

In the actuarial literature, capital allocation methods are either based on a univariate risk measure combined with an allocation principle, or a multivariate risk indicator with a multivariate ruin probability. The aim is to minimize this multivariate ruin probability. In this paper, we focus on an allocation technique by minimizing some of the risk indicators.

The literature on capital allocation methods is very rich. Several principles have been proposed over the last twenty years. The most important and most studied are Shapley method, Aumann-Shapley method and Euler's method.

The Shapley method is based on cooperative game theory. It is described in detail in Denault's paper (2001) [11]. Denault proved that this method, originally used to allocate the total cost between players in coalition games contexts, can be easily adapted to solve the problem of the overall risk allocation between segments.

Tasche devoted two papers [26] and [27] to describing Euler's method. Euler's method is also found in the literature under the name of gradient method. It is based on the idea of allocating capital according to the infinitesimal marginal impact of each risk. This impact corresponds to the increase obtained on the overall risk, which yields an infinitely small increment in a marginal risk. Euler's method is well known in the actuarial literature. Several papers analyze its properties (RORAC compatibility [7], [28], Coherence [6],...) under different assumptions (Tasche (2004) [25], Balog (2011) [2]). Its fame is due to the existence of economic arguments that may justify its use to develop allocation rules.

Finally, Aumann-Shapley method is a continuous generalization of Shapley method. Its principle is based on the value introduced by Aumann and Shapley in game theory. Denault (2001) [11] analyzes this method and its application to capital allocation.

These three capital allocation principles rely on different risk measures. The coherence of the allocation method depends on the properties of the selected risk measure. Several papers deal with capital allocation coherence based on the properties of the risk measure used. We quote as examples, Fischer (2003) [13], Bush and Dorfleitner (2008) [6], and Kalkbrener (2009) [20].

Other techniques have been proposed more recently for building optimal allocation methods, one by minimizing some multivariate ruin probabilities, as in Cai and Li (2007) [8], other by minimizing some new multivariate risk indicators, as in Cénac et al. [9], [10]. The latter defined three types of indicators, which take into account both the ruin severity at the branch level, and the impact of the dependence structure on this local severity. In the one-period case, these indicators can be considered as special cases of a general indicator family introduced in Dhaene et al. (2012) [12]. Allocation by minimizing these indicators was studied in bivariate dimension by Cénac et al. (2014) [9], and

in higher dimension by Maume-Deshamps et al. (2016) [21]. Its behavior and asymptotic behavior for some special distributions' families were also analyzed in [21]. The impact of dependence on the allocation composition was studied in the same paper. In the present paper we focus on the coherence properties of this kind of allocation.

Allocation by minimizing some real multivariate risk indicators can be used in the more general framework of modeling systemic risk. In reinsurance, the allocation can be used to find the optimal stop-loss level in some concrete treaties. This allocation technique can also help to measure the performance of calculating the groups' capital requirement in the Swiss Solvency Test (SST), which provides a consistent framework for both the legal branch and the group solvency capital requirement.

The article is organized as follows. In the first section, we introduce the optimal allocation method by minimizing multivariate risk indicators. Using an axiomatic approach, we define in Section 2 some coherence properties for the allocation method in a multivariate context. The third section is devoted to the coherence properties of this optimal allocation. Section 4 is a discussion on the best allocation method choice for a group insurance.

3.1 Optimal allocation presentation

In a multivariate risk framework, we consider a vectorial risk process $X^p = (X_1^p, \dots, X_d^p)$, where X_k^p corresponds to the losses of the k^{th} business line during the p^{th} period. We denote by R_k^p the reserve of the k^{th} line at time p , so: $R_k^p = u_k - \sum_{l=1}^p X_k^l$, where $u_k \in \mathbb{R}^+$ is the initial capital of the k^{th} business line. $u = u_1 + \dots + u_d$ is the initial capital of the group and d is the number of business lines.

Cénac et al. (2012) [10] defined the two following multivariate risk indicators, for d risks and n periods, given penalty functions g_k , $k \in \{1, \dots, d\}$:

– the indicator I :

$$I(u_1, \dots, u_d) = \sum_{k=1}^d \mathbb{E} \left(\sum_{p=1}^n g_k(R_k^p) \mathbb{1}_{\{R_k^p < 0\}} \mathbb{1}_{\{\sum_{j=1}^d R_j^p > 0\}} \right),$$

– the indicator J :

$$J(u_1, \dots, u_d) = \sum_{k=1}^d \mathbb{E} \left(\sum_{p=1}^n g_k(R_k^p) \mathbb{1}_{\{R_k^p < 0\}} \mathbb{1}_{\{\sum_{j=1}^d R_j^p < 0\}} \right),$$

$g_k : \mathbb{R}^- \rightarrow \mathbb{R}^+$ are C^1 , convex functions with $g_k(0) = 0$, $g_k(x) \geq 0$, $k = 1, \dots, d$, g_k are decreasing functions on \mathbb{R}^- . These functions represent the cost that each branch has to pay when it becomes insolvent while the group is solvent for the I indicator, or while the group is also insolvent in the case of the J indicator.

Cénac et al. (2012) [10] proposed to allocate some capital u by minimizing these indicators. The idea is to find an allocation vector (u_1, \dots, u_d) that minimizes the indicator such as $u = u_1 + \dots + u_d$, where u is the initial capital that needs to be shared among all branches.

The indicator I represents the expected sum of penalty amounts of local ruins, knowing that the group remains solvent. In the case of the indicator J , the local ruin severities are taken into account only in the case of group insolvency.

By using stochastic optimization algorithms, we may estimate the minimum of these risk indicators. Cénac et al. (2012) [10] propose a Kiefer-Wolfowitz version of the mirror algorithm as a convergent algorithm under general assumptions to find an optimal allocation which minimizes the indicator I . This algorithm is effective in solving the optimal allocation problem, especially, for a large number of business lines, and for allocation over several periods.

3.1.1 Definitions and notations

Since the new regulation, such as Solvency 2, require a one year allocation strategy, in this paper we focus on a single period ($n = 1$). An annual allocation seems to be a more efficient decision for an insurer; during a year of operation, this allows the firm to dynamically respond to the changes that occurred in its risk portfolio.

The following notations are used:

- u is the initial capital of the firm.
- $\mathcal{U}_u^d = \{v = (v_1, \dots, v_d) \in [0, u]^d, \sum_{i=1}^d v_i = u\}$ is the set of possible allocations of capital u .
- For all $i \in \{1, \dots, d\}$ let $\alpha_i = \frac{u_i}{u}$ then $\sum_{i=1}^d \alpha_i = 1$ if $(u_1, \dots, u_d) \in \mathcal{U}_u^d$.
- $\mathbb{1}_u^d = \{\alpha = (\alpha_1, \dots, \alpha_d) \in [0, 1]^d, \sum_{i=1}^d \alpha_i = 1\}$ is the set of possible allocation percentages.
- X_k corresponds to the losses of the k^{th} branch during one period. It is a non negative random variable in our context.
- For $(u_1, \dots, u_d) \in \mathcal{U}_u^d$, $R^k = u_k - X_k$ is the reserve of the k^{th} business line at the end of the period, u_k represents the part of capital allocated to the k^{th} branch.
- The aggregate sum of risks is: $S = \sum_{i=1}^d X_i$, and let $S^{-i} = \sum_{j=1; j \neq i}^d X_j$ for all $i \in \{1, \dots, d\}$.
- F_Z is the cumulative distribution function of a random variable Z , \bar{F}_Z is its survival function and f_Z its probability density function.

Definition 3.1.1 (Optimal allocation). Let \mathbf{X} be a non negative random vector of \mathbb{R}^d , $u \in \mathbb{R}^+$ and $\mathcal{K}_{\mathbf{X}} : \mathcal{U}_u^d \rightarrow \mathbb{R}^+$ a multivariate risk indicator associated to \mathbf{X} and u . An optimal allocation of the capital u for the risk vector \mathbf{X} is defined by:

$$(u_1, \dots, u_d) \in \underset{(v_1, \dots, v_d) \in \mathcal{U}_u^d}{\arg \inf} \{ \mathcal{K}_{\mathbf{X}}(v_1, \dots, v_d) \}.$$

For risk indicators of the form $\mathcal{K}_{\mathbf{X}}(v) = \mathbb{E}[S(\mathbf{X}, \mathbf{v})]$, with a scoring function $S : \mathbb{R}^+ \times \mathbb{R}^+ \rightarrow \mathbb{R}^+$, this definition can be seen as an extension in a multivariate framework of the concept of elicibility. Elicibility is presented for example in Gneiting (2011) [16], and studied recently for univariate risk measures, by Bellini and Bignozzi (2015)[3], Ziegel (2014)[29] and Steinwart et al. (2014)[24], for examples.

Optimal allocations can be seen as vector-valued risk measures in the sense defined by Jouini et al. (2004) [19], and can also be used to construct set-valued risk measures as defined in Hamel and Heyde (2010) [17].

Assumptions Throughout this paper, we will use the following assumptions:

- H1** $\mathcal{K}_{\mathbf{X}}$ admits a unique minimum in \mathcal{U}_u^d . In this case, we denote by $A_{X_1, \dots, X_d}(u) = (u_1, \dots, u_d)$ the optimal allocation of u on the d risky branches in \mathcal{U}_u^d .
- H2** The functions g_k are differentiable and such that for all $k \in \{1, \dots, d\}$, $g'_k(u_k - X_k)$ admits a moment of order one, and (X_k, S) has a joint density distribution denoted by $f_{(X_k, S)}$.
- H3** The d risks have the same penalty function $g_k = g, \forall k \in \{1, \dots, d\}$.

The first assumption is verified when the indicator is strictly convex, this is particularly true if at least one function g_k is strictly convex; and the joint density $f_{(X_k, S)}$ support contains $[0, u]^2$ (see [10]).

3.1.2 Optimality conditions

In this section, we focus on the optimality condition for the indicators I and J .

For an initial capital u , and an optimal allocation minimizing I , we seek $u^* \in \mathbb{R}_+^d$ such that:

$$I(u^*) = \inf_{v_1 + \dots + v_d = u} I(v), \quad v \in \mathbb{R}_+^d.$$

Under assumption H2, the risk indicators I and J are differentiable, and in this case, we may calculate the following gradients:

$$\begin{aligned} (\nabla I(v))_i &= \sum_{k=1}^d \int_{v_k}^{+\infty} g_k(v_k - x) f_{X_k, S}(x, u) dx + \mathbb{E}[g'_i(v_i - X_i) \mathbb{1}_{\{X_i > v_i\}} \mathbb{1}_{\{S \leq u\}}] \\ \text{and, } (\nabla J(v))_i &= \sum_{k=1}^d \int_{v_k}^{+\infty} g_k(v_k - x) f_{X_k, S}(x, u) dx + \mathbb{E}[g'_i(v_i - X_i) \mathbb{1}_{\{X_i > v_i\}} \mathbb{1}_{\{S \geq u\}}]. \end{aligned}$$

Under H1 and H2, using Lagrange multipliers, we obtain an optimality condition verified by the unique solution to this optimization problem:

$$\mathbb{E}[g'_i(u_i - X_i) \mathbb{1}_{\{X_i > u_i\}} \mathbb{1}_{\{S \leq u\}}] = \mathbb{E}[g'_j(u_j - X_j) \mathbb{1}_{\{X_j > u_j\}} \mathbb{1}_{\{S \leq u\}}], \quad \forall (i, j) \in \{1, \dots, d\}^2. \quad (3.1)$$

A natural choice for penalty functions is the ruin severity: $g_k(x) = |x|$. In that case, and if the joint density $f_{(X_k, S)}$ support contains $[0, u]^2$, for at least one $k \in \{1, \dots, d\}$, our optimization problem has a unique solution.

We may write the indicators as follows:

$$\begin{aligned} I(u_1, \dots, u_d) &= \sum_{k=1}^d \mathbb{E} \left(|R^k| \mathbb{1}_{\{R^k < 0\}} \mathbb{1}_{\{\sum_{i=1}^d R^i \geq 0\}} \right) \\ &= \sum_{k=1}^d \mathbb{E} \left((X_k - u_k) \mathbb{1}_{\{X_k > u_k\}} \mathbb{1}_{\{\sum_{i=1}^d X_i \leq u\}} \right) = \sum_{k=1}^d \mathbb{E} \left((X_k - u_k)^+ \mathbb{1}_{\{S \leq u\}} \right), \end{aligned}$$

and,

$$\begin{aligned} J(u_1, \dots, u_d) &= \sum_{k=1}^d \mathbb{E} \left(|R^k| \mathbb{1}_{\{R^k < 0\}} \mathbb{1}_{\{\sum_{i=1}^d R^i \leq 0\}} \right) \\ &= \sum_{k=1}^d \mathbb{E} \left((X_k - u_k) \mathbb{1}_{\{X_k > u_k\}} \mathbb{1}_{\{\sum_{i=1}^d X_i \geq u\}} \right) = \sum_{k=1}^d \mathbb{E} \left((X_k - u_k)^+ \mathbb{1}_{\{S \geq u\}} \right). \end{aligned}$$

Respective components of the gradient of these indicators are

$$K_I - \mathbb{P} \left(X_1 > u_1, \sum_{j=1}^d X_j \leq u \right), \dots, K_I - \mathbb{P} \left(X_d > u_d, \sum_{j=1}^d X_j \leq u \right),$$

and,

$$K_J - \mathbb{P} \left(X_1 > u_1, \sum_{j=1}^d X_j \geq u \right), \dots, K_J - \mathbb{P} \left(X_d > u_d, \sum_{j=1}^d X_j \geq u \right),$$

where,

$$K_I = K_J = \sum_{k=1}^d \int_{u_k}^{+\infty} (x - u_k) f_{X_k, S}(x, u) dx.$$

Using Lagrange multipliers to solve our convex optimization problem under the only constraint $u_1 + u_2 + \dots + u_d = u$, the following optimality conditions are obtained from 3.1 in the special case where $g_k(x) = |x|$:

$$\mathbb{P}(X_i > u_i, S \leq u) = \mathbb{P}(X_j > u_j, S \leq u), \quad \forall (i, j) \in \{1, 2, \dots, d\}^2. \quad (3.2)$$

For the J indicator, this condition becomes

$$\mathbb{P}(X_i > u_i, S \geq u) = \mathbb{P}(X_j > u_j, S \geq u), \quad \forall (i, j) \in \{1, 2, \dots, d\}^2. \quad (3.3)$$

Some explicit and semi-explicit formulas for the optimal allocation can be obtained with this optimality condition. Our problem reduces to the study of this allocation depending on the nature of the distributions of the risk X_k and on the form of dependence between them.

3.2 Coherence of a capital allocation in a multivariate context

In his article [11], Denault introduced the notion of a coherent allocation, fixing four axioms that must be verified by a principle of capital allocation in order to be qualified as coherent. Denault's definition can be used only for allocation methods driven by univariate risk measures, especially coherent ones, according to the criteria defined by Artzner et al. (1999) [1]. Our optimal capital allocation is not directly derived from a univariate risk measure, even if it is obtained by minimizing a multivariate risk indicator. Brunnermeier and Cheridito (2014) [5] fixed four axioms that should be satisfied by allocation used to implement capital requirements or systemic risk charges. These axioms are also based on univariate risk measures. They are therefore not directly used in our context.

In this section, we reformulate coherence axioms in a more general multivariate context. We also define other coherence properties and justify for each one why it is a desirable property from an economic point of view.

3.2.1 Coherence

We follow Denault's idea to define a coherent capital allocation in a multivariate context. We define risk measure free counterparts to the four Denault's axioms and define a fifth axiom that seems important to describe the presence of perfectly correlated risks on capital allocation.

Definition 3.2.1 (Coherence). A capital allocation $(u_1, \dots, u_d) = A_{X_1, \dots, X_d}(u)$ of an initial capital $u \in \mathbb{R}^+$ is coherent if it satisfies the following properties:

1. **Full allocation:** All of the capital $u \in \mathbb{R}^+$ must be allocated between the branches:

$$\sum_{i=1}^d u_i = u.$$

2. **Symmetry:** If the joint distribution of the vector (X_1, \dots, X_d) is unchanged by permutation of the risks X_i and X_j , then the allocation remains also unchanged by this permutation, and the i^{th} and j^{th} business lines both make the same contribution to the risk capital: if

$$\begin{aligned} & (X_1, \dots, X_{i-1}, X_i, X_{i+1}, \dots, X_{j-1}, X_j, X_{j+1}, \dots, X_d) \\ & \quad \underline{\underline{\mathcal{L}}} \\ & (X_1, \dots, X_{i-1}, X_j, X_{i+1}, \dots, X_{j-1}, X_i, X_{j+1}, \dots, X_d), \end{aligned}$$

then $u_i = u_j$.

3. **Riskless allocation:** For a deterministic risk $X = c$, where the constant $c \in \mathbb{R}^+$:

$$A_{X, X_1, \dots, X_d}(u) = (c, A_{X_1, \dots, X_d}(u - c)).$$

This property means that the presence of a deterministic risk has no impact on allocation to the risky branches.

4. **Sub-additivity:** $\forall M \subseteq \{1, \dots, d\}$, let $(u^*, u_1^*, \dots, u_r^*) = A_{\sum_{i \in M} X_i, X_j \in \{1, \dots, d\} \setminus M}(u)$, where $r = d - \text{card}(M)$ and $(u_1, \dots, u_d) = A_{X_1, \dots, X_d}(u)$:

$$u^* \leq \sum_{i \in M} u_i.$$

This property means that the allocation takes into account the diversification gain. It is related to the *no undercut* property defined by Denault, which has no sense in our context.

5. **Comonotonic additivity:** For $r \leq d$ comonotonic risks,

$$A_{X_{i \in \{1, \dots, d\} \setminus CR}, \sum_{k \in CR} X_k}(u) = (u_{i \in \{1, \dots, d\} \setminus CR}, \sum_{k \in CR} u_k),$$

where $(u_1, \dots, u_d) = A_{X_1, \dots, X_d}(u)$ is the allocation of u on the d risks (X_1, \dots, X_d) and CR denotes the set of the r comonotonic risk indexes.

The concept of comonotonic random variables is related to Hoeffding and Fréchet's works (1940) [18] and (1951) [15]. Here we use the definition of comonotonic risks as it was first mentioned in the actuarial literature in Borch (1962) [4].

A vector of random variables (X_1, X_2, \dots, X_n) is comonotonic if and only if there exists a random variable Y and non-decreasing functions $\varphi_1, \dots, \varphi_n$ such that:

$$(X_1, \dots, X_n) =_d (\varphi_1(Y), \dots, \varphi_n(Y)).$$

3.2.2 Other desirable properties

We also define some desirable properties that an allocation should naturally satisfy. These properties are based on the ideas presented by Artzner et al. (1999) [1] for coherent risk measures and on the axiomatic characterization of coherent capital allocations given by Kalkbrener (2009) [20].

Definition 3.2.2 (Positive homogeneity). An allocation is positively homogeneous, if for any $\alpha \in \mathbb{R}^+$, it satisfies:

$$A_{\alpha X_1, \dots, \alpha X_d}(\alpha u) = \alpha A_{X_1, \dots, X_d}(u).$$

In other words, a capital allocation method is positively homogeneous, if it is insensitive to cash changes.

Definition 3.2.3 (Translation invariance). An allocation is invariant by translation, if for all $(a_1, \dots, a_d) \in \mathbb{R}^d$ such that $u > \sum_{k=1}^d a_k$, it satisfies:

$$A_{X_1+a_1, \dots, X_d+a_d}(u) = A_{X_1, \dots, X_d}\left(u - \sum_{k=1}^d a_k\right) + (a_1, \dots, a_d).$$

The translation invariance property shows the impact of a deterministic translation of the risk vector (X_1, \dots, X_d) . If one increases a risk by a constant amount (a_1, \dots, a_n) , then the allocation is the sum of this constant and the allocation of the remaining capital $u - \sum_{i=1}^d a_i$.

Definition 3.2.4 (Continuity). An allocation is continuous, if for all $i \in \{1, \dots, d\}$:

$$\lim_{\epsilon \rightarrow 0} A_{X_1, \dots, (1+\epsilon)X_i, \dots, X_d}(u) = A_{X_1, \dots, X_i, \dots, X_d}(u).$$

This property reflects the fact that a small change to the risk of a business line, have only limited effect on the capital part that we attribute to it.

Let us recall the definition of first-order stochastic dominance, as it is presented in Shaked and Shanthikumar (2007)[23]. For random variables X and Y , Y first-order stochastically dominates X if and only if:

$$\bar{F}_X(x) \leq \bar{F}_Y(x), \quad \forall x \in \mathbb{R}^+,$$

and in this case we denote: $X \leq_{st} Y$.

This definition is also equivalent to the following one:

$$X \leq_{st} Y \Leftrightarrow \mathbb{E}[u(X)] \leq \mathbb{E}[u(Y)], \text{ for all } u \text{ non-decreasing function}$$

Definition 3.2.5 (Monotonicity). An allocation satisfies the monotonicity property, if for $(i, j) \in \{1, \dots, d\}^2$:

$$X_i \leq_{st} X_j \Rightarrow u_i \leq u_j.$$

The monotonicity is a natural requirement, it reflects the fact that if X_j is riskier than X_i then, it is natural to allocate more capital to the j^{th} branch. An allocation that satisfies Monotonicity and Translation invariance properties can be considered as a monetary allocation in the sense used by Föllmer et al. (2010) [14] to define monetary risk measures.

The RORAC compatibility property defined by Tasche (2007) [26] loses its meaning in our risk measure free allocation method.

3.3 Coherence of the optimal allocation

In what follows, we show that minimizing the indicator I gives a capital allocation satisfying the coherence axioms of Definition 3.2.1, except sub-additivity. We also prove homogeneity, translation invariance, monotonicity and continuity for this allocation. The same holds for the allocation by minimizing J .

3.3.1 Coherence

Firstly, the *full allocation* axiom is verified by construction.

Proposition 3.3.1 shows that the optimal allocation satisfies the symmetry property.

Proposition 3.3.1 (Symmetry). *Under H1, if for $(i, j) \in \{1, 2, \dots, d\}^2$, $i \neq j$, the couples (X_i, S^{-i}) and (X_j, S^{-j}) are identically distributed and $g_i = g_j$, then:*

$$u_i = u_j.$$

Proof. Let $(i \neq j) \in \{1, 2, \dots, d\}^2$ be such that (X_i, S^{-i}) and (X_j, S^{-j}) have the same distribution and the same penalty function $g_i = g_j = g$. If $u_i \neq u_j$, we may assume $i < j$, and denote:

$$(u_1, \dots, u_i, \dots, u_j, \dots, u_d) = A_{X_1, \dots, X_i, \dots, X_j, \dots, X_d}(u),$$

then,

$$I(u_1, \dots, u_i, \dots, u_j, \dots, u_d) = \inf_{v \in \mathcal{U}_u^d} I(v) = \inf_{v \in \mathcal{U}_u^d} \sum_{k=1}^d \mathbb{E} \left(g_k(v_k - X_k) \mathbb{1}_{\{X_k > v_k\}} \mathbb{1}_{\{S \leq u\}} \right).$$

On the other hand, and since $g_i = g_j = g$ and $(X_i, S^{-i}) \stackrel{\mathcal{L}}{=} (X_j, S^{-j})$, then:

$$\begin{aligned} I(u_1, \dots, u_{i-1}, u_j, u_{i+1}, \dots, u_{j-1}, u_i, u_{j+1}, \dots, u_d) &= \sum_{k=1, k \neq i, k \neq j}^d \mathbb{E} \left(g_k(u_k - X_k) \mathbb{1}_{\{X_k > u_k\}} \mathbb{1}_{\{S \leq u\}} \right) \\ &\quad + \mathbb{E} \left(g(u_i - X_i) \mathbb{1}_{\{X_i > u_i\}} \mathbb{1}_{\{S \leq u\}} \right) \\ &\quad + \mathbb{E} \left(g(u_j - X_j) \mathbb{1}_{\{X_j > u_j\}} \mathbb{1}_{\{S \leq u\}} \right) \\ &= I(u_1, \dots, u_i, \dots, u_j, \dots, u_d). \end{aligned}$$

From H1, I admits a unique minimum in \mathcal{U}_u^d , we deduce that:

$$(u_1, \dots, u_i, \dots, u_j, \dots, u_d) = (u_1, \dots, u_{i-1}, u_j, u_{i+1}, \dots, u_{j-1}, u_i, u_{j+1}, \dots, u_d).$$

We conclude that $u_i = u_j$. □

Corollary 3.3.2. *Under Assumptions H1 and H3, if (X_1, \dots, X_d) is an exchangeable random vector, then the allocation by minimizing I and J indicators is the same and is given by:*

$$A_{X_1, \dots, X_d}(u) = \left(\frac{u}{d}, \frac{u}{d}, \dots, \frac{u}{d} \right).$$

The following proposition shows that the optimal allocation verifies the Riskless allocation axiom.

Proposition 3.3.3 (Riskless Allocation). *Under Assumptions H1 and H3, and for 1-homogeneous* penalty functions, for any $c \in [0, u]$:*

$$A_{c, X_1, \dots, X_d}(u) = (c, A_{X_1, \dots, X_d}(u - c)),$$

where $(c, A_{X_1, \dots, X_d}(u - c))$ is the concatenated vector of c and the vector $A_{X_1, \dots, X_d}(u - c)$.

*. we recall f is a 1-homogeneous function if for all $\lambda \in \mathbb{R}$ and $x \in D_f$, $f(\lambda x) = \lambda f(x)$.

Remark that if all risks are constant, the uniqueness of the minimum is lost and therefore assumptions of Proposition 3.3.3 are not satisfied. This is also the case if $c > u$, the lack of uniqueness of the minimum makes Proposition 3.3.3 not applicable.

Proof. The presence of a discrete distribution makes the indicator I not differentiable, so we cannot use neither the gradient, nor the optimality condition obtained in the case of existence of joined densities.

Let $(u^*, u_1^*, \dots, u_d^*) = A_{c, X_1, \dots, X_d}(u)$ and $(u_1, \dots, u_d) = A_{X_1, \dots, X_d}(u - c)$.

We denote $S = \sum_{i=1}^d X_i$, and the common penalty function $g = g_k, \forall k \in \{1, \dots, d\}$.

We distinguish between three possibilities:

– *Case 1: $u^* < c$*

In this case,

$$\begin{aligned} I(u^*, u_1^*, \dots, u_d^*) &= \inf_{v \in \mathcal{U}_u^{d+1}} I(v) = \inf_{v \in \mathcal{U}_u^{d+1}} \sum_{k=0}^d \mathbb{E} \left(g(v_k - X_k) \mathbb{1}_{\{X_k > v_k\}} \mathbb{1}_{\{S \leq u-c\}} \right) \\ &= \mathbb{E} \left(g(u^* - c) \mathbb{1}_{\{S \leq u-c\}} \right) + \sum_{k=1}^d \mathbb{E} \left(g(u_k^* - X_k) \mathbb{1}_{\{X_k > u_k^*\}} \mathbb{1}_{\{S \leq u-c\}} \right), \end{aligned}$$

for all $k \in \{1, \dots, d\}$ we put, for example, $\alpha_k = \alpha = \frac{u^* - c}{d} < 0$, and since the function g is convex and $g(0) = 0$, it satisfies for all real $0 < \beta < 1$, $g(\beta x) \leq \beta g(x), \forall x \in \mathbb{R}^-$. Then:

$$\begin{aligned} I(u^*, u_1^*, \dots, u_d^*) &\geq \mathbb{E} \left(d \cdot g \left(\frac{u^* - c}{d} \right) \mathbb{1}_{\{S \leq u-c\}} \right) + \sum_{k=1}^d \mathbb{E} \left(g(u_k^* - X_k) \mathbb{1}_{\{X_k > u_k^*\}} \mathbb{1}_{\{S \leq u-c\}} \right) \\ &= \mathbb{E} \left(d \cdot g(-(-\alpha)_+) \mathbb{1}_{\{S \leq u-c\}} \right) + \sum_{k=1}^d \mathbb{E} \left(g(-(X_k - u_k^*)_+) \mathbb{1}_{\{S \leq u-c\}} \right) \\ &= \sum_{k=1}^d \mathbb{E} \left([g(-(X_k - u_k^*)_+) + g(-(-\alpha_k)_+)] \mathbb{1}_{\{S \leq u-c\}} \right), \end{aligned}$$

$x \rightarrow g(-(x)_+)$ is also a 1-homogeneous convex function, then:

$$I(u^*, u_1^*, \dots, u_d^*) \geq \sum_{k=1}^d \mathbb{E} \left(g(-(X_k - (u_k^* + \alpha_k))_+) \mathbb{1}_{\{S \leq u-c\}} \right),$$

we remark that $\sum_{k=1}^d (u_k^* + \alpha_k) = u - c$, then $(u_1^* + \alpha, \dots, u_d^* + \alpha) \in \mathcal{U}_{u-c}^d$.

So,

$$\begin{aligned} I(u^*, u_1^*, \dots, u_d^*) &\geq \sum_{k=1}^d \mathbb{E} \left(g((u_k^* + \alpha_k) - X_k) \mathbb{1}_{\{X_k > u_k^* + \alpha_k\}} \mathbb{1}_{\{S \leq u-c\}} \right) \\ &\geq \inf_{v \in \mathcal{U}_{u-c}^d} \sum_{k=1}^d \mathbb{E} \left(g(v_k - X_k) \mathbb{1}_{\{X_k > v_k\}} \mathbb{1}_{\{S \leq u-c\}} \right) \\ &= \sum_{k=1}^d \mathbb{E} \left(g(u_k - X_k)^+ \mathbb{1}_{\{S \leq u-c\}} \right) \\ &= I(c, u_1, \dots, u_d), \end{aligned}$$

then,

$$I(u^*, u_1^*, \dots, u_d^*) \geq I(c, u_1, \dots, u_d).$$

That is contradictory with the uniqueness of the minimum on the set \mathcal{U}_u^{d+1} .

– *Case 2: $u^* > c$*

We have :

$$I(u^*, u_1^*, \dots, u_d^*) = \sum_{k=1}^d \mathbb{E} \left(g(u_k^* - X_k) \mathbb{1}_{\{X_k > u_k^*\}} \mathbb{1}_{\{S \leq u-c\}} \right),$$

and,

$$I(c, u_1, \dots, u_d) = \sum_{k=1}^d \mathbb{E} \left(g(u_k - X_k) \mathbb{1}_{\{X_k > u_k\}} \mathbb{1}_{\{S \leq u-c\}} \right).$$

Let $\alpha = \frac{u^* - c}{d} > 0$, we remark that, $(u_1^* + \alpha, \dots, u_d^* + \alpha) \in \mathcal{U}_{u-c}^d$, and that the penalty function g is decreasing on \mathbb{R}^- because $g''(x) \geq 0, \forall x \in \mathbb{R}^-$ and $g'(0^+) = 0$. Then,

$$\begin{aligned} I(c, u_1, \dots, u_d) &= \sum_{k=1}^d \mathbb{E} \left(g(-(X_k - u_k)_+) \mathbb{1}_{\{S \leq u-c\}} \right) \\ &= \inf_{v \in \mathcal{U}_{u-c}^d} \sum_{k=1}^d \mathbb{E} \left(g(v_k - X_k) \mathbb{1}_{\{X_k > v_k\}} \mathbb{1}_{\{S \leq u-c\}} \right) \\ &\leq \sum_{k=1}^d \mathbb{E} \left(g(-(X_k - (u_k^* + \alpha))_+) \mathbb{1}_{\{S \leq u-c\}} \right) \\ &< \sum_{k=1}^d \mathbb{E} \left(g(-(X_k - u_k^*)_+) \mathbb{1}_{\{S \leq u-c\}} \right) \\ &= \sum_{k=1}^d \mathbb{E} \left(g(u_k^* - X_k) \mathbb{1}_{\{X_k > u_k^*\}} \mathbb{1}_{\{S \leq u-c\}} \right) = I(u^*, u_1^*, \dots, u_d^*). \end{aligned}$$

That is contradictory with the fact that $I(u^*, u_1^*, \dots, u_d^*) = \inf_{v \in \mathcal{U}_u^{d+1}} I(v)$.

We deduce that the only possible case is the third one $u^* = c$.

– *Case 3: $u^* = c$*

The uniqueness of the minimum implies that:

$$\begin{aligned} (u^*, u_1^*, \dots, u_d^*) &= (c, u_1^*, \dots, u_d^*) = \arg \min_{\mathcal{U}_{u-c}^{d+1}} \sum_{k=1}^d \mathbb{E} [g(u_k - X_k) \mathbb{1}_{\{X_k > u_k\}} \mathbb{1}_{\{S \leq u-c\}}] \\ &= \arg \min_{\mathcal{U}_{u-c}^d} \sum_{k=1}^d \mathbb{E} [g(u_k - X_k) \mathbb{1}_{\{X_k > u_k\}} \mathbb{1}_{\{S \leq u-c\}}] \\ &= (c, u_1, \dots, u_d). \end{aligned}$$

Finally, we have proven that:

$$(u^*, u_1^*, \dots, u_d^*) = (c, u_1, \dots, u_d).$$

This property is presented under Assumption H3 of identical penalty functions, but can be easily generalized. □

Lemma 3.3.4 is related to subadditivity. It will be used in the proof of the comonotonic additivity property.

Lemma 3.3.4. *Under Assumptions H1, H2 and H3, and for all $(i, j) \in \{1, \dots, d\}^2$,*

– *if $A_{X_1, \dots, X_{i-1}, X_i+X_j, X_{i+1}, \dots, X_{j-1}, X_{j+1}, \dots, X_d}(u).e_i < A_{X_1, \dots, X_d}(u).(e_i + e_j)$, then:*

$$\forall k \in \{1, \dots, d\} \setminus \{i, j\}, \quad A_{X_1, \dots, X_{i-1}, X_i+X_j, X_{i+1}, \dots, X_{j-1}, X_{j+1}, \dots, X_d}(u).e_k > A_{X_1, \dots, X_d}(u).e_k,$$

– *if $A_{X_1, \dots, X_{i-1}, X_i+X_j, X_{i+1}, \dots, X_{j-1}, X_{j+1}, \dots, X_d}(u).e_i > A_{X_1, \dots, X_d}(u).(e_i + e_j)$, then:*

$$\forall k \in \{1, \dots, d\} \setminus \{i, j\}, \quad A_{X_1, \dots, X_{i-1}, X_i+X_j, X_{i+1}, \dots, X_{j-1}, X_{j+1}, \dots, X_d}(u).e_k < A_{X_1, \dots, X_d}(u).e_k,$$

– *if $A_{X_1, \dots, X_{i-1}, X_i+X_j, X_{i+1}, \dots, X_{j-1}, X_{j+1}, \dots, X_d}(u).e_i = A_{X_1, \dots, X_d}(u).(e_i + e_j)$, then:*

$$\forall k \in \{1, \dots, d\} \setminus \{i, j\}, \quad A_{X_1, \dots, X_{i-1}, X_i+X_j, X_{i+1}, \dots, X_{j-1}, X_{j+1}, \dots, X_d}(u).e_k = A_{X_1, \dots, X_d}(u).e_k,$$

where $x.e_i$ is the dot product of the vector $x \in \mathbb{R}^d$ and the i^{th} component of the canonical basis of \mathbb{R}^d .

Proof. In order to simplify notations, and without loss of generality, we assume $i = d - 1$ and $j = d$. Let $(u_1, \dots, u_{d-1}, u_d) = A_{X_1, \dots, X_d}(u)$ and $(u_1^*, \dots, u_{d-2}^*, u_{d-1}^*) = A_{X_1, \dots, X_{d-2}, X_{d-1} + X_d}(u)$. The optimality conditions (3.1) are:

$$\mathbb{E}[g'_i(u_i - X_i) \mathbb{1}_{\{X_i > u_i\}} \mathbb{1}_{\{S \leq u\}}] = \mathbb{E}[g'_j(u_j - X_j) \mathbb{1}_{\{X_j > u_j\}} \mathbb{1}_{\{S \leq u\}}] = \lambda \quad \forall (i, j) \in \{1, \dots, d\}^2,$$

and

$$\mathbb{E}[g'_i(u_i^* - X_i) \mathbb{1}_{\{X_i > u_i^*\}} \mathbb{1}_{\{S \leq u\}}] = \mathbb{E}[g'_i(u_{d-1}^* - (X_{d-1} + X_d)) \mathbb{1}_{\{X_{d-1} + X_d > u_{d-1}^*\}} \mathbb{1}_{\{S \leq u\}}] = \lambda^* \quad \forall i \in \{1, \dots, d-2\}.$$

Now, we suppose that $u_{d-1}^* > u_d + u_{d-1}$. In this case there exists $k \in \{1, \dots, d-2\}$ such that $u_k^* < u_k$, and since $x \rightarrow g'(-x)_+$ is non increasing on \mathbb{R}^+ , then:

$$\mathbb{E}[g'_i(u_k - X_k) \mathbb{1}_{\{X_k > u_k\}} \mathbb{1}_{\{S \leq u\}}] = \lambda < \mathbb{E}[g'_i(u_k^* - X_k) \mathbb{1}_{\{X_k > u_k^*\}} \mathbb{1}_{\{S \leq u\}}] = \lambda^*$$

we deduce that for all $k \in 1, \dots, d-2$: $u_k^* < u_k$.

The proof is the same if we assume that $u_{d-1}^* < u_d + u_{d-1}$, and the results follow. \square

Proposition 3.3.5 (Comonotonic additivity). *Let CR be the set of the comonotonic risk indexes and $r = |CR|$. Under Assumption H2, and for $g_k(x) = |x|$, for all $k \in \{1, \dots, d\}$*

$$A_{X_{i \in \{1, \dots, d\} \setminus CR}, \sum_{k \in CR} X_k}(u) = (u_{i \in \{1, \dots, d\} \setminus CR}, \sum_{k \in CR} u_k),$$

where $(u_1, \dots, u_d) = A_{X_1, \dots, X_d}(u)$ is the optimal allocation of u on the d risks (X_1, \dots, X_d) , $A_{X_{i \in \{1, \dots, d\} \setminus CR}, \sum_{k \in CR} X_k}(u)$ is the optimal allocation of u on the $n-d+1$ risks $(X_{i \in \{1, \dots, d\} \setminus CR}, \sum_{k \in CR} X_k)$.

Proof. For $(i, j) \in \{1, \dots, d\}^2$, if X_i and X_j are comonotonic risks, then, there exists a non decreasing non negative function h such that $X_i = h(X_j)$, and remark that h is increasing under Assumption H2. Let f be the function $x \rightarrow f(x) = x + h(x)$, so that $X_i + X_j = f(X_j)$.

We denote $(u_1, \dots, u_d) = A_{X_1, \dots, X_d}(u)$ and $(u_1^*, \dots, u_{d-1}^*) = A_{X_{i \in \{1, \dots, d\} \setminus \{i, j\}}, X_i + X_j}(u)$, then,

$$A_{X_{i \in \{1, \dots, d\} \setminus \{i, j\}}, X_i + X_j}(u).e_{d-1} = u_{d-1}^* \quad \text{and} \quad A_{X_1, \dots, X_d}(u).(e_i + e_j) = u_i + u_j.$$

From the optimality condition (3.2)

$$\mathbb{P}(X_i \geq u_i, S \leq u) = \mathbb{P}(X_j \geq u_j, S \leq u),$$

we deduce that $u_i = h(u_j)$ and that $u_i + u_j = f(u_j)$.

If there exists $k \in \{1, \dots, d\} \setminus \{i, j\}$, such that $u_k^* < u_k$, then $\forall k \in \{1, \dots, d\} \setminus \{i, j\}$:

$$\mathbb{P}(X_k \geq u_k^*, S \leq u) > \mathbb{P}(X_k \geq u_k, S \leq u),$$

so,

$$\begin{aligned} \mathbb{P}(X_i + X_j \geq u_{d-1}^*, S \leq u) &= \mathbb{P}(X_j \geq f^{-1}(u_{d-1}^*), S \leq u) \\ &= \mathbb{P}(X_k \geq u_k^*, S \leq u) \\ &> \mathbb{P}(X_k \geq u_k, S \leq u) \\ &= \mathbb{P}(X_j \geq u_j, S \leq u), \end{aligned}$$

finally, we deduce that: $f^{-1}(u_{d-1}^*) < u_j$, then $u_{d-1}^* < f(u_j) = u_i + u_j$ and,

$\sum_{k \in \{1, \dots, d\} \setminus \{i, j\}} u_k^* < \sum_{k \in \{1, \dots, d\} \setminus \{i, j\}} u_k$ which is absurd.

In the same way, the case $u_k < u_k^*$ for $k \in \{1, \dots, d\} \setminus \{i, j\}$ leads to the contradiction.

Using Lemma 3.3.4, and under Assumption H3, we deduce $u_k = u_k^*$ for all $k \in \{1, \dots, d\} \setminus \{i, j\}$.

The additivity property for two comonotonic risks can be trivially generalized to several comonotonic risks. \square

Concerning the sub-additivity property, we have not yet managed to build a demonstration for this property. However, simulations using the optimization algorithm presented in Cénac et al. (2012) [10], seem to confirm the sub-additivity of the allocation by minimizing the indicators I and J , even for classic examples of non sub-additivity of the risk measure VaR.

Remark 3.3.1. *The previous properties have been demonstrated for the optimal allocation by minimizing I , they can be demonstrated with the same arguments for the optimal allocation by minimization of the indicator J .*

3.3.2 Other desirable properties

In this section, we show that the optimal allocation by minimization of I and J satisfies some desirable properties. We consider the allocation by minimizing the multivariate risk indicator I , the proofs are almost the same in the case of the indicator J .

Proposition 3.3.6 (Positive homogeneity). *Under Assumption H1, and for 1-homogeneous penalty functions g_k , $k \in \{1, \dots, d\}$, for any $\alpha \in \mathbb{R}^+$:*

$$A_{\alpha X_1, \dots, \alpha X_d}(\alpha u) = \alpha A_{X_1, \dots, X_d}(u).$$

Proof. Since penalty functions are convex and 1-homogeneous, then, for any $\alpha \in \mathbb{R}^+$:

$$\begin{aligned} A_{\alpha X_1, \dots, \alpha X_d}(\alpha u) &= \arg \min_{(u_1^*, \dots, u_d^*) \in \mathcal{U}_{\alpha u}^d} \sum_{k=1}^d \mathbb{E}[g_k(u_k^* - \alpha X_k) \mathbb{1}_{\{\alpha X_k > u_k^*\}} \mathbb{1}_{\{\alpha S \leq \alpha u\}}] \\ &= \arg \min_{(u_1^*, \dots, u_d^*) \in \mathcal{U}_{\alpha u}^d} \sum_{k=1}^d \alpha \mathbb{E}[g_k\left(\frac{u_k^*}{\alpha} - X_k\right) \mathbb{1}_{\{X_k > \frac{u_k^*}{\alpha}\}} \mathbb{1}_{\{S \leq u\}}] \\ &= \arg \min_{(u_1^*, \dots, u_d^*) \in \mathcal{U}_{\alpha u}^d} \sum_{k=1}^d \mathbb{E}[g_k\left(\frac{u_k^*}{\alpha} - X_k\right) \mathbb{1}_{\{X_k > \frac{u_k^*}{\alpha}\}} \mathbb{1}_{\{S \leq u\}}] \\ &= \alpha \arg \min_{(u_1, \dots, u_d) \in \mathcal{U}_u^d} \sum_{k=1}^d \mathbb{E}[g_k(u_k - X_k) \mathbb{1}_{\{X_k > u_k\}} \mathbb{1}_{\{S \leq u\}}] \\ &= \alpha A_{X_1, \dots, X_d}(u). \end{aligned}$$

□

Proposition 3.3.7 (Translation invariance). *Under Assumptions H1, H2 and for all $(a_1, \dots, a_d) \in \mathbb{R}^d$, such that $u > \sum_{k=1}^d a_k$ and the joint density $f_{(X_k, S)}$ support contains $[0, u - \sum_{k=1}^d a_k]^2$, for all $k \in \{1, \dots, d\}$,*

$$A_{X_1+a_1, \dots, X_d+a_d}(u) = A_{X_1, \dots, X_d}\left(u - \sum_{k=1}^d a_k\right) + (a_1, \dots, a_d).$$

Proof. We denote by (u_1^*, \dots, u_d^*) the optimal allocation $A_{X_1+a_1, \dots, X_d+a_d}(u)$, and by (u_1, \dots, u_d) the optimal allocation $A_{X_1, \dots, X_d}\left(u - \sum_{k=1}^d a_k\right)$.

Using the optimality condition (3.1), (u_1^*, \dots, u_d^*) is the unique solution in \mathcal{U}_u^d of the following equations system:

$$\mathbb{E}[g'_i(u_i^* - (X_i + a_i)) \mathbb{1}_{\{X_i + a_i > u_i^*\}} \mathbb{1}_{\{S + a \leq u\}}] = \mathbb{E}[g'_j(u_j^* - (X_j + a_j)) \mathbb{1}_{\{X_j + a_j > u_j^*\}} \mathbb{1}_{\{S + a \leq u\}}], \quad \forall (i, j) \in \{1, \dots, d\}^2,$$

where $a = \sum_{k=1}^d a_k$. Then, (u_1^*, \dots, u_d^*) satisfies also:

$$\mathbb{E}[g'_i(u_i^* - a_i - X_i) \mathbb{1}_{\{X_i > u_i^* - a_i\}} \mathbb{1}_{\{S \leq u - a\}}] = \mathbb{E}[g'_j(u_j^* - a_j - X_j) \mathbb{1}_{\{X_j > u_j^* - a_j\}} \mathbb{1}_{\{S \leq u - a\}}], \quad \forall (i, j) \in \{1, \dots, d\}^2.$$

Since, $(u_1^* - a_1, \dots, u_d^* - a_d) \in \mathcal{U}_{u-a}^d$, and from the solution uniqueness of the optimality condition (3.1) for the allocation $A_{X_1, \dots, X_d}(u - a)$, we deduce that: $u_k^* = u_k + a_k$ for all $k \in \{1, \dots, d\}$. □

Proposition 3.3.8 (Continuity). *Under Assumptions H1 and H2, and if $\forall k \in \{1, \dots, d\}$, $\exists \epsilon_0 > 0$ such that:*

$$\forall \epsilon, |\epsilon| < \epsilon_0, \quad \mathbb{E}[\sup_{v \in [0, u]} |g'_k(v - (1 + \epsilon)X_k)|] < +\infty,$$

then, if (X_1, \dots, X_d) is a continuous random vector, for all $i \in \{1, \dots, d\}$:

$$\lim_{\epsilon \rightarrow 0} A_{X_1, \dots, (1+\epsilon)X_i, \dots, X_d}(u) = A_{X_1, \dots, X_i, \dots, X_d}(u).$$

Proof. Let (u_1, \dots, u_d) be the optimal allocation of u

$$(u_1, \dots, u_d) = A_{X_1, \dots, X_i, \dots, X_d}(u),$$

then (u_1, \dots, u_d) is the unique solution in \mathcal{U}_u^d of Equations system (3.1):

$$\mathbb{E}[g'_i(u_i - X_i) \mathbb{1}_{\{X_i > u_i\}} \mathbb{1}_{\{S \leq u\}}] = \mathbb{E}[g'_j(u_j - X_j) \mathbb{1}_{\{X_j > u_j\}} \mathbb{1}_{\{S \leq u\}}], \quad \forall (i, j) \in \{1, \dots, d\}^2.$$

For $\epsilon \in \mathbb{R}$, let $(u_1^\epsilon, \dots, u_d^\epsilon)$ be the optimal allocation of u

$$(u_1^\epsilon, \dots, u_d^\epsilon) = A_{X_1, \dots, X_{i-1}, (1+\epsilon)X_i, X_{i+1}, \dots, X_d}(u),$$

then $(u_1^\epsilon, \dots, u_d^\epsilon)$ is the unique solution in \mathcal{U}_u^d of the following equations system:

$$\mathbb{E}[g'_i(u_i^\epsilon - (1+\epsilon)X_i) \mathbb{1}_{\{(1+\epsilon)X_i > u_i^\epsilon\}} \mathbb{1}_{\{S + \epsilon X_i \leq u\}}] = \mathbb{E}[g'_j(u_j^\epsilon - X_j) \mathbb{1}_{\{X_j > u_j^\epsilon\}} \mathbb{1}_{\{S + \epsilon X_i \leq u\}}], \quad \forall (i, j) \in \{1, \dots, d\}^2.$$

Since \mathcal{U}_u^d is a compact set of $(\mathbb{R}^+)^d$, we may consider a convergent subsequence $(u_1^{\epsilon_k}, \dots, u_d^{\epsilon_k})$ of $(u_1^\epsilon, \dots, u_d^\epsilon)$.

Since the penalties functions satisfy:

$$\exists \epsilon_0 > 0, \quad \forall \epsilon, |\epsilon| < \epsilon_0, \quad \mathbb{E}[\sup_{v \in [0, u]} |g'_k(v - (1 + \epsilon)X_k)|] < +\infty,$$

we use Lebesgue's dominated convergence Theorem to get:

$$\mathbb{E}[g'_i(\lim_{\epsilon \rightarrow 0} u_i^{\epsilon_k} - X_i) \mathbb{1}_{\{X_i > \lim_{\epsilon \rightarrow 0} u_i^{\epsilon_k}\}} \mathbb{1}_{\{S \leq u\}}] = \mathbb{E}[g'_j(\lim_{\epsilon \rightarrow 0} u_j^{\epsilon_k} - X_j) \mathbb{1}_{\{X_j > \lim_{\epsilon \rightarrow 0} u_j^{\epsilon_k}\}} \mathbb{1}_{\{S \leq u\}}], \quad \forall (i, j) \in \{1, \dots, d\}^2,$$

thereby $(\lim_{\epsilon \rightarrow 0} u_1^{\epsilon_k}, \dots, \lim_{\epsilon \rightarrow 0} u_d^{\epsilon_k})$ is a solution of Equation (3.1), because $\sum_{l=1}^d \lim_{\epsilon \rightarrow 0} u_l^{\epsilon_k} = \lim_{\epsilon \rightarrow 0} \sum_{l=1}^d u_l^\epsilon = u$, $(\lim_{\epsilon \rightarrow 0} u_1^{\epsilon_k}, \dots, \lim_{\epsilon \rightarrow 0} u_d^{\epsilon_k}) \in \mathcal{U}_u^d$.

From the solution uniqueness of (3.1) in \mathcal{U}_u^d , we deduce that: $\lim_{k \rightarrow \infty} u_i^{\epsilon_k} = u_i$ for all $i \in \{1, \dots, d\}$.

For all convergent subsequence of $(u_1^\epsilon, \dots, u_d^\epsilon)$ the limit point is (u_1, \dots, u_d) , we deduce that:

$$\lim_{\epsilon \rightarrow 0} (u_1^\epsilon, \dots, u_d^\epsilon) = (u_1, \dots, u_d).$$

□

Proposition 3.3.9 (Monotonicity). *Under Assumption H2, and for $(i, j) \in \{1, \dots, d\}^2$, such that $g_i = g_j$:*

$$X_i \leq_{st} X_j \Rightarrow u_i \leq u_j.$$

Proof. Let (u_1, \dots, u_d) be the optimal allocation $A_{X_1, \dots, X_i, \dots, X_d}(u)$ and $g = g_i = g_j$ be the common penalty function for X_i and X_j , under Assumption H2, the optimality condition (3.1) is written as follows:

$$\mathbb{E}[g'(u_i - X_i) \mathbb{1}_{\{X_i > u_i\}} \mathbb{1}_{\{S \leq u\}}] = \mathbb{E}[g'(u_j - X_j) \mathbb{1}_{\{X_j > u_j\}} \mathbb{1}_{\{S \leq u\}}].$$

Now if $X_i \leq_{st} X_j$, and since, $x \mapsto -g'(-(u_i - x)_+) \mathbb{1}_{\{S \leq u\}}$ is a non decreasing function on \mathbb{R}^+ , then:

$$\mathbb{E}[g'(u_i - X_j) \mathbb{1}_{\{X_j > u_j\}} \mathbb{1}_{\{S \leq u\}}] \leq \mathbb{E}[g'(u_i - X_i) \mathbb{1}_{\{X_i > u_i\}} \mathbb{1}_{\{S \leq u\}}].$$

We deduce that:

$$\mathbb{E}[g'(u_i - X_j) \mathbb{1}_{\{X_j > u_j\}} \mathbb{1}_{\{S \leq u\}}] \leq \mathbb{E}[g'(u_j - X_j) \mathbb{1}_{\{X_j > u_j\}} \mathbb{1}_{\{S \leq u\}}],$$

and since, g' is a non decreasing function, and the distributions are all continuous, that implies: $u_j \geq u_i$. □

By combining all the properties demonstrated in this section, we show that in the case of penalty functions $g_k(x) = |x| \forall k \in \{1, \dots, d\}$, and for a continuous random vector (X_1, \dots, X_d) , such that the joint density $f_{(X_k, S)}$ support contains $[0, u]^2$, for at least one $k \in \{1, \dots, d\}$, the optimal allocation by minimization of the indicators I and J is a symmetric riskless full allocation. It satisfies the properties of comonotonic additivity, positive homogeneity, translation invariance, monotonicity, and continuity. These properties are therefore desirable from an economic point of view. As they are satisfied by the proposed optimal allocation, this allocation method may be used for the economic capital allocation between the different branches of a group, in terms of their actual participation in the overall risk, taking into account both their marginal distributions and their dependence structure.

3.4 Discussion: What could be the best choice for a capital allocation?

In Section 3, we explained why the optimal allocation can be considered coherent from an economic point of view. Now, the most practical question is to define the best allocation method choice for an insurer.

The first goal of the Solvency 2 norms is the insurers' protection, while the ORSA approach is based on the minimization of risk at both the local and global levels. The classical methods of risk allocation give weight to each business line in the group risk. The optimal allocation is based on a global risk optimization. From this point of view, the capital allocation by minimizing a risk indicator seems more in line with the ORSA goals. Indeed, this method constitutes a second risk management level, after the solvency capital requirement determination.

The best allocation method choice depends finally on the risk aversion of the insurer. If the SCR is considered as the only risk management level, the classical methods of risk allocation are sufficient. If the insurer accepts to enhance his security level as it is the ORSA aim, the allocation by minimization of a multivariate risk indicator can be a good practical answer.

Conventional capital allocation methods are based on a chosen univariate risk measure and their properties are derived from those of this risk measure. It seems more coherent in a multivariate framework to use directly a multivariate risk indicator, not only for risk measurement, but also for capital allocation.

Another important criterion in the choice of the allocation method is the nature of the capital. The allocation of an investment capital may be different from that of a solvency capital.

Conclusion

In this article, we have shown that the capital allocation method by minimization of some multivariate risk indicators can be considered as coherent from an economic point of view. This method also illustrates the importance of the choice of the risky business in the portfolio and its impact on the management of the overall company's capital.

Finally, the choice of a capital allocation method remains a complex and crucial exercise; some methods may be better suited to deal with specific issues, others can lead to dangerously wrong financial decisions. In the case of the proposed optimal capital allocation, the risk management is at the heart of the allocation process, and thus when the company allocates its capital and reduces its overall risk at the same time. In the case of risk aversion this is reflected in the choice of the multivariate risk indicator to be minimized. That is why we think that from a risk management point of view, this method can be considered as more flexible.

Acknowledgment

We are grateful to Corina Constantinescu for valuable suggestions on our paper. We also thank the editor and the reviewer for her/his valuable comments which helped to improve the quality of the manuscript.

Bibliography

- [1] P. Artzner, F. Delbaen, J-M. Eber, and D. Heath. Coherent measures of risk. *Mathematical Finance*, 9(3):203–228, 1999. 4, 39, 44, 47, 130, 131, 177
- [2] D. Balog. Capital allocation in financial institutions: the euler method. IeHAS discussion papers, Institute of Economics, Hungarian Academy of Sciences, June 2011. 126
- [3] F. Bellini and V. Bignozzi. On elicitable risk measures. *Quantitative Finance*, 15(5):725–733, 2015. 6, 40, 41, 42, 43, 44, 128
- [4] K. Borch. Equilibrium in a reinsurance market. *Econometrica*, (30):424–444, 1962. 9, 131, 167
- [5] M.K. Brunnermeier and P. Cheridito. Measuring and allocating systemic risk. *Available at SSRN 2372472*, 2014. 130
- [6] A. Buch and G. Dornflleitner. Coherent risk measures, coherent capital allocations and the gradient allocation principle. *Insurance: Mathematics and Economics*, 42(1):235 – 242, 2008. 126
- [7] A. Buch, G. Dornflleitner, and M. Wimmer. Risk capital allocation for rorac optimization. *Journal of Banking & Finance*, 35(11):3001–3009, August 2011. 126
- [8] J. Cai and H. Li. Dependence properties and bounds for ruin probabilities in multivariate compound risk models. *Journal of Multivariate Analysis*, 98(4):757 – 773, 2007. 126, 144
- [9] P. Cénac, S. Loisel, V. Maume-Deschamps, and C. Prieur. Risk indicators with several lines of business: comparison, asymptotic behavior and applications to optimal reserve allocation. *Annales de l’ISUP*, 58(3), 2014. 17, 27, 126, 144, 148, 151, 162
- [10] P. Cénac, V. Maume-Deschamps, and C. Prieur. Some Multivariate risk indicators: Minimization by using a Kiefer-Wolfowitz approach to the mirror stochastic algorithm . *Statistics and Risk Modeling*, 29(1):47–71, 2012. 29, 126, 127, 128, 136, 143, 144, 145, 146
- [11] M. Denault. Coherent allocation of risk capital. *Journal of risk*, 4:1–34, January 2001. 126, 130, 144
- [12] J. Dhaene, E.A. Valdez, A. Tsanakas, and S. Vanduffel. Optimal capital allocation principles. *The Journal of Risk and Insurance*, 79(1):1–28, 2012. 17, 126, 144, 145
- [13] T. Fischer. Risk capital allocation by coherent risk measures based on one-sided moments. *Insurance: Mathematics and Economics*, 32(1):135 – 146, 2003. 126
- [14] H. Föllmer, A. Schied, and T.J. Lyons. Stochastic finance. an introduction in discrete time. *The Mathematical Intelligencer*, 26(4):67–68, 2004. 131
- [15] M. Fréchet. Sur les tableaux de corrélation dont les marges sont données. *Annales de l’université de Lyon, Section.A, Series 3.*, (14):53–77, 1951. 9, 131, 167
- [16] T. Gneiting. Making and Evaluating Point Forecasts . *Journal of the American Statistical Association*, 106(494):746–762, June 2011. 6, 40, 41, 42, 43, 44, 128
- [17] A. H. Hamel and F. Heyde. Duality for set-valued measures of risk. *SIAM Journal on Financial Mathematics*, 1(1):66–95, 2010. 53, 128
- [18] W. Hoeffding. Masstabinvariante Korrelationstheorie. *Schriften des mathematischen Instituts und des Instituts für angewandte Mathematik der Universität Berlin*, (5):179–233, 1940. 9, 131, 167
- [19] E. Jouini, M. Meddeb, and N. Touzi. Vector-valued coherent risk measures. *Finance and Stochastics*, 8(4):531–552, 2004. 4, 39, 47, 53, 128, 144, 177

-
- [20] M. Kalkbrenner. An axiomatic characterization of capital allocations of coherent risk measures. *Quantitative Finance*, 9(8):961–965, 2009. 126, 131
- [21] V. Maume-Deschamps, D. Rullière, and K. Said. Impact of dependence on some multivariate risk indicators. *Methodology and Computing in Applied Probability*, 2016. 127, 143
- [22] V. Maume-Deschamps, D. Rullière, and K. Said. On a capital allocation by minimization of some risk indicators. *European Actuarial Journal*, pages 1–20, 2016. 40, 47, 125, 143, 144, 145, 146, 156
- [23] M. Shaked and J.G. Shanthikumar. *Stochastic Orders*. Springer Series in Statistics, 2007. 8, 131, 156
- [24] I. Steinwart, C. Pasin, R. C. Williamson, and S. Zhang. Elicitation and Identification of Properties. *Workshop and Conference Proceedings*, 34:1–45, 2014. 40, 42, 128
- [25] D. Tasche. Allocating portfolio economic capital to sub-portfolios. In *Economic Capital: A Practitioner Guide, Risk Books*, pages 275–302. Risk Books, 2004. 126
- [26] D. Tasche. Euler allocation: Theory and practice. Technical Report arXiv:0708.2542, Aug 2007. 126, 131, 144
- [27] D. Tasche. Capital allocation to business units and sub-portfolios: the euler principle. Papers 0708.2542, arXiv.org, 2008. 126
- [28] G. Zanjani. An economic approach to capital allocation. *Journal of Risk and Insurance*, 77(3):523–549, 2010. 126
- [29] J.F. Ziegel. Coherence and elicibility. *arXiv preprint arXiv:1303.1690*. 40, 44, 128

Chapitre 4

Impact de la dépendance sur l'allocation du capital par minimisation d'indicateurs multivariés de risque

Impact of dependence on some multivariate risk indicators

Ce chapitre est composé de l'article Maume-Deschamps et al. (2016) [22] qui porte le même titre, accepté pour publication dans le journal *Methodology and Computing in Applied Probability*, publié en ligne depuis le 29 février 2016.

<http://link.springer.com/article/10.1007/s11009-016-9489-4>

abstract

The minimization of some multivariate risk indicators may be used as an allocation method, as proposed in Cénac et al. (2012) [6]. The aim of capital allocation is to choose a point in a simplex, according to a given criterion. In Maume-Deschamps et al. (2016) [23], it is proved that the proposed allocation technique satisfies a set of coherence axioms. In the present one, we study the properties and asymptotic behavior of the allocation for some distribution models. We also analyze the impact of the dependence structure on the allocation using some copulas.

Keywords : Multivariate risk indicators, dependence modeling, sub-exponential distributions, risk theory, optimal capital allocation, copulas..

Introduction

Natural phenomena, financial events and risks are usually modeled through random vectors or processes. In these fields, considering dependencies between random variables is necessary. In actuarial science, this issue has led to the development of multivariate risk theories. Indeed, an insurance company is generally exposed to several risks which cannot be assumed to be independent. It is therefore necessary to adopt a multivariate approach that takes into account both the marginal structures of risks and their dependence structure.

Multivariate risk theory is based on dependence modeling and includes multivariate ruin probabilities and multivariate risk measures. In the univariate case, ruin probability has been widely studied since the beginning of the 20th century (Lundberg (1903) [20], Cramer (1930) [10]). In multivariate contexts, several definitions are proposed for ruin probability. Hult and Lindskog (2006) [17] defined a multivariate ruin probability based on a notion of ruin sets. Cai and Li (2007) [4] defined different finite-time ruin probabilities. Risk measure theory was also enriched by some multivariate risk measure definitions. Jouini et al. (2004) [19] defined vector-valued coherent risk measures. Cousin and Di Bernardino (2013) [9] introduced some multi-dimensional extensions of usual univariate risk measures such as the Value-at-Risk. These measures are multi-dimensional valued, and thus cannot take benefit from a full order. Dhaene et al. (2012) introduced real valued family for risk measures for mono-periodic multivariate processes ([13]). In multi-periodic contexts, Cénac et al. (2012) [6] defined new multivariate risk indicators as sums of expected local ruin amounts using penalty functions.

Capital allocation may be seen as a direct application of multivariate risk theory. Allocating a capital in actuarial contexts means distributing a positive quantity representing an allocation capital in a simplex. For the insurance industry the calculation of the regulatory economic capital, which is called the Solvency Capital Requirement, is well controlled and its methodology is almost imposed by the supervisory authorities of the sector. Nevertheless, the allocation of this capital may be considered as an internal exercise for each company, and constitutes a management choice whose success is a key factor for firm performance optimization. It can also be seen as an indicator of its good governance.

Several allocation methods were proposed in the actuarial literature. Some are designed by Euler or Shapley principles (see as examples Tasche (2007) [27] for the Euler allocation principle and Denaux (2001) [12] for the Shapley allocation principle). Others are based on the minimization of some ruin probabilities or multivariate risk indicators. In this context, Cénac et al. (2012) [6] proposed a capital allocation by minimizing some multivariate risk indicators. They presented a numerical optimization algorithm to find the optimal allocation in some general cases. In [5] properties and asymptotic behavior of the allocation are studied in some bivariate cases. The impact of dependence on capital allocation was studied for certain cases of Euler's method by Bargès et al. (2009) [2] and Cossette et al. (2012) [8].

In this paper, we give a further study of the allocation method by minimizing some multivariate risk indicators. The idea of minimizing multivariate risk indicators seems from our point of view more suited for the ORSA* approach of Solvency 2 European norms. In a recent paper [23], we show that this method satisfies a set of coherence axioms. In the present paper, we study the impact of marginal distributions and dependence structure on the optimal allocation. We generalize the results presented in bivariate cases by Cénac et al. (2014) [5]. Moreover, we study the impact of the dependence structure on the allocation using some parametric copulas. Our main results are explicit formulas for various classes of multivariate processes, as well as the limit behavior of the allocation as the capital goes to infinity.

The paper is organized as follows. In the first section, we recall the allocation method by minimizing multivariate risk indicators. Section 2 is a study of the independence case and the impact of the

*. Own Risk and Solvency Assessment

marginals on the capital allocation. We present some explicit formulas obtained for some particular models, and we discuss the asymptotic behavior of the optimal allocation for these models. The impact of the dependence structure on the allocation is studied in Sections 3, 4 and 5. Section 3 presents the dependence case with asymptotic independence. The asymptotic dependence case is studied in Section 4 through the analysis of two models. Section 5 is devoted to the perfect dependence case.

4.1 Optimal allocation

In this section, we recall one capital allocation principle that consists in minimizing some risk indicators. This method was introduced in Cénac et al. (2012) [6]. It is based on the minimization of some multivariate risk indicators.

In a multivariate risk framework, we consider an insurance group composed of d branches or business lines. We denote by u the initial capital of the group. Let X_p be a vectorial risk process $X_p = (X_p^{(1)}, \dots, X_p^{(d)})$, where $X_p^{(k)}$ corresponds to the losses of the k^{th} business line during the p^{th} period. We denote by $R_p^{(k)}$ the reserve of the k^{th} line at time p , so: $R_p^{(k)} = u_k - \sum_{\ell=1}^p X_\ell^{(k)}$, where $u_k \in \mathbb{R}^+$ is the initial capital of the k^{th} business line, then $u = u_1 + \dots + u_d$.

Cénac et al. (2012) [6] defined the two following multivariate risk indicators, given penalty functions g_k :

– The indicator I :

$$I(u_1, \dots, u_d) = \sum_{k=1}^d \mathbb{E} \left(\sum_{p=1}^n g_k(R_p^{(k)}) \mathbb{1}_{\{R_p^{(k)} < 0\}} \mathbb{1}_{\{\sum_{j=1}^d R_p^{(j)} > 0\}} \right),$$

– The indicator J :

$$J(u_1, \dots, u_d) = \sum_{k=1}^d \mathbb{E} \left(\sum_{p=1}^n g_k(R_p^{(k)}) \mathbb{1}_{\{R_p^{(k)} < 0\}} \mathbb{1}_{\{\sum_{j=1}^d R_p^{(j)} < 0\}} \right),$$

$g_k : \mathbb{R}^- \rightarrow \mathbb{R}^+$ are C^1 , convex functions with $g_k(0) = 0$, $g_k(x) \geq 0$ for $x < 0$, $k = 1, \dots, d$. They represent the cost that each branch has to pay when it becomes insolvent while the group is solvent for the I indicator, or while the group is also insolvent in the case of the J indicator.

These multivariate risk indicators are a generalization of the mono-periodic indicator family introduced in Dhaene et al.(2012)[13].

The indicator I represents the expected sum of penalty amounts of local ruins, knowing that the group remains solvent. In the case of the indicator J , the local ruin severities are taken into account only in the case of group insolvency.

The idea of the optimal allocation is to allocate some capital u by minimizing these indicators. This is finding an allocation vector (u_1, \dots, u_d) that minimizes the indicator such as $u = u_1 + \dots + u_d$, where u is the initial capital that needs to be shared among all branches. In our paper [23], we explained why this capital allocation method can be considered as economically coherent.

By using a stochastic optimization algorithm, one can estimate the minimum of these risk indicators. Cénac et al. (2012) [6] propose a Kiefer-Wolfowitz version of the mirror algorithm as a convergent algorithm under general assumptions to find optimal allocation minimizing the indicator I . This algorithm is effective to solve the optimal allocation problem, especially for a large number of business lines, and for allocation over several periods.

Since the solvency capital requirement (SCR) is calculated for a one year time horizon only, in this paper, we have focused on the case of mono-periodic allocation.

The risk X_k corresponds to the losses of the k^{th} branch during one period. It is a positive random variable in our context. We denote by u the initial capital of the firm, it represents the allocation capital and it can be, as an example, the SCR or another investment capital. The quantity u_k represents the portion of capital allocated to the k^{th} branch, then $\sum_{i=1}^d u_i = u$. The simplex $\mathcal{U}_u^d = \{v = (v_1, \dots, v_d) \in [0, u]^d, \sum_{i=1}^d v_i = u\}$ is the set of all possible allocations. For all $i \in \{1, \dots, d\}$ we denote $\alpha_i = \frac{u_i}{u}$, then, $\sum_{i=1}^d \alpha_i = 1$. For $(u_1, \dots, u_d) \in \mathcal{U}_u^d$, we also define the reserve of the k^{th} business line at the end of the period as $R^{(k)} = u_k - X_k$. The aggregate sum of risks is $S = \sum_{i=1}^d X_i$, and let $S^{(-i)} = \sum_{j=1; j \neq i}^d X_j$ for all $i \in \{1, \dots, d\}$. Finally, F_Z is the distribution function of a random variable Z , \bar{F}_Z is its survival function and f_Z its density function.

Now we focus on the optimality condition for the indicators I and J if they are strictly convex.

For an initial capital u , and an optimal allocation minimizing the multivariate risk indicator I , we look for $u^* \in \mathbb{R}_+^d$ such that:

$$I(u^*) = \inf_{v_1 + \dots + v_d = u} I(v), \quad v \in \mathbb{R}_+^d.$$

We assume that the functions g_k are differentiable and for all $k \in \{1, \dots, d\}$, the derivative $g'_k(u_k - X_k)$ admits a moment of order one, and the random vector (X_k, S) has a joint density distribution denoted by $f_{(X_k, S)}$. In order to ensure the existence and uniqueness of the indicators' minimum in \mathcal{U}_u^d , we assume that for at least one $k \in \{1, \dots, d\}$, g_k is convex. We also assume that the joint density $f_{(X_k, S)}$ support contains $[0, u]^2$ (see [6]). In this case, the indicators I and J are strictly convex and admit a unique minimum. Under these assumptions, the risk indicators I and J are differentiable, and in this case, we can calculate the following gradients:

$$\begin{aligned} (\nabla I(v))_i &= \sum_{k=1}^d \int_{v_k}^{+\infty} g_k(v_k - x) f_{X_k, S}(x, u) dx + \mathbb{E}[g'_i(v_i - X_i) \mathbb{1}_{\{X_i > v_i\}} \mathbb{1}_{\{S \leq u\}}] \\ \text{and, } (\nabla J(v))_i &= \sum_{k=1}^d \int_{v_k}^{+\infty} g_k(v_k - x) f_{X_k, S}(x, u) dx + \mathbb{E}[g'_i(v_i - X_i) \mathbb{1}_{\{X_i > v_i\}} \mathbb{1}_{\{S \geq u\}}]. \end{aligned}$$

Using the Lagrange multipliers method, we obtain an optimality condition verified by the unique solution to this optimization problem:

$$\mathbb{E}[g'_i(u_i - X_i) \mathbb{1}_{\{X_i > u_i\}} \mathbb{1}_{\{S \leq u\}}] = \mathbb{E}[g'_j(u_j - X_j) \mathbb{1}_{\{X_j > u_j\}} \mathbb{1}_{\{S \leq u\}}], \quad \forall j \in \{1, \dots, d\}^2. \quad (4.1)$$

A natural choice for penalty functions is the ruin severity: $g_k(x) = |x|$.

In Maume-Deschamps et al. (2016) [23], we showed that in the case of penalty functions $g_k(x) = |x| \forall k \in \{1, \dots, d\}$, and for continuous random vector (X_1, \dots, X_d) such that the joint density $f_{(X_k, S)}$ support contains $[0, u]^2$, for at least one $k \in \{1, \dots, d\}$, the optimal allocation by minimization of the indicators I and J is a symmetric riskless full allocation. It also satisfies the properties of comonotonic additivity, positive homogeneity, translation invariance, monotonicity, and continuity.

We may write the indicators as follows:

$$\begin{aligned} I(u_1, \dots, u_d) &= \sum_{k=1}^d \mathbb{E} \left(|R^{(k)}| \mathbb{1}_{\{R^{(k)} < 0\}} \mathbb{1}_{\{\sum_{i=1}^d R^{(i)} \geq 0\}} \right) \\ &= \sum_{k=1}^d \mathbb{E} \left((X_k - u_k) \mathbb{1}_{\{X_k > u_k\}} \mathbb{1}_{\{\sum_{i=1}^d X_i \leq u\}} \right) = \sum_{k=1}^d \mathbb{E} \left((X_k - u_k)^+ \mathbb{1}_{\{S \leq u\}} \right), \end{aligned}$$

and,

$$\begin{aligned} J(u_1, \dots, u_d) &= \sum_{k=1}^d \mathbb{E} \left(|R^{(k)}| \mathbb{1}_{\{R^{(k)} < 0\}} \mathbb{1}_{\{\sum_{i=1}^d R^{(i)} \leq 0\}} \right) \\ &= \sum_{k=1}^d \mathbb{E} \left((X_k - u_k) \mathbb{1}_{\{X_k > u_k\}} \mathbb{1}_{\{\sum_{i=1}^d X_i \geq u\}} \right) = \sum_{k=1}^d \mathbb{E} \left((X_k - u_k)^+ \mathbb{1}_{\{S \geq u\}} \right). \end{aligned}$$

In the special case where $g_k(x) = |x|$, 4.1 becomes:

$$\mathbb{P}(X_i > u_i, S \leq u) = \mathbb{P}(X_j > u_j, S \leq u), \forall (i, j) \in \{1, 2, \dots, d\}^2. \quad (4.2)$$

For the J indicator, this condition can be written:

$$\mathbb{P}(X_i > u_i, S \geq u) = \mathbb{P}(X_j > u_j, S \geq u), \forall (i, j) \in \{1, 2, \dots, d\}^2. \quad (4.3)$$

Some explicit and semi-explicit formulas for the optimal allocation can be obtained with this optimality condition. Our problem reduces to the study of this allocation depending on the nature of the distributions of the risk X_k and on the form of dependence between them.

From now on, we will call *I-allocation* the optimal capital allocation by minimizing the multivariate risk indicator I , and analogously the capital allocation by minimizing the indicator J will be called *J-allocation*.

In the next sections, we focus on the study of the impact of dependence on the allocation of capital. We first examine the case of independence for exponential marginal laws and some sub-exponential marginal laws. Then, we study different types of dependence using some copulas and numerical illustrations. We distinguish the cases of asymptotic independence, asymptotic dependence and perfect dependence. In fact, the asymptotic behavior is useful to be analyzed from a practical point of view, it is natural to separate the kind of dependence that takes into account the asymptotical dependence from the asymptotic independence models. The perfect dependence case is also important to be examined; it is extremely interesting from an actuarial point of view to have an idea on the allocation behavior in presence of comonotonic risks.

In order to separate these families we recall the definition of the upper tail dependence coefficient as presented in Joe (1997) [18], for bivariate random variables (X, Y) of a continuous marginal distributions

$$\lambda_U = \lim_{u \rightarrow 1^-} \mathbb{P}(Y > F_Y^{-1}(u) | X > F_X^{-1}(u)),$$

by Sklar's Theorem, there exist a unique copula $C(., .) : [0, 1] \times [0, 1] \rightarrow \mathbb{R}$ such that $F_{X,Y}(x, y) = C(F_X(x), F_Y(y))$ for all $(x, y) \in \mathbb{R}^2$. The upper tail dependence coefficient can be expressed in terms of copula,

$$\lambda_U = \lim_{u \rightarrow 1^-} \frac{1 - 2u + C(u, u)}{1 - u},$$

when the limit exists.

Analogously, the lower tail dependence coefficient is defined as

$$\lambda_L = \lim_{u \rightarrow 0^+} \mathbb{P}(Y \leq F_Y^{-1}(u) | X \leq F_X^{-1}(u)) = \lim_{u \rightarrow 0^+} \frac{C(u, u)}{u},$$

when the limit exists. We shall consider only the upper tail dependence.

Coles et al. (1999) [7] provides an asymptotically simple equivalent version of λ_U function of the dependence structure C . De Luca and Rieviccio (2009) [11] proposed a multidimensional extension to the upper tail dependence coefficient, defined for a d -random vector (X_1, \dots, X_d) as

$$\lambda_U^i = \lim_{u \rightarrow 1^-} \mathbb{P}(F_{X_i}(X_i) > u | F_{X_j}(X_j) > u, j = \{1, \dots, d\}, j \neq i).$$

The impact of dependence on the capital allocation will be studied in four cases. We start in the next section by examining the impact of marginals nature on the allocation in the independence context. The second case will be the dependence with asymptotic independence case $\lambda_U = 0$. The case of dependence with asymptotic dependence $\lambda_U > 0$ will be the subject of section 4, and the perfect dependence case, where $\lambda_U = 1$, will be studied in the last section.

4.2 Independence models

In this section we generalize the results presented in dimension 2 by Cénac et al. (2014) in the first section of their paper [5] to higher dimension.

The results presented here give explicit forms to the optimal allocation for some specific distributions. This could be used as a benchmark to test optimization algorithms convergence.

The asymptotic behavior of capital allocation is useful from a practical point of view for insurers. It presents an important information in the ruin theory context and gives an estimate of ruin probability. The solvency capital requirement level is set as the 99.5%-quantile, the study of the asymptotic behavior gives an idea on the case of capital exceeding that level. We devote then a part of this sub-section to the asymptotic behavior study. We also get some asymptotic results, when the capital u goes to infinity. We study both the exponential and the sub-exponential cases, and we determine the difference between between asymptotic allocations for exponential and Pareto distributions cases.

We consider from now on that the penalty functions are identical and equal to the severity of local ruin $g_k(x) = g(x) = |x|, \forall k \in \{1, 2, \dots, d\}$. The optimality conditions for minimizing the multivariate risk indicators I and J are given respectively by Equations (4.2) and (4.3).

Recall that $\alpha_i = \frac{u_i}{u} \in [0, 1]$, so that when $u \rightarrow +\infty$, we may consider convergent subsequences in the proofs below. By abuse of notation, we consider $\lim_{u \rightarrow +\infty} \alpha_i$. In fact, we consider a convergent subsequence and get the existence of the limit by obtaining the uniqueness of the limit point.

4.2.1 Independent exponentials

Assume X_1, X_2, \dots, X_d are independent exponential random variables with respective parameters $0 < \beta_1 < \beta_2 < \dots < \beta_d$. Remark that in the particular case where $\beta_1 = \beta_2 = \dots = \beta_d$, the optimal allocation is given by $u_1 = \dots = u_d = u/d$.

Proposition 4.2.1 (*I*-allocation, independent exponentials case).

The allocation minimizing the risk indicator I is the unique solution in \mathcal{U}_u^d , of the following equations system:

$$h(\beta_i \alpha_i) - h(\beta_j \alpha_j) - \sum_{\ell=1}^d A_\ell h(\beta_\ell) [h(\alpha_i \cdot (\beta_i - \beta_\ell)) - h(\alpha_j \cdot (\beta_j - \beta_\ell))] = 0, \forall (i, j) \in \{1, 2, \dots, d\}^2, \quad (4.4)$$

where h is the function defined by $h(x) = \exp(-u \cdot x)$, and A_ℓ denotes the constants $A_\ell = \prod_{j=1, j \neq \ell}^d \frac{\beta_j}{\beta_j - \beta_\ell}$, for $\ell = 1, \dots, d$.

When the capital u goes to infinity, the asymptotic optimal allocation satisfies:

$$\lim_{u \rightarrow \infty} \left(\frac{u_1}{u}, \frac{u_2}{u}, \dots, \frac{u_d}{u} \right) = \left(\frac{1}{\beta_i} \right)_{i=1,2,\dots,d}.$$

Proof.

If $X_i \sim \mathcal{E}(\beta_i)$ are independent exponential random variables, then $S^{(-i)} = \sum_{j=1; j \neq i}^d X_j$ have a generalized Erlang distribution with parameters $(\beta_1, \beta_2, \dots, \beta_{i-1}, \beta_{i+1}, \dots, \beta_d)$, so we write:

$$\begin{aligned} \mathbb{P}\left(X_i > u_i, \sum_{j=1}^d X_j \leq u\right) &= \mathbb{P}(X_i > u_i) - \mathbb{P}\left(X_i > u_i, \sum_{j=1}^d X_j > u\right) \\ &= \bar{F}_{X_i}(u_i) - \bar{F}_{X_i}(u) - \int_{u_i}^u \bar{F}_{S^{(-i)}}(u-s) f_{X_i}(s) ds \\ &= h(\beta_i \alpha_i) - h(\beta_i) - \sum_{\ell=1}^d A_\ell h(\beta_\ell) h(\alpha_i \cdot (\beta_i - \beta_\ell)) + \sum_{\ell=1}^d A_\ell h(\beta_i) \\ &= h(\beta_i \alpha_i) - \sum_{\ell=1}^d A_\ell h(\beta_\ell) h(\alpha_i \cdot (\beta_i - \beta_\ell)), \end{aligned}$$

because, $\bar{F}_{X_i}(x) = e^{-\beta_i x}$, $\bar{F}_{S^{(-i)}}(x) = \sum_{\ell=1, \ell \neq i}^d \left(\prod_{j=1, j \neq \ell, j \neq i}^d \frac{\beta_j}{\beta_j - \beta_\ell} \right) e^{-\beta_\ell x}$ and $\sum_{\ell=1}^d A_\ell = 1$.

The optimal allocation is the unique solution in \mathcal{U}_u^d , of the following equations system:

$$\mathbb{P}\left(X_i > u_i, \sum_{k=1}^d X_k \leq u\right) = \mathbb{P}\left(X_j > u_j, \sum_{k=j}^d X_k \leq u\right), \forall (i, j) \in \{1, 2, \dots, d\}^2,$$

which leads to (4.4).

Equations system (4.4) is equivalent to:

$$\forall (i, j) \in \{1, 2, \dots, d\}^2, \quad h(\beta_i \alpha_i) \left[1 - \sum_{\ell=1}^d A_\ell h((1 - \alpha_i) \beta_\ell)\right] = h(\beta_j \alpha_j) \left[1 - \sum_{\ell=1}^d A_\ell h((1 - \alpha_j) \beta_\ell)\right]. \quad (4.5)$$

Remark that for all $i \in \{1, 2, \dots, d\}$, $\limsup_{u \rightarrow \infty} \frac{u_i}{u} < 1$ because if this result was not satisfied then there would exist $i \in \{1, \dots, d\}$ such that $\lim_{u \rightarrow \infty} \frac{u_i}{u} = \lim_{u \rightarrow \infty} \alpha_i = 1$, taking if necessary a convergent subsequence of α_i . For all $j \neq i$, $\lim_{u \rightarrow \infty} \frac{u_j}{u} = \lim_{u \rightarrow \infty} \alpha_j = 0$, and Equations system (4.2) cannot be satisfied in this case.

Equations system (4.5) is equivalent to:

$$\forall (i, j) \in \{1, 2, \dots, d\}^2, \quad h(\beta_i \alpha_i - \beta_j \alpha_j) = \frac{1 - \sum_{\ell=1}^d A_\ell h((1 - \alpha_j) \beta_\ell)}{1 - \sum_{\ell=1}^d A_\ell h((1 - \alpha_i) \beta_\ell)},$$

the right side of the last equations system tends to 1 when u tends to ∞ , therefore, we deduce that $\lim_{u \rightarrow \infty} h(\beta_i \alpha_i - \beta_j \alpha_j) = 1$ and consequently:

$$\forall (i, j) \in \{1, 2, \dots, d\}^2, \quad \lim_{u \rightarrow \infty} \alpha_i = \frac{\beta_j}{\beta_i} \lim_{u \rightarrow \infty} \alpha_j,$$

then, for all $i \in \{1, 2, \dots, d\}$:

$$\lim_{u \rightarrow \infty} \alpha_i = \frac{1}{\beta_i} \frac{1}{\sum_{j=1}^d \frac{1}{\beta_j}}.$$

□

The resulting system for the optimal allocation is a system of nonlinear equations, which can be solved numerically.

Remark 4.2.1. Based on the above result, we can conclude that asymptotically:

- since $\beta_i < \beta_j$ we have $\alpha_i > \alpha_j$, this means that we allocate more capital to the most risky business line.
- α_i is a decreasing function of β_i .
- α_j is an increasing function of β_i for $j \neq i$.

In order to illustrate the speed of the convergence to the asymptotic allocation, we present in Figure 4.1 a numerical illustration in a bivariate case, for the I -allocation.

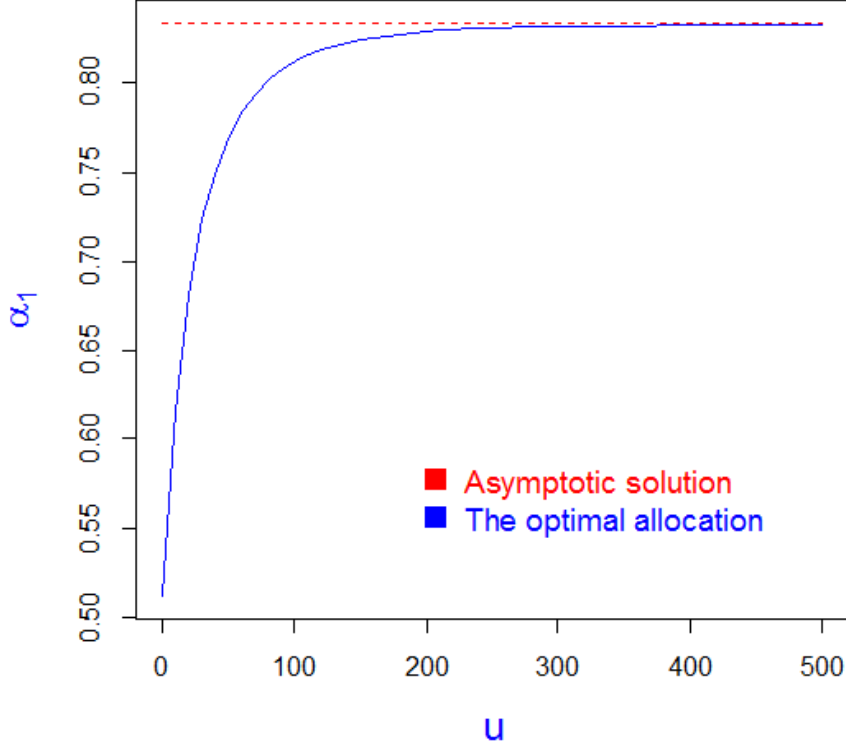


Figure 4.1: Speed of convergence of allocation toward the asymptotic allocation in the bivariate independent exponentials case. Allocation percentage α_1 as a function of initial capital u . $\beta_1 = 0.05$, $\beta_2 = 0.25$

For a tolerated error of 1%, the convergence level is near $u = 100$ which is less than the 99.5%-quantile of X_1 . In our numerical illustration, the value of this quantile is 105.97.

Proposition 4.2.2 (J -allocation, independent exponentials case).

The allocation minimizing the risk indicator J is the unique solution in \mathcal{U}_u^d , of the following equations system:

$$\forall (i, j) \in \{1, 2, \dots, d\}^2, \quad \sum_{\ell=1}^d A_{\ell} h(\beta_{\ell}) [h(\alpha_i \cdot (\beta_i - \beta_{\ell})) - h(\alpha_j \cdot (\beta_j - \beta_{\ell}))] = 0. \quad (4.6)$$

When the capital u goes to infinity, the optimal allocation minimizing the risk indicator J is the following:

$$\lim_{u \rightarrow \infty} \frac{u_1}{u} = 1 \text{ and } \lim_{u \rightarrow \infty} \frac{u_j}{u} = 0 \quad \forall j \in \{2, 3, \dots, d\}.$$

Proof.

The proof of (4.6) is similar to that of Proposition 4.2.1.

Equations system (4.6) is equivalent to the following one:

$$\forall (i, j) \in \{1, 2, \dots, d\}^2, \quad \sum_{\ell=1}^d A_\ell \cdot h((1 - \alpha_i) \cdot (\beta_\ell - \beta_i)) = \sum_{\ell=1}^d A_\ell \cdot h(\alpha_j \cdot (\beta_j - \beta_\ell) + \beta_\ell - \beta_i).$$

If $\overline{\lim} \frac{u_1}{u} < 1$, as u goes to $+\infty$ in the equations of the previous system for $i = 1$ we get, taking if necessary a convergent subsequence,

$$\forall j \in \{2, 3, \dots, d\}, \quad \lim_{u \rightarrow \infty} \sum_{\ell=1}^d A_\ell \cdot h(\alpha_j \cdot (\beta_j - \beta_\ell) + \beta_\ell - \beta_1) = A_1. \quad (4.7)$$

The first terms of these equations can be decomposed into three parts as follows:

$$\begin{aligned} \lim_{u \rightarrow \infty} \sum_{\ell=1}^d A_\ell \cdot h(\alpha_j \cdot (\beta_j - \beta_\ell) + \beta_\ell - \beta_1) &= A_1 \cdot \lim_{u \rightarrow \infty} h(\alpha_j \cdot (\beta_j - \beta_1)) \\ &+ \sum_{\ell=2}^j A_\ell \cdot \lim_{u \rightarrow \infty} h(\alpha_j \cdot (\beta_j - \beta_\ell) + \beta_\ell - \beta_1) \\ &+ \sum_{\ell=j+1}^d A_\ell \cdot \lim_{u \rightarrow \infty} h(\alpha_j \cdot (\beta_j - \beta_\ell) + \beta_\ell - \beta_1). \end{aligned}$$

For all, $j > 1$, $\sum_{\ell=2}^j A_\ell \cdot \lim_{u \rightarrow \infty} h(\alpha_j \cdot (\beta_j - \beta_\ell) + \beta_\ell - \beta_1) = 0$, because $\beta_\ell - \beta_1 > 0$ and $\beta_j - \beta_\ell \geq 0$ for $\ell \in$

$\{2, 3, \dots, j\}$. Moreover, $\sum_{\ell=j+1}^d A_\ell \cdot \lim_{u \rightarrow \infty} h(\alpha_j \cdot (\beta_j - \beta_\ell) + \beta_\ell - \beta_1) = 0$, because for all $\ell \in \{j+1, j+2, \dots, d\}$,

$\alpha_j \cdot (\beta_j - \beta_\ell) + \beta_\ell - \beta_1 = (\beta_\ell - \beta_j)(1 - \alpha_j) + \beta_j - \beta_1 > 0$. So that 4.7 leads to $\forall j \in \{1, \dots, d\}$ $\lim_{u \rightarrow \infty} h(\alpha_j \cdot (\beta_j - \beta_1)) = 1$. We deduce that $\forall j \in \{2, 3, \dots, d\}$; $\alpha_j \stackrel{u \rightarrow \infty}{\sim} o(\frac{1}{u})$. This contradicts the necessary condition: $\lim_{u \rightarrow \infty} \sum_{\ell=1}^d \alpha_\ell = 1$. \square

4.2.2 Some distributions of the sub-exponential family

In most cases of risk distributions, we cannot give explicit or semi-explicit optimal allocation formulas, the difficulty comes from the lack of a simple form of the risks sum S and its joint distribution with each risk X_j . In this section, we present asymptotic results (as u goes to infinity) for the optimal allocation in the case of some distributions of the sub-exponential family. In this way, we generalize results of [5] to higher dimension.

We recall that a non negative random variable belongs to the sub-exponential family definition if its survival function satisfies,

$$\frac{\overline{F^{*2}}(x)}{\overline{F}(x)} \xrightarrow{x \rightarrow +\infty} 2,$$

where $\overline{F^{*2}}$ is the convolution of \overline{F} .

In Asmussen (2000) [1], it is proven that sub-exponential distributions satisfy also the following relation, for all $d \in \mathbb{N}^*$:

$$\frac{\overline{F^{*d}}(x)}{\overline{F}(x)} \xrightarrow{x \rightarrow +\infty} d,$$

where $\overline{F^{*d}}$ is the d^{th} convolution of \overline{F} .

We shall use the following theorem proved in Cénac and al.

Theorem 4.2.3 (Sub-exponential distributions [5]). *Let X be a random variable with sub-exponential distribution F_X , Y a random variable with support \mathbb{R}^+ , independent of X , and $(u, v) \in (\mathbb{R}^+)^2$, such that:*

- there exists $0 < \kappa_1 < \kappa_2 < 1$ such that for u large enough, $\kappa_1 \leq \frac{v}{u} \leq \kappa_2$,
 - $\frac{\bar{F}_X(y)}{\bar{F}_X(x)} \stackrel{x \rightarrow +\infty}{\asymp} O(1)$, if $y = \Theta(x)$.
- Then,

$$\lim_{u \rightarrow \infty} \frac{\mathbb{P}(X \geq v, X + Y \geq u)}{\bar{F}_X(u)} = 1.$$

The asymptotic behavior

Here we examine the asymptotic behavior of the optimal allocation by minimizing the indicators I and J in the cases of some sub-exponential distributions.

In what follows, (u_1, \dots, u_d) denotes the optimal allocation of u associated to the risk indicator I .

Lemma 4.2.4. *Let (X_1, X_2, \dots, X_d) be continuous positive and independent random variables whose support is \mathbb{R}^+ .*

Assume that $\forall (i, j) \in \{1, 2, \dots, d\}^2$:

1. $\bar{F}_{X_i}(x) \stackrel{x \rightarrow +\infty}{\asymp} \Theta(\bar{F}_{X_j}(x))$,
2. $\bar{F}_{X_i}(s) \stackrel{s \rightarrow +\infty}{\asymp} o(\bar{F}_{X_i}(t))$, if $t = o(s)$,

then, there exist $\kappa_1 > 0$ and $\kappa_2 < 1$ such that,

$$\kappa_1 \leq \frac{u_\ell}{u} \leq \kappa_2 \quad \forall \ell \in \{1, 2, \dots, d\}, \forall u \in \mathbb{R}^+. \quad (4.8)$$

Note that the first condition of Lemma 4.2.4 is not satisfied for exponential distributions. However, Pareto distributions satisfy the hypothesis of Lemma 4.2.4.

Proof. Taking if necessary a convergent subsequence, we assume that $\exists i \in \{1, \dots, d\}$ such that: $\frac{u_i}{u} \stackrel{u \rightarrow +\infty}{\rightarrow} 1$ or $\frac{u_i}{u} \stackrel{u \rightarrow +\infty}{\rightarrow} 0$, the first case implies that for all $j \neq i$, $\frac{u_j}{u} \stackrel{u \rightarrow +\infty}{\rightarrow} 0$, then, it is sufficient to prove that the existence of an $i \in \{1, \dots, d\}$ such that $\frac{u_i}{u} \stackrel{u \rightarrow +\infty}{\rightarrow} 0$ is impossible.

Let us assume the existence of $i \in \{1, \dots, d\}$, such that: $\frac{u_i}{u} \stackrel{u \rightarrow +\infty}{\rightarrow} 0$.

Then, $\exists j \in \{1, \dots, d\} \setminus i$ such that $\lim_{u \rightarrow +\infty} \frac{u_j}{u} \in]0, 1]$, therefore, $u_j \stackrel{u \rightarrow +\infty}{\rightarrow} +\infty$ and $u_i \stackrel{u \rightarrow +\infty}{\asymp} o(u_j)$.

Using Assumptions (1) and (2), we deduce that:

$$\frac{\bar{F}_{X_j}(u_j)}{\bar{F}_{X_i}(u_i)} \stackrel{u \rightarrow +\infty}{\rightarrow} 0. \quad (4.9)$$

The optimality condition (4.2) can also be written for all $j \neq i$ as follows:

$$\bar{F}_{X_i}(u_i) - \mathbb{P}(X_i > u_i, S > u) = \bar{F}_{X_j}(u_j) - \mathbb{P}(X_i > u_j, S > u). \quad (4.10)$$

That presents a trivial contradiction if u_i remains bounded.

Now, assume that $u_i \rightarrow +\infty$. Recall that $S^{(-i)} = \sum_{k=1, k \neq i}^d X_k$, then:

$$\mathbb{P}(X_i > u_i, S > u) = \mathbb{P}(X_i > u_i, S^{(-i)} > \sqrt{u}, S > u) + \mathbb{P}(X_i > u_i, S^{(-i)} < \sqrt{u}, S > u).$$

We have:

$$\mathbb{P}(X_i > u_i, S^{(-i)} > \sqrt{u}, S > u) \leq \mathbb{P}(X_i > u_i) \mathbb{P}(S^{(-i)} > \sqrt{u}) \stackrel{u \rightarrow +\infty}{\asymp} o(\bar{F}_{X_i}(u_i)).$$

*. For $(x, y) \in \mathbb{R}^{2+}$, we shall denote $y = \Theta(x)$ if there exist $0 < C_1 \leq C_2 < \infty$, such that for x large enough, $C_1 \leq \frac{y}{x} \leq C_2$.

Using assumption (2) and since $u_i = o(u)$,

$$\mathbb{P}\left(X_i > u_i, S^{(-i)} < \sqrt{u}, S > u\right) \leq \bar{F}_{X_i}(u - \sqrt{u}) \stackrel{u \rightarrow +\infty}{\equiv} o(\bar{F}_{X_i}(u_i)).$$

We deduce that:

$$\mathbb{P}(X_i > u_i, S > u) \stackrel{u \rightarrow +\infty}{\equiv} o(\bar{F}_{X_i}(u_i)). \quad (4.11)$$

We remark also that:

$$\mathbb{P}(X_j > u_j, S > u) \stackrel{u \rightarrow +\infty}{\equiv} O(\bar{F}_{X_j}(u_j)) \stackrel{u \rightarrow +\infty}{\equiv} o(\bar{F}_{X_i}(u_i)). \quad (4.12)$$

Equation (4.10) leads to:

$$1 - \underbrace{\frac{\mathbb{P}(X_i > u_i, S > u)}{\bar{F}_{X_i}(u_i)}}_{T_1} = \underbrace{\frac{\bar{F}_{X_j}(u_j)}{\bar{F}_{X_i}(u_i)}}_{T_2} - \underbrace{\frac{\mathbb{P}(X_j > u_j, S > u)}{\bar{F}_{X_i}(u_i)}}_{T_3}.$$

Now, relations: (4.11), (4.9), and (4.12), imply that T_1 , T_2 , and T_3 , go to zero, and this is a contradiction. \square

Lemma 4.2.5. *Let X_1, \dots, X_d be continuous, positive and independent random variables such that the support of the density of (X_i, S) is $(\mathbb{R}^+)^2$. Let (u_1, \dots, u_d) be the optimal allocation of u associated to the risk indicator I . We assume:*

1. *there exist $0 < \kappa_1 < \kappa_2 < 1$ such that for all $i = 1, \dots, d$ and for all $u \in \mathbb{R}^+$,*

$$\kappa_1 \leq \frac{u_i}{u} \leq \kappa_2,$$

2. *for all $i = 1, \dots, d$, if $y = y(x)$ is such that*

$$0 < \liminf_{x \rightarrow \infty} \frac{y}{x} \leq \limsup_{x \rightarrow \infty} \frac{y}{x} < 1$$

then

$$\frac{\bar{F}_{X_i}(x)}{\bar{F}_{X_i}(y)} \xrightarrow{x \rightarrow \infty} 0.$$

Then, for all $i, j = 1, \dots, d$,

$$\frac{\bar{F}_{X_i}(u_i)}{\bar{F}_{X_j}(u_j)} \xrightarrow{u \rightarrow \infty} 1.$$

Assumptions of Lemma 4.2.5 are satisfied for distributions of exponential type (see remark (4.2.2) below). Its application gives another proof to Proposition 4.2.1. In contrast, Lemma 4.2.5 cannot be used for Pareto distributions.

Proof. Following the lines of the proof of Lemma 4.2.4, take $0 < \gamma < 1 - \kappa_2$,

$$\mathbb{P}(X_i > u_i, S > u) = \mathbb{P}(X_i > u_i, S^{(-i)} > \gamma u, S > u) + \mathbb{P}(X_i > u_i, S^{(-i)} > \gamma u, S > u).$$

As before,

$$\mathbb{P}(X_i > u_i, S^{(-i)} > \gamma u, S > u) \leq \mathbb{P}(X_i > u_i) \mathbb{P}(S^{(-i)} > \gamma u) \stackrel{u \rightarrow +\infty}{\equiv} o(\bar{F}_{X_i}(u_i)).$$

On the other hand,

$$\begin{aligned} \mathbb{P}(X_i > u_i, S^{(-i)} > \gamma u, S > u) &\leq \mathbb{P}(X_i > u - \gamma u) \\ &= \bar{F}_{X_i}((1 - \gamma)u) \stackrel{u \rightarrow +\infty}{\equiv} o(\bar{F}_{X_i}(u_i)) \\ &\text{because } 0 < \frac{\kappa_1}{1 - \gamma} \leq \frac{u_i}{(1 - \gamma)u} \leq \frac{\kappa_2}{1 - \gamma} < 1. \end{aligned}$$

So that, $\mathbb{P}(X_i > u_i, S > u) \stackrel{u \rightarrow +\infty}{\underset{\equiv}{\sim}} o(\bar{F}_{X_i}(u_i))$ and the same computation gives $\mathbb{P}(X_j > u_j, S > u) \stackrel{u \rightarrow +\infty}{\underset{\equiv}{\sim}} o(\bar{F}_{X_j}(u_j))$. Now, u_i and u_j satisfy Equation (1) and thus,

$$1 + o(1) \stackrel{u \rightarrow +\infty}{\underset{\equiv}{\sim}} \frac{\bar{F}_{X_j}(u_j)}{\bar{F}_{X_i}(u_i)} + o(1) \frac{\bar{F}_{X_j}(u_j)}{\bar{F}_{X_i}(u_i)}.$$

This implies that $\frac{\bar{F}_{X_j}(u_j)}{\bar{F}_{X_i}(u_i)}$ is bounded from above and thus

$$\frac{\bar{F}_{X_j}(u_j)}{\bar{F}_{X_i}(u_i)} \stackrel{u \rightarrow +\infty}{\underset{\equiv}{\sim}} 1 + o(1).$$

□

Remark 4.2.2. We remark that the hypothesis of Lemma 4.2.5 are satisfied for distribution of exponential type, that is distributions verifying:

$$\bar{F}_{X_i}(x) = \Theta(e^{-\mu_i x}).$$

Indeed, in this case, we have $0 < \underline{\lim} \frac{u_i}{u} \leq \overline{\lim} \frac{u_i}{u} < 1$, $\forall i \in \{1, \dots, d\}$. In fact, if $u_i \stackrel{u \rightarrow +\infty}{\underset{\equiv}{\sim}} o(u)$, then, $\exists j \in \{1, \dots, d\} \setminus \{i\}$ such that $\frac{u_j}{u} \rightarrow \kappa \in]0, 1]$, so $u_i \stackrel{u \rightarrow +\infty}{\underset{\equiv}{\sim}} o(u_j)$.

Since $\mu_i, \mu_j > 0$, $\mu_i u_i \stackrel{u \rightarrow +\infty}{\underset{\equiv}{\sim}} o(\mu_j u_j)$. As in the proof of Lemma 4.2.4, we get:

$$\mathbb{P}(X_i > u_i, S > u) \stackrel{u \rightarrow +\infty}{\underset{\equiv}{\sim}} o(\bar{F}_{X_i}(u_i))$$

and,

$$\mathbb{P}(X_j > u_j, S > u) \stackrel{u \rightarrow +\infty}{\underset{\equiv}{\sim}} o(\bar{F}_{X_j}(u_j)).$$

From the optimality condition,

$$\frac{\bar{F}_{X_i}(u_i)}{e^{-\mu_i u_i}} - \underbrace{\frac{\mathbb{P}(X_i > u_i, S > u)}{e^{-\mu_i u_i}}}_{T_1} = \frac{\bar{F}_{X_j}(u_j)}{e^{-\mu_i u_i}} - \underbrace{\frac{\mathbb{P}(X_j > u_j, S > u)}{e^{-\mu_i u_i}}}_{T_3},$$

which is absurd because as $u \rightarrow +\infty$:

- $\bar{F}_{X_i}(u_i) = \Theta(e^{-\mu_i u_i})$,
- $T_1 = o(1)$,
- $T_2 = \frac{\bar{F}_{X_j}(u_j)}{e^{-\mu_j u_j}} e^{-\mu_j u_j + \mu_i u_i} \rightarrow 0$, since $\mu_i u_i = o(\mu_j u_j)$,
- and $T_3 = o(1) e^{-\mu_j u_j + \mu_i u_i} \rightarrow 0$.

Proposition 4.2.6 (Asymptotic I -allocation, independent sub-exponential case).

Under conditions of Lemma 4.2.4, and if, for all $i \in \{1, \dots, d\}$, F_{X_i} is a sub-exponential distribution, such that:

$$\frac{\bar{F}_{X_i}(y)}{\bar{F}_{X_i}(x)} \stackrel{x \rightarrow +\infty}{\underset{\equiv}{\sim}} O(1), \text{ for } 0 < \kappa_1 \leq \frac{y}{x} \leq \kappa_2 < 1.$$

Then, by minimizing the I indicator, u_i and u_j satisfy:

$$\bar{F}_{X_i}(u_i) - \bar{F}_{X_i}(u) \stackrel{u \rightarrow +\infty}{\underset{\equiv}{\sim}} \bar{F}_{X_j}(u_j) - \bar{F}_{X_j}(u) + o(\bar{F}_{X_i}(u)). \quad (4.13)$$

Proposition 4.2.6 is applicable in the case of Pareto distributions. So, we will use it for determining the optimal asymptotic allocation, for independent risks of Pareto distributions in the next subsection.

Proof. The proof of this Proposition is a direct application of Theorem 4.2.3 and Lemma 4.2.4. □

Now, we focus on the asymptotic optimal allocation by minimizing the risk indicator J , and we study the case of the sub-exponential distribution family.

Proposition 4.2.7 (Asymptotic J -allocation, independent sub-exponential case).

Let (X_1, X_2, \dots, X_d) be continuous positive and independent random variables, such that there exists $i \in \{1, \dots, d\}$ with a sub-exponential distribution, the optimal capital allocation by minimizing the J indicator (u_1, \dots, u_d) verifies, for all $j \neq i$:

$$\lim_{u \rightarrow \infty} \frac{\mathbb{P}(X_j \geq u_j, S \geq u)}{\bar{F}_{X_i}(u)} = 1.$$

Proof. The solution to (4.3) satisfies :

$$\forall j \in \{1, 2, \dots, d\}, \quad \frac{\mathbb{P}(X_i > u_i, S \geq u)}{\mathbb{P}(X_i > u)} = \frac{\mathbb{P}(X_j > u_j, S \geq u)}{\mathbb{P}(X_i > u)}.$$

When u goes to $+\infty$, and using Theorem 4.2.3, we obtain Proposition 4.2.7. □

Application to Pareto independent distributions

We consider d independent random variables (X_1, X_2, \dots, X_d) of Pareto distributions, with parameters $(a, b_i)_{\{i=1,2,\dots,d\}}$ respectively, such that $b_1 > b_2 > \dots > b_d > 0$. Therefore, these distributions will be characterized by densities and survival functions of the following forms:

$$f_{X_i}(x) = \frac{a}{b_i} \left(1 + \frac{x}{b_i}\right)^{-a-1},$$

and

$$\bar{F}_{X_i}(x) = \left(1 + \frac{x}{b_i}\right)^{-a}.$$

Proposition 4.2.8 (Asymptotic I -allocation, independent Pareto case).

Asymptotically, the unique solution to (4.2) satisfies:

$$\forall (i, j) \in \{1, 2, \dots, d\}^2, \quad \left(\frac{\lim_{u \rightarrow \infty} \alpha_i}{b_i}\right)^{-a} - \left(\frac{\lim_{u \rightarrow \infty} \alpha_j}{b_j}\right)^{-a} = \left(\frac{1}{b_i}\right)^{-a} - \left(\frac{1}{b_j}\right)^{-a}.$$

Proof. Follows from Proposition 4.2.6. □

Proposition 4.2.9 (Asymptotic J -allocation, independent Pareto case).

The unique solution to (4.3) satisfies:

$$\lim_{u \rightarrow \infty} \alpha_1 = 1 \text{ and } \lim_{u \rightarrow \infty} \alpha_i = 0, \forall i \in \{2, 3, \dots, d\}.$$

Proof. We suppose that $\exists \in \{1, \dots, d\}$ such that $0 < \lim_{u \rightarrow \infty} \alpha_j < 1$.

From Theorem 4.2.3 :

$$\lim_{u \rightarrow \infty} \frac{\mathbb{P}(X_j \geq u_j, S \geq u)}{\bar{F}_{X_j}(u)} = 1.$$

On the other hand, and applying Proposition 4.2.7 in the Pareto distributions case, we get for $i \in \{1, \dots, d\} \setminus \{j\}$:

$$\frac{\mathbb{P}(X_j \geq u_j, S \geq u)}{\bar{F}_{X_j}(u)} = \frac{\mathbb{P}(X_j \geq u_j, S \geq u)}{\bar{F}_{X_i}(u)} \cdot \frac{\bar{F}_{X_i}(u)}{\bar{F}_{X_j}(u)} \underset{u \rightarrow +\infty}{\sim} \frac{\bar{F}_{X_i}(u)}{\bar{F}_{X_j}(u)} = \left(\frac{1 + \frac{u}{b_i}}{1 + \frac{u}{b_j}}\right)^{-a},$$

then, for $i \in \{1, \dots, d\} \setminus \{j\}$:

$$\lim_{u \rightarrow \infty} \frac{\mathbb{P}(X_j \geq u_j, S \geq u)}{\bar{F}_{X_j}(u)} = \left(\frac{b_j}{b_i}\right)^{-a} \neq 1.$$

That is absurd. We deduce that $\forall i \in \{1, 2, \dots, d\}$:

$$\lim_{u \rightarrow \infty} \alpha_i \in \{0, 1\},$$

and since $\sum_{i=1}^d \alpha_i = 1$, then, there is a unique i such that $\lim_{u \rightarrow \infty} \alpha_i = 1$ and for all $j \neq i$ $\lim_{u \rightarrow \infty} \alpha_j = 0$. Let us recall the definition of the order stochastic dominance, as it is presented in Shaked and Shanthikumar (2007)[26]. For random variables X and Y , Y first-order stochastically dominates X if and only if:

$$\bar{F}_X(x) \leq \bar{F}_Y(x), \quad \forall x \in \mathbb{R}^+,$$

and in this case we denote: $X \preceq_{st} Y$.

Now, clearly $X_d \preceq_{st} \dots \preceq_{st} X_2 \preceq_{st} X_1$ because $b_1 > b_2 > \dots > b_d > 0$.

In [23] (Proposition 3.10) we have proved that the optimal capital allocation satisfies the monotonicity property. We deduce that $\alpha_1 \geq \dots \geq \alpha_d$, and thus $\lim_{u \rightarrow \infty} \alpha_1 = 1$. \square

4.2.3 Comparison between the asymptotic behaviors in exponential and sub-exponential cases

Let us consider the following risk indicator:

$$I_{loc}(u_1, \dots, u_d) = \sum_{k=1}^d \mathbb{E} \left((X_k - u_k) \mathbb{1}_{\{X_k > u_k\}} \right) = \sum_{k=1}^d \mathbb{E} \left((X_k - u_k)^+ \right).$$

With this indicator, the impact of each branch on the group is not taken into account. This indicator thus takes only *local* effects into account. If, for any $k = 1, \dots, d$, the random vector (X_k, S) admits a density whose support contains $[0, u]^2$, then the optimality condition associated to I_{loc} gives:

$$\mathbb{P}(X_i > u_i) = \mathbb{P}(X_j > u_j).$$

We remark that this corresponds to the asymptotic result for indicator I of Lemma 4.2.5, where $\frac{\bar{F}_{X_j}(u_j)}{\bar{F}_{X_i}(u_i)} \underset{u \rightarrow +\infty}{=} 1 + o(1)$. In other words, in the case of *exponential type* and independent risks, the group effect is asymptotically negligible. This behavior is also clear in Proposition 4.2.1, where we found the asymptotic capital allocation for I for independent exponential distributions as:

$$u_i = \frac{\frac{1}{\beta_i}}{\sum_{j=1}^d \frac{1}{\beta_j}} u, \quad \text{for all } i = 1, \dots, d,$$

as for I_{loc} .

Proposition 4.2.6 shows that for independent sub-exponential distributions the asymptotic behavior is different. Indeed, in Equation (4.13), the terms $\bar{F}_{X_i}(u)$ and $\bar{F}_{X_j}(u)$ lead to take into account the group effect. In the Pareto distributions case for example, recall that Proposition 4.2.8 gives that the asymptotic behavior of the optimal allocation for I is described by :

$$\left(\frac{\lim_{u \rightarrow \infty} \alpha_i}{b_i} \right)^{-a} - \left(\frac{\lim_{u \rightarrow \infty} \alpha_j}{b_j} \right)^{-a} = \left(\frac{1}{b_i} \right)^{-a} - \left(\frac{1}{b_j} \right)^{-a},$$

whereas the optimal allocation for I_{loc} for independent Pareto distributions is given by $\alpha_i = \frac{b_i}{\sum_{\ell=1}^d b_\ell}$, for all $i \in \{1, \dots, d\}$.

For optimal allocation by minimizing the J indicator, the asymptotic behavior is identical for the two families of distributions, the riskiest branch is considered as first responsible of the overall ruin, and

thus, the optimal asymptotic solution is to allocate the entire capital u to this business line.

Finally, the nature of the marginal risk distribution has an impact on the asymptotic behavior of the optimal allocation.

In the rest of this paper, we study the impact of the dependence on the optimal capital allocation using some copulas (see Nelsen [24] for review on copulas). The idea is to determine the optimal allocation as a function of the copula parameters in each case. We focus on the indicator I , the same kind of calculations can be done for the indicator J .

4.3 Asymptotic independence models

In this section, We focus on the asymptotic independence case. It is a kind of dependence of which vanishes for high values of the random variables. In order to illustrate this case, we choose the FGM copula model, well studied in the actuarial literature. To visualize the impact of dependence on the capital allocation, we will limit our study to the bivariate model. The results can be generalized in the multivariate context using the multivariate form of FGM copula.

4.3.1 Impact of dependence in FGM Model

Let X_1 and X_2 be two risks of marginal exponential distributions $X_i \sim \exp(\beta_i)$ and FGM bivariate dependence structure with $-1 \leq \theta \leq 1$ as parameter (see Nelsen [24], Example 3.12, Section 3.2.5). We assume that $\beta_1 < \beta_2$ and $\beta_2 \neq 2\beta_1$.

In this model, the Pearson correlation coefficient is given by $\rho_P = \frac{\theta}{4}$, its Spearman's correlation coefficient is $\rho = \frac{\theta}{3}$. We recall that Spearman's rho is the concordance measure defined for continuous bivariate distributions with copula C as dependence structure by

$$\rho_S = 12 \int \int_{[0,1]^2} uv dC(u, v) - 3 = 12 \int \int_{[0,1]^2} C(u, v) dudv - 3.$$

The bivariate distribution function is:

$$F_{X_1, X_2}(x_1, x_2) = (1 - e^{-\beta_1 x_1})(1 - e^{-\beta_2 x_2}) + \theta(1 - e^{-\beta_1 x_1})(1 - e^{-\beta_2 x_2})e^{-\beta_1 x_1} e^{-\beta_2 x_2}.$$

The fact that $\rho \in [-1/3, 1/3]$ reduces the capacity of FGM copula to take into account the strong dependence. The related copula for the FGM distribution is

$$C_\theta(u, v) = uv[1 + \theta(1 - u)(1 - v)],$$

and the upper tail dependence coefficient is

$$\lambda_U = 0,$$

the FGM model is then an asymptotically independent model.

Proposition 4.3.1 (*I*-allocation, FGM Model).

For the indicator I, the optimal allocation of a capital u is given by $(\alpha_1 u, (1 - \alpha_1)u)$ such that $\alpha_1 = u_1/u$ is the unique solution in $[0, 1]$ of the following equation:

$$\begin{aligned} h(\alpha_1) - h(\gamma(1 - \alpha_1)) &= \frac{1 + \theta}{\gamma - 1} [h(\gamma) + \gamma h(1) - (1 + \gamma)h(\gamma + \alpha_1 - \gamma\alpha_1)] \\ &+ \frac{\theta}{\gamma - 1} [h(2\gamma) + \gamma h(2) - (1 + \gamma)h(2\gamma + 2\alpha_1 - 2\gamma\alpha_1)] \\ &- \frac{\theta}{\gamma - 2} [2h(\gamma) + \gamma h(2) - (2 + \gamma)h(\gamma + 2\alpha_1 - \gamma\alpha_1)] \\ &- \frac{\theta}{2\gamma - 1} [h(2\gamma) + 2\gamma h(1) - (1 + 2\gamma)h(2\gamma + \alpha_1 - 2\gamma\alpha_1)], \end{aligned} \quad (4.14)$$

where, h is the function $h(x) = \exp(-\beta_1 ux)$, and $\gamma = \beta_2/\beta_1$.

Proof. The proof is postponed to Appendix 4.6.1. □

Remark 4.3.1. *The case where $\theta = 0$, is the independent exponential distributions model considered in Proposition 4.2.1.*

Equation (4.14) gives the behavior of the optimal allocation with respect to θ . It may be solved numerically. Figure 4.2 presents an illustration of the optimal allocation variation with respect to the dependence parameter of the FGM copula.

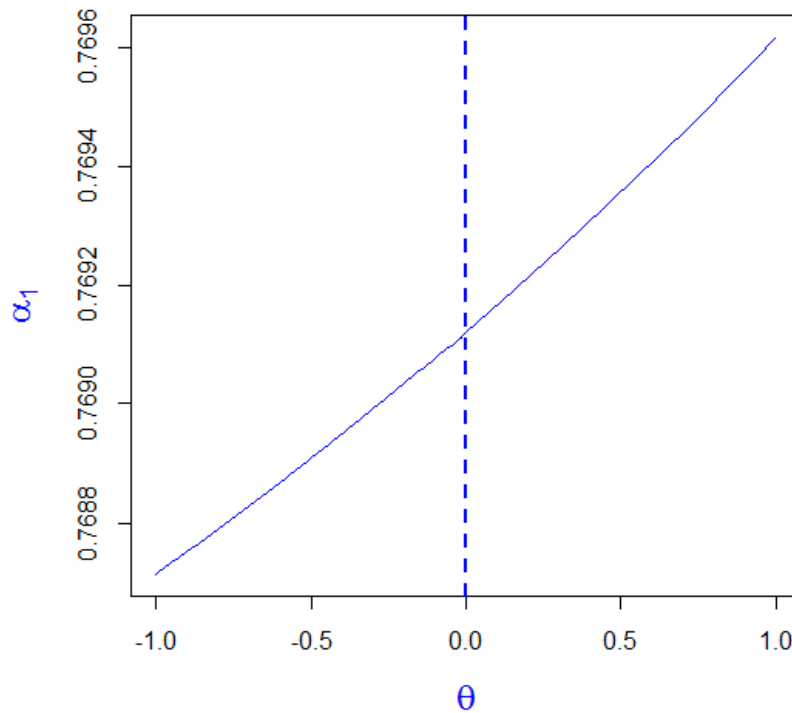


Figure 4.2: Impact of dependence on the optimal allocation. Allocation percentage α_1 as a function of the dependence parameter θ . FGM Model : $\beta_1 = 0.05$, $\beta_2 = 0.25$, and $u = 50$.

For the illustration parameters and for a capital level $u = 50$, we remark that α_1 is an increasing function of θ , this can be verified analytically using the implicit function theorem. It is important to remark that the optimal allocation is also a function of the capital u . The variations of α_1 with respect to the dependence parameter depends on u . Figure 4.3 gives the obtained result for an allocation with the same distributions parameters but with $u = 100$ as allocation capital.

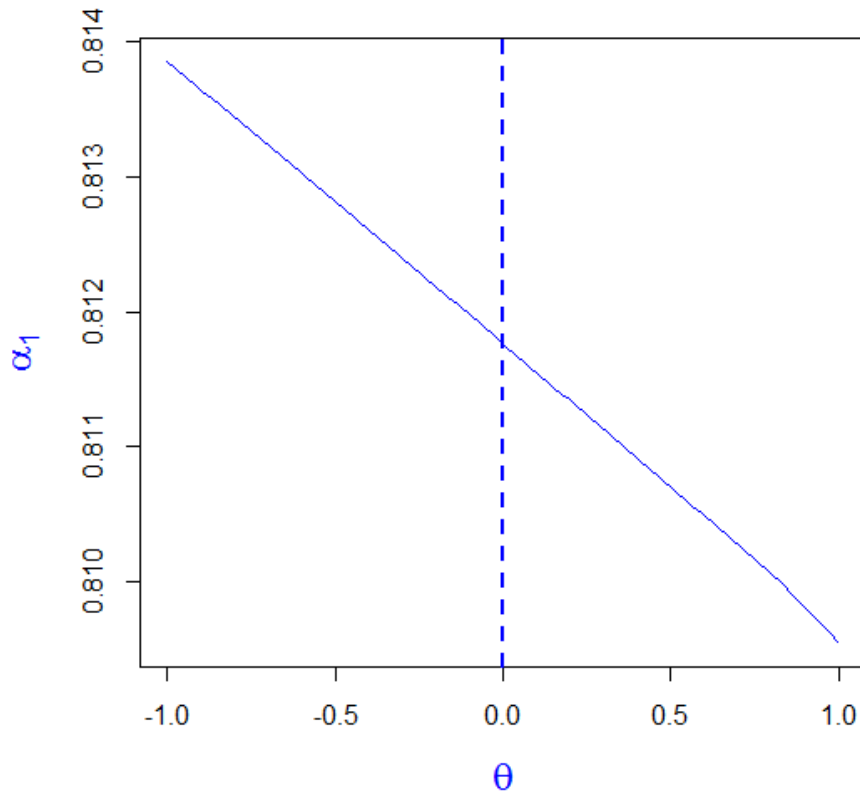


Figure 4.3: Impact of dependence on the optimal allocation. Allocation percentage α_1 as a function of the dependence parameter θ . FGM Model : $\beta_1 = 0.05$, $\beta_2 = 0.25$, and $u = 100$.

The optimal allocation depends on the initial capital. As an example, with the same parameters, α_1 is a decreasing function of θ for u smaller than a certain threshold u^* , and it is an increasing function of θ for u larger than the threshold u^* . In the case of our numerical illustrations ($\beta_1 = 0.05$, $\beta_2 = 0.25$), u^* is approximately 53. This behavior can be proved analytically in details using the derivative function. The variation range size of α_1 is very small, and its size is function of the distributions parameters and the allocation capital. That is due to the dependence structure, the FGM copula present only a light dependence. For survival Clayton copula as example, which has a stronger dependence structure, the variation range size of α_1 is more important (see subsection 4.4.2).

4.3.2 Impact of capital level in FGM model

In order to analyse the impact of the capital level on the dependence impact on the capital allocation, we present in Figure 4.4 the evolution of α_1 function of the dependence parameter for three different capital allocation levels.

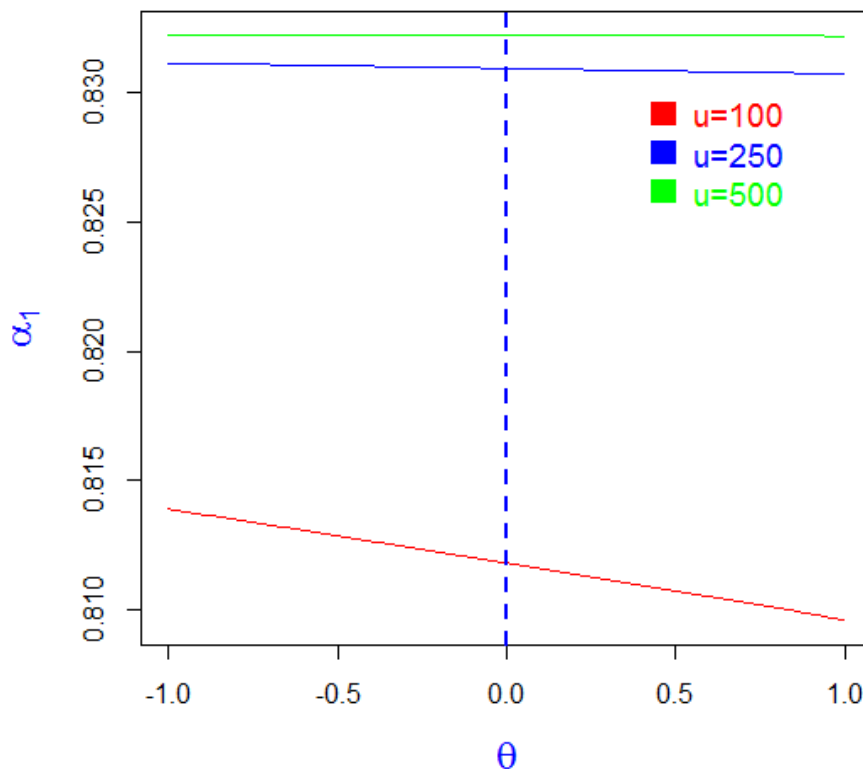


Figure 4.4: Impact of capital level in FGM model. Allocation percentage α_1 as a function of the dependence parameter θ . $\beta_1 = 0.05$, $\beta_2 = 0.25$, and $u = 100, 250, 5000$.

The dependence impact is less important for higher capital allocation level.

Finally, the capital allocation by minimizing the multivariate risk indicators takes into account the weak dependence, like FGM copula. The impact of dependence depends on the capital level. Its more important for small capital levels.

4.4 Asymptotic dependence models

This section is devoted to the study of the dependence models with asymptotic dependence. In our context, the upper tail dependence is the only interesting quantity, we shall consider models with a strict positive upper tail dependence coefficient $\lambda_U > 0$.

Using two different models, Marshall-Olkin and Common mixture, we analyze the impact of asymptotic dependence on the asymptotic behavior of the optimal allocation. The impact of the asymptotic dependence on the speed of convergence of the allocation when u goes to infinity is also considered.

The Marshall-Olkin model is a special kind of dependence structure, because the marginals depend on the degree of dependence, so when we change the dependence level, we change at the same time the marginals. With the common mixture model, we will construct a correlated Pareto model in order to compare it with the independent Pareto model. Common mixture model will also be used to study the impact of dependence on the asymptotic behavior.

4.4.1 Marshall-Olkin Model

Let $Y_i \sim \exp(\lambda_i)$, with $i = 0, 1, 2$ be three independent random variables.

We construct two random variables with common shock: $X_i = \min(Y_i, Y_0)$ for $i = 1, 2$. X_i 's have

exponential marginal distributions of parameters $\beta_i = \lambda_i + \lambda_0$ (see e.g. Nelsen [24] section 3.1.1.). This dependence construction model has as Pearson correlation coefficient $\rho_P = \frac{\lambda_0}{\lambda_0 + \lambda_1 + \lambda_2}$. The joint distribution function is given by:

$$\begin{aligned}\bar{F}_{X_1, X_2}(x_1, x_2) &= \mathbb{P}(X_1 > x_1, X_2 > x_2) = \mathbb{P}(Y_1 > x_1, Y_2 > x_2, Y_0 > \max(x_1, x_2)) \\ &= e^{-\lambda_1 x_1} e^{-\lambda_2 x_2} e^{-\lambda_0 \max(x_1, x_2)} \\ &= e^{-(\lambda_0 + \lambda_1)x_1} e^{-(\lambda_0 + \lambda_2)x_2} e^{\lambda_0 \min(x_1, x_2)} \\ &= \bar{F}_{X_1}(x_1) \bar{F}_{X_2}(x_2) e^{\lambda_0 \min(x_1, x_2)},\end{aligned}$$

and the joint density function is the following:

$$f_{X_1, X_2}(x_1, x_2) = \begin{cases} f_{X_1, X_2}^1(x_1, x_2) = \beta_1 e^{-\beta_1 x_1} (\beta_2 - \lambda_0) e^{-(\beta_2 - \lambda_0)x_2} & si \quad x_1 > x_2 \\ f_{X_1, X_2}^2(x_1, x_2) = (\beta_1 - \lambda_0) e^{-(\beta_1 - \lambda_0)x_1} \beta_2 e^{-\beta_2 x_2} & si \quad x_1 < x_2 \\ f_{X_1, X_2}^0(x_1, x_2) = \lambda_0 e^{-\beta_1 x} e^{-\beta_2 x} e^{\lambda_0 x} & si \quad x_1 = x_2 = x \end{cases}.$$

Embrechts et al. (2001) [14] defined a multivariate extension of the bivariate Marshall-Olkin model. In the bivariate case, this construction leads to a copula given by

$$C(u_1, u_2) = \min(u_1^{1-\lambda_0/\beta_1} u_2, u_1 u_2^{1-\lambda_0/\beta_2}).$$

The Spearman's rho for Marshall-Olkin copulas is

$$\rho_S = \frac{1}{1 + \frac{2}{3} \frac{\lambda_1 + \lambda_2}{\lambda_0}},$$

Since $\rho_S \in]0, 1[$, the Marshall-Olkin copulas model only the positive dependence. The Marshall-Olkin copulas have upper tail dependence, it is given by

$$\lambda_U = \min\left(\frac{\lambda_0}{\beta_1}, \frac{\lambda_0}{\beta_2}\right) = \frac{\lambda_0}{\max(\lambda_1, \lambda_2) + \lambda_0},$$

Marshall-Olkin model is then an asymptotic dependence model.

Proposition 4.4.1 (*I*-allocation, Marshall-Olkin Model).

We suppose that $\lambda_1 < \lambda_2$ and that $\lambda_0 \neq \lambda_2 - \lambda_1$. The optimal allocation of a capital u minimizing the indicator I is given by $(\alpha_1 u, (1 - \alpha_1)u)$, such that $\alpha_1 = u_1/u$ is the unique solution in $[0, 1]$ of the following equation:

$$\begin{aligned}g(\beta_2(1 - \alpha_1)) - g(\beta_1 \alpha_1) + \frac{\beta_1}{\beta_1 - \lambda_2} g((\beta_1 - \lambda_2)\alpha_1 + \lambda_2) + \frac{\lambda_2}{\beta_1 - \lambda_2} g((\lambda_2 - \beta_1)(1 - \alpha_1) + \beta_1) \\ - \frac{\lambda_1}{\lambda_1 - \beta_2} g(\beta_2) = \frac{\lambda_2}{\beta_1 - \lambda_2} g(\beta_1) + g(\lambda_s/2) \left[\frac{\beta_1}{\beta_1 - \lambda_2} - \frac{\lambda_1}{\lambda_1 - \beta_2} \right],\end{aligned}$$

where, $\lambda_s = \lambda_0 + \lambda_1 + \lambda_2$, and g is the function $g(x) = \exp(-ux)$.

Proof. The proof is postponed to Appendix 4.6.2. □

Remark 4.4.1. In the case $\lambda_0 = 0$ which is the independence case, we find again Equation (4.4) given by Proposition 4.2.1.

We can consider λ_0 as a dependence parameter in this model. Figure 4.5 presents an illustration of the variation of α_1 as a function of λ_0 .

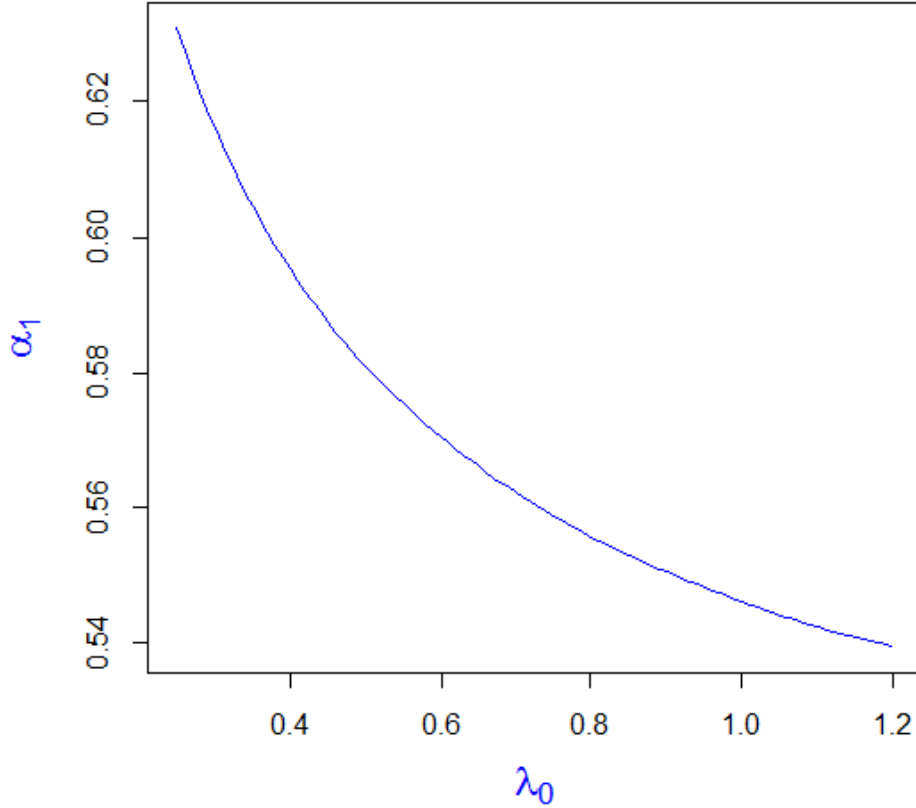


Figure 4.5: Impact of dependence on the optimal allocation in MO model. Allocation percentage α_1 as a function of the dependence parameter λ_0 . $\beta_1 = 0.05$, $\beta_2 = 0.25$ and $u = 50$.

One can notice that α_1 is a decreasing function of λ_0 . This is coherent with the increase of α_1 as a function of $\gamma = \frac{\beta_2}{\beta_1} = \frac{\lambda_2 + \lambda_0}{\lambda_1 + \lambda_0}$ demonstrated in [5] in the independence case, since the two risks are independent, conditionally to Y_0 .

Remark that in the Marshall-Olkin model, the dependence construction changes the marginals, contrary to the FGM model or any archimedean copulas as example, where by construction the marginals are all the time the same, and the dependence effect is limited to the copula.

4.4.2 Common Mixture Model

This method of multivariate models construction is presented in detail in Joe (1997) [18]. It is based on the choice of a random variable Θ of support S_Θ , and independent random variables Y_i to construct random variables X_i conditionally independent knowing Θ , such that:

$$\bar{F}_{X_i|\Theta=\theta}(x_i) = (\bar{F}_{Y_i}(x_i))^\theta.$$

This construction provides the marginal distributions and the joint distribution by integration with respect to the law of Θ , as presented in Marceau (2013) [21].

We are interested here to the case of bivariate exponential mix model, and we assume that the moment generating function of Θ , M_Θ exists, then:

$$f_{X_1,S}(x_1, s) = \int_{\theta \in S_\Theta} \beta_1 \theta e^{-\beta_1 \theta x_1} \beta_2 \theta e^{-\beta_2 \theta (s-x_1)} dF_\Theta(\theta) = \beta_1 \beta_2 \frac{d^2 M_\Theta(t)}{dt^2} \Big|_{t=-(\beta_2(s-x_1)+\beta_1 x_1)}.$$

Correlated Pareto

For this model, we shall visualize the impact of dependence and compare it with Pareto independence model previously studied.

Let (X_1, \dots, X_d) be a mixture of exponential distributions such that for all $i \in \{1, 2, \dots, d\}$, $X_i \sim \mathcal{E}(\beta_i \theta)$, with $(\beta_1 < \beta_2 \dots < \beta_d)$, and $\theta \sim \Gamma(a, b)$. Therefore, X_i 's survival functions is

$$\bar{F}_{X_i}(x) = \int_0^\infty \bar{F}_{X_i|\Theta=\theta} f_\Theta(\theta) d\theta = \int_0^\infty e^{-\beta_i \theta x} f_\Theta(\theta) d\theta = \left(1 + \frac{\beta_i x}{b}\right)^{-a},$$

consequently, X_i is Pareto distributed with parameters $(a, \frac{b}{\beta_i})$. The X_i 's are conditionally independents. Computations are done by conditioning on the random variable θ and then integrating the formulas found for the case of independent exponential distributions. This model has been studied in

e.g.[25, 28]. Recall that $A_\ell = \prod_{j=1, j \neq \ell}^d \frac{\beta_j}{\beta_j - \beta_\ell}$, $\ell = 1, \dots, d$.

Proposition 4.4.2 (*I*-allocation, correlated Pareto case).

The optimal allocation minimizing the multivariate risk indicator I is the unique solution in \mathcal{U}_u^d , of the following equations system:

$$\forall (i, j) \in \{1, 2, \dots, d\}^2,$$

$$s(\beta_i \alpha_i) - s(\beta_j \alpha_j) - \sum_{\ell=1}^d A_\ell [s(\alpha_i \beta_i + (1 - \alpha_i) \beta_\ell) - s(\alpha_j \beta_j + (1 - \alpha_j) \beta_\ell)] = 0, \quad (4.15)$$

where s is the function defined by $s(x) = (1 + x \frac{u}{b})^{-a}$ and $\alpha_i = \frac{u_i}{u}$ for all $i \in \{1, \dots, d\}$.

When the capital u goes to infinity, the optimal allocation by minimization of the risk indicator I is the unique solution in \mathcal{U}_u^d of the following equations system:

$$\forall (i, j) \in \{1, 2, \dots, d\}^2,$$

$$(\beta_i \alpha_i)^{-a} - (\beta_j \alpha_j)^{-a} - \sum_{\ell=1}^d A_\ell [(\alpha_i \beta_i + (1 - \alpha_i) \beta_\ell)^{-a} - (\alpha_j \beta_j + (1 - \alpha_j) \beta_\ell)^{-a}] = 0. \quad (4.16)$$

Proof.

For (4.15), it suffices to integrate Equations System (4.4), multiplied by the density function of θ .

We divide Equations System (4.15) by $s(1)$, and let u go to $+\infty$ to get Equations System (4.16). \square

Proposition 4.4.2 shows the impact of the dependence related to the mixture. Indeed, in the case of independent Pareto distributions, of parameters $(a, \frac{b}{\beta_i})_{i=1, \dots, d}$, the asymptotic optimal allocation for the indicator I is given by Proposition 4.2.8 as the solution of the equations system:

$$\forall (i, j) \in \{1, 2, \dots, d\}^2, \quad (\beta_i \alpha_i)^{-a} - (\beta_j \alpha_j)^{-a} = (\beta_i)^{-a} - (\beta_j)^{-a}.$$

Each equation in this system depends only on two risks, unlike the mixture case, where the equations of Equations System (4.16), depend on all the risks.

Proposition 4.4.3 (*J*-allocation, correlated Pareto case).

The optimal allocation minimizing the multivariate risk indicator J is the unique solution in \mathcal{U}_u^d , of the following equations system:

$$\forall (i, j) \in \{1, 2, \dots, d\}^2, \quad \sum_{\ell=1}^d A_\ell [s(\alpha_i \beta_i + (1 - \alpha_i) \beta_\ell) - s(\alpha_j \beta_j + (1 - \alpha_j) \beta_\ell)] = 0. \quad (4.17)$$

When the capital u goes to infinity, the optimal allocation by minimization of the risk indicator J , is the unique solution in \mathcal{U}_u^d of the following equations system:

$$\forall (i, j) \in \{1, 2, \dots, d\}^2, \quad \sum_{\ell=1}^d A_\ell [(\alpha_i \beta_i + (1 - \alpha_i) \beta_\ell)^{-a} - (\alpha_j \beta_j + (1 - \alpha_j) \beta_\ell)^{-a}] = 0. \quad (4.18)$$

Proof.

It suffices to integrate Equations System (4.6), multiplied by the density function of θ .

We divide Equations system (4.17) by $s(1)$, and we let u go to $+\infty$ to get Equations system (4.18). \square

Proposition 4.4.3 shows that for the indicator J , the asymptotic behaviour of the optimal capital allocation takes into account the mixture effect. Indeed, for independent Pareto distributions, we have proved in Proposition 4.2.9, that we allocate the entire capital u to the riskiest branch X_1 , while the asymptotic optimal allocation in the correlated Pareto distributions case is the solution of Equations system (4.18).

Mixture effect

In order to simplify the previous model, we assume without losing generality that $b = 1$. So, $\Theta \sim \Gamma(\gamma, 1)$ and $Y_i \sim \exp(\beta_i)$. The marginals are *Pareto* $(\gamma, 1/\beta_i)$ distributions, and the survival bivariate distribution is given by

$$\bar{F}_{X_1, X_2}(x_1, x_2) = \left(\frac{1}{1 + \beta_1 x_1 + \beta_2 x_2} \right)^\gamma = \left(\bar{F}_{X_1}(x_1)^{-1/\gamma} + \bar{F}_{X_1}(x_1)^{-1/\gamma} - 1 \right)^{-\gamma},$$

that is the Survival Clayton copula of dependence parameter $\theta = 1/\gamma$ so that the upper tail dependence coefficient is:

$$\lambda_U = \lambda_L^{Clayton} = 2^{-\gamma},$$

where $\lambda_L^{Clayton}$ is the lower tail dependence coefficient of Clayton copula. This dependence model have then upper tail dependence. Proposition 4.4.4 gives the optimal allocation for the considered model.

Proposition 4.4.4 (*I*-allocation, common mixture model).

For $\Theta \sim \text{Gamma}(\gamma, 1)$, the optimal allocation of u between the two branches is given by: $(\alpha_1 u, (1 - \alpha_1)u)$ such that $\alpha_1 = u_1/u$ is the unique solution in $[0, 1]$ of the following equation:

$$h(\beta_1 \alpha_1) - h(\beta_2(1 - \alpha_1)) + \frac{\beta_1}{\beta_2 - \beta_1} h(\beta_1) + \frac{\beta_2}{\beta_2 - \beta_1} h(\beta_2) - \frac{\beta_1 + \beta_2}{\beta_2 - \beta_1} h(\beta_1 \alpha_1 + \beta_2(1 - \alpha_1)) = 0 \quad (4.19)$$

where, h is the function: $h(x) = \frac{1}{(1+x)^\gamma}$.

In this model, we can consider γ as a dependence parameter. So, we can easily study the variation of the optimal allocation according to γ , the lower γ , the stronger the dependence. The independence case corresponds to $\gamma = +\infty$.

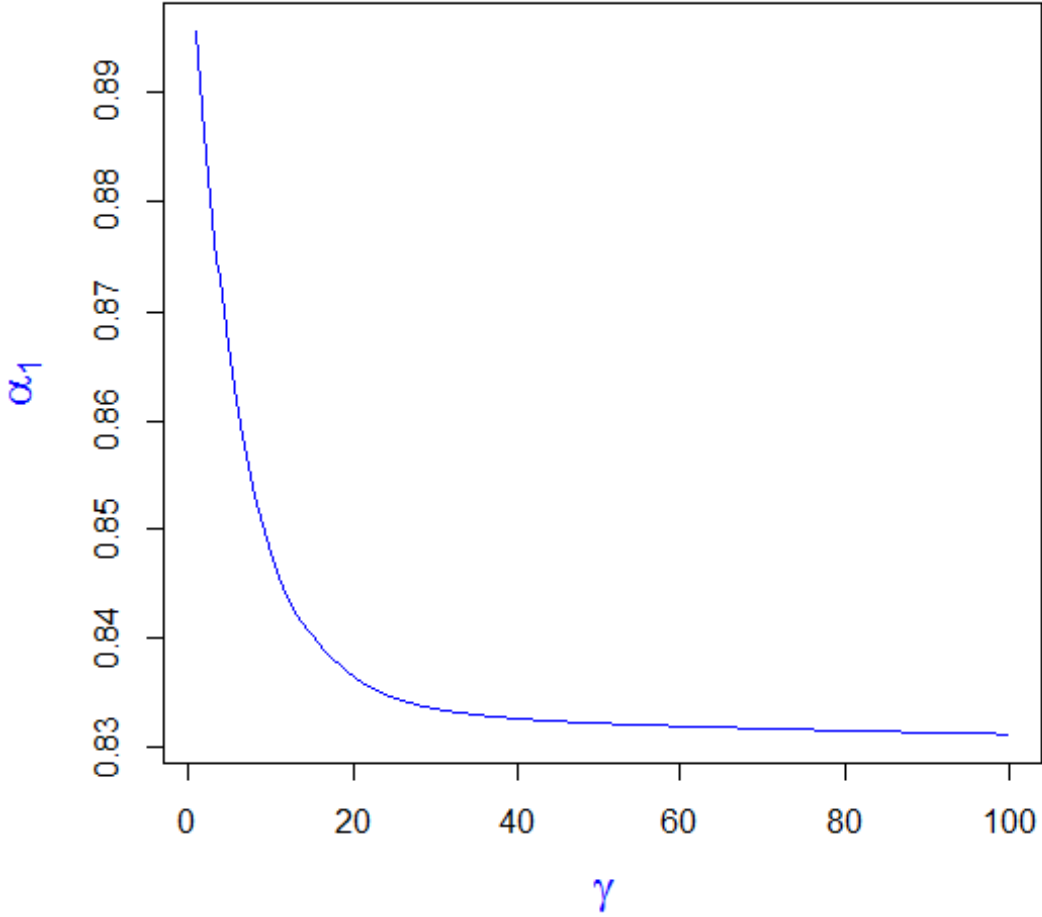


Figure 4.6: Impact of mixture on the optimal allocation. Allocation percentage α_1 as a function of the mixture parameter γ . $\beta_1 = 0.01$, $\beta_2 = 0.05$ and $u = 100$.

Figure 4.6 above presents an illustration of the variation of α_1 , function of γ , we note that the choice of γ is a key factor in determining the composition of the optimal allocation. The percentage allocated to the riskiest branch is an increasing function of dependence $1/\gamma$, it is not the same behavior obtained previously for the FGM model. The presence of dependence increases the percentage of the capital allocated to the riskiest branch.

We also analyze the asymptotic behavior of the capital allocation for this mixture model. Proposition 4.4.5 gives the asymptotic solution.

Proposition 4.4.5 (Asymptotic I -allocation, common mixture model).

When the capital u goes to infinity, the optimal allocation is $(\alpha_1 u, (1 - \alpha_1)u)$ such that α_1 is the unique solution in $[0, 1]$ of the following equation:

$$g(\beta_1 \alpha_1) - g(\beta_2 (1 - \alpha_1)) + \frac{\beta_1}{\beta_2 - \beta_1} g(\beta_1) + \frac{\beta_2}{\beta_2 - \beta_1} g(\beta_2) - \frac{\beta_1 + \beta_2}{\beta_2 - \beta_1} g(\beta_1 \alpha_1 + \beta_2 (1 - \alpha_1)) = 0 \quad (4.20)$$

where g is the function: $g(x) = \frac{1}{x^\gamma}$.

In Figure 4.7 we observe the joint impact of capital level and dependence effect on the capital allocation; It presents the percentage of the capital allocated to the riskiest branch in function of the capital levels, for different values of γ . Figure 4.7 is also an illustration of the impact of dependence on the asymptotic behavior of the optimal capital allocation; the weaker the dependence, the quicker the convergence.

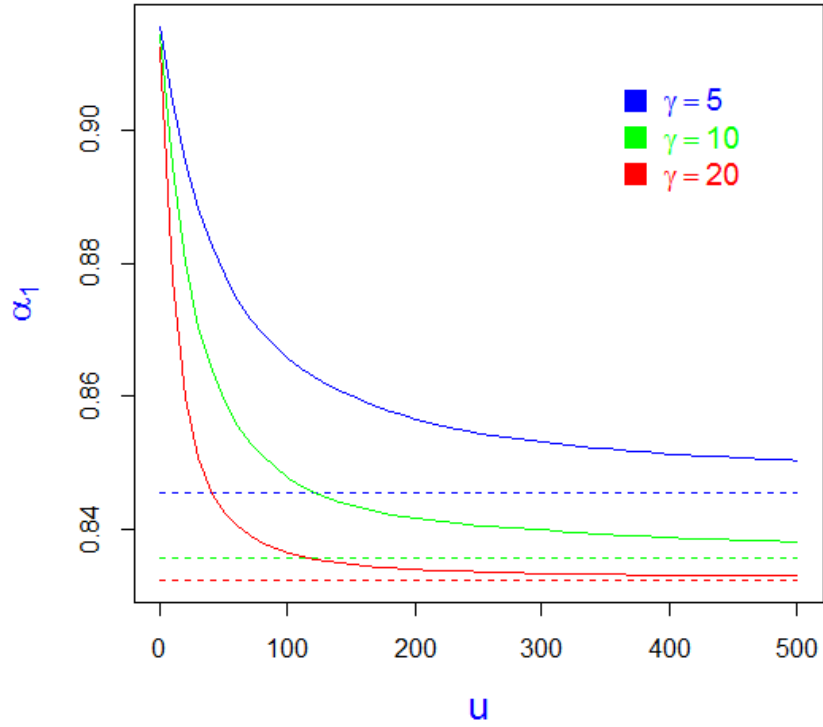


Figure 4.7: Dependence impact on the speed of convergence to the asymptotic case. Allocation percentage α_1 as a function of initial capital u . $\beta_1 = 0.01$, $\beta_2 = 0.05$.

In Figure 4.8, we compare the evolution of the capital allocation for $u = 100$ with the asymptotic allocation as functions of γ .

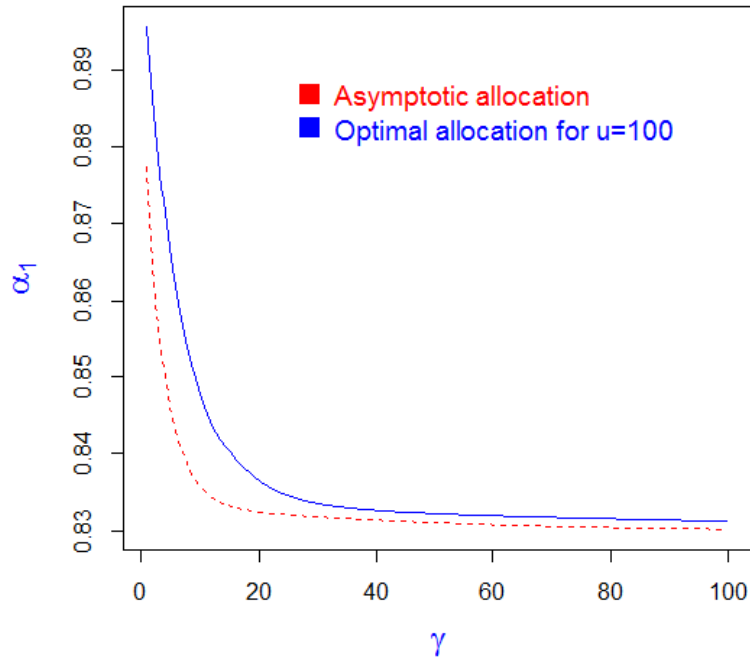


Figure 4.8: Speed of convergence to the asymptotic allocation. Allocation percentage α_1 as a function of the mixture parameter γ . α_1 as function of γ . $\beta_1 = 0.01$, $\beta_2 = 0.05$, $u = 100$.

The difference between the capital allocation and the asymptotic allocation is a decreasing function of γ , the speed of convergence is then a decreasing function of the dependence level $1/\gamma$.

Finally, the percentage of the capital allocated to the riskiest branch for all the models studied in this

section is an increasing function of the dependence level. The impact of dependence on the asymptotic behavior of the optimal allocation was also observed, the speed of convergence is a decreasing function of the dependence.

4.5 Perfect dependence

We devote this section to the perfect dependence case. The perfect dependence corresponds to the comonotonic risks case. In particular, in the comonotonic case, the dependence and concordance measures equal to their maximums $\rho = \rho_S = 1$, and the upper tail dependence coefficient is $\lambda_U = 1$. The concept of comonotonic random variables is related to the studies of Hoeffding (1940) [16] and Fréchet (1951) [15]. Here we use the definition of comonotonic risks as it was first mentioned in the actuarial literature in Borch (1962) [3].

A vector of random variables (X_1, X_2, \dots, X_n) is comonotonic if and only if there exists a random variable Y and non-decreasing functions $\varphi_1, \dots, \varphi_n$ such that:

$$(X_1, \dots, X_n) \stackrel{d}{=} (\varphi_1(Y), \dots, \varphi_n(Y)).$$

In the case where the risks X_1, \dots, X_d are comonotonic, we may give explicit formulas for the optimal allocation minimizing the multivariate risk indicators I and J , and for some risk models. We use the existence of a uniform random variable U such that $X_i = F_{X_i}^{-1}(U)$ for all $i \in \{1, \dots, d\}$, and $S = \sum_{i=1}^d F_{X_i}^{-1}(U) = \varphi(U)$, where $\varphi(t) = \sum_{i=1}^d F_{X_i}^{-1}(t)$, φ is a non-decreasing function. The main result of this section is given below.

Proposition 4.5.1 (Optimal allocation, comonotonic case).

Let X_1, \dots, X_d be comonotonic risks, with increasing distribution functions and support containing $[0, u]$. The optimal allocations for indicators I , J and I_{loc} coincide, they are given by $(u_1, \dots, u_d) \in \mathcal{U}_u^d$ and

$$F_{X_i}(u_i) = F_{X_j}(u_j) \quad \forall i, j = 1, \dots, d,$$

where I_{loc} is the risk indicator defined in (4.2.3) by

$$I_{loc}(u_1, \dots, u_d) = \sum_{k=1}^d \mathbb{E} \left((X_k - u_k) \mathbb{1}_{\{X_k > u_k\}} \right) = \sum_{k=1}^d \mathbb{E} \left((X_k - u_k)^+ \right).$$

In addition, the allocation percentages $\alpha_i = \frac{u_i}{u}$ are independent of the capital level u for all $i \in \{1, \dots, d\}$ if and only if there exist positive constants a_{ij} such that

$$X_i \stackrel{L}{=} a_{ij} X_j, \quad \forall (i, j) \in \{1, \dots, d\}^2.$$

Proof. Let us denote: $w_i = F_{X_i}(u_i)$, $v = \varphi^{-1}(u)$, $M_i = \max(w_i, v)$. The indicators I and J may be rewritten for $(u_1, \dots, u_d) \in \mathcal{U}_u^d$:

$$\begin{aligned} I(u_1, \dots, u_d) &= \sum_{i=1}^d \mathbb{E} \left((F_{X_i}^{-1}(U) - u_i) \mathbb{1}_{\{U \geq w_i, U \leq v\}} \right) \\ J(u_1, \dots, u_d) &= \sum_{i=1}^d \mathbb{E} \left((F_{X_i}^{-1}(U) - u_i) \mathbb{1}_{\{U \geq M_i\}} \right). \end{aligned}$$

We remark that since $(u_1, \dots, u_d) \in \mathcal{U}_u^d$, and F_{X_i} is increasing for all $i \in \{1, \dots, d\}$, we cannot have that $w_i < v$ for all i , so that, I is not trivially equal to 0. We use Lagrange multiplier to get that the minimum of I and J are reached in \mathcal{U}_u^d respectively for:

- $\mathbb{P}(U \geq w_i, U \leq v) = \mathbb{P}(U \geq w_j, U \leq v)$, for $i, j = 1, \dots, d$,
- $\mathbb{P}(U \geq M_i) = \mathbb{P}(U \geq M_j)$, for $i, j = 1, \dots, d$.

These equality are achieved if and only if $w_i = w_j = v$ or in other words if $F_{X_i}(u_i) = F_{X_j}(u_j)$. We remark that the minimum of I is then 0.

In another hand, if $\alpha_i = \frac{u_i}{u}$ is independent of u for all $i \in \{1, \dots, d\}$, and since $F_{X_i}(\alpha_i u) = F_{X_j}(\alpha_j u)$ and F_{X_i} is an increasing function for all $i \in \{1, \dots, d\}$, then $F_{X_j}^{-1} \circ F_{X_i}$ are a linear functions for all $(i, j) \in \{1, \dots, 2\}^2$

$$X_i \stackrel{\mathcal{L}}{=} \frac{\alpha_i}{\alpha_j} X_j, \quad \forall (i, j) \in \{1, \dots, 2\}^2.$$

In the other way, if there exist a non negative constants a_{ij} such that

$$X_i \stackrel{\mathcal{L}}{=} a_{ij} X_j, \quad \forall (i, j) \in \{1, \dots, 2\}^2,$$

then the optimal allocation satisfies $u_i = a_{ij} u_j$ for all $(i, j) \in \{1, \dots, 2\}^2$, and from that we deduce

$$\alpha_i = \frac{1}{1 + \sum_{\ell=1, \ell \neq i}^d a_{\ell i}},$$

for all $i \in \{1, \dots, d\}$ which is independent of u . □

The following three corollaries are direct applications of Proposition 4.5.1 to some particular cases.

Example 4.5.2 (Comonotonic exponential model).

For comonotonic risks of exponential distributions $X_i \sim \exp(\beta_i)$, the optimal allocation by minimization of the two risk indicators is:

$$\forall i \in \{1, \dots, d\}, \quad u_i = \frac{1/\beta_i}{\sum_{j=1}^d 1/\beta_j} u.$$

It is noticeable that in this particular case, the optimal allocation for comonotonic risks coincides with the asymptotic allocation in the independence case (see Proposition 4.2.1). In fact, in the asymptotic case, the impact of dependence becomes negligible, and the optimal allocation is proportional to the risk level of each branch, which is also the same behavior in the comonotonic case.

Example 4.5.3 (Comonotonic log-normal model).

For comonotonic risks of log-normal distributions $X_i \sim LN(\mu_i, \sigma)$, the optimal allocation by minimization of the two risk indicators is:

$$\forall i \in \{1, \dots, d\}, \quad u_i = \frac{\exp(\mu_i)}{\sum_{\ell=1}^d \exp(\mu_\ell)} u.$$

Example 4.5.4 (Comonotonic Pareto model).

For comonotonic risks of Pareto distributions of the same shape parameter γ : $X_i \sim Pa(\gamma, \lambda_i)$, the optimal allocation by minimization of the two risk indicators is:

$$\forall i \in \{1, \dots, d\}, \quad u_i = \frac{\lambda_i}{\sum_{\ell=1}^d \lambda_\ell} u.$$

Contrary to the exponential case, this result does not coincide with the asymptotic behavior obtained for independent Pareto distributions (see Proposition 4.2.8).

Conclusion

In this article, we have studied the capital allocation based on the level of the group capital. It gives an idea of the capital level impact on the sensitivity of its allocation between branches. The comparison between the asymptotic optimal allocation, in the case of sub-exponential and exponential distributions, underscores the impact of the risks nature on the behavior of the allocation for a very large capital. Explicit and semi-explicit formulas for different models could allow practical applications.

Our empirical study on the bivariate FGM model indicates that for large enough capital u the percentage allocated to the riskiest branch is a decreasing function of the dependence strength. Remark that for Marshall-Olkin and correlated Pareto models, the dependence structure and the marginal laws are both functions of the shock parameter (λ_0) or the mixture parameter (γ). So that, for these models we can assess a conclusion on the behavior of the capital allocation with respect to λ_0 or γ but we have to keep in mind that their behaviors not only reflect the impact of dependence but also the impact of the marginal laws.

Compared to the classical risk allocation methods, the allocation by minimizing multivariate risk indicators takes into account the capital level. This seems to be a more acceptable behavior in a capital allocation operation.

The risk aggregation transforms the global risk portfolio in a univariate risk. At the opposite, our allocation is based on a multivariate analysis. The goal of this paper is to highlight the importance of the dependence modeling in the process of an allocation capital operation. The capital allocation is also sensitive to the risks nature, we tried in this paper to make more evident the impact of distributions' nature on the allocation behavior. Since the capital allocation is an important financial decision, especially for groups, insurers must be very careful in their risk and dependence modeling choices to get an allocation that reduces negative impacts of business lines ruin.

The various behaviors observed depending on dependence strength, capital, and marginal laws could help practitioners in selecting a model in coherence with expert advices.

Bibliography

- [1] S. Asmussen. *Ruin probabilities. Advanced series on statistical science & applied probability, Volume 2.* 2000. 151
- [2] M. Bargès, H. Cossette, and E. Marceau. Tvar-based capital allocation with copulas. *Insurance: Mathematics and Economics*, 45(3):348–361, 2009. 144
- [3] K. Borch. Equilibrium in a reinsurance market. *Econometrica*, (30):424–444, 1962. 9, 131, 167
- [4] J. Cai and H. Li. Dependence properties and bounds for ruin probabilities in multivariate compound risk models. *Journal of Multivariate Analysis*, 98(4):757 – 773, 2007. 126, 144
- [5] P. Cénac, S. Loisel, V. Maume-Deschamps, and C. Prieur. Risk indicators with several lines of business: comparison, asymptotic behavior and applications to optimal reserve allocation. *Annales de l'ISUP*, 58(3), 2014. 17, 27, 126, 144, 148, 151, 162
- [6] P. Cénac, V. Maume-Deschamps, and C. Prieur. Some Multivariate risk indicators: Minimization by using a Kiefer-Wolfowitz approach to the mirror stochastic algorithm . *Statistics and Risk Modeling*, 29(1):47–71, 2012. 29, 126, 127, 128, 136, 143, 144, 145, 146
- [7] S. Coles, J. Heffernan, and J. Tawn. Dependence measures for extreme value analyses. *Extremes*, 2(4):339–365, 1999. 147
- [8] H. Cossette, M. Mailhot, and E. Marceau. Tvar-based capital allocation for multivariate compound distributions with positive continuous claim amounts. *Insurance: Mathematics and Economics*, 50(2):247–256, 2012. 144
- [9] A. Cousin and E. Di Bernardino. On multivariate extensions of value-at-risk. *Journal of Multivariate Analysis*, 119:32–46, 2013. 144
- [10] H. Cramér. *On the mathematical theory of risk.* Centraltryckeriet, 1930. 144
- [11] G. De Luca and G. Riveccio. Multivariate tail dependence coefficients. *Statistical methods for the analysis of large data-sets, Book of short papers*, 2009. 147
- [12] M. Denault. Coherent allocation of risk capital. *Journal of risk*, 4:1–34, January 2001. 126, 130, 144
- [13] J. Dhaene, E.A. Valdez, A. Tsanakas, and S. Vanduffel. Optimal capital allocation principles. *The Journal of Risk and Insurance*, 79(1):1–28, 2012. 17, 126, 144, 145
- [14] Paul Embrechts, Filip Lindskog, and Alexander McNeil. Modelling dependence with copulas. *Rapport technique, Département de mathématiques, Institut Fédéral de Technologie de Zurich, Zurich*, 2001. 161
- [15] M. Fréchet. Sur les tableaux de corrélation dont les marges sont données. *Annales de l'université de Lyon, Section.A, Series 3.*, (14):53–77, 1951. 9, 131, 167
- [16] W. Hoeffding. Masstabinvariante Korrelationstheorie. *Schriften des mathematischen Instituts und des Instituts für angewandte Mathematik der Universität Berlin*, (5):179–233, 1940. 9, 131, 167
- [17] H. Hult and F. Lindskog. Heavy-tailed insurance portfolios: buffer capital and ruin probabilities. Technical Report 1441, School of ORIE, Cornell University, January 2006. 144
- [18] H. Joe. *Multivariate models and multivariate dependence concepts.* CRC Press, 1997. 8, 147, 162
- [19] E. Jouini, M. Meddeb, and N. Touzi. Vector-valued coherent risk measures. *Finance and Stochastics*, 8(4):531–552, 2004. 4, 39, 47, 53, 128, 144, 177

- [20] F. Lundberg. *1. Approximerad framställning af sannolikhetsfunktionen: 2. Återförsäkring af kollektivrisker. Akademisk afhandling.* Almqvist & Wiksells, 1903. 144
- [21] E. Marceau. *Modélisation et évaluation quantitative des risques en actuariat.* Springer, 2013. 162
- [22] V. Maume-Deschamps, D. Rullière, and K. Said. Impact of dependence on some multivariate risk indicators. *Methodology and Computing in Applied Probability*, 2016. 127, 143
- [23] V. Maume-Deschamps, D. Rulliere, and K. Said. On a capital allocation by minimization of some risk indicators. *European Actuarial Journal*, pages 1–20, 2016. 40, 47, 125, 143, 144, 145, 146, 156
- [24] R.B. Nelsen. *An introduction to copulas.* Springer, 2007. 70, 77, 92, 111, 157, 161
- [25] D. Oakes. Bivariate survival models induced by frailties. *Journal of the American Statistical Association*, 84(406):487–493, 1989. 163
- [26] M. Shaked and J.G. Shanthikumar. *Stochastic Orders.* Springer Series in Statistics, 2007. 8, 131, 156
- [27] D. Tasche. Euler allocation: Theory and practice. Technical Report arXiv:0708.2542, Aug 2007. 126, 131, 144
- [28] H-C. Yeh. The frailty and the Archimedean structure of the general multivariate Pareto distributions. *Bulletin Institute of Mathematics Acedemia Sinica*, 2(3):713–729, 2007. 163

4.6 Appendix: Proofs

4.6.1 Proof of Proposition 4.3.1

Proof. The bivariate density function is the following:

$$f_{X_1, X_2}(x_1, x_2) = (1 + \theta)f(x_1, x_2, \beta_1, \beta_2) + \theta f(x_1, x_2, 2\beta_1, 2\beta_2) \\ - \theta f(x_1, x_2, 2\beta_1, \beta_2) - \theta f(x_1, x_2, \beta_1, 2\beta_2)f_{X_1, X_2}(x_1, x_2),$$

where f is the function $f(x, t, a, b) = abe^{-ax}e^{-bt}$.

We use the equality: $f_{X_1, S=X_1+X_2}(x_1, s) = f_{X_1, X_2}(x_1, s - x_1)\mathbb{1}_{\{s \geq x_1\}}$ for all $s \geq x_1$, to find the expression of $F_{X_1, S}(x_1, s)$, using a double integration:

$$F_{X_1, S}(x_1, s) = \int_0^s \int_0^{x_1} f_{X_1, X_2}(x, t - x)\mathbb{1}_{\{t \geq x\}} dx dt = \int_0^{x_1} \int_x^s f_{X_1, X_2}(x, t - x) dt dx \\ = (1 + \theta)F(x_1, s, \beta_1, \beta_2) + \theta F(x_1, s, 2\beta_1, 2\beta_2) - \theta F(x_1, s, 2\beta_1, \beta_2) - \theta F(x_1, s, \beta_1, 2\beta_2),$$

where F is the following function:

$$F(x_1, s, a, b) = \int_0^{x_1} \int_x^s abe^{-(a-b)x}e^{-bt} dt dx = 1 - e^{-ax_1} + \frac{a}{b-a}e^{-bs} - \frac{a}{b-a}e^{-bs+(b-a)x_1}.$$

The same way and by the symmetry of the FGM model:

$$F_{X_2, S}(x_2, s) = (1 + \theta)F(x_2, s, \beta_2, \beta_1) + \theta F(x_2, s, 2\beta_2, 2\beta_1) - \theta F(x_2, s, 2\beta_2, \beta_1) - \theta F(x_2, s, \beta_2, 2\beta_1).$$

Using $\mathbb{P}(X_i > u_i, S \leq u) = \mathbb{P}(S \leq u) - \mathbb{P}(X_i \leq u_i, S \leq u)$, the optimal allocation is the unique solution in \mathcal{U}_u^2 of the equation: $F_{X_1, S}(u_1, u) = F_{X_2, S}(u_2, u)$. Then, the optimal allocation is determined by α_1 the solution of the equation: $F_{X_1, S}(\alpha_1 u, u) = F_{X_2, S}((1 - \alpha_1)u, u)$.

Since,

$$F_{X_1, S}(\alpha_1 u, u) = 1 - h(\alpha_1) + \frac{1 + \theta}{\gamma - 1}[h(\gamma) - h(\gamma + \alpha_1 - \gamma\alpha_1)] + \frac{\theta}{\gamma - 1}[h(2\gamma) - h(2\gamma + 2\alpha_1 - 2\gamma\alpha_1)] \\ - \frac{2\theta}{\gamma - 2}[h(\gamma) - h(\gamma + 2\alpha_1 - \gamma\alpha_1)] - \frac{\theta}{2\gamma - 1}[h(2\gamma) - h(2\gamma + \alpha_1 - 2\gamma\alpha_1)],$$

and,

$$\begin{aligned}
F_{X_2,S}((1-\alpha_1)u, u) &= 1 - h(\gamma(1-\alpha_1)) + \frac{(1+\theta)\gamma}{1-\gamma}[h(1) - h(\gamma + \alpha_1 - \gamma\alpha_1)] \\
&\quad + \frac{\theta\gamma}{1-\gamma}[h(2) - h(2\gamma + 2\alpha_1 - 2\gamma\alpha_1)] \\
&\quad - \frac{2\gamma\theta}{1-2\gamma}[h(1) - h(2\gamma + \alpha_1 - 2\gamma\alpha_1)] - \frac{\theta\gamma}{2-\gamma}[h(2) - h(\gamma + 2\alpha_1 - \gamma\alpha_1)],
\end{aligned}$$

we deduce from that the equation presented in the proposition 4.3.1. \square

4.6.2 Proof of Proposition 4.4.1

Proof. The joint distribution function is given by:

$$\begin{aligned}
F_{X_1,S}(x_1, s) &= \int_0^s \int_0^{x_1} f_{X_1,S}(x, t-x) \mathbb{1}_{\{t>x\}} dx dt \\
&= \int_0^s \int_0^{x_1} f_{X_1,S}^1(x, t-x) \mathbb{1}_{\{2x>t>x\}} dx dt + \int_0^s \int_0^{x_1} f_{X_1,S}^2(x, t-x) \mathbb{1}_{\{t>2x\}} dx dt \\
&\quad + \int_0^s \int_0^{x_1} f_{X_1,S}^0(x, t-x) \mathbb{1}_{\{t=2x\}} dx dt.
\end{aligned}$$

we distinguish between two cases:

Case $s > 2x_1$: in this case,

$$\begin{aligned}
\int_0^s \int_0^{x_1} f_{X_1,S}^1(x, t-x) \mathbb{1}_{\{2x>t>x\}} dx dt &= \int_{2x_1}^s \int_0^{x_1} f_{X_1,S}^1(x, t-x) \mathbb{1}_{\{2x>t>x\}} dx dt \\
&\quad + \int_0^{2x_1} \int_0^{x_1} f_{X_1,S}^1(x, t-x) \mathbb{1}_{\{2x>t>x\}} dx dt \\
&= \int_0^{2x_1} \int_0^{x_1} f_{X_1,S}^1(x, t-x) \mathbb{1}_{\{2x>t>x\}} dx dt \\
&= \int_0^{2x_1} \int_{t/2}^{\min(x_1,t)} f_{X_1,S}^1(x, t-x) dx dt \\
&= \int_0^{x_1} \int_{t/2}^t f_{X_1,S}^1(x, t-x) dx dt + \int_{x_1}^{2x_1} \int_{t/2}^{x_1} f_{X_1,S}^1(x, t-x) dx dt,
\end{aligned}$$

and,

$$\int_0^s \int_0^{x_1} f_{X_1,S}^2(x, t-x) \mathbb{1}_{\{t>2x\}} dx dt = \int_0^{x_1} \int_0^s f_{X_1,S}^2(x, t-x) \mathbb{1}_{\{t>2x\}} dt dx = \int_0^{x_1} \int_{2x}^s f_{X_1,S}^2(x, t-x) dt dx,$$

and,

$$\int_0^s \int_0^{x_1} f_{X_1,S}^0(x, t-x) \mathbb{1}_{\{t=2x\}} dx dt = \frac{\lambda_0}{\lambda_0 + \lambda_1 + \lambda_2} (1 - e^{-(\lambda_0 + \lambda_1 + \lambda_2)x_1}),$$

then, we deduce the explicit expression of $F_{X_1,S}(x_1, s)$:

$$\begin{aligned}
F_{X_1,S}(x_1, s) &= \frac{2\beta_1\lambda_2}{(\beta_1 - \lambda_2)(\beta_1 + \lambda_2)} (1 - e^{-(\beta_1 + \lambda_2)x_1}) - \frac{\lambda_2}{\beta_1 - \lambda_2} (1 - e^{-\beta_1 x_1}) - \frac{\beta_1}{\beta_1 - \lambda_2} (e^{-\beta_1 x_1} - e^{-(\beta_1 + \lambda_2)x_1}) \\
&\quad + \frac{\lambda_1}{\lambda_1 + \beta_2} (1 - e^{-(\lambda_1 + \beta_2)x_1}) - \frac{\lambda_1}{\lambda_1 - \beta_2} e^{-\beta_2 s} + \frac{\lambda_1}{\lambda_1 - \beta_2} e^{-(\lambda_1 - \beta_2)x_1 - \beta_2 s} \\
&\quad + \frac{\lambda_0}{\lambda_0 + \lambda_1 + \lambda_2} (1 - e^{-(\lambda_0 + \lambda_1 + \lambda_2)x_1}).
\end{aligned}$$

Case $2x_1 > s > x_1$:

$$\begin{aligned} \int_0^s \int_0^{x_1} f_{X_1,S}^1(x, t-x) \mathbb{1}_{\{2x > t > x\}} dx dt &= \int_0^s \int_{t/2}^{\min(x_1, t)} f_{X_1,S}^1(x, t-x) dx dt \\ &= \int_0^{x_1} \int_{t/2}^t f_{X_1,S}^1(x, t-x) dx dt + \int_{x_1}^s \int_{t/2}^{x_1} f_{X_1,S}^1(x, t-x) dx dt, \end{aligned}$$

and,

$$\int_0^s \int_0^{x_1} f_{X_1,S}^2(x, t-x) \mathbb{1}_{\{t > 2x\}} dx dt = \int_0^s \int_0^{t/2} f_{X_1,S}^2(x, t-x) dx dt,$$

and,

$$\int_0^s \int_0^{x_1} f_{X_1,S}^0(x, t-x) \mathbb{1}_{\{t=2x\}} dx dt = \frac{\lambda_0}{\lambda_0 + \lambda_1 + \lambda_2} (1 - e^{-(\lambda_0 + \lambda_1 + \lambda_2)s/2}),$$

then, we deduce also in this case, the explicit expression of $F_{X_1,S}(x_1, s)$:

$$\begin{aligned} F_{X_1,S}(x_1, s) &= \frac{2\beta_1\lambda_2}{(\beta_1 - \lambda_2)(\beta_1 + \lambda_2)} (1 - e^{-(\beta_1 + \lambda_2)s/2}) - \frac{\lambda_2}{\beta_1 - \lambda_2} (1 - e^{-\beta_1 x_1}) \\ &\quad - \frac{\beta_1}{\beta_1 - \lambda_2} (e^{-\beta_1 x_1} - e^{-(\beta_1 - \lambda_2)x_1 - \lambda_2 s}) + \frac{\lambda_1}{\lambda_1 - \beta_2} (1 - e^{-\beta_2 s}) \\ &\quad - \frac{2\lambda_1\beta_2}{(\lambda_1 - \beta_2)(\lambda_1 + \beta_2)} (1 - e^{-(\lambda_1 + \beta_2)s/2}) + \frac{\lambda_0}{\lambda_0 + \lambda_1 + \lambda_2} (1 - e^{-(\lambda_0 + \lambda_1 + \lambda_2)s/2}). \end{aligned}$$

We remark that $\lambda_0 + \lambda_1 + \lambda_2 = \lambda_1 + \beta_2 = \lambda_2 + \beta_1$, and we suppose that $\lambda_1 > \lambda_2$. Using the monotony property, we deduce that $1 > \alpha_1 > 1/2$, then $2\alpha_1 u > u > \alpha_1 u$. So, for $u_1 = \alpha_1 u$, and $g(x) = \exp(-xu)$:

$$F_{X_1,S}(\alpha_1 u, u) = 1 - g(\beta_1 \alpha_1) - \frac{\lambda_1}{\lambda_1 - \beta_2} g(\beta_2) + \frac{\beta_1}{\beta_1 - \lambda_2} g((\beta_1 - \lambda_2)\alpha_1 + \lambda_2) + g(\lambda_s/2) \left[\frac{\lambda_1}{\lambda_1 - \beta_2} - \frac{\beta_1}{\beta_1 - \lambda_2} \right],$$

and,

$$F_{X_2,S}((1 - \alpha_1)u, u) = 1 - g(\beta_2(1 - \alpha_1) + \frac{\lambda_2}{\lambda_2 - \beta_1} [g((\lambda_2 - \beta_1)(1 - \alpha_1) + \beta_1) - g(\beta_1)]).$$

That is sufficient to get 4.4.1. □

Conclusion

Conclusion et perspectives

La gestion des risques en assurance évolue d'une façon permanente et rapide ces dernières années. La théorie multivariée du risque a fait son apparition suite à la nécessité de la prise en compte de la dépendance. L'approche multivariée trouve son importance dans plusieurs problématiques qui impactent directement les résultats des compagnies d'assurance, comme l'allocation du capital et la recherche d'un programme de réassurance optimal. Dans ce contexte, nous avons essayé dans le cadre de cette thèse de comprendre une partie des problématiques actuarielles de nature multivariée et de proposer des réponses sous forme de mesures de risque multidimensionnelles et méthodologies d'allocation du capital.

Dans le premier chapitre de ce travail, nous nous sommes basés sur l'importance de la propriété de l'élicitabilité des mesures de risque pour proposer différentes constructions de mesures expectiles multivariées. Les propriétés des mesures obtenues dépendent de la définition choisie. Cette décision est laissée à l'utilisateur et doit répondre à ses besoins d'utilisation pratique, à la nature de son portefeuille de risques et à ses pratiques de modélisation de la dépendance. Nous avons ensuite étudié les propriétés de cohérence de la construction à l'aide de matrices semi-définies positives à coefficients positifs. La même étude peut être menée pour toutes les autres méthodologies. L'analyse de la robustesse par rapport à une distance comme Wasserstein sera nécessaire pour vérifier la stabilité des mesures obtenues. Dans le cas général, l'expectile multivarié ne peut être obtenu explicitement, nous nous sommes donc appuyés sur des outils d'approximation stochastique pour fournir une première solution pratique qui permet de déterminer des valeurs approchées des vecteurs expectiles. Des résultats sous forme de systèmes d'équations explicites peuvent être obtenus pour des cas particuliers de modèles de distributions multivariées.

La richesse de la théorie des mesures de risque univariées nous encourage à continuer la recherche dans ce domaine pour développer d'autres familles de mesures multidimensionnelles, qui répondent à des besoins économiques précis tels que la quantification du risque systémique. Le concept de cohérence n'est pas totalement caractérisé dans le cas des mesures vectorielles, nous essayerons dans la suite de formaliser une axiomatisation de la propriété de la cohérence qui généralise la définition univariée introduite par Artzner et al. (1999) [3]. D'un autre côté, nous nous intéresserons de près aux mesures sous forme d'ensembles, dont la cohérence était introduite par Jouini et al. (2004) [71]. Des mesures vectorielles peuvent être obtenues par ajout de critère d'optimisation aux ensembles mesures.

Le deuxième chapitre de la thèse a été consacré à l'étude du comportement asymptotique des mesures expectiles multivariées dans un environnement de variations régulières. Nous avons utilisé la relation existante entre la mesure spectrale d'un vecteur aléatoire et sa fonction de dépendance des queues droites pour caractériser le comportement asymptotique des expectiles. Cette étude nous a permis de visualiser la régularité des composantes du vecteur expectile dans le cas de l'équivalence des queues marginales et de proposer des estimateurs statistiques dans les cas des modèles Fréchet indépendant asymptotiquement et Fréchet comonotone.

Ce travail s'est focalisé sur le cas du domaine d'attraction de Fréchet qui englobe les distributions à queues lourdes, très présentes dans la modélisation des risques actuariels. Cette analyse requiert également le traitement des cas d'appartenance des distributions marginales aux autres domaines d'attraction Gumbel et Weibull. Un travail de recherche dans cet objectif est en cours de finalisation,

et ses premiers résultats confirment que le comportement des composantes de l'expectile asymptotique reste régulier dans ces deux cas également. Durant cette partie, nous avons imposé des hypothèses d'existence des limites, qui sont vérifiées numériquement dans de nombreux cas mais restent en attente de démonstration théorique. La construction des estimateurs des expectiles multivariés en cas d'indépendance asymptotique a été faite à l'aide des outils de la théorie des valeurs extrêmes dans sa version univariée. L'utilisation de la littérature récente de la théorie multivariée des valeurs extrêmes sera plus pertinente vu la nature de ces mesures. L'estimation de l'expectile asymptotique dans le cas général sera possible si on peut estimer efficacement des intégrales des fonctions de la dépendance des queues. Une autre dernière perspective, qui nous semble intéressante pour améliorer cette partie, sera l'analyse asymptotique dans un environnement à variations régulières cachées. Ce concept dernièrement introduit par Resnick (2002) [94] commence à attirer la curiosité académique grâce à ses importants champs d'applications possibles.

Les normes Solvabilité 2 proposent une méthode standard pour la détermination des capitaux de Solvabilité réglementaires, mais l'allocation de ces capitaux reste un exercice libre qui entre complètement dans les choix internes de l'assureur. Nous avons donc choisi d'étudier la problématique de l'allocation du capital qui concerne particulièrement les groupes d'assurance et les assureurs multi-branches. Nous nous sommes concentrés sur une approche par minimisation du risque global de ruine du groupe qui diffère des différentes techniques classiques d'allocation du risque, fondées sur l'idée de la détermination de l'impact marginal de chaque risque sur le groupe.

Le troisième chapitre de cette thèse est composé d'un article qui présente cette approche multivariée d'allocation et étudie sa cohérence. Il introduit une caractérisation axiomatique de la cohérence économique d'une méthodologie d'allocation, indépendamment des mesures utilisées pour déterminer le capital à allouer dans la phase agrégation. Une section de ce chapitre est consacrée à la discussion du meilleur choix possible de méthode pour réaliser cette importante opération en assurance. Nous sommes convaincus qu'une allocation par minimisation du risque est en conformité avec l'objectif Solvabilité 2 de limiter la probabilité de la ruine. Cette méthode d'allocation constitue une réponse multivariée à une problématique de la même nature.

Une méthode d'allocation du capital doit être sensible aux distributions des risques, mais aussi à celle de la structure de dépendance. Partant de ce constat, on s'est proposé, dans le chapitre quatre de cette thèse, de traiter l'impact de la dépendance sur l'allocation optimale par minimisation d'indicateurs multivariés de risque. Nous avons ainsi étudié d'abord le cas de l'indépendance pour isoler l'impact des distributions marginales, pour des modèles exponentiel et sous-exponentiel, sur la composition de l'allocation et son comportement asymptotique. L'impact de la dépendance a été ensuite étudié pour les cas de l'indépendance asymptotique et de dépendance asymptotique à l'aide d'une sélection particulière de modèles de dépendance et de mélange. L'allocation optimale est entièrement déterminée dans le cas comonotone. Ce chapitre confirme que la méthode d'allocation par minimisation du risque prend en compte tous les aspects de la modélisation multivariée des risques. Elle a la particularité d'être sensible au niveau du capital de l'allocation, et l'aversion au risque est traduite par le choix de l'indicateur minimisé.

En ce qui concerne les perspectives en matière de construction de méthodes d'allocation optimale, adaptées à la nature du capital à allouer et à l'aversion au risque de l'assureur, il peut être intéressant d'étendre les résultats théoriques obtenus à une classe plus grande d'indicateurs de risque multivariés, et d'établir dans ce cadre des stratégies qui minimisent le rapport risque/rentabilité. On peut notamment étudier les différents choix possibles des fonctions de pénalités afin de mieux modéliser la différence des natures de risques. Enfin, cette méthode qui présente des garanties théoriques satisfaisantes mériterait d'être exploitée, dans le monde actuariel, sur des données réelles.

Bibliographie

- [1] B. Abdous and R. Theodorescu. Note on the spatial quantile of a random vector. *Statistics & Probability Letters*, 13(4) :333–336, 1992. 54
- [2] C. Acerbi. Spectral measures of risk : a coherent representation of subjective risk aversion. *Journal of Banking & Finance*, 26(7) :1505–1518, 2002. 6
- [3] P. Artzner, F. Delbaen, J-M. Eber, and D. Heath. Coherent measures of risk. *Mathematical Finance*, 9(3) :203–228, 1999. 4, 39, 44, 47, 130, 131, 177
- [4] S. Asmussen. *Ruin probabilities. Advanced series on statistical science & applied probability, Volume 2.* 2000. 151
- [5] A.A. Balkema and L. De Haan. Residual life time at great age. *The Annals of probability*, pages 792–804, 1974. 10
- [6] D. Balog. Capital allocation in financial institutions : the euler method. Iehas discussion papers, Institute of Economics, Hungarian Academy of Sciences, June 2011. 126
- [7] M. Bargès, H. Cossette, and E. Marceau. Tvar-based capital allocation with copulas. *Insurance : Mathematics and Economics*, 45(3) :348–361, 2009. 144
- [8] B. Basrak. *The sample autocorrelation function of non-linear time series.* PhD thesis, University of Groningen, 2000. 90, 91
- [9] J. Beirlant, P. Vynckier, and J.L. Teugels. Tail index estimation, pareto quantile plots regression diagnostics. *Journal of the American statistical Association*, 91(436) :1659–1667, 1996. 14
- [10] F. Bellini and V. Bignozzi. On elicitable risk measures. *Quantitative Finance*, 15(5) :725–733, 2015. 6, 40, 41, 42, 43, 44, 128
- [11] F. Bellini and E. Di Bernardino. Risk management with expectiles. *The European Journal of Finance*, pages 1–20, 2015. 40, 45, 46, 87, 104, 110
- [12] F. Bellini, B. Klar, A. Müller, and E.R. Gianin. Generalized quantiles as risk measures. *Insurance : Mathematics and Economics*, 54 :41–48, 2014. 6, 40, 46
- [13] N. Bingham, C. M Goldie, and J. Teugels. *Regular variation*, volume 27. Cambridge university press, 1989. 87, 89, 114
- [14] J.R Blum. Multidimensional stochastic approximation methods. *The Annals of Mathematical Statistics*, pages 737–744, 1954. 40, 64
- [15] K. Borch. Equilibrium in a reinsurance market. *Econometrica*, (30) :424–444, 1962. 9, 131, 167
- [16] M.K. Brunnermeier and P. Cheridito. Measuring and allocating systemic risk. *Available at SSRN 2372472*, 2014. 130
- [17] A. Buch and G. Dorffleitner. Coherent risk measures, coherent capital allocations and the gradient allocation principle. *Insurance : Mathematics and Economics*, 42(1) :235–242, 2008. 39
- [18] A. Buch and G. Dorffleitner. Coherent risk measures, coherent capital allocations and the gradient allocation principle. *Insurance : Mathematics and Economics*, 42(1) :235 – 242, 2008. 126
- [19] A. Buch, G. Dorffleitner, and M. Wimmer. Risk capital allocation for rorac optimization. *Journal of Banking & Finance*, 35(11) :3001–3009, August 2011. 126
- [20] J. Cai and H. Li. Dependence properties and bounds for ruin probabilities in multivariate compound risk models. *Journal of Multivariate Analysis*, 98(4) :757 – 773, 2007. 126, 144

- [21] A. Charpentier and J. Segers. Tails of multivariate archimedean copulas. *Journal of Multivariate Analysis*, 100(7) :1521–1537, 2009. 101
- [22] P. Chaudhuri. On a geometric notion of quantiles for multivariate data. *Journal of the American Statistical Association*, 91(434) :862–872, 1996. 54
- [23] P. Cénac, S. Loisel, V. Maume-Deschamps, and C. Prieur. Risk indicators with several lines of business : comparison, asymptotic behavior and applications to optimal reserve allocation. *Annales de l'ISUP*, 58(3), 2014. 17, 27, 126, 144, 148, 151, 162
- [24] P. Cénac, V. Maume-Deschamps, and C. Prieur. Some Multivariate risk indicators : Minimization by using a Kiefer-Wolfowitz approach to the mirror stochastic algorithm . *Statistics and Risk Modeling*, 29(1) :47–71, 2012. 29, 126, 127, 128, 136, 143, 144, 145, 146
- [25] S. Coles, J. Heffernan, and J. Tawn. Dependence measures for extreme value analyses. *Extremes*, 2(4) :339–365, 1999. 147
- [26] H. Cossette, M. Mailhot, and E. Marceau. Tvar-based capital allocation for multivariate compound distributions with positive continuous claim amounts. *Insurance : Mathematics and Economics*, 50(2) :247–256, 2012. 144
- [27] A. Cousin and E. Di Bernardino. On multivariate extensions of value-at-risk. *Journal of Multivariate Analysis*, 119 :32–46, 2013. 40
- [28] A. Cousin and E. Di Bernardino. On multivariate extensions of value-at-risk. *Journal of Multivariate Analysis*, 119 :32–46, 2013. 144
- [29] A. Cousin and E. Di Bernardino. On multivariate extensions of conditional-tail-expectation. *Insurance : Mathematics and Economics*, 55 :272–282, 2014. 40
- [30] H. Cramér. *On the mathematical theory of risk*. Centraltryckeriet, 1930. 144
- [31] M. Csörgő and L. Viharos. Estimating the tail index. *Asymptotic Methods in Probability and Statistics (B. Szyszkowicz, ed.)*, page 833–881, 1998. 14
- [32] J. Danielsson and C.G. de Vries. *Extreme returns, tail estimation, and value-at-risk*. University of Iceland, Institute of Economic Studies, 1997. 11
- [33] A. Daouia, S. Girard, and G. Stupfler. Estimation of tail risk based on extreme expectiles. In *Workshop Extremes-Copulas-Actuarial science*, 2016. 46
- [34] L. De Haan and A. Ferreira. *Extreme value theory : an introduction*. Springer Science & Business Media, 2007. 87, 90, 116
- [35] L. De Haan and S. Resnick. Limit theory for multivariate sample extremes. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 40(4) :317–337, 1977. 90
- [36] G. De Luca and G. Riveccio. Multivariate tail dependence coefficients. *Statistical methods for the analysis of large data-sets, Book of short papers*, 2009. 147
- [37] A.L.M. Dekkers and L. De Haan. On the estimation of the extreme-value index and large quantile estimation. *The Annals of Statistics*, pages 1795–1832, 1989. 13
- [38] A.L.M. Dekkers, J.H.J. Einmahl, and L. De Haan. A moment estimator for the index of an extreme-value distribution. *The Annals of Statistics*, pages 1833–1855, 1989. 13
- [39] F. Delbaen, F. Bellini, V. Bignozzi, and J.F. Ziegel. Risk measures with the cxls property. *arXiv preprint arXiv :1411.0426*, 2014. 43
- [40] M. Denault. Coherent allocation of risk capital. *Journal of risk*, 4 :1–34, 2001. 28, 39
- [41] M. Denault. Coherent allocation of risk capital. *Journal of risk*, 4 :1–34, January 2001. 126, 130, 144
- [42] M. Denuit and A. Charpentier. *Mathématiques de l'assurance non-vie*. Economica, 2005. 10
- [43] J. Dhaene, E.A. Valdez, A. Tsanakas, and S. Vanduffel. Optimal capital allocation principles. *The Journal of Risk and Insurance*, 79(1) :1–28, 2012. 17, 126, 144, 145
- [44] E. Di Bernardino. *Modélisation de la dépendance et mesures de risque multidimensionnelles*. PhD thesis, Citeseer, 2013. 40

- [45] E. Di Bernardino, JM. Fernández-Ponce, F. Palacios-Rodríguez, and MR. Rodríguez-Griñolo. On multivariate extensions of the conditional value-at-risk measure. *Insurance : Mathematics and Economics*, 61 :1–16, 2015. 40
- [46] M. Duflo. Random iterative models, volume 34 of applications of mathematics, 1997. 64
- [47] P. Embrechts. *Extremes and integrated risk management*. Risk Books, 2000. 10
- [48] P. Embrechts, C. Klüppelberg, and T. Mikosch. *Modelling extremal events*, volume 33. Springer Science & Business Media, 1997. 14, 87, 117
- [49] P. Embrechts, G. Puccetti, L. Rüschendorf, R. Wang, and A. Beleraj. An academic response to basel 3.5. *Risks*, 2(1) :25–48, 2014. 39
- [50] Paul Embrechts, Filip Lindskog, and Alexander McNeil. Modelling dependence with copulas. *Rapport technique, Département de mathématiques, Institut Fédéral de Technologie de Zurich, Zurich*, 2001. 161
- [51] E. Emmer, M. Kratz, and D. Tasche. What is the best risk measure in practice? A comparison of standard measures. December 2013. 40, 41, 45, 46
- [52] T. Fischer. Risk capital allocation by coherent risk measures based on one-sided moments. *Insurance : Mathematics and Economics*, 32(1) :135 – 146, 2003. 126
- [53] T. Fissler and J.F. Ziegel. Higher order elicibility and osband’s principle. *arXiv preprint arXiv :1503.08123*, 2015. 46, 47
- [54] H. Föllmer, A. Schied, and T.J. Lyons. Stochastic finance. an introduction in discrete time. *The Mathematical Intelligencer*, 26(4) :67–68, 2004. 131
- [55] P. Fraysse. *Estimation récursive dans certains modeles de déformation*. PhD thesis, Université Sciences et Technologies-Bordeaux I, 2013. 65
- [56] M. Fréchet. Sur les tableaux de corrélation dont les marges sont données. *Annales de l’université de Lyon, Section.A, Series 3.*, (14) :53–77, 1951. 9, 131, 167
- [57] R. Frongillo and I. Kash. Vector-valued property elicitation. *Preprint*, 2014. 47
- [58] M. Garrido. *Modélisation des évènements rares et des quantiles extrêmes, méthodes de sélection de modèle pour les queues de distribution*. PhD thesis, Thèse de doctorat, Université J. Fourier, 2002. 10
- [59] B. Gnedenko. Sur la distribution limite du terme maximum d’une serie aleatoire. *Annals of mathematics*, pages 423–453, 1943. 9, 114
- [60] T. Gneiting. Making and Evaluating Point Forecasts . *Journal of the American Statistical Association*, 106(494) :746–762, June 2011. 6, 40, 41, 42, 43, 44, 128
- [61] A. H. Hamel and F. Heyde. Duality for set-valued measures of risk. *SIAM Journal on Financial Mathematics*, 1(1) :66–95, 2010. 53, 128
- [62] B.M. Hill. A simple general approach to inference about the tail of a distribution. *The annals of statistics*, 3(5) :1163–1174, 1975. 12, 26
- [63] W. Hoeffding. Masstabinvariante Korrelationstheorie. *Schriften des mathematischen Instituts und des Instituts für angewandte Mathematik der Universität Berlin*, (5) :179–233, 1940. 9, 131, 167
- [64] J.R.M. Hosking, J. R Wallis, and E.F. Wood. Estimation of the generalized extreme-value distribution by the method of probability-weighted moments. *Technometrics*, 27(3) :251–261, 1985. 11
- [65] J.R.M. Hosking and J.R. Wallis. Parameter and quantile estimation for the generalized pareto distribution. *Technometrics*, 29(3) :339–349, 1987. 11
- [66] H. Hult and F. Lindskog. Heavy-tailed insurance portfolios : buffer capital and ruin probabilities. Technical Report 1441, School of ORIE, Cornell University, January 2006. 144
- [67] W. Hürlimann. On higher-degree bivariate stop-loss transforms, with applications. *Blätter der DGVMF*, 25(3) :485–502, 2002. 52

- [68] A.F. Jenkinson. The frequency distribution of the annual maximum (or minimum) values of meteorological elements. *Quarterly Journal of the Royal Meteorological Society*, 81(348) :158–171, 1955. 9
- [69] H. Joe. *Multivariate models and multivariate dependence concepts*. CRC Press, 1997. 8, 147, 162
- [70] H. Joe, H. Li, and A.K. Nikoloulopoulos. Tail dependence functions and vine copulas. *Journal of Multivariate Analysis*, 101(1) :252–270, 2010. 101
- [71] E. Jouini, M. Meddeb, and N. Touzi. Vector-valued coherent risk measures. *Finance and Stochastics*, 8(4) :531–552, 2004. 4, 39, 47, 53, 128, 144, 177
- [72] W.K. Newey and J.L. Powell. Asymmetric Least Squares Estimation and Testing. *Econometrica*, 55(4) :819–847, July 1987. 6, 44
- [73] M. Kalkbrenner. An axiomatic characterization of capital allocations of coherent risk measures. *Quantitative Finance*, 9(8) :961–965, 2009. 126, 131
- [74] O. Kallenberg. Random measures, 3rd. Edition, Akademie Verlag, Berlin, Academic Press, New York, 1983. 90
- [75] J. Kiefer, J. Wolfowitz, et al. Stochastic estimation of the maximum of a regression function. *The Annals of Mathematical Statistics*, 23(3) :462–466, 1952. 64
- [76] R. Kiesel, R. Rühlicke, G. Stahl, and J. Zheng. Conceptualizing robustness in risk management. Available at SSRN 2065723, 2012. 46
- [77] C. Klüppelberg, G.I. Kuhn, and L. Peng. Semi-parametric models for the multivariate tail dependence function—the asymptotically dependent case. *Scandinavian Journal of Statistics*, 35(4) :701–718, 2008. 8, 91
- [78] M. Kratz and S.I. Resnick. The qq-estimator and heavy tails. *Stochastic Models*, 12(4) :699–724, 1996. 14
- [79] N.S. Lambert, D.M. Pennock, and Y. Shoham. Eliciting properties of probability distributions. In *Proceedings of the 9th ACM Conference on Electronic Commerce*, pages 129–138. ACM, 2008. 44, 46, 47
- [80] J. Lelong. *Etude asymptotique des algorithmes stochastiques et calcul du prix des options Parisiennes*. PhD thesis, Ecole des Ponts ParisTech, 2007. 68
- [81] H. Li and Y. Sun. Tail dependence for heavy-tailed scale mixtures of multivariate distributions. *Journal of Applied Probability*, pages 925–937, 2009. 92
- [82] F. Lundberg. *1. Approximerad framställning af sannolikhetsfunktionen : 2. Återförsäkring af kollektivrisiker. Akademisk afhandling*. Almqvist & Wiksells, 1903. 144
- [83] T. Mao and F. Yang. Risk concentration based on expectiles for extreme risks under fgm copula. *Insurance : Mathematics and Economics*, 64 :429–439, 2015. 46, 87
- [84] E. Marceau. *Modélisation et évaluation quantitative des risques en actuariat*. Springer, 2013. 162
- [85] V. Maume-Deschamps, D. Rullière, and K. Said. Impact of dependence on some multivariate risk indicators. *Methodology and Computing in Applied Probability*, 2016. 127, 143
- [86] V. Maume-Deschamps, D. Rullière, and K. Said. Multivariate extensions of expectiles risk measures. working paper or preprint, September 2016. 39
- [87] V. Maume-Deschamps, D. Rullière, and K. Said. On a capital allocation by minimization of some risk indicators. *European Actuarial Journal*, pages 1–20, 2016. 40, 47, 125, 143, 144, 145, 146, 156
- [88] A.J. McNeil and J. Nešlehová. Multivariate archimedean copulas, d-monotone functions and ℓ_1 -norm symmetric distributions. *The Annals of Statistics*, pages 3059–3097, 2009. 101
- [89] T. Mikosch. Modeling dependence and tails of financial time series. *Extreme Values in Finance, Telecommunications, and the Environment*, pages 185–286, 2003. 89, 91

- [90] R.B. Nelsen. *An introduction to copulas*. Springer, 2007. 70, 77, 92, 111, 157, 161
- [91] D. Oakes. Bivariate survival models induced by frailties. *Journal of the American Statistical Association*, 84(406) :487–493, 1989. 163
- [92] K.H. Osband. *Providing Incentives for Better Cost Forecasting*. University of California, Berkeley, 1985. 42, 43, 46
- [93] J. Pickands. Statistical inference using extreme order statistics. *the Annals of Statistics*, pages 119–131, 1975. 10, 11, 13
- [94] S. Resnick. Hidden regular variation, second order regular variation and asymptotic independence. *Extremes*, 5(4) :303–336, 2002. 178
- [95] S. Resnick. *Heavy-tail phenomena : probabilistic and statistical modeling*. Springer Science & Business Media, 2007. 89, 90
- [96] S. Resnick and C. Stărică. Smoothing the hill estimator. *Advances in Applied Probability*, pages 271–293, 1997. 14
- [97] S.I. Resnick. *Extreme values, regular variation and point processes*. Springer, 2013. 88
- [98] H. Robbins and S. Monro. A stochastic approximation method. *The annals of mathematical statistics*, pages 400–407, 1951. 40, 64
- [99] R.T. Rockafellar and S. Uryasev. The fundamental risk quadrangle in risk management, optimization and statistical estimation . *Surveys in Operations Research and Management Science*, (18) :33–53, March 2013. 44
- [100] N. S. Lambert. Elicitation and Evaluation of Statistical Forecasts . July 2013. 42
- [101] S. Sarykalin, G. Serraino, and S. Uryasev. Value-at-risk vs. conditional value-at-risk in risk management and optimization. *Tutorials in Operations Research. INFORMS, Hanover, MD*, pages 270–294, 2008. 39
- [102] M. Shaked and J.G. Shanthikumar. *Stochastic Orders*. Springer Series in Statistics, 2007. 8, 131, 156
- [103] A. Sklar. *Fonctions de répartition à n dimensions et leurs marges*. Université Paris 8, 1959. 7
- [104] I. Steinwart, C. Pasin, R. C. Williamson, and S. Zhang. Elicitation and Identification of Properties. *Workshop and Conference Proceedings*, 34 :1–45, 2014. 40, 42, 128
- [105] D. Tasche. Expected shortfall and beyond. *Journal of Banking & Finance*, 26(7) :1519–1533, 2002. 6
- [106] D. Tasche. Allocating portfolio economic capital to sub-portfolios. In *Economic Capital : A Practitioner Guide, Risk Books*, pages 275–302. Risk Books, 2004. 126
- [107] D. Tasche. Euler allocation : Theory and practice. Technical Report arXiv :0708.2542, Aug 2007. 126, 131, 144
- [108] D. Tasche. Capital allocation to business units and sub-portfolios : the euler principle. Papers 0708.2542, arXiv.org, 2008. 126
- [109] R. Wang and J. F. Ziegel. Elicitable distortion risk measures : A concise proof. *Statistics & Probability Letters*, 100 :172–175, 2015. 40, 44
- [110] S. Wang. A class of distortion operators for pricing financial and insurance risks. *Journal of risk and insurance*, pages 15–36, 2000. 6
- [111] S. Weber. Distribution-invariant risk measures, information, and dynamic consistency. *Mathematical Finance*, 16(2) :419–441, 2006. 43
- [112] I. Weissman. Estimation of parameters and large quantiles based on the k largest observations. *Journal of the American Statistical Association*, 73(364) :812–815, 1978. 117
- [113] C. Weng and Y. Zhang. Characterization of multivariate heavy-tailed distribution families via copula. *Journal of Multivariate Analysis*, 106 :178–186, 2012. 92
- [114] H-C. Yeh. The frailty and the Archimedean structure of the general multivariate Pareto distributions. *Bulletin Institute of Mathematics Academia Sinica*, 2(3) :713–729, 2007. 163

- [115] G. Zanjani. An economic approach to capital allocation. *Journal of Risk and Insurance*, 77(3) :523–549, 2010. 126
- [116] J. F. Ziegel. Coherence and elicibility. *Mathematical Finance*, 2014. 40, 44, 128