Université de Montréal

On the design of customized risk measures in insurance, the problem of capital allocation and the theory of fluctuations for Lévy processes

par

Hassan Omidi Firouzi

Département de mathématiques et de statistique
Faculté des arts et des sciences

Thèse présentée à la Faculté des études supérieures
en vue de l'obtention du grade de
Philosophiæ Doctor (Ph.D.)
en mathématiques

December 16, 2014

Université de Montréal
Faculté des études supérieures

Cette thèse intitulée

On the design of customized risk measures in insurance, the problem of capital allocation and the theory of fluctuations for Lévy processes

présentée par

Hassan Omidi Firouzi

a été évaluée par un jury composé des personnes suivantes :

Benjamin Avanzi
(président-rapporteur)

Manuel Morales, Mélina Mailhot
(directeur de recherche)

Ghislain Léveillé
(membre du jury)

Julien Trufin
(examinateur externe)

Jean-Yves Potvin
(représentant du doyen de la FAS)

Thèse acceptée le
Date d'acceptation

RÉSUMÉ

Dans cette thèse, nous étudions quelques problèmes fondamentaux en mathématiques financières et actuarielles, ainsi que leurs applications. Cette thèse est constituée de trois contributions portant principalement sur la théorie de la mesure de risques, le problème de l'allocation du capital et la théorie des fluctuations. Dans le chapitre 2, nous construisons de nouvelles mesures de risque cohérentes et étudions l'allocation de capital dans le cadre de la théorie des risques collectifs. Pour ce faire, nous introduisons la famille des "mesures de risque entropique cumulatifs" (*Cumulative Entropic Risk Measures*). Le chapitre 3 étudie le problème du portefeuille optimal pour le *Entropic Value at Risk* [2, 46] dans le cas où les rendements sont modélisés par un processus de diffusion à sauts (*Jump-Diffusion*). Dans le chapitre 4, nous généralisons la notion de "statistiques naturelles de risque" (*natural risk statistics*) [53] au cadre multivarié. Cette extension non-triviale produit des mesures de risque multivariées construites à partir des données financiéres et de données d'assurance. Le chapitre 5 introduit les concepts de *"drawdown"* et de la "*vitesse d'épuisement*" (*speed of depletion*) dans la théorie de la ruine. Nous étudions ces concepts pour des modeles de risque décrits par une famille de processus de Lévy spectrallement négatifs.

Mots-clés: mesures de risque cohérentes et convexes, allocation de capital, mesures de risque multivariées construites à partir des données, processus càdlàg, problème de portefeuille optimal, processus Jump-Diffusion, processus de Lévy spectrallement négatifs, drawdown, vitesse d'épuisement

SUMMARY

The aim of this thesis is to study fundamental problems in financial and insurance mathematics particularly the problem of measuring risk and its application within financial and insurance frameworks. The main contributions of this thesis can be classified in three main axes: the theory of risk measures, the problem of capital allocation and the theory of fluctuation. In Chapter 2, we design new coherent risk measures and study the associated capital allocation in the context of collective risk theory. We introduce the family of *Cumulative Entropic Risk Measures*. In Chapter 3, we study the optimal portfolio problem for the *Entropic Value at Risk* coherent risk measure [2, 46] for particular return models which are based on relevant cases of Jump-Diffusion models. In Chapter 4, we extending the notion of natural risk statistics [53] to the multivariate setting. This non-trivial extension will endow us with multivariate data-based risk measures that are bound to have applications in finance and insurance. In Chapter 5, we introduce the concepts of *drawdown* and *speed of depletion* to the ruin theory literature and study them for the class of spectrally negative Lévy risk processes.

Keywords: Coherent and Convex Risk Measure, Capital Allocation, Multivariate Data-Based Risk Measures, Càdlàg Process, Optimal Portfolio Problem, Jump-Diffusion Processes, Spectrally Negative Lévy Process, Drawdown, Speed of Depletion

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

No one walks alone on the journey of life. The writing of this thesis has been one of the most significant academic challenges I have ever had to face. This research project would not have been completed without the support, patience and guidance of many people. The author would like to express his deepest gratitude to his supervisor Prof. Manuel Morales for his excellent guidance, patience, motivation, enthusiasm, and providing the author with an excellent atmosphere for carrying out research. Deepest appreciation goes to my co supervisor Prof. Mélina Mailhot for her invaluable assistance and support during my studies at the University of Montreal. Special thanks to Prof. Louis-Pierre Arguin for whose knowledge and guidance during my PhD. The author takes this opportunity to express his deepest appreciation to Prof. Andrew Luong, Prof. Hélène Guérin without whose knowledge and assistance this study would not have been successful. Special appreciation to Prof. Hirbod Assa, who as a great friend, was always willing to help and give his best suggestions and guidance during my studies at the University of Montreal.

The author would also like to convey thanks to the staff in the Department of Mathematics and Statistics at the University of Montreal for their invaluable assistance. The author wishes to use this occasion to express his sincere gratitude to *Le Centre de recherches mathématiques (CRM), L'Institut des sciences mathématiques (ISM), the Institute of Mathematical Finance of Montreal (IFM2), La Direction des relations internationales (DRI) de l'Université de Montréal and Le Fonds de recherche du Québec Nature et technologies (FQRNT)* for providing the financial means.

The author wishes to express his love and appreciation to his beloved family for their endless love, emotional support and encouragement through the duration of his studies.

Last but not the least, the author would like to dedicate this thesis to his uncle (Shamshir Ali Omidi Firouzi), who has been a constant source of support and encouragement during his life. I am so grateful to have you in my life.

INTRODUCTION

MOTIVATION AND RESEARCH PROBLEMS

Studying and evaluating the risk faced by a financial or an insurance institution has always represented challenge from both theoretical and practical points of view. Indeed, the good health of financial and insurance companies relies on hedging a given level of risk by identifying how much money should be kept aside in order to face worse case scenarios. This requires methods to readily quantify financial risks. In [6], the authors paved the mathematical ground to measure the risk of a company by using a solid mathematical construction. The authors defined the concept of a risk measure as a real-valued function on the space of random variables representing possible outcomes of a financial position. Such a function assigns to each financial model a numerical value quantifying the level of risk associated with it. They devised an axiomatic structure for such measures that is compatible with the risk management point of view in which risks are evaluated through weighted sums of a set of possible scenarios. A large amount of scientific work has appeared since their seminal work which not only generalizes the theory of risk measures but also brings the theory into the context of different applications.

We find a fair amount of literature concerned with the practical aspects of the theory of risk measures. Two important financial and insurance problems which make use of this theory of risk measures are: 1) the problem of capital allocation, and 2) the optimal portfolio problem. The problem of capital allocation investigates different solutions in which one can allocate the total risk of a company, to each of its lines of business or departments. In other words, it seeks to understand how each component of a company is responsible for the total risk. On the other hand, the optimal portfolio problem is at the heart of Modern portfolio theory (MPT). It concerns itself with a portfolio optimization problem where one tries to minimize risk of a given portfolio, composed of different stocks, given a fixed required level of return on investment. These are two problems that we address in this thesis under new perspectives.

Indeed, although, the theory and applications of risk measures has been extensively studied for random variables, there are open frontiers when it comes to a generalized framework that takes into account the dynamical features of market evolution. A space of suitable random variables endows us with static models operating on a given point in time but that overlooks the time evolution of financial assets. Such a question can be framed within a sound theory of risk measures for stochastic processes. This is the framework in which we look at the problem of capital allocation in this thesis. We know that the random behavior of financial phenomena is modeled by a stochastic process representing the financial fluctuation of system over time. The problem of capital allocation for stochastic processes needs to be well studied and explored in the framework of the theory of risk measures defined on space of stochastic processes. This requires advanced notions and techniques from different fields of mathematics like functional and convex analysis. This makes the generalization of the notions of coherent and convex risk measures to the space of stochastic processes very challenging to say the least (see [24] for thorough discussion on such an axiomatic generalization).

We have then that designing risk measures in the space of stochastic processes becomes a challenge not only at the theoretical level but as far as practical applications are concerned. Problems like capital allocation turns out to be a more difficult endeavor under such a framework. Studying the problem of capital allocation for risk measures on space of stochastic processes requires a representation of a risk measure analog to the one already available for random variables. This can be done in different ways. To the best of our knowledge, the theory of coherent and convex risk measures on space of stochastic processes based on the axiomatic definition proposed in [24] has not been extensively studied and it represents a promising avenue of research.

After all the above considerations, the first question that motivates the current work is to investigate a new class of coherent risk measures on the space of stochastic processes having the following properties:

- They satisfy in the axiomatic definition proposed in [24] and also capture the risk associated with the path-properties of a financial or an insurance model.
- They have smooth explicit representations which enable us to easily compute without having to make use of the advanced mathematical tools.

- They can be enough tractable to apply them in practice. For instance, applying them to the problem of capital allocation.

We address this problem by tackling these challenges discussed above in particular cases of interest.

For instance, in this thesis we are interested in proposing a class of risk measures capturing the risk associated with the path-properties of a financial or an insurance model. This takes us to the realm of ruin theory where traditionally the risk reserve of an insurance company can be modeled by a spectrally negative Lévy processes (SNLP). Now, since this class of processes accepts a Laplace transform, we focus our attention to risk measures which have closed-form expressions and have a representation in terms of Laplace transform of the underlying model. There are at least two good reasons to do this. First, through characterization theorems for Lévy processes, it is apparent that Laplace transforms of SNLP provide a description of the path-properties for this class of processes which motivates our approach of constructing risk measures based on such notions. Second, we are ultimately looking for a coherent risk measure defined on the space of stochastic processes. Our construction allows us to build up such measures by using risk measures for random variables as a benchmark that can be extended from the one-dimensional law of the process to the law of the process itself. As we discuss in our contribution, we carry this out by multiplying a benchmark risk measure to a suitable weight function and integrating over a fixed period of time. Interesting enough, this construction allows us to address the other two features that we seek in our risk measure. In fact, this method of designing risk measures will also simplify the problem of capital allocation for processes and avoids the use of advanced techniques in mathematics to numerically compute the resulting expressions.

Another direction that we explore in this thesis and that it relates to the main goal of designing new risk measures defined on space of stochastic processes is the study of new quantities in collective risk theory. In fact, before the actual design of a risk measure one needs to define quantities of interest that describe the risky features of a given model. In collective risk theory one traditional example of such quantity is the ruin time and its associated ruin probability. In this thesis, we turn our interest to the definition of alternative quantities that describe path-dependent features of the process in question and in turn of the associated risk of the insurance model.

In [93] we find the novel concept of speed of market crash in finance. This study motivated us to introduce and propose new path-related concepts in collective risk theory which have not yet been studied in an insurance management context. The ultimate goal of having expressions for such quantities relates to the core of this thesis in the sense that one can devise risk measures through the integrating procedure discussed in a previous application by using the new quantity's probability functions as weights functions in the construction. Although this is the ultimate motivation, in this part of the contribution we concern ourselves only with the introduction and study of quantities like the speed of depletion and drawdowns which turned to be a challenge in its own. In an insurance context, these concepts have not been studied and yet they seem to possess a potential in insurance risk management. These quantities enable us to know how an insurance reserve is affected by drawdowns and how fast and frequent these drawdowns occur. This is problem constitutes a second contribution in this thesis.

A third problem which is close to the problem of designing coherent risk measures is to devise a class of risk measures for data sets instead. Traditionally, risk measures are defined on a space of models but in practice the underlying model is unknown leaving us only with observed data sets. These data sets can be univariate or multivariate depending on the application at hand. We seek to extend the notion of univariate data-based risk measures to a multivariate setting. We have then to define suitable vector-valued risk measures on the space of random vectors. In our contribution we start by looking at the theory of risk measures in multivariate setting, which is a challenge in its own, in order to produce a multivariate data-based axiomatic risk measure. Studying risk measures in multivariate setting has only recently been studied. In this framework, a risk measure on space of random vectors assigns a set of vectors instead of a real single value to a random vector. To the best of our knowledge unfortunately, these vector-valued risk measures defined on the space of random vectors are not tractable enough to apply them into practice. We circumvent this problem by proposing another alternative method. To do this, we propose to use multivariate data sets instead of using random vectors and define a multivariate data-based risk measure. In [53], the authors propose a new risk measure defined on a set of data. This so-called natural risk statistics also satisfies in an axiomatic definition which are based on data than random variables. By inspiring from the definition of natural risk statistics and its properties, we propose to define a class of risk measures defined on the multivariate data sets. This risk measures satisfy an axiomatic definition based on a novel couple ordering that we propose and that suits our goals.

In all we investigate three questions dealing, at different levels, with the main objective put forward at the beginning of this thesis which is designing suitable risk measures in insurance and finance.

CONTRIBUTIONS AND STRUCTURE OF THE THESIS

This thesis is based on four independent research articles that deal with both theoretical and practical aspects of risk measures and collective risk theory. These articles are presented in chapters 2, 3 ,4 and 5. An introductory Chapter 1 is included to give a brief summary of the main definitions and results of the theory of risk measures as well as the fluctuation theory for Lévy processes.

We now provide a brief account of the content of each of the main chapters.

In Chapter 2, which is based on the paper [11] entitled *On the Capital Allocation Problem for a New Coherent Risk Measure in Collective Risk Theory*, we deal with a definition of suitable risk measures on the space of bounded càdlàg process having moment generating functions. The objective of this contribution is two-fold. First, we introduce and study new coherent risk measures on the space of stochastic processes having moment generating functions based on a recently introduced risk measure for random variables presented in [2]. These so-called *cumulative* risk measures, are defined as a weighted integral of a given coherent risk measure over a finite-time interval. Second, we study the capital allocation problem in the context of collective risk theory for these risk measures. This line of research is relevant since the theory of risk measures on the space of bounded stochastic processes lacks tractable examples that could be used in practical applications. In fact, we study the problem of capital allocation in an insurance context for these tailor-made risk measures.

The first objective is accomplished, we have shown that our *Cumulative Entropic Risk Measure* is coherent and that it has a representation that allows us to tackle the capital allocation problem. We have indeed shown that the capital allocation problem for processes has a unique solution determined by the *Euler allocation* method, as presented in [83], under some assumptions. Furthermore, we applied this result for a proposed model for the net-loss claim process associated to an insurance company with n different departments. In our model we let the aggregate claims process, $(X_t^i)_{0\leq t\leq T}$, to be a linear combination of m independent spectrally positive Lévy processes for $1 \leq i \leq n$.

Chapter 3, which is based on the article [74] entitled *Optimal Portfolio Problem Using Entropic Value at Risk: When the Underlying Distribution is Non-Elliptical*, deals with the optimal portfolio problem. Due to both practical and theoretical objections that can be made about the framework of modern portfolio problem, in this chapter, we propose a new stochastic model for the asset returns that is based on *Jumps-Diffusion* **(J-D)** distributions [76, 80]. This family of distributions is more compatible with stylized features of asset returns, not to mention that it allows for a straight-forward statistical inference from readily available data. We also propose to use a new coherent risk measure, so-called, *Entropic Value at Risk(EVaR)* as presented in [2], in the optimization problem. For certain models, including a jump-diffusion distribution, this risk measure yields an explicit formula for the objective function so that the optimization problem can be solved without resorting to numerical approximations.

In Chapter 4, which is based on the paper [69] entitled *Data-Based Natural Risk Statistics*, we deal with definitions of suitable multivariate risk measures for data sets. In this chapter we propose a non-trivial extension of the concept of *natural risk statistics* to the multivariate setting. The notion of *natural risk statistics* was introduced as a technique to measure risk from data as opposed to measuring risk from predetermined models [53]. In other words, these are axiomatic risk measures defined, not on the space of random variables or processes, but rather on the space of sequences. These new risk measures have been recently introduced and their full potential has not yet been explored.

The challenge to redefine the concept of natural risk statistics for multivariate data sets is to define a right set of axioms that will allow for a coherent representation of such risk measures. This is a non-trivial problem. First, one must come with a reasonable way of ordering vectors. A second objective is to study these data-based risk measures as estimators of model-based risk measures and look into the problem of sensitivity to model mismatch. We propose a couple ordering for data vectors in $\mathbb{R}^n \times \mathbb{R}^n$ that yields a natural axiomatic definition for multivariate data-based risk measures. We find a representation for these risk measures. We also characterize these multivariate data-based risk measures via acceptance sets which complete the extension of known results from the univariate to the multivariate setting. The notion of data-based risk measures has gone unnoticed for the past couple of years but now attention is turning to the problem of risk measure estimators or *natural risk statistics*.

Chapter 5, which is based on the article [15] entitled *On the Depletion Problem for an Insurance Risk Process: New Non-ruin Quantities in Collective Risk Theory*, deals with the new quantities which are not ruin related yet they capture important features of an insurance position. This chapter elaborates on alternative non-ruin quantities that measure the riskiness associated with large claims in an insurance reserve. The field of risk theory has traditionally focused on ruin-related quantities that can naturally be interpreted in terms of risk. Although it is true that there are still many challenging questions, ruin-related quantities do not seem to capture path-dependent properties of the reserve process. In this chapter, we aim at presenting the probabilistic properties of *drawdowns* and the speed at which an insurance reserve depletes as a consequence of the risk exposure of the company. These new quantities are not ruin related yet they capture important features of an insurance position and we believe it can lead to the design of a meaningful risk measures. Studying *drawdowns* and *speed of depletion* for Lévy insurance risk processes represents a novel and challenging concept in insurance mathematics. Indeed, drawdowns and speed of depletion are quantities that do not depend on the level but on path properties of the model, which explains how fast the process can drop. In other words, how faster the claims are depleting the reserve than premiums can be collected. This type of quantities has never been proposed before as measures of riskiness in collective insurance risk theory. Drawdowns have been studied for diffusion processes in a financial setting as presented in [93]. However, in insurance, we need expressions for processes with jumps.

We have obtained expressions for many drawdown-related quantities in different cases of Lévy insurance risk processes for which they can be calculated, in particular for the classical Cramér-Lundberg model. In this model, we assume aggregate claims are modeled by a compound Poisson process with exponentially distributed severities.

Chapter 1

PRELIMINARIES

The main objective which we investigate in this thesis is designing a class of risk measures on space of stochastic processes. To achieve this, the main idea we put forward in this thesis, to achieve the main goal, requires recalling basic notions in different theories and making use of them simultaneously. These theories are 1) the theory of Lévy processes, 2) the theory of fluctuation for Lévy processes and 3) the theory of risk measures. In this chapter, we aim at providing main and important notions tying with one of these theories.

The outline of this chapter is as follows. In Section 1.1, we provide basic definitions and properties of Lévy processes. In the other words, in Subsection 1.1.1, we recall the definition of a Lévy process and provide one of the most important characterization theorem in probability theory so called Lévy-Khintchine Theorem which provides a representation for characteristic function for a Lévy process. Section 1.1 will continue in Subsection 1.1.2 by recalling an important subclass of Lévy processes so called subordinators. We use this class of processes in Chapter 2 to apply in the problem of capital allocation. In Subsections 1.1.3 and 1.1.4, we study some distributional properties for Lévy processes which we make us in Chapter 5, as well as provide some examples for Lévy processes. In Section 1.2, we will provide main notions in collective risk theory as well as the theory of fluctuations for Lévy processes. In Subsection 1.2.1, we recall the main problem in collective risk theory so called ruin problem and provide important theorems related to a particular Lévy process in hand. We finish Section 1.2 with studying the theory of fluctuation for Lévy processes and show how theses two theories are connected in Subsection 1.2.2. We also provide different important theorems available in this context. In Section 1.3, we recall main definitions and theorems for risk measures as well as providing different classes of risk measures which we use in Chapter 2 and Chapter 4 of this thesis. In Subsection 1.3.1, we provide the definition of convex and coherent risk measures as well as their representations. In Subsection 1.3.2, we recall a definition of capital allocation which we widely use in Chapter 2 of this thesis. Section 1.3 will continue in Subsections 1.3.3 and 1.3.4 by giving the definition of two important classes of risk measures which motivated us to study multivariate data-based risk measures in Chapter 4.

1.1. Lévy processes

Lévy processes is an important class of stochastic processes that are used in mathematical finance and insurance. Having càdlàg (continuous from right and having limit from left) paths along with stationary and independent increments makes this family of processes tangible to apply as mathematical models for financial and insurance phenomena.

Lévy processes will appear in this thesis as a core and body of our work. On one hand, we are interested in designing a new class of risk measures on the space of bounded càdlàg processes that can capture the risk associated with the path-properties of the insurance model. These risk measures have a mathematical closed form for processes which have Laplace transform or moment generating function. Another application of Lévy processes ties with introducing the new concept in insurance called drawdowns. We generalize this financial concept to insurance by using Lévy processes.

1.1.1. Lévy Processes and Infinitely Divisible Distributions

There are different ways to study behavior of a random variable (or a stochastic process). One of such ways which provides the distributional properties of a random variable is to study its characteristic function (or in particular, Laplace transform). In fact, characteristic functions or Laplace transforms characterize probability distributions. We can also find different applications of characteristic functions (Laplace transforms) in probability theory. For instance, we can see applications of characteristic functions (or Laplace transforms if exist) in studying the first passage times for Lévy processes (or generally stochastic processes). Therefore, knowing characteristic function (or Laplace transform) for a Lévy process is of a great importance. One of the most important theorems in probability theory is called Lévy-Khintchine Theorem. This theorem provides characteristic functions for a class of distributions so called infinitely divisible distributions (or somehow characterizes this class of distributions). We will see in the sequel of this chapter that there is a one to one correspondence between Lévy processes and the class of infinitely divisible distributions. Hence, using Lévy-Khintchine Theorem we can obtain characteristic function for Lévy processes as well.

In this subsection, we will recall important definitions and theorem related to infinitely divisible distributions as well as definition of Lévy processes and their relation with this class of distributions. Most of the definitions and results about Lévy processes and their properties in this section are taken from [5, 81].

In the following definition, we introduce a class of probability distributions which can be presented in terms of convolution functions of independent and identically distributed (i.i.d.) distributions.

Definition 1.1.1. *([5]) Let $\mathcal{P}(\mathbb{R})$ denote the set of all Borel probability measures on $\mathbb{R}$. Then a Borel probability measure $\mu \in \mathcal{P}(\mathbb{R})$ is infinitely divisible if it has a convolution n-th root in $\mathcal{P}(\mathbb{R})$. i.e., there exists a probability measure $\mu^{\frac{1}{n}} \in \mathcal{P}(\mathbb{R})$ for which $\mu = (\mu^{\frac{1}{n}})^{*^n}$, for each $n \in \mathbb{N}$.*

We can specialize Definition 1.1.1 for a random variable and investigate when a random variable can be infinitely divisible. A random variable X (see Definition A.1.4) is infinitely divisible if its law P_X (see Definition A.1.5) is infinitely divisible, e.g. $X =^d Y_1^{(n)} + \cdots + Y_n^{(n)}$, where $=^d$ means equality in distribution and $Y_1^{(n)}, \ldots, Y_n^{(n)}$ are i.i.d. random variables, for each $n \in \mathbb{N}$. Note that, the characteristic function of X (see Definition A.1.7) can be written as $\phi_X(u) = (\phi_{Y_1^{(n)}}(u))^n$.

Poisson distribution and normal distributions are two main examples of infinitely divisible distributions. In order to provide a characterization for this class of distributions we need to know some basic notions. One of them is Lévy measure which appears in the characterization theorem, Lévy-Khintchine Theorem.

Definition 1.1.2. *([5]) A measure ν (see Definition A.1.2) defined on $\mathbb{R}\backslash\{0\}$ is called a Lévy measure if*

$$\int_{\mathbb{R}\backslash\{0\}} (|y|^2 \wedge 1)\nu(dy) < \infty, \tag{1.1.1}$$

where $|.|$ is the absolute value function.

Lévy measures are one of the key components in representations for Lévy processes. They includes useful information about the structure of Lévy processes and infinitely divisible distributions (see Propositions 1.1.1, 1.1.2 and 1.1.3).

The following theorem characterizes infinitely divisible probability measures on $\mathbb{R}$ using characteristic functions.

Theorem 1.1.1. (Lévy-Khintchine) *([5, 63]) A Borel probability measure μ on $\mathbb{R}$ is infinitely divisible if there exists a value $b \in \mathbb{R}$, $\sigma \geq 0$ and a Lévy measure*

ν on $\mathbb{R}\backslash\{0\}$ such that for all $u \in \mathbb{R}$,

$$\phi_\mu(u) = \exp\left\{ibu - \frac{1}{2}\sigma^2 u^2 + \int_{\mathbb{R}\backslash\{0\}} \left(e^{iuy} - 1 - iuy\mathbb{1}_{(|y|<1)}(y)\right)\nu(dy)\right\} \quad (1.1.2)$$

Conversely, any mapping of the form (1.1.2) is the characteristic function of an infinitely divisible probability measure on $\mathbb{R}$.

Proof. For a proof, we refer to Subsection 1.2.4 of [5]. □

The triple (b, σ, ν) is called the characteristics of the infinitely divisible probability measure μ, and $\eta := \log \phi_\mu$ is called the Lévy symbol or Characteristic exponent.

In the following, we recall the definition of an example of infinitely divisible random variables so called α-stable random variable. We apply this type of random variable in Chapter 2 and Chapter 5. We also provide a theorem characterizing these random variables which mainly uses Lévy-Khintchin Theorem. For more details, read Subsection 1.2.5 of [5].

Definition 1.1.3. (α-stable random variable) *([5]) A random variable X is said to be α-stable if there exist real-valued sequences $(\sigma n^{\frac{1}{\alpha}}, n \in \mathbb{N})$ and $(d_n, n \in \mathbb{N})$ such that*

$$X_1 + X_2 + \cdots + X_n \stackrel{d}{=} \sigma n^{\frac{1}{\alpha}} X + d_n \quad (1.1.3)$$

where $X_1, X_2, \ldots, X_n$ are independent copies of X . In particular, X is said to be strictly stable if each $d_n = 0$.

Theses random variables are heavy tailed (except normal distribution) in the sense that the tails of their distributions decay slowly enough to zero (see [73] for more information). It follows immediately from (1.1.3) that all stable random variables are infinitely divisible. The characteristics in the Lévy-Khintchine formula are given by the following theorem.

Theorem 1.1.2. *([63, 81]) If X is an α-stable real-valued random variable, then its characteristics in Theorem 1.1.1 must take one of the two following forms:*

(1) when $\alpha = 2, \nu = 0$, so $X \sim N(b, \sigma)$;

(2) when $\alpha \neq 2, \sigma = 0$ and

$$\nu(dx) = \frac{c_1}{x^{\alpha+1}}\mathbb{1}_{(0,\infty)}(dx) + \frac{c_2}{|x|^{\alpha+1}}\mathbb{1}_{(-\infty,0)}(dx),$$

where $c_1 \geq 0,\ c_2 \geq 0$ and $c_1 + c_2 > 0$.

Proof. A proof can be found in [81], page 80. □

In the sequel of this section we provide a concise description of a class of processes (Lévy processes) which have a one to one correspondence with infinitely divisible distributions. In probability theory, a Lévy process, named after the French mathematician Paul Lévy, is any continuous-time stochastic process that starts at 0, admits càdlàg modification and has "stationary independent increments", (this phrase will be explained below). The most well-known examples are the Wiener process and the Poisson process.

Definition 1.1.4. *([5]) A Lévy process $X = (X_t)_{t \geq 0}$ is a stochastic process satisfying the following:*

(1) $X(0) = 0$ (a.s.);

(2) X has independent and stationary increments;

(3) X is stochastically continuous, i.e. for all $a > 0$ and for all $s \geq 0$,

$$\lim_{t \to s} P(|X_t - X_s| > a) = 0.$$

Lévy-Ito decomposition Theorem (see Theorem A.3.1) provides a characterization for Lévy processes. It turns out form Theorem A.3.1 that a Lévy process is an independent sum of a Brownian motion with drift, an independent compound Poisson processes and a square integrable jump process. Independent structure of Lévy processes along with having both continuous and jump parts in its nature, make this class of processes more suitable to model dynamical financial and insurance phenomena.

It follows from the Definition 1.1.4 that each Lévy process X_t is infinitely divisible. Referring to [5] we have the following theorem.

Theorem 1.1.3. *([5]) If X is a Lévy process, then X_t is infinitely divisible for each $t \geq 0$. Furthermore,*

$$\phi_{X_t}(u) = e^{t\eta(u)}, \;\; \forall u \in \mathbb{R}^d, t \geq 0,$$

where η is the Lévy symbol of $X(1)$.

As we mentioned earlier, there is an one to one correspondence between Lévy processes and infintely divisible distributions. Theorem 1.1.3 along with the following theorem show this correspondence. For a comprehensive discussion on this theorem we refer to Subsection 1.4.1 in [5].

Theorem 1.1.4. *([5]) If μ is an infinitely divisible distribution on $\mathbb{R}$ with Lévy symbol η, then there exists a Lévy process X such that μ is the law of $X(1)$.*

There are two important subclasses of the class of Lévy processes for which we can find a large amount of literature studying their path properties as well as their distributional properties (see for instance [60, 61, 62, 63]). In this thesis we also focus on these classes of processes which model risk and surplus of insurance companies. Referring to Lévy-Ito decomposition (Theorem A.3.1), a Lévy process has both continuous and jump parts. These jumps can be either upward or downwards. i.e., the process can have positive jumps or negative jumps.

Definition 1.1.5. *([63]) A Lévy process having just negative jumps is called a* spectrally negative Lévy process *(SNLP). A Lévy process which has only positive jumps is called a* spectrally positive Lévy process *(SPLP).*

One of the simplest classes of processes which is included in the class of SPLP is the class of subordinators. Elements of this class have increasing paths and do not have Brownian motion part in their Lévy-Ito decomposition. In the following subsection, we are interested to provide the mathematical definition of these processes as well as study their Laplace transforms (they exist).

1.1.2. Subordinators

In this part we will provide a particular example of spectrally positive Lévy processes so called subordinators. A subordinator is a Lévy process which is increasing almost surely. Such processes can be thought of as a random model of time evolution [5].

Definition 1.1.6. *([5]) The process $X = (X_t)_{t\geq 0}$ is called a subordinator if we have*

$$X_t \geq 0 \text{ } a.s. \text{ } for \text{ } each \text{ } t > 0,$$

and

$$X(t_1) \leq X(t_2) \text{ } a.s. \text{ } whenever \text{ } t_1 \leq t_2.$$

The following theorem, which is taken from [5], provides Lévy-Khintchine characterization for subordinators.

Theorem 1.1.5. *([5]) A Lévy process $X = (X_t)_{t\geq 0}$ be a subordinator if and only if its Lévy symbol takes the form*

$$\eta(u) = ibu + \int_{(0,\infty)} (e^{iuy} - 1)\lambda(dy), \tag{1.1.4}$$

for some $b \geq 0$ and a Lévy measure λ satisfying

$$\lambda(-\infty, 0) = 0 \text{ } and \text{ } \int_{(0,\infty)} (y \wedge 1)\lambda(dy) < \infty.$$

We call the pair (b, λ) the characteristics of the subordinator X.

We then obtain the following expression for the Laplace transform of the distribution:

$$\mathbb{E}(e^{-uX_t}) = e^{-t\psi(u)},$$

where

$$\psi(u) = -\eta(iu) = bu + \int_0^\infty (1 - e^{-uy})\lambda(dy),$$

for each $u > 0$. The function $\psi(u)$ is called the Laplace exponent of the subordinator.

For some application of subordinators in mathematical insurance, see [49, 55]. An example of subordinators which is of interest in this thesis is gamma subordinator. In fact, having Laplace transform for this process enables us to apply it in the new coherent risk measure introduced in Chapter 2 of this thesis. In the following we provide a brief description with properties of this subordinator.

Example 1.1.1. *(Gamma Subordinators) ([5]) Let $X = (X_t)_{t\geq 0}$ be a gamma process with parameters $a, b > 0$, so that each X_t has density*

$$f_{X_t}(x) = \frac{b^{at}}{\Gamma(at)} x^{at-1} e^{-bx},$$

for $x \geq 0$; then it is easy to verify that, for each $u \geq 0$

$$\int_0^\infty e^{-ux} f_{X_t}(x)dx = \left(1 + \frac{u}{b}\right)^{-at} = \exp\left[-ta \log\left(1 + \frac{u}{b}\right)\right]. \tag{1.1.5}$$

From here it is a straightforward exercise to show that

$$\int_0^\infty e^{-ux} f_{X_t}(x)dx = \exp\left[-t \int_0^\infty (1 - e^{-ux}) a x^{-1} e^{-bx} dx\right]. \tag{1.1.6}$$

From this we see that $X = (X_t)_{t\geq 0}$ is a subordinator with $b = 0$ and $\lambda(dx) = ax^{-1}e^{-bx}dx$. Moreover, $\psi(u) = a \log(1 + u/b)$.

1.1.3. Examples of Lévy processes

In this subsection we give two well know examples of Lévy processes. In fact, referring to Theorem A.3.1 we see theses examples are two important parts in Lévy-Ito decomposition. Moreover, different applications of these examples in mathematical modeling in finance and insurance can be also found. We apply these examples in Chapter 2 of this thesis as well.

Example 1.1.2. *(Brownian Motion With Drift) ([5]) Let $(B_t)_{t\geq 0}$ be a standard Brownian motion in $\mathbb{R}$. Then, the process $(C_t)_{t\geq 0}$ with $C_t = bt + \sigma B_t$ is a Lévy process with its characteristic function given by*

$$\phi_{B_t}(u) = \exp\{-\frac{1}{2} t\eta_c(u)\},$$

where $\eta_c(u)$ is Lévy symbol of $C(1)$ of the form

$$\eta_c(u) = ibu - \frac{1}{2}\sigma^2 u^2,$$

For $b \in \mathbb{R}$ and $\sigma \geq 0$. In fact a Lévy process has a continuous sample paths if and only if it is of the form $C(t)$. For some application of Brownian motion in mathematical insurance, see the article [49].

Example 1.1.3. *(Compounded Poisson Process) ([5]) Let $\{Z(n) : n \in \mathbb{N}\}$ be a sequence of i.i.d. random variables in $\mathbb{R}$ with law μ_Z. Let N be a Poisson process of intensity λ , which is independent of $Z(n)$. The compound Poisson process Y is defined as follows:*

$$Y(t) = Z(1) + \cdots + Z(N(t)).$$

Then, Y has Lévy symbol of

$$\eta_Y(u) = \int_{\mathbb{R}} (e^{iuy} - 1)\lambda\mu_Z(dy).$$

If $\mu_Z = \delta_1$ where δ_1 is the Dirac delta measure, then Y is said to be Poisson process (see [5]). Classical risk model can be viewed as an application of compounded Poisson process in mathematical modeling in insurance (see [50]).

1.1.4. Distributional properties for Lévy processes

As we mentioned earlier, the class of Lévy processes is the main core and body of this thesis. We apply this class of processes in different cases. Since we deal with a subclass of Lévy processes with important distributional properties in this thesis, knowing distributional properties for Lévy processes helps us to understand much better this class of processes. Studying the variational behavior as well as verifying when these processes have finite moments are two important problems which we need to consider in this subsection. Lévy processes are not in general of finite variation and they do not have finite moments. In the following, we recall some important results which under some assumptions classify the bounded variation Lévy processes. The main components in Lévy processes which play important roles are Lévy measure ν and continuous component $\sigma \geq 0$.

Proposition 1.1.1. *([63]) Let $(X_t)_{t\geq 0}$ be a Lévy process with triplet (b, σ, ν).*

(1) If $\nu(\mathbb{R}) < \infty$, then almost all paths of X_t have a finite number of jumps on every compact interval.

(2) If $\nu(\mathbb{R}) = \infty$, then almost all paths of X_t have an infinite number of jumps on every compact interval.

Proposition 1.1.2. *([5]) Let $(X_t)_{t\geq 0}$ be a Lévy process with triplet (b, σ, ν).*

(1) If $\sigma = 0$ and $\int_{|x|\leq 1} |x|\nu(dx) < \infty$, then almost all paths of X_t have finite variation.

(2) If $\sigma \neq 0$ or $\int_{|x|\leq 1} |x|\nu(dx) = \infty$, then almost all paths of X_t have infinite variation.

It can be seen from Proposition 1.1.2 that a Brownian motion with drift is a Lévy process with infinite variation as $\sigma > 0$. In the next proposition we will recall necessary and sufficient conditions for a Lévy process $(X_t)_{t\geq 0}$ to have a finite moment.

Proposition 1.1.3. *([5]) Let $(X_t)_{t\geq 0}$ be a Lévy process and $n \in \mathbb{N}$. $\mathbb{E}(|X_t|^n) < \infty$ for all $t > 0$ if and only if $\int_{|x|\geq 1} |x|^n \nu(dx) < \infty$.*

The reason we should care about big jumps to figure out the moment properties of a Lévy process is that both the Brownian motion part and the martingale pure jump part have finite moments at any order (see Chapter 2 of [5]) while the compound Poisson part with big jumps, bigger than 1, has not necessarily finite moments.

It can be seen from Propositions 1.1.1, 1.1.2 and 1.1.3 that activity of jumps of a Lévy process depends on all the jumps of the process, the variation of a Lévy process depends on the small jumps and the Brownian motion part and the moment properties of a Lévy process depend on the big jumps.

1.2. Fluctuation Theory for Lévy Process and Collective Risk Theory

So far, we paved the way to study the theory of fluctuation by introducing the class of Lévy processes and their properties. Fluctuation theory which recently found its application in insurance context, mainly focuses on studying the path behavior of stochastic processes (in particular Lévy processes). Studying problems like the problem of ruin for a spectrally negative Lévy process, which models the surplus process of an insurance company, can be found in the realm of this theory. In the other words, ruin problem has been studied as a particular case of so called exit problem in the theory of fluctuation for Lévy processes (we refer to [17, 63]). Known expressions in ruin theory can be derived as applications of the results of theory of fluctuation. To better understand this theory and main problems in it, we recall important results related to exit problems in the sequel of this section. We start by providing a flavor of ruin problem for the reader and show how this problem can be connected to an exit problem in the theory of fluctuation. We also study the theory of fluctuation by introducing new path-related quantities for the class of spectarlly negative Lévy process in Chapter 5 of this thesis.

Ruin problem for an insurance company which was introduced in collective risk theory has been extensively studied in the past decades. We can find in the related literature a range of examples of Lévy processes for which the ruin problem has been studied and interesting results have been derived (see [55, 58, 63, 72]).

1.2.1. Collective Risk Theory and Ruin Problem

In this subsection we provide basic concepts and notations widely used in collective risk theory. We give a mathematical definition of ruin probability and Gerber-Shiu function for the risk reserve of an insurance company. We also provide some results related to ruin probability for a classical risk reserve process studied in [55].

Consider the following general model for the risk reserve of an insurance company,

$$R_t = x + c\,t - X_t\ , \quad t \geq 0\ , \tag{1.2.1}$$

where the aggregate claims process X is a spectrally positive Lévy process with zero drift, with $X(0) = 0$. Moreover, x is the initial reserve level and c is a constant premium rate defined as

$$c = (1+\theta)\mathbb{E}[X_1] \tag{1.2.2}$$

where $\theta > 0$ is the security loading factor.

Then the associated ruin time is

$$\tau_x := \inf\{t \geq 0 \mid X_t - c\,t \geq x\}\ , \tag{1.2.3}$$

and the infinite-horizon ruin probability can be defined by

$$\Psi(x) := \mathbb{P}_x(\tau_x < \infty)\ , \tag{1.2.4}$$

where $\mathbb{P}_x$ is short-hand notation for $\mathbb{P}(\,\cdot \mid X(0) = x)$.

Ruin probability in model presented in (1.2.1) has been discussed in different articles. For instance, we refer to [55, 58, 63, 72] for a thorough discussion.

In the sequel, we provide some important results related to different cases of risk models in (1.2.1).

Theorem 1.2.1. *([7]) If R_t is a surplus process in (1.2.1) based upon a compound Poisson aggregate claim process, X_t, with $c = (1+\theta)\mathbb{E}[X_1]$, that θ is security loading factor, then for $x \geq 0$,*

$$\Psi(x) = \frac{e^{-Kx}}{\mathbb{E}[e^{-KR(\tau)}|\tau < \infty]} \tag{1.2.5}$$

Where K is the smallest positive root of the Lundberg's fundamental equation, $M_{X_t}(k) = e^{kct}$.

Corollary 1.2.1. *Using the fact that $R(\tau) < 0$ when $\tau < \infty$ and one application of theorem 1.2.1 yields a nice approximation for ruin probability. That is,*

$$\Psi(x) \cong e^{-Kx}.$$

Now, we are going to give the important theorem taken from [55] studying a ladder-height decomposition for the ruin probability for the risk process of the form

$$R_t = x + ct - S(t) + \eta Z(t). \tag{1.2.6}$$

Here S is a subordinator with Lévy measure $d\nu$ and Z is a Lévy motion with no positive jumps and zero drift.

Theorem 1.2.2. *([55]) Let R be a risk process as in (1.2.6) and denote by $Y_t = R_t - x$. Then its associated ruin probability $\Psi(u) = \mathbb{P}[sup_{t\geq 0}\{-Y_t > x\}]$ satisfies the equation*

$$1 - \Psi(x) = \frac{\theta}{\theta+1}\sum_{n=0}^{\infty}(\frac{1}{1+\theta})^n M^{\star n} \star G^{\star(n+1)}(x), \; x \geq 0, \tag{1.2.7}$$

where M is the distribution with Laplace transform given by

$$\xi_M(s) = \int_0^{\infty} e^{-sx} dM(x) = \frac{\psi_S(s)}{s\mathbb{E}(S_1)}, \tag{1.2.8}$$

where ψ_S is the Laplace exponent of S and G is the distribution with Laplace transform given by

$$\xi_G(s) = \int_0^{\infty} e^{-sx} dG(x) = \frac{cs}{\psi_{ct+\eta Z}(s)}. \tag{1.2.9}$$

The following example recalls results for a particular case of the model model (1.2.6).

Example 1.2.1. *[38, 50] study model (1.2.6) when $Z = W$ is a Brownian motion with zero drift and variance σ^2, and S is a compound Poisson process, that is,*

$$R_t = x + ct - S_t + \sigma W_t, \; t \geq 0.$$

The corresponding functions G and M in Theorem 1.2.2 for this process are:

$$G(u) = 1 - e^{-(\frac{c}{\sigma^2})u}, \; u > 0,$$

and $M(u) = \frac{1}{\beta}\int_0^u [1 - F(t)]dt$, where $\beta = \int_0^{\infty}[1 - F(t)]dt < \infty$.

Another historic contribution to collective risk theory was made by actuarial scientists Hans U. Gerber and Elias S.W. Shiu in their article [50], where the expected discounted penalty function (EDPF) comes to light. The problem of ruin can be studied as a particular case of this function. The expected discounted penalty function (or called Gerber-Shiu function) is defined by

$$\eta(x) = \mathbb{E}_x[e^{-\delta\tau} w(R_{\tau^-}, |R_\tau|) I(\tau < \infty)], \; x \geq 0, \tag{1.2.10}$$

where $\delta \geq 0$ is the discounting force of interest, τ is the time of ruin, R_{τ^-} is the surplus immediately before ruin and the bounded function $w(u,v)$ is called a penalty function depending on the amount of surplus prior to ruin u and the amount of deficit at ruin v. In equation (1.2.10), $\mathbb{E}_x$ is shorthand for $\mathbb{E}(.|X_0 = x)$.

([50]) Let $f(u,v,t|x)$ denote the joint probability density function of R_{τ^-}, $|R_\tau|$and τ, then we can rewrite the function $\eta(x)$ as follows:

$$\eta(x) = \int_0^\infty \int_0^\infty \int_0^\infty w(u,v) e^{-\delta t} f(u,v,t|x) dt du dv. \tag{1.2.11}$$

For $u_0 > 0$ and $v_0 > 0$, if $w(u,v)$ is a "generalized" density function with mass 1 for $(u,v) = (u_0, v_0)$ and 0 for other values of (u,v), then

$$\eta(x) = \int_0^\infty e^{-\delta t} f(u_0, v_0, t|x) dt = f(u_0, v_0|x).$$

The expected discounted penalty, considered as a function of the initial surplus, satisfies a certain renewal equation, which has a probabilistic interpretation. Explicit answers are obtained for different cases. For instance, when the initial surplus is zero, initial surplus is very large, and for arbitrary initial surplus if the claim distribution follows an exponential distribution or a mixture of exponentials. For more discussion, we refer to the [50].

We refer to [19, 49, 50, 66] for results on EDPF for different risk models.

1.2.2. Fluctuation Theory for Lévy Processes

In this subsection, we discuss fluctuation theory and give some important results related to the first passage time of a Lévy process. We introduce some notions and results that are needed in the rest of the thesis. It is worthwhile to know the connection between fluctuation theory and collective risk theory. In fact, by defining the first passage time, important concept in fluctuation theory, we see that the time of ruin in collective risk theory defined in (1.2.3) for the general risk model (1.2.1) is in fact a first passage time for a Lévy process. To show this connection, first we need to provide notations and definitions available in the theory of fluctuation.

Let $X = (X_t)_{t \geqslant 0}$ be a spectrally negative Lévy process defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geqslant 0}, \mathbb{P})$. Since X has no positive jumps, the expectation $\mathbb{E}\left[e^{sX_t}\right]$ exists for all $s \geqslant 0$ and it is given by $\mathbb{E}\left[e^{sX_t}\right] = e^{t\psi(s)}$ where $\psi(s)$ is of the form

$$\psi(s) = a\,s + \frac{1}{2}\sigma^2 s^2 + \int_0^\infty (e^{-x\,s} - 1 + s\,x\mathbb{1}_{\{x<1\}})\,\nu(dx)\ , \tag{1.2.12}$$

where $a \in \mathbb{R}$, $\sigma \geq 0$ and ν is the Lévy measure associated with the process $-X$ (for a thorough account on Lévy process see [17, 63]). For the right inverse of ψ, we shall write Φ on $[0, \infty)$. Formally, for each $q \geqslant 0$,

$$\Phi(q) := \sup\{s \geqslant 0 : \psi(s) = q\}\ . \tag{1.2.13}$$

Notice that since X is a spectrally negative Lévy process X, we have that $\Phi(q) > 0$ for $q > 0$ (see [63]).

In the sequel, we consider the following stopping times so called *first passage times*. Afterward, we show how we can rewrite the time of ruin presented in equation (1.2.3) as a first passage time.

$$\tau_x^+ = \inf\{t > 0 : X_t > x\}. \tag{1.2.14}$$

for $x > 0$ and

$$\tau_x^- = \inf\{t > 0 : X_t < x\}, \tag{1.2.15}$$

for $x \in \mathbb{R}$. If we consider $Y_t = ct - X_t$ in equation (1.2.1), then by replacing Y_t in equation (1.2.3) we get

$$\tau := \inf\{t \geq 0 \mid Y_t \leq -x\}\ ,$$

where is in form of equation (1.2.15).

In the following theorem we recall Laplace transform for the first passage time τ_x^+.

Theorem 1.2.3. *([63]) Let $(X_t)_{t\geq 0}$ be a spectrally negative Lévy process and τ_x^+ is given in (1.2.14). Then,*

$$\mathbb{E}(e^{-q\tau_x^+}\mathbb{1}_{(\tau_x^+<\infty)}) = e^{-\Phi(q)x}, \tag{1.2.16}$$

where $\Phi(q)$ is given in (1.2.13).

Corollary 1.2.2. *From Theorem 1.2.3 we have that $\mathbb{P}(\tau_x^+ < \infty) = e^{-\Phi(0)x}$ which is one if and only if $\mathbb{E}(X_1) \geq 0$.*

In this sequel of this subsection we provide the definition of the so-called *q-scale functions* $W^{(q)}$ which is a key notion in the analysis of path properties for spectrally negative Lévy processes. Moreover, we provide some important results related to the Laplace transform of one- and two-sided exit problem for spectrally negative Lévy processes.

Theorem 1.2.4. *([63]) Let $(X_t)_{t\geq 0}$ be a spectrally negative Lévy process. Then for every $q \geqslant 0$, there exists a strictly increasing and continuous function $W^{(q)} : \mathbb{R} \longrightarrow [0, \infty)$ such that $W^{(q)}(y) = 0$ for all $y < 0$ satisfying*

$$\int_0^\infty e^{-\lambda y} W^{(q)}(y)dy = \frac{1}{\psi(\lambda) - q}, \qquad \lambda > \Phi(q). \tag{1.2.17}$$

Furthermore, $W^{(q)}$ has left and right derivatives on $(0, \infty)$ which may not agree for all spectrally negative Lévy processes. If $\sigma > 0$ then the left and right derivatives agree and $W^{(q)} \in C^1(0, \infty)$.

Notice that for $q = 0$, equation (1.2.17) defines the so-called *scale function* and we simply write W. Although, q-scale functions are one of the most important key notions to study path properties of Lévy processes, but unfortunately, there are few cases of processes for which we have explicit representation for associated q-scale functions. We refer to Subsection 5.5.1 on page 108 for an example of a q-scale function. In this case, we consider a compound Poisson process which jump distributions follow an exponential distribution. Studying q-scale functions can be reduced to studying W using the following result. In the following corollary we recall a representation for $W^{(q)}$ in terms of W.

Corollary 1.2.3. *([63]) Let $(X_t)_{t\geq 0}$ be a spectrally negative Lévy process. Then*

$$W^{(q)}(x) = \sum_{n\geq 0} q^n W^{*(n+1)}(x), \tag{1.2.18}$$

where $W^{*(n)}$ is the n-th convolution of W with itself.

In the following we recall a theorem providing the Laplace transforms for one- and two-sided exit problem.

Theorem 1.2.5. *([63]) Let $(X_t)_{t\geq 0}$ be a spectrally negative Lévy process and $W^{(q)}$ the q-scale function defined in (1.2.17). Also consider τ_x^+ and τ_x^- as the stopping times defined in (1.2.14) and (1.2.15) for $x \in \mathbb{R}$ respectively. Then the following statements are true.*

(1) For any $x \in \mathbb{R}$ and $q \geq 0$,

$$\mathbb{E}_x(e^{-q\tau_0^-} \mathbb{1}_{(\tau_0^- < \infty)}) = Z^{(q)}(x) - \frac{q}{\Phi(q)} W^{(q)}(x), \tag{1.2.19}$$

where $Z^{(q)}(x) = 1 + q \int_0^x W^{(q)}(y)dy$ for $x \in \mathbb{R}$.

(2) For any $x \leq a$ and $q \geq 0$,

$$\mathbb{E}_x(e^{-q\tau_a^+} \mathbb{1}_{(\tau_0^- > \tau_a^+)}) = \frac{W^{(q)}(x)}{W^{(q)}(a)}. \tag{1.2.20}$$

(3) For any $x \leq a$ and $q \geq 0$,

$$\mathbb{E}_x(e^{-q\tau_0^-}\mathbb{1}_{(\tau_0^- < \tau_a^+)}) = Z^{(q)}(x) - Z^{(q)}(a)\frac{W^{(q)}(x)}{W^{(q)}(a)}. \tag{1.2.21}$$

This theorem provides important results related to Laplace transforms of exit problems. One of the important results which we can derive from this theorem is the ruin probability for a SNLP. Setting $q = 0$ in Theorem 1.2.5 yields to the following corollary.

Corollary 1.2.4. *Under the same assumptions in Theorem 1.2.5 we have*

(1) For any $x \in \mathbb{R}$,

$$\mathbb{P}_x(\tau_0^- < \infty) = 1 - (\mathbb{E}(X_1) \vee 0)W(x), \tag{1.2.22}$$

where $(\mathbb{E}(X_1) \vee 0)$ is the maximum between $\mathbb{E}(X_1)$ and 0.

(2) For any $x \leq a$,

$$\mathbb{P}_x(\tau_a^+ < \tau_0^-) = \frac{W(x)}{W(a)}. \tag{1.2.23}$$

Proof. (1.2.22) and (1.2.23) can be derived by replacing $q = 0$ in (1.2.19) and (1.2.20) respectively. □

It can be seen from equation (1.2.22) when $\mathbb{E}(X_1) \leq 0$, claim sizes have negative means, the ruin surely happens. Setting $x = a$ in equation (1.2.23) yields to $\mathbb{P}_x(\tau_a^+ < \tau_0^-) = 1$ which means the process creeps upward at the initial value a before going below 0. This is compatible with the path behavior of a SNLP. As a particular case if we set $x = a = 0$, then we have $\mathbb{P}(\tau_0^+ < \tau_0^-) = 1$. This means that a SNLP with initial value 0, this process creeps upward than downward at 0. This is also compatible with path behavior of a SNLP. Consider a Brownian motion as an example for this case.

1.3. Risk Measures

As a part of this thesis, we are to design a risk measure for the class of bounded càdlàg processes. Knowing the basic definitions and properties of risk measures leads to bring this section into this chapter.

Article [6] and book [45] gave a mathematically construction of a risk measure. The authors introduced the concept of a coherent risk measure based on a preference relation on a subset of L^∞ (see Definition A.2.3). In this section L^∞ represents the space of uncertain payoff values of financial positions.

1.3.1. Coherent and Convex Risk Measures on the Space L^∞

In this subsection we bring the axiomatic definition for convex and coherent risk measures on the space of bounded random variables. This axiomatic definition implies to study the theory of risk measure in the context of convex analysis. Referring to standard textbook on convex analysis, Fenchel's Theorem states that if f is a convex and lower semi-continuous function, then it meets a representation in terms of its conjugate (see [79]). Convex and coherent risk measures defined on L^p spaces also meet such representations, so called robust representations, under some limit constraints (we refer to [45]). In the sequel of this section, we recall important theorems and representations for both convex and coherent risk measures. In the following definitions and theorems we assume the argument of a risk measure is "profit" than "loss".

Definition 1.3.1. *([45]) A mapping $\rho : L^\infty \to R$ is called a Convex risk measure if it satisfies the following conditions for all $X, Y \in L^\infty$.*

(1) Monotonocity: if $X \leq Y$ then $\rho(X) \geq \rho(Y)$.
(2) Cash invariance: If $m \in \mathbb{R}$, then $\rho(X + m) = \rho(X) - m$.
(3) Convexity: $\rho(\lambda X + (1-\lambda)Y) \leq \lambda\rho(X) + (1-\lambda)\rho(Y)$, for $0 \leq \lambda \leq 1$.

In Definition 1.3.1 if the argument of risk measure ρ considered to be "loss", then conditions (1) and (2) will be replaced by $\rho(X) \leq \rho(Y)$ and $\rho(X+m) = \rho(X)+m$.

Definition 1.3.2. *A convex risk measure ρ is called a coherent risk measure if it satisfies*

- *Positive Homogeneity: if $\lambda \geq 0$, then $\rho(\lambda X) = \lambda\rho(X)$.*

If a risk measure ρ is positively homogeneous, then it is normalized, i.e., $\rho(0) = 0$. Under the assumption of positive homogeneity, convexity is equivalent to

- Subadditivity: $\rho(X + Y) \leq \rho(X) + \rho(Y)$.

The financial meaning of conditions above are clear:

- Monotonocity: That is, if portfolio Y always has better values than portfolio X under almost all scenarios then the risk of Y should be less than the risk of X.
- Cash Invariance: The value m is just adding cash to your portfolio X, which acts like an insurance: the risk of $X + m$ is less than the risk of X, and the difference is exactly the added cash m. In particular, if $m = \rho(X)$ then $\rho(X + \rho(X)) = 0$.
- Positive Homogeneity: Doubling your portfolio, doubles your risk.
- Subadditivity: The risk of two portfolios together cannot get any worse than adding the two risks separately.

Example 1.3.1. *One of the most important risk measures which is widely used in financial risk management is value at risk (VaR). Let X show profit of a financial position. VaR is defined as follows.*

$$VaR_\beta(X) = \inf\{m \in \mathbb{R} \mid \mathbb{P}(X + m < 0) \leq \beta\}. \tag{1.3.1}$$

([45]) In financial terms, $VaR_\beta(X)$ is the smallest amount of capital which, if added to X and invested in the risk-free asset, keeps the probability of a negative outcome below the level λ.

Value at Risk is neither a convex nor a coherent risk measure as it does not meet subadditivity property. But, in the following example we recall an upper bound coherent risk measure for VaR.

Example 1.3.2. *Let X show profit of a financial position. The Conditional Value at Risk at level $\beta \in (0, 1)$ of X is given by*

$$CVaR_\beta(X) = \frac{1}{\beta}\int_0^\beta VaR_\lambda(X)d\lambda, \tag{1.3.2}$$

where $VaR_\lambda(X)$ is defined by equation (1.3.1) *(see [45] for a thorough discussion).*

So far, we have provided axiomatic definition of convex and coherent risk measures along with the financial interpretation of axioms. In the sequel of this section we are interested in recalling robust representations which these risk measures meet. Before going further, we provide an important set of financial positions which an underlined risk measure induced. This set helps us introduce robust representations for convex or coherent risk measures.

Definition 1.3.3. *([45]) A risk measure ρ induces the class*

$$\mathcal{A}_\rho := \{X \in L^\infty | \rho(X) \leq 0\}, \tag{1.3.3}$$

of positions which are acceptable in the sense that they do not require additional capital. The class $\mathcal{A}_\rho$ will be called the acceptance set of ρ.

As we mentioned earlier in this section, Convex and coherent risk measures meet explicit representations in terms of their conjugate functions. These representations help apply available results on risk measures to practice. For instance, these representations enable to study practical problems using coherent risk measures and derive interesting results (see [27]). For instance, we apply these representations to study the problem of capital allocation in Chapter 2 of this thesis. This part of Chapter 1 is devoted to recall main definitions and representation theorems for convex and coherent risk measures. For more information see [45, 9].

Definition 1.3.4. *([45]) A coherent risk measure $\rho : L^\infty \to \mathbb{R}$ is said to have the* **Fatou property** *if for any bounded sequence X_n in L^∞ converging to $X \in L^\infty$ almost surely(in probability) , we have*

$$\rho(X) \leq \liminf \rho(X_n). \tag{1.3.4}$$

Fatou Property enables us to study convex and coherent risk measures using their penalty functions (which are also called conjugate functions in convex analysis context). In fact, this property which is equivalent with continuity from above for a risk measure plays the role of property so called lower semi continuity for convex functions in convex analysis.

Now,let us denote by $\mathcal{M}_1 := \mathcal{M}_1(\Omega, \mathcal{F}, P)$ the set of all probability measures on $(\Omega, \mathcal{F})$ which are absolutely continuous with respect to P. The following theorem characterizes those convex risk measures on L^∞ that can be represented by a penalty function concentrated on probability measures, and hence on $\mathcal{M}_1(P)$ as presented in [45].

Theorem 1.3.1. *([45]) For a convex risk measure $\rho : L^\infty \to \mathbb{R}$ the following are equivalent:*

(1) ρ is a convex risk measure with the Fatou property

(2) The acceptance set $\mathcal{A}_\rho$ of ρ is weak closed in L^∞ i.e., $\mathcal{A}_\rho$ is closed with respect to the topology $\sigma(L^\infty, L^1)$ and convex.*

(3) ρ is a convex risk measure which is continuous from above i.e. for any bounded and decreasing sequence X_n converging to X, $\rho(X) = \lim \rho(X_n)$.

(4) ρ can be represented by:

$$\rho(X) = \sup_{Q \in \mathcal{M}_1(P)} (E_Q(-X) - \alpha_{min}(Q)) \; \forall X \in L^\infty, \tag{1.3.5}$$

where, $\alpha_{min}(Q) = \sup_{\mathcal{A}_\rho} E_Q(-X)$.

We point out that penalty function in Theorem 1.3.1 is not unique for a given convex risk measure. These Penalty functions α_{min} are not only hard to find for convex risk measures but also difficult to deal with in practical problems. In the next corollary we specialize Theorem 1.3.1 for coherent risk measures where it shows for coherent risk measures α_{min} is zero.

Corollary 1.3.1. *([45]) A coherent risk measure on L^∞ can be represented by a set $\mathcal{M} \subset \mathcal{M}_1(P)$ if and only if the equivalent conditions of Theorem 1.3.1 are satisfied. In this case:*

$$\rho(X) = \sup_{Q \in \mathcal{M}} (E_Q(-X)) \; \forall X \in L^\infty. \tag{1.3.6}$$

We see from Corollary 1.3.1 that penalty function α_{min} in Theorem 1.3.1 is vanished for a coherent risk measure . As the set A_ρ for the coherent risk measure ρ is a cone and also ρ has positive homogeneous property, we come up with only two values for α_{min} which are 0 and ∞. This is the reason why α_{min} is vanished for coherent risk measures.

Remark 1.3.1. *If we identify a subset of absolutely continuous measures $\mathcal{M}$ with the set of its Radon-Nikodym derivatives i.e. $D_\sigma = \{f \in L^1_+(\Omega) | \exists Q \in \mathcal{M}, f = \frac{dQ}{dP}\}$ where*

$$L^1_+(\Omega, \mathcal{F}, \mathbb{P}) = \{f \in L^1(\Omega, \mathcal{F}, \mathbb{P}) \mid f \geq 0,\ \mathbb{E}(f) = 1\}, \tag{1.3.7}$$

then we can rewrite the representation of ρ by:

$$\rho(X) = \sup_{f \in D_\sigma} (E_P(-fX))\ \forall X \in L^\infty.$$

1.3.2. Capital Allocation

Risk measures find different applications in mathematical finance and insurance. One of the problem for which we see the theory of risk measures has been widely used is the problem of capital allocation. We also study this problem in Chapter 2 under a class of coherent risk measures when we model the reserve of an insurance company by spectrally negative Lévy processes. We also provide explicit results for this problem. Let ρ be a risk measure on L^∞ and let $X = X_1 + \cdots + X_d$ be the cash flow produced over a unit time period by the portfolio of a company consisting of several departments, so that X_i is the cash flow produced by the i^{th} department. The problem of the capital allocation is: how is the total risk $\rho(X)$ allocated between the departments? There have been proposed different allocation methods to solve this problem. For a comparison of different combinations of risk measures and allocation methods we refer to [88]. For class of coherent risk measures [36, 42] propose the following axiomatic definition for capital allocation which we use as benchmark for capital allocation in Chapter 2 of this thesis.

Definition 1.3.5. *([27]) Let ρ be a coherent risk measure on L^∞ and let $X = X_1 + \cdots + X_d$ be the cash flow produced over a unit time period by the portfolio of a company consisting of several departments, so that X_i is the cash flow produced by the i^{th} department. a capital allocation between $X_1, \ldots, X_d$ is a vector $(x_1, ..., x_d) \in \mathbb{R}^d$ such that*

$$\sum_{i=1}^d x_i \;=\; \rho\left(\sum_{i=1}^d X_i\right), \tag{1.3.8}$$

$$\sum_{i=1}^{d} h_i x_i \quad \leq \quad \rho\left(\sum_{i=1}^{d} h_i X_i\right), \; h = (h_1, \ldots, h_d) \in \mathbb{R}_+^d. \tag{1.3.9}$$

The first condition is called the *full allocation* property the second condition is called the *linear diversification property* of capital allocation. If we set $h_i = 1$ and $h_j = 0$ for $j = 1, \ldots, n, \; j \neq i$ in inequality (1.3.9), then $x_i \leq \rho(X_i)$. This means that we expect i-th department in an insurance company has smaller risk contribution in the total risk of the company than when it considered as a separate sector. This is compatible with diversification property of capital allocation.

1.3.3. Set-Valued Risk Measures

In previous subsection we have provided the definition and robust representation for coherent and convex risk measures defined on the space L^∞ where it is the space of all real -valued risky portfolios. Now, consider the case where the risky portfolio is an $\mathbb{R}^d$-valued random variable. In [56], the authors propose an axiomatic definition for set-valued coherent risk measures defined on the space of $L_d^\infty(\mathbb{R}^d)$, the space of all equivalence classes of (essentially) bounded $\mathbb{R}^d$-valued random variables. They define (d, n) -coherent risk measures as set-valued maps from $L_d^\infty(\mathbb{R}^d)$ into subsets of $\mathbb{R}^d$, $\mathcal{P}(\mathbb{R}^n)$, satisfying some axioms for $n \leq d$ (see [56]). As we will see in the following, dealing with vector-valued risk measures turns out to be a difficult task. This inspired us to look for an alternative way to extend the theory of risk measures to vector-valued random variables. The structure proposed in [56] inspired us to study vector-valued risk measures defined on a set of data instead of studying them for multidimensional random variables. In Chapter 4 of this thesis we propose an axiomatic definition for vector-valued data-based risk measures which enables us to evaluate the possible risk of multivariate data. We consider these risk measures defined on multivariate data as alternative risk measures for vector-valued risk measures. To give the reader a flavor of this structure, we need to provide the proposed definition for vector-valued risk measures presented in [56].

In order to provide the definition of (d, n) -coherent risk measures we need to recall the portfolio ordering for $L_d^\infty(\mathbb{R}^d)$. The following discussion is taken from [56].

Let $(\Omega, \mathcal{F}, \mathbb{P})$ and K be a probability space and a closed convex cone of $\mathbb{R}^d$ respectively such that

$$\mathbb{R}_+^d \subset K, \qquad K \neq \mathbb{R}^d,$$

and

$$\forall i = n+1, \ldots, d: \quad -\mathbf{1}^i + \alpha \mathbf{1}^1 \text{ and } \mathbf{1}^i - \beta \mathbf{1}^1 \in K \quad \text{for some } \alpha, \beta > 0, \tag{1.3.10}$$

where $\mathbf{1}^i$ is the i^{th} canonical basis vector defined by $\mathbf{1}^i_j = 1$ if $i = j$, and zero otherwise. The condition (1.3.10) is called the substitutability condition. The closed convex cone K induces the partial ordering on $\mathbb{R}^d$ by $x \succeq 0$ if and only if $x \in K$. This induces the partial ordering on $L^\infty_d(\mathbb{R}^d)$ in the following way.

$$X \succeq 0 \quad \text{if and only if} \quad X \in K, \tag{1.3.11}$$

for $X \in L^\infty_d(\mathbb{R}^d)$.

For each $x \in \mathbb{R}^n$ we consider the following notation for the translation of x into $\mathbb{R}^d$ for $n \leq d$.

$$\forall x \in \mathbb{R}^n, \qquad \tilde{x} := (x, \mathbf{0}) \in \mathbb{R}^d,$$

where the last $d - n$ components of $\tilde{x}$ are zero.

Definition 1.3.6. *([56]) A (d,n)-coherent risk measure is a set-valued map $R : L^\infty_d(\mathbb{R}^d) \longrightarrow \mathcal{P}(\mathbb{R}^n)$ satisfying the following axioms:*

(1) $\forall X \in L^\infty_d(\mathbb{R}^d)$, $R(X)$ is closed, and $\mathbf{0} \in R(\mathbf{0}) \neq \mathbb{R}^n$;

(2) $\forall X \in L^\infty_d(\mathbb{R}^d)$: $X \succeq 0$ $\mathbb{P} - a.s.$ $\Longrightarrow R(\mathbf{0}) \subset R(X)$;

(3) $\forall X, Y \in L^\infty_d(\mathbb{R}^d)$, $R(X) + R(Y) \subset R(X+Y)$;

(4) $\forall t > 0$ and $X \in L^\infty_d(\mathbb{R}^d)$, $R(tX) = tR(X)$;

(5) $\forall x \in \mathbb{R}^n$ and $X \in L^\infty_d(\mathbb{R}^d)$, $R(X + \tilde{x}) = \{-x\} + R(X)$.

In Definition 1.3.6 the summation between two sets $R(X)$ and $R(Y)$ is the Minkowski sum which is defined as follows.

$$R(X) + R(Y) = \{a + b : \quad a \in R(X),\ b \in R(Y)\}.$$

1.3.4. Natural Risk Statistics

In this subsection, we give the definition of the concept so called natural risk statistics and a representation of it. Multivariate data-based risk measures which we propose in Chapter 4 of this thesis are alternatives for vector-valued risk measures. These data-based risk measures meet an axiomatic construction which is a natural development of a construction of the natural risk statistics propose in [53]. Natural risk statistics, as defined in [53], is an alternative to coherent risk measures. This type of risk measure is defined on $\mathbb{R}^n$, as the space of data with length n. Before moving on further, note that in the definition of natural risk statistics, the argument of risk measure is "loss" instead of "profit" or "pay-off".

Definition 1.3.7. *([53]) A function $\rho : \mathbb{R}^n \to \mathbb{R}$ is a natural risk statistics if,*

(1) Positive homogeneity and translation invariance:

$$\rho(a\tilde{x} + b\mathbf{1}) = a\rho(\tilde{x}) + b,\ \forall \tilde{x} \in \mathbb{R}^n,\ a \geq 0,\ b \in \mathbb{R}$$

where $\mathbf{1} = (1, \ldots, 1)^T \in \mathbb{R}^n$.

(2) Monotonicity:

$$\rho(\tilde{x}) \le \rho(\tilde{y}),\ if\ \tilde{x} \le \tilde{y},$$

where $\tilde{x} = (x_1, \ldots, x_n) \le \tilde{y} = (y_1, \ldots, y_n)$ *if and only if* $x_i \le y_i,\ i = 1, \ldots, n$.

(3) Comonotonic subadditivity:

$$\rho(\tilde{x} + \tilde{y}) \le \rho(\tilde{x}) + \rho(\tilde{y}),\ if\ \tilde{x}\ and\ \tilde{y}\ are\ comonotonic,$$

where $\tilde{x}$ *and* $\tilde{y}$ *are comonotonic if and only if* $(x_i - x_j)(y_i - y_j) \ge 0$, *for any* $i \ne j$.

(4) Permutation invariance:

$$\rho((x_1, \ldots, x_n)) = \rho((x_{i_1}, \ldots, x_{i_n}))$$

for any permutation $(i_1, \ldots, i_n)$.

The following theorem, proved in [53], gives a robust representation for natural risk statistics.

Theorem 1.3.2. *([53]) Let* $x_{(1)}, \ldots, x_{(n)}$ *be the order statistics of the observation* $\tilde{x}$ *with* $x_{(n)}$ *being the largest. Then* ρ *is a natural risk statistic if and only if there exists a set of weights* $\mathcal{W} = \{\tilde{w} = (w_1, \ldots, w_n)\} \subset \mathbb{R}^n$ *with each* $\tilde{w} \in \mathcal{W}$ *satisfying* $\sum_{i=1}^n w_i = 1$ *and* $w_i \ge 0,\ \forall 1 \le i \le n$, *such that*

$$\rho(\tilde{x}) = \sup_{\tilde{w} \in \mathcal{W}} \sum_{i=1}^{n} w_i x_{(i)},\ \forall \tilde{x} \in \mathbb{R}^n. \tag{1.3.12}$$

Chapter 2

ON THE CAPITAL ALLOCATION PROBLEM FOR A NEW COHERENT RISK MEASURE IN COLLECTIVE RISK THEORY

Abstract

In this chapter we introduce a new coherent cumulative risk measure on a subclass in the space of càdlàg processes. This new coherent risk measure turns out to be tractable enough within a class of models where the aggregate claims is driven by a spectrally positive Lévy process. Moreover, this risk measure is well-suited to address the problem of capital allocation in an insurance context. Indeed, we show that the capital allocation problem for this risk measure has a unique solution determined by the Euler allocation method. Some examples are provided.

Keywords. Capital allocation, Euler allocation method, Coherent risk measures, Lévy insurance processes, Risk measures on the space of stochastic processes.

This chapter is a joint research work with Hirbod Assa and Manuel Morales; see [11] [1].

2.1. Introduction

Collective risk theory has built upon the pioneering work of Filip Lundberg [30] and it now comprises a substantial body of knowledge that concerns itself

[1]In this project, my contribution was to identify the main idea of the research problem where we were interested to introduce a class of coherent risk measures on the space of stochastic processes . After discussing with Prof. Assa about the possible extensions of research projects he has already done in his thesis [9], we came up with a general idea we put forward in [11]. I prepared the mathematical ground for the project by recalling important required theorems and definitions in the context. I proved the results in the project and applied them to important examples in actuarial context. I also contributed in this project by writing up the first draft of it, then polishing it with Prof. Morales to get the final version.

with the study the riskiness of an insurer's reserve as measured by the ruin probability and related quantities [7]. A large amount of literature now exists on such insolvency measures for a wide variety of models, the latest being the so-called Lévy insurance risk models [19] and [20].

Traditionally, collective risk theory focuses on the insurer's ability to manage the solvency of its reserve through the control of initial investment x. The mathematical tool often cited for such task is the probability of ruin. It is indeed a measure of the likelihood that an insurer's reserve would eventually be insufficient to cover its liabilities in the long run.

More precisely, consider the following general model for the risk reserve of an insurance company,

$$R_t = x + c\,t - X_t\ , \quad t \geq 0\ , \tag{2.1.1}$$

where the aggregate claims process X is a spectrally positive Lévy process with zero drift, with $X(0) = 0$ and jump measure denoted by ν. Moreover, x is the initial reserve level and c is a constant premium rate defined as

$$c = (1 + \theta)\mathbb{E}[X_1] \tag{2.1.2}$$

where $\theta > 0$ is the security loading factor.

Then the associated ruin time is

$$\tau_x := \inf\{t \geq 0 \mid X_t - c\,t \geq x\}\ , \tag{2.1.3}$$

and the *infinite-horizon ruin probability* can be defined by

$$\Psi(x) := \mathbb{P}_x(\tau_x < \infty)\ , \tag{2.1.4}$$

where $\mathbb{P}_x$ is short-hand notation for $\mathbb{P}(\,\cdot \mid X(0) = x)$.

Much of the literature in collective risk theory studies the problem of deriving expressions and reasonable approximations for the probability of ruin as a function of the initial reserve level x. This problem is addressed within an ever-growing set of models for the aggregate claims process. See [7] for a thorough account on the so-called ruin theory.

Naturally, the ruin probability Ψ quantifies the solvency of the net-loss process $Y_t := X_t - ct$ as a function of the initial reserve level x. In fact, we can define a risk measure $\rho_\beta : \mathcal{X} \longrightarrow [0,1]$ on a suitable model space $\mathcal{X}$ (say the space of bounded càdlàg stochastic processes $\mathcal{R}^\infty$). Let $Y_t = X_t - ct$ be the net-loss process associated with the reserve process (2.1.1), then

$$\rho_\beta(Y) \longmapsto a := \inf\{x \geq 0 \mid \Psi(x) \leq \beta\}\ , \tag{2.1.5}$$

where ψ is the associated ruin probability (2.1.4) and $\beta \in (0,1)$ represents a given tolerance to ruin. We refer to [25] for a particular case dealing with Value at Risk.

One can interpret a as the smallest initial level for which the process R has an acceptable risk level, i.e. its associated ruin probability is less or equal to a tolerable figure β. Such risk measures have been recently studied (see [84]) and although they exhibit interesting properties, they lack the tractability of an efficient risk management tool. In fact, any meaningful risk management application, such as capital allocation, would be hard to implement using (2.1.5). Recently, other risk measures have been studied such as one based on the concept of *area in red*, which is a measure of how large the overall deficit of the company can be (see [65] for details). These new notions turn out to exhibit some interesting properties, yet the issue with these risk measures remains, they are very difficult to implement in a risk management problem such as capital allocation.

In this chapter, we recover this idea of measuring the risk of an insurance risk process but with a view towards an application in the capital allocation problem. In fact, we reverse-engineer a risk measure with the sole purpose of addressing the non-trivial problem of capital allocation in a collective risk theory context. That is, we look at the aggregate loss from a number of, potentially dependent, lines of business and we give a way to allocate a portion of the overall risk to each component. More precisely, we define a coherent risk measure on a suitable subspace of the space of càdlàg processes, R^p_L, as a mapping $\rho : R^p_L \longrightarrow \mathbb{R}_+$ (the precise description of the subspace in question, R^p_L, is given in Defintion 2.2.1 on page 35). We then give a definition of what we mean by capital allocation in this context and give an explicit solution for it. Unlike (2.1.5), this measure is tractable enough and allows for a solution of the capital allocation problem within a suitable space of stochastic processes. This is carried out within the framework given by the theory of coherent and convex risk measures for stochastic processes. There is indeed a fair amount of research on the question of how to define risk measures for stochastic processes. Among previous works on these issues we find, for instance, [24] and [25] where the authors work out risk measures on the space of random processes modeling the evolution of a certain financial position or [26] where they develop risk measures in a dynamic fashion. Indeed there is now a comprehensive theory of risk measures on the space of stochastic processes that draws from convex analysis, probability and the general theory of stochastic processes in order to build a mathematical framework for the quantification of economic risks in a dynamic fashion.

On the other hand, applications of risk measures defined on the space of stochastic processes are less abundant. In particular, the problem of capital allocation in such a framework is far from trivial. On the space of random variables, this problem is well-studied and it has been addressed in different ways. For instance, in [82], the author applies risk measures to study the problem of capital allocation for random variables in a general framework while in [37] we find a cooperative game theory study of the fair allocation principle for coherent risk measures. In [37, 82], the authors propose the Euler principle for allocating the required capital under some technical assumptions and, in [88], a comparison of different combinations of risk measures and allocation methods can be found. Another approach can be found in [86], where they propose a sensitivity analysis framework for internal risk models for the class of distortion risk measures. As a special case they study the problem of capital allocation for this class of risk measures under their sensitivity framework.

Unfortunately, when it comes to the space of stochastic processes, it is a different story. The problem of capital allocation for coherent risk measures on the space of stochastic processes turns out to be a more difficult task. A comprehensive understanding of the problem of capital allocation for coherent risk measures in this setting requires advanced notions and techniques from functional analysis as well as convex analysis. In fact, a formal treatment of the problem of capital allocation for coherent risk measures requires studying the weak subgradient set associated to the risk measure [10]. As it turns out, in order to get a good understanding of the sub-gradient set and its properties, we need a robust representation of the underlying risk measure which, in turn, requires studying of the dual space of R_L^p. Now, the fact that, we have a sophisticated topological structure to deal with for the dual space of R_L^p, makes it difficult to characterize the sub-gradient set and to give a solution to the problem of capital allocation. To the best of our knowledge, this problem has not been thoroughly studied for risk measures defined on the space of stochastic processes. We can only cite, [10], where the author discusses the problem of capital allocation for risk measures defined on the space of cádlág processes. Or, [65] where the authors study the capital allocation problem for a new risk measure that, as it turned out, it does not satisfy an axiomatic definition of coherent risk measures defined on stochastic processes proposed in [24]. As a drawback, the resulting solution of the capital allocation problem, does not follow an axiomatic definition of capital allocation. Moreover, neither the proposed risk measure, nor the capital allocation solution, have an explicit formula.

We circumvent all these issues by giving an ad-hoc smooth explicit representation of a class of risk measures. Such construction leads naturally to an explicit solution for the capital allocation problem without having to make use of the advanced machinery from functional analysis.

The contribution of this chapter is then two-fold. First, based on [2] and [9], we design a new risk measure on the space of bounded càdlàg processes that can capture the risk associated with the path-properties of an insurance model. We do this by extending the notion of *Entropic Value at Risk*, first introduced in [2], to a suitable space of stochastic processes. Second, we explore the capital allocation problem using this new risk measure in an insurance context and we show that the Euler allocation method is the only method to allocate the requiring capital for this risk measure.

The outline of the chapter is as follows. In Section 2.2, we introduce the notion of *Cumulative Entropic Value at Risk* (CEVaR_β) as a coherent risk measure on the space of bounded stochastic processes and we explore some of its relevant features. In Section 2.3, we explore the capital allocation problem and give a theorem which characterizes the capital allocation set for these measures. In fact, we show that for the CEVaR_β risk measure the Euler allocation method is the only way to allocate the risk capital. Finally, in Section 2.4, we show some results for CEVaR_β and provide some examples.

2.2. Cumulative Entropic Risk Measures

Let $(\Omega, \mathcal{F}, \mathbb{P}, \bar{\mathcal{F}})$ be a filtered probability space. We consider the space $\mathcal{R}^p$ of stochastic processes on $[0, T]$ that are càdlàg, adapted and such that $X^* := \sup_{[0,T]} |X_t| \in L^p(\Omega, \mathcal{F})$, with $1 \le p \le \infty$. Furthermore, assume that $L^1(\Omega, \mathcal{F}, \mathbb{P})$ has a countable dense subset. Notice that, for any $1 \le p \le \infty$, the space $\mathcal{R}^p$ endowed with the norm $||X||_{\mathcal{R}^p} = ||X^*||_{L^p}$, is a Banach space.

In [24] and [25] the authors developed the theory of convex risk measures on the space of $\mathcal{R}^p$ ($\rho : \mathcal{R}^p \longrightarrow \mathbb{R}_+$). It is within this framework that we develop our approach. We start by defining a subset of $\mathcal{R}^p$ that serves our purposes and with which we will work for the rest of the chapter.

Definition 2.2.1. *We define the subspace $\mathcal{R}^p_L$ containing the processes $X \in \mathcal{R}^p$ with the following property.*

$$m_t(s) = \mathbb{E}[\exp(-s\,X_t)] < \infty\ , \qquad s \ge 0\ ,$$

for $t \in [0, T]$.

The idea we put forward in this chapter is to use a *cumulative risk measure* based on the *Entropic Value at Risk* that was defined in [2]. That is, following

[10], we measure the risk of a random process $X \in \mathcal{R}_L^p$ by defining a *cumulative risk measure* $\rho : \mathcal{R}_L^p \longrightarrow \mathbb{R}_+$ as follows. Let ρ_0 be a given risk measure on $L^p(\Omega, \mathcal{F}, \mathbb{P})$, i.e. $\rho_0 : L^p(\Omega, \mathcal{F}, \mathbb{P}) \longrightarrow \mathbb{R}$, and let $\omega : [0,T] \longrightarrow \mathbb{R}^+$ be a suitable weight function, i.e. $\int_0^T \omega(t)dt = 1$. Then we can define a cumulative risk measure $\rho : \mathcal{R}_L^p \longrightarrow \mathbb{R}_+$ based on ρ_0 as the weighted aggregate risk of a random process $X \in \mathcal{R}_L^p$. More precisely,

$$\rho(X) := \int_0^T \rho_0(X_t)\omega(t)dt \ . \tag{2.2.1}$$

An interesting and meaningful choice for a weight function mentioned above can be a density function for a random time which captures important moments associated to the underlined process $X = (X_t)_{t \geq 0}$. Such constructions were proposed and studied in [10]. The features of such measures inherently depend on the choice of base risk measure ρ_0. In fact, if the risk measure ρ_0 is coherent then ρ in (2.2.1) is coherent as well.

Theorem 2.2.1. *Let $L^p(\Omega, \mathcal{F}, \mathbb{P})$ be the space of financial positions with finite $||.||_p$ and ρ_0 be a coherent risk measure on $L^p(\Omega, \mathcal{F}, \mathbb{P})$. Then the risk measure $\rho : \mathcal{R}_L^p \longrightarrow \mathbb{R}_+$, given in (2.2.1), is a coherent risk measure on the space $\mathcal{R}_L^p$.*

Proof. First we show the positive homogeneity and translation invariance properties of ρ. For $\lambda > 0$ and $m \in \mathbb{R}$ we have,

$$\rho(\lambda X + m) = \int_0^T \rho_0(\lambda X_t + m)\omega(t)dt = \lambda\ \rho(X) - m \int_0^T \omega(t)dt \ ,$$

which shows the positive homogeneity and translation invariance properties since $\int_0^T \omega(t)dt = 1$.

As for monotonicity, if $X_t \leq Y_t$ a.s., then $\rho_0(X_t) \geq \rho_0(Y_t)$ for $t \in [0,T]$. Now, since ω is a positive real valued function, we have $\rho_0(X_t)\omega(t) \geq \rho_0(Y_t)\omega(t)$ for any $t \in [0,T]$ as well. This implies that $\rho(X) \geq \rho(Y)$ which proves the monotonicity.

Now using the subadditivity property of ρ_0 and since ω is a positive function we have,

$$\rho_0(X_t + Y_t)\omega(t) \leq \rho_0(X_t)\omega(t) + \rho_0(Y_t)\omega(t) \ ,$$

for $t \in [0,T]$. This directly implies the subadditivity property of ρ. i.e.,

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

□

In this chapter, we propose to use the *Entropic Value at Risk* measure (EVaR_β) as our measure ρ_0 in (2.2.1). This yields an interesting family of risk measures on the space of bounded stochastic processes. Following [2, 46] we now give a first definition.

Definition 2.2.2. *Let X be a random variable in $L^\infty(\Omega, \mathcal{F}, \mathbb{P})$ with*

$$\mathbb{E}[\exp(-s\,X)] < \infty \;, \qquad s > 0 \;. \tag{2.2.2}$$

Then the **Entropic Value at Risk***, denoted by $EVaR_\beta$, is given by*

$$EVaR_\beta(X) := \inf_{s>0} \frac{\ln \mathbb{E}[\exp(-s\,X)] - \ln \beta}{s}, \tag{2.2.3}$$

for risk level $\beta \in (0,1)$.

We can generalize this definition for random variables in $L^p(\Omega, \mathcal{F}, \mathbb{P})$ for $p \geq 1$, as long as they meet (2.2.2). The following key result for EVaR$_\beta$ can be found in [2].

Theorem 2.2.2. *The risk measure $EVaR_\beta$ from Definition 2.2.2 is a coherent risk measure. Moreover, for any $X \in L^\infty(\Omega, \mathcal{F}, \mathbb{P})$ having property (2.2.2), its dual representation has the form*

$$EVaR_\beta(X) = \sup_{f \in \mathcal{D}} \mathbb{E}_{\mathbb{P}}(-f\,X) \;, \tag{2.2.4}$$

where $\mathcal{D} = \{f \in L^1_+(\Omega, \mathcal{F}, \mathbb{P}) \mid \mathbb{E}_{\mathbb{P}}[f \ln(f)] \leq -\ln \beta\}$ and

$$L^1_+(\Omega, \mathcal{F}, \mathbb{P}) = \{f \in L^1(\Omega, \mathcal{F}, \mathbb{P}) \mid \mathbb{E}_{\mathbb{P}}(f) = 1\}. \tag{2.2.5}$$

For the proof we refer to [2].

If we use the risk measure (2.2.3) in our general definition of a cumulative risk measure (2.2.1), we naturally obtain a risk measure on the space $\mathcal{R}^p_L$ that would inherit some of the key features of the original risk measure.

We now formally introduce the concept of *Cumulative Entropic Value at Risk*, denoted by CEVaR_β, on the space $\mathcal{R}^p_L$.

Definition 2.2.3. *Let X be a stochastic process in $\mathcal{R}^p_L$ and let $EVaR_\beta$ be the risk measure in Definition 2.2.2. Then, for a given weight function $\omega : [0,T] \longrightarrow \mathbb{R}^+$ (i.e. $\int_0^T \omega(t)dt = 1$), the* **Cumulative Entropic Value at Risk***, denoted by $CEVaR_\beta$, is defined by*

$$CEVaR_\beta(X) = \int_0^T EVaR_\beta(X_t)\omega(t)dt \;. \tag{2.2.6}$$

The main advantage of using (2.2.3) as our based measure is that the resulting cumulative risk measure (2.2.6) is tractable enough for a wide family of collective risk models. This comes from the fact that the expectation appearing in (2.2.3) is merely the Laplace exponent of the random variable X_t (for $t \geq 0$). In collective risk theory, many of the models used for insurance reserves have closed-form Laplace transforms, in particular the so-called Lévy insurance risk processes. If the aggregate claims process is driven by a spectrally negative Lévy processes then a cumulative entropic risk measure based on the EVAR_β is an natural choice to work with in risk management applications.

The risk measure in Definition 2.2.3 belongs to the general framework of axiomatic risk measures on the space of stochastic processes developed in [24]. We now study some of its properties.

Corollary 2.2.1. *The risk measure $CEVaR_\beta$, given in Definition 2.2.3, is a coherent risk measure on the space $\mathcal{R}^p_L$.*

Proof. Since EVaR_β is of the form (2.2.1) with a coherent base risk measure ρ_0, it follows that EVaR_β is a coherent risk measure as a special case of Theorem 2.2.1. □

Now, one can notice that in Definition 2.2.3 the weight function ω plays an important role. Different choices of weight functions would result in different cumulative Entropic risk measures. One can naturally think of ω as a density function that distributes a probability mass over the interval $[0, T]$. Interesting choices would be to use the density function f_τ of a suitable stopping time $\tau \in [0, T]$, like the first passage time or ruin time. This would penalize certain regions of the interval $[0, T]$ according to whether a certain meaningful event is more or less likely to occur over these regions.

For tractability purposes, in this chapter, we use a uniform weight function, i.e. we consider $\omega(t) = \frac{1}{T}$. In the remaining of the chapter we will be working with the following subfamily of CEVaR,

$$CEVaR_\beta(X) = \frac{1}{T}\int_0^T EVaR_\beta(X_t)dt \; . \tag{2.2.7}$$

Now, the object of our interest in this chapter is to apply the *CEVaR* in (2.2.7) within an insurance context where the aggregate claims are modeled by a spectrally positive Lévy processes (or equivalently, the surplus process is modeled by a spectrally negative Lévy process). The following proposition enables us to include a subclass of spectrally negative Lévy processes having m.g.f in the space $\mathcal{R}^p$ for some $p \geq 1$. This would enable us to use *CEVaR* with this class of processes.

Proposition 2.2.1. *Let X be a càdlàg Lévy process with $X_0 = 0$ and let p be a real number in $[1, \infty)$. Then, the following are equivalent.*

(1) X is L^p-integrable.

(2) $X_t^* = \sup_{0 \le s \le t} |X_s|$ *is* L^p*-integrable.*

PROOF. This is a special case of Theorem 25.18 in [81]. □

Remark 2.2.1. *It can be seen from Proposition 2.2.1 that the class of* L^1*-integrable spectrally negative Lévy processes having m.g.f is in the space* $\mathcal{R}^1$*. Thus, we can model the surplus process of an insurance company using the elements of this class of processes and apply CEVaR.*

2.2.1. Examples

The *Cumulative Entropic Risk Measure* introduced in Definition 2.2.3 has the advantage of being tractable enough for a large family of processes which have moment generating function and that can be used as models for the net-loss process in (2.1.1). Here we discuss a few examples and compute expressions for the CEVaR in (2.2.7) for some Lévy insurance risk models. More examples will be provided in Section 5.5 when we set up our surplus model in Subsection 2.3.2.

2.2.1.1. *Brownian Motion with Drift*

Let $Y_t = \mu t + \sigma W_t$ be a Brownian motion with drift parameter μ and scale parameter σ for $\mu < 0, \sigma > 0$. Such a process is used in collective risk theory as the net-loss process in (2.1.1) for an approximation to the classical Cramer-Lundberg model ([51]). The moment generating function of $Z_t = -Y_t$ is

$$\mathbb{E}(e^{-sZ_t}) = \mathbb{E}(e^{sY_t}) = e^{\mu ts + \frac{1}{2}\sigma^2 s^2 t}\,.$$

By direct substitution in (2.2.3) and differentiation with respect to s we have, for $t \in [0, T]$,

$$s^* = \sqrt{\frac{-2\ln\beta}{\sigma^2 t}}.$$

By direct substitution in (2.2.3) we then have

$$EVaR_\beta(Z_t) = \mu t + \frac{1}{2}\sigma\sqrt{-2t\ln\beta} - \frac{\sigma\ln\beta}{\sqrt{-2\ln\beta}}\sqrt{t}\,. \tag{2.2.8}$$

Direct substitution and integration of (2.2.8) into (2.2.7) results in

$$CEVaR_\beta(Z) = \frac{\mu T}{2} - \frac{4\sigma\ln\beta}{3\sqrt{-2\ln\beta}}\sqrt{T}\,.$$

2.2.1.2. *Gamma Subordinator*

Let $Y_t = \mu t + X_t$ be a gamma process with parameters $a, b > 0$, with drift parameter $\mu < 0$ and mean $\mathbb{E}(Y_t) = (\mu + \frac{a}{b})t$. The moment generating function of $Z_t = -Y_t$ is

$$\mathbb{E}(e^{-sZ_t}) = \mathbb{E}(e^{sY_t}) = e^{\mu ts - ta\ln(1-s/b)}.$$

In this case, to obtain $\mathrm{EVaR}_\beta(Z_t)$ we need to find s^* from the following equation.

$$\frac{tas}{b-s} + at\ln(1-\frac{s}{b}) + \ln\beta = 0\ , \qquad s \geq 0\ .$$

The above equation is obtained by applying the moment generating function of Z_t in the definition EVaR_β and by straight-forward differentiation with respect to s. Unlike the previous examples, there is no closed-form expression for the solution of this equation. But once s^* is obtained numerically, we can calculate $\mathrm{EVaR}_\beta(Z_t)$ by direct substitution s^* in (2.2.3). $\mathrm{CEVaR}_\beta(Z)$ can be obtained by direct integration of $\mathrm{EVaR}_\beta(Z_t)$ over $[0, T]$.

2.3. Capital Allocation

We now study the problem of capital allocation in an insurance context with the coherent risk measure *CEVaR* that we introduced in Section 2.2. A discussion of the problem of capital allocation for *CEVaR*, which is a risk measure defined on $\mathcal{R}_L^p$, must start with an analysis of this problem for *EVaR*, which is a risk measure on a subspace of $L^\infty(\Omega, \mathcal{F}, \mathbb{P})$.

Finding the capital allocation for a risk measure on the space of stochastic processes typically requires knowledge of its robust representation and its subgradient set (see [9] for a detailed account on this problem). This robust representation is typically a hard problem in the space $\mathcal{R}_L^p$ that normally requires advanced functional analysis tools. In the case of *EVaR* we propose to tackle the problem of finding the capital allocation for CEVaR by finding first the capital allocation for EVaR and then use the linear relation between EVaR and CEVaR to get the capital allocation for CEVaR.

We first give some definitions that will be needed throughout this section.

Definition 2.3.1. *Let ρ be a coherent risk measure defined on $L^\infty(\Omega, \mathcal{F}, \mathbb{P})$. Now let $\mathcal{D} \subset L^1_+$ be the largest set for which the following robust representation holds true for ρ.*

$$\rho(X) = \underset{f\in\mathcal{D}\subset L^1_+}{sup}\ \mathbb{E}_{\mathbb{P}}(-fX)\ \forall X \in L^\infty(\Omega, \mathcal{F}, \mathbb{P})\ , \tag{2.3.1}$$

where L^1_+ is the set defined in (1.3.7). The set $\mathcal{D}$ is called the determining set of ρ (see [45]).

The following definition is taken from [27].

Definition 2.3.2. *Let ρ be a coherent risk measure defined on $L^\infty(\Omega, \mathcal{F}, \mathbb{P})$ with determining set $\mathcal{D} \subset L^1_+$. Let $X \in L^\infty(\Omega, \mathcal{F}, \mathbb{P})$. A function $f \in \mathcal{D}$ is called* *an extreme function for X if $\rho(X) = \mathbb{E}_{\mathbb{P}}(fX) \in (-\infty, \infty)$. The set of extreme functions will be denoted by $\chi_{\mathcal{D}}(X)$.*

The following result is taken from [27] and gives conditions for the set of extreme functions defined above to be non-empty.

Proposition 2.3.1. *Let $\mathcal{D} \subset L^1_+$ be the determining set of a given coherent risk measure ρ on $L^\infty(\Omega, \mathcal{F}, \mathbb{P})$. Now consider the following set.*

$$L^1(\mathcal{D}) := \{X \in L^\infty(\Omega, \mathcal{F}, \mathbb{P}) \mid \lim_{n \longrightarrow \infty} \sup_{f \in \mathcal{D}} \mathbb{E}_{\mathbb{P}}[f\,|X|\,\mathbb{I}_{\{|X|>n\}}] = 0\}. \tag{2.3.2}$$

If the determining set $\mathcal{D}$ is weakly compact and $X \in L^1(\mathcal{D})$, then the set of extreme functions for X is not empty, i.e.$\chi_{\mathcal{D}}(X) \neq \emptyset$.

Now, we turn our attention to the concept of **capital allocation**. Consider a vector of risks $X = (X^1, \ldots, X^d)$, such that $X^i \in L^\infty(\Omega, \mathcal{F}, \mathbb{P})$ for $i = 1, \ldots, d$, are random variables representing the cash flow or risk exposure of a portfolio consisting of d risky positions or departments.

Given a coherent risk measure ρ on $L^\infty(\Omega, \mathcal{F}, \mathbb{P})$, we now look at the problem of how to allocate the total risk of the portfolio $\rho\left(X^1 + \cdots + X^d\right)$ among the different departments such that the individual risk of each one of them is properly measured.

The following formal definition of capital allocation was proposed by [36] and [42] and it is the one we set out to study in this chapter. In fact, the following gives a mathematical definition of capital allocation for a coherent risk measure.

Definition 2.3.3. *Consider a coherent risk measure ρ on $L^\infty(\Omega, \mathcal{F}, \mathbb{P})$ and a vector of risks $X = (X^1, \ldots, X^d)$ such that $X^i \in L^\infty(\Omega, \mathcal{F}, \mathbb{P})$ for $i = 1, \ldots, d$. A fair capital allocation for X is a vector $(K^1, ..., K^d) \in \mathbb{R}^d$ such that*

(1)

$$\sum_{i=1}^{d} K^i = \rho\left(\sum_{i=1}^{d} X^i\right),$$

(2)

$$\sum_{i=1}^{d} h^i K^i \leq \rho\left(\sum_{i=1}^{d} h^i X^i\right), \qquad \forall\, h = (h^1, \ldots, h^d) \in \mathbb{R}^d_+ .$$

The first condition is called the *full allocation* property and it simply states the fact that the total risk of the whole portfolio should be the aggregated risks of each department. The second condition is called the *linear diversification property* of capital allocation. In fact, this condition has a one to one correspondence with the positive homogeneity and subadditivity of a coherent risk measure ρ (see [57]). Since we work in this chapter with a coherent risk measure it is somehow natural to adopt this definition of capital allocation.

The following is an interesting result characterizing the set of possible such capital allocations and it is adapted from [27].

Theorem 2.3.1. *Let $\mathcal{D} \subset L^1_+$ be the determining set of a given coherent risk measure ρ on $L^\infty(\Omega, \mathcal{F}, \mathbb{P})$ and let $X = (X^1, \ldots, X^d)$ be a vector such that $X^i \in L^\infty(\Omega, \mathcal{F}, \mathbb{P})$ for $i = 1, \ldots, d$. Consider the following set*

$$G = \overline{\{(\mathbb{E}_{\mathbb{P}}(-f\, X^1), \ldots, \mathbb{E}_{\mathbb{P}}(-f\, X^d)) \mid f \in \mathcal{D}\}} \subset \mathbb{R}^d\,. \tag{2.3.3}$$

The set $U \subset \mathbb{R}^d$ of capital allocations for $X = (X^1, \ldots, X^d)$, satisfying Definition 2.3.3, is convex and bounded and it has the form

$$U = \underset{x \in G}{argmax} < e, x > , \tag{2.3.4}$$

where $< \cdot, \cdot >$ is the inner product in $\mathbb{R}^d$, $e = (1, \ldots, 1)$ and argmax is the set of points of G for which $< e, x >$ attains its maximum value.

If moreover, $X^1, \ldots, X^d \in L^1(\mathcal{D})$ and $\mathcal{D}$ is weakly compact, then U can be identified to be

$$U = \left\{ \Big(\mathbb{E}_{\mathbb{P}}(-f\, X^1), \ldots, \mathbb{E}_{\mathbb{P}}(-f\, X^d) \Big) \mid f \in \chi_{\mathcal{D}} \left(\sum_{i=1}^{d} X^i \right) \right\}. \tag{2.3.5}$$

Proof. In [27], the author provides a proof of the theorem for coherent utility functions. The result follows by noticing that, for a given coherent risk measure ρ, if we set $\rho^*(X) := -\rho(-X)$ we obtain a coherent utility function and the result in [27] holds. So, from $\rho(X) = -\rho^*(-X)$ the results for the statement of our theorem holds. □

The set $G \subset \mathbb{R}^d$ in Theorem 2.3.1 is called the *generator* for X and ρ (see [27]). The following corollary gives a condition on G for the uniqueness of the capital allocation.

Corollary 2.3.1. *Under the conditions of Theorem 2.3.1. If moreover, $G \subset \mathbb{R}^d$ is strictly convex (i.e. its interior is non-empty and its border contains no interval), then there is a unique capital allocation satisfying Definition 2.3.3.*

Proof. See [27] for a proof in terms of coherent utility functions. □

2.3.1. CEVaR and the Capital Allocation Problem

Our main goal in this chapter is to apply cumulative entropic risk measure in a capital allocation problem. So far, we have discussed key notions of the capital allocation problem for a risk measure on $L^\infty(\Omega, \mathcal{F}, \mathbb{P})$. In this subsection, we apply these results in order to give an answer to the problem of capital allocation for *CEVaR* which is a risk measure on $\mathcal{R}_L^p$. Notice that this is a somewhat more complicated problem since there is a dynamic component to this problem. Here, this is overcome by the cumulative property of CEVaR_β. We start by extending Definition 2.3.3 to the more general notion of capital allocation with respect to a coherent risk measure on the space $\mathcal{R}_L^p$. The following definition is taken from [18].

Definition 2.3.4. *Let $\left(X_t^1,\dots,X_t^d\right)_{t\in[0,T]}$ be d random processes in $\mathcal{R}_L^p$ representing d financial positions or departments. Moreover, consider a coherent risk measure $\rho:\mathcal{R}_L^p\longrightarrow\mathbb{R}_+$ defined on the space $\mathcal{R}_L^p$. A fair capital allocation for $\left(X_t^1,\dots,X_t^d\right)_{t\in[0,T]}$ with respect to ρ is a vector $(L^1,...,L^d)\in\mathbb{R}^d$ such that*

(1)

$$\sum_{i=1}^d L^i=\rho\left(\sum_{i=1}^d X^i\right)\ ,$$

(2)

$$\sum_{i=1}^d h^iL^i\le\rho\left(\sum_{i=1}^d h^iX^i\right)\ ,\qquad\forall\, h=(h^1,\dots,h^d)\in\mathbb{R}_+^d\ ,$$

where $\sum_{i=1}^d X^i$ denotes the process $\left(\sum_{i=1}^d X_t^i\right)_{t\in[0,T]}$.

In this section, we show how a capital allocation satisfying Definition 2.3.4 can be obtained when using CEVaR as risk measure. In fact we show that the border of the set $\mathcal{D}$ in the robust representation (2.3.1) for EVaR_β is not a convex set so leads to the fact that the border of the set G in (2.3.3) is not a convex set as well. This yields to the uniqueness of the capital allocation for *EVaR* (Corollary 2.3.1) as well as for *CEVaR*.

Theorem 2.3.2. *Let D be the determining set in the robust representation* (2.2.4) *for EVaR_β. Then the convex combination of elements of $\partial\mathcal{D}$ is never an element of $\partial\mathcal{D}$ where $\partial\mathcal{D}=\{f\in L_+^\infty:\mathbb{E}_\mathbb{P}(f\ln(f))=-\ln\beta\}$.*

Proof. It is sufficient to show that for any $\lambda\in[0,1]$ and any two functions f and g in $\partial\mathcal{D}$, the function $\lambda f+(1-\lambda)g$ is not in $\partial\mathcal{D}$. Define the function H on the space of positive real line taking real values as follows.

$$H(x):=x\ln x,$$

for all $x\in\mathbb{R}^+$.

It is clear that the function H is strictly convex on the positive real line. Since, $H'(x)=\ln x+1$ and $H''(x)=\frac{1}{x}>0$ for all $x\in\mathbb{R}_+$. Now again we define a new function K on $[0,1]$ with its values in $\mathbb{R}$ by using the composition function

$H(\lambda f + (1-\lambda)g)$ as follows.

$$K(\lambda) = E_{\mathbb{P}}(H(\lambda f + (1-\lambda)g)),$$

for the fixed functions f and g in $\partial\mathcal{D}$. Notice that we use a slight abuse of notation, here $H(\lambda f + (1-\lambda)g)$ is to be understood point-wise. That is, for $x \in \mathbb{R}$, the function $H(\lambda f + (1-\lambda)g) \longrightarrow H(\lambda f(x) + (1-\lambda)g(x))$.

If we take the first and second derivatives for the function K, we see that this function is strictly convex too. $K'(\lambda) = E_{\mathbb{P}}((f-g)(H'(\lambda f + (1-\lambda)g))$ and $K''(\lambda) = E_{\mathbb{P}}((f-g)^2(H''(\lambda f + (1-\lambda)g)) = \frac{(f-g)^2}{\lambda f + (1-\lambda)g} > 0$. Now, considering $K(0) = E_{\mathbb{P}}(H(f))$ and $K(1) = E_{\mathbb{P}}(H(g))$ along with the strictly convexity of the function K, we come up with the inequality

$$K(\lambda) = E_{\mathbb{P}}(H(\lambda f + (1-\lambda)g)) < -\ln\beta \;\forall\lambda \in (0,1).$$

This proves our assertion. $\square$

Now, we are going to characterize the capital allocation satisfying Definition 2.3.4 with respect to CEVaR_β given by (2.2.6). Notice that this seems to be a more complicated problem since CEVaR_β is a risk measure defined on the space of stochastic processes $\mathcal{R}_L^p$. However, this is possible thanks to the cumulative property of CEVaR_β. In fact, this enables us to study the problem of capital allocation for CEVaR_β through studying the same problem for EVaR_β. A characterization theorem for the problem of capital allocation for coherent risk measures on space of random variables has been proved by [37] in the context of game theory. But in the following theorem, for the sake of completeness, we provide a different proof for the capital allocation for EVaR_β using different approach.

Theorem 2.3.3. *Let $(X_t^1, \dots, X_t^d)_{0\le t\le T}$ be a vector such that each $(X_t^i)_{0\le t\le T} \in \mathcal{R}_L^p$ (for $i = 1, \dots, d$) represents the cash-flow or risk exposure from one risk position or department at time $t \in [0,T]$. We denote by $X_t = \sum_{i=1}^d X_t^i$ the portfolio-wide cash-flow produced at time $t \in [0,T]$. Furthermore, define the function $f_\rho^t(u_1, \dots, u_d) = \rho(\sum_{i=1}^d u_i X_t^i)$ where ρ is EVaR_β as defined in Definition 2.2.2. Then, the capital allocation satisfying Definition 2.3.4 over the period $[0,T]$, with respect to CEVaR_β as defined in (2.2.6), is determined uniquely for $i = 1, \dots, d$ by*

$$L^i = \int_0^T K_t^i \omega(t) dt\ , \tag{2.3.6}$$

where K_t^i is

$$K_t^i = \frac{d\rho}{dh}(X_t + hX_t^i)|_{h=0} = \frac{\partial}{\partial u_i} f_\rho^t(1, \dots, 1) \qquad 1 \le i \le d, \qquad t \in [0,T]\ . \tag{2.3.7}$$

PROOF. From the definition of $CEVaR_\beta$ it is clear that

$$CEVaR_\beta\left((\sum_{i=1}^{d} X_t^i)_{t\in[0,T]}\right) = \int_0^T EVaR_\beta\left(\sum_{i=1}^{d} X_t^i\right)\omega(t)dt = \int_0^T EVaR_\beta(X_t)\omega(t)dt.$$

Because of the linear property of integral, studying the problem of capital allocation, both existence and uniqueness of capital allocation, for $CEVaR_\beta$ can be reduced to study this problem for $EVaR_\beta$. Now, we show that K_t^i provided in the theorem is the capital allocation for $EVaR_\beta$ and the vector $(X_t^1,\ldots,X_t^d)$ for a fixed $t\in[0,T]$. For this first of all we show that the possible capital allocations for the vector $(X_t^1,\ldots,X_t^d)$ are those belonging to the following set.

$$A_t := \{x\in G \mid f_\rho^t(V) - f_\rho^t(e) \geq < x, V-e >,\ \forall\ V\in\mathbb{R}^d\}, \tag{2.3.8}$$

where G is given by (2.3.3), $e=(1,\ldots,1)$ and $< .,. >$ is the inner product in $\mathbb{R}^d$. To show this, we assume that $x\in G$ is a capital allocation for the vector $(X_t^1,\ldots,X_t^d)$. Thus, we have $f_\rho^t(e) =< e, x >$. Then, for every $V\in\mathbb{R}^d$ we have $f_\rho^t(V) - f_\rho^t(e) \geq < x, V-e >$. Therefore, $x\in A_t$. Now, assume that $x\in A_t$, then $f_\rho^t(V) - f_\rho^t(e) \geq < x, V-e >$ for all $V\in\mathbb{R}^d$. By replacing $V=2e$ and $V=\frac{1}{2}e$ into the last inequality we get $f_\rho^t(e) =< e, x >$. Therefore, the set of capital allocation associated to the vector $(X_t^1,\ldots,X_t^d)$ is A_t. Since, the risk measure EVaR_β is positive homogeneous, i.e., for all $\lambda>0$ we have $\mathrm{EVaR}_\beta(\lambda X_t) = \lambda\,\mathrm{EVaR}_\beta(X_t)$, we deduce that the function f_ρ above is a homogeneous function. So, by Euler's theorem on homogeneous functions we have $f_\rho^t(u_1,\ldots,u_n) = \sum_{i=1}^d u_i \frac{\partial}{\partial u_i} f_\rho^t(u_1,\ldots,u_d)$ which implies

$$f_\rho^t(1,\ldots,1) = \sum_{i=1}^{d} \frac{\partial}{\partial u_i} f_\rho^t(1,\ldots,1)\ . \tag{2.3.9}$$

for $(u_1,\ldots,u_n)=(1,\ldots,1)$. Because $EVaR_\beta(X_t)$ has smooth explicit representation so partial derivative $\frac{\partial}{\partial u_i} f_\rho^t(1,\ldots,1)$ exists for all $i=1,\ldots,d$. Now, referring to [27] we can see that $x=\left(\frac{\partial}{\partial u_i} f_\rho^t(1,\ldots,1)\right)_{1\leq i\leq d}$ belongs in G as well as in A_t. This shows that gradient of f_ρ^t evaluated at $e=(1,\ldots,1)$ is one choice for capital allocation of the vector $(X_t^1,\ldots,X_t^d)$. To show the uniqueness we refer to any classical convex analysis textbook to see when the gradient exists then the set A_t is a singleton [79]. i.e., the gradient of f_ρ^t evaluated at $e=(1,\ldots,1)$ is the only possible allocation for the vector $(X_t^1,\ldots,X_t^d)$. Now to get the capital allocation for $CEVaR_\beta$ we just need to multiply K_t^i by $\omega(t)$ and take integral w.r.t t. This finishes the proof. $\square$

Corollary 2.3.2. *Under the assumptions of Theorem 2.3.3, the capital allocated to the department i^{th} for the cumulative risk measure $CEVaR_\beta$ given by (2.2.7)* *is*

$$L^i = \frac{1}{T}\int_0^T K_t^i dt \ ,$$

where K_t^i *is given by equation* (2.3.7).

Proof. By replacing $\omega(t) = \frac{1}{T}$ into equation (2.3.6) we get the required capital for the department i^{th}. □

Theorem 2.3.3, gives us a solution to the problem of capital allocation for stochastic processes over a finite time period $[0, T]$. Interesting enough, unlike other solutions to this problem, this capital allocation can be readily computed for a large family of processes. Now, we turn our attention to an application of our results.

2.3.2. Capital Allocation for Insurance Lévy Risk Processes

We now apply Corollary 2.3.2 to give an answer to the capital allocation problem for an insurance risk process. We consider here an insurance company consisting of n departments. For each department, we let R_t^i be a risk reserve process of the form (2.1.1). In other words, $R_t^i = x^i - Y_t^i$ where $Y_t^i = X_t^i - c^i t$ denotes the net-loss claim process related to the i-th department. We recall that x^i is the initial reserve, c^i is the loaded premium and X_t^i is a model for the aggregate claims while the index i refers to one of the n departments. In order to allow for a more rich description of an insurance company, we think of the aggregate claims process X_t^i as the aggregate amount paid out by the department i which is composed of fractions of m independent classes of claims. That is, let $W_t^1, \ldots, W_t^m$ be m independent spectrally positive Lévy process modeling aggregate claims of m different types. One can think for instance of claims associated with car accidents, home damage, medical insurance, etc. Then, the aggregate claims X_t^i paid out by the i-th department would be a linear combination of some of these W_t^j claims processes. For example, consider aggregate claims produced by a car insurance contract. We suppose that one department will pay out property damage coverage (a fraction of the aggregate claims from the contract) while another department will pay out third-party liability costs (another fraction of the aggregate claims from the contract).

Mathematically, we let $W_t^1, \ldots, W_t^m$ be m independent spectrally positive Lévy processes having moment generating function(m.g.f) for $j = 1, \ldots, m$. Now, we let each X_t^i to be a linear combination of some, or all, of the $W_t^1, \ldots, W_t^m$, i.e.

$$X = \begin{pmatrix} X_t^1 \\ X_t^2 \\ \vdots \\ X_t^n \end{pmatrix} = \begin{pmatrix} a_{11} & \ldots & a_{1m} \\ a_{21} & \ldots & a_{2m} \\ \vdots & \ldots & \vdots \\ a_{n1} & \ldots & a_{nm} \end{pmatrix} \begin{pmatrix} W_t^1 \\ W_t^2 \\ \vdots \\ W_t^m \end{pmatrix}, \tag{2.3.10}$$

where a_{ij}'s are non-negative real numbers for $1 \leq i \leq n$ and $1 \leq j \leq m$.

We point out that we chose this structure because it admits a neat solution for the capital allocation problem through Theorem 2.3.3. One can always fall back on the more simple case where each department pays out one, and only one, type of claims as oppose to paying fractions of different types of claims. This would correspond to having $n = m$ and a diagonal matrix in (2.3.10) with all elements in the diagonal equal to one yielding $X_t^i = W_t^i$ for all i. We also point out that this construction endows the processes R^i's with a dependence structure through the aggregate claims X^i's. The next result is one of the main contribution of our chapter.

Theorem 2.3.4. *Consider n risk processes such that $(R_t^i)_{0 \leq t \leq T} \in \mathcal{R}_L^p$, for $i = 1, \ldots, n$. Now, let such $R_t^i = x^i + c^i t - X_t^i$ where the aggregate risk processes X_t^i be those defined in (2.3.10). Then the capital allocation that satisfies Definition 2.3.4 over the time period $[0, T]$, for each net-gross process $c^i t - X_t^i$ and with respect to the risk measure CEVaR$_\beta$ in (2.2.7) is*

$$L^i = \frac{1}{T} \int_0^T K_t^i dt - c^i \frac{T}{2}, \tag{2.3.11}$$

where

$$K_t^i = t \sum_{j=1}^m a_{ij} \phi_j'(s^* \sum_{k=1}^n a_{kj}), \qquad t \in [0, T], \tag{2.3.12}$$

and $\mathbb{E}(e^{sW_1^j}) = e^{\phi_j(s)}$ for $s \geq 0$, $\phi_j'(0) \geq 0$, $1 \leq i \leq n$ and $1 \leq j \leq m$.

Proof. First we want to find the capital allocation with respect to the risk measure EVaR_β before applying Theorem 2.3.3. For any coherent risk measure ρ defined on $L^\infty(\Omega, \mathcal{F})$, we have, by the cash-invariant property, that, for each $t \in [0, T]$

$$\rho(\sum_{i=1}^n Y_t^i) = \rho(\sum_{i=1}^n -X_t^i) - \sum_{i=1}^n c^i\, t\,.$$

That is, in order to find the capital allocation (at $t \in [0, T]$) in this setting with respect to a coherent risk measure (in particular for EVaR_β), we just need to find the capital allocation for each claim process X_t^i.

For a given coherent risk measure ρ on $L^{\infty}(\Omega,\mathcal{F})$, let us define the function $f^t_\rho(u_1,\ldots,u_n) := \rho(\sum_{i=1}^n -u_i X^i_t)$ for $t\in[0,T]$. Taking into account the structure of the processes $X^1_t,\ldots,X^n_t$, we can write, for $t\in[0,T]$

$$\begin{aligned} f^t_\rho(u_1,\ldots,u_n) &= \rho(\sum_{i=1}^n -u_i X^i_t) \\ &= \rho\left(-\left((\sum_{j=1}^m u_1 a_{1j} W^j_t) + (\sum_{j=1}^m u_2 a_{2j} W^j_t) + \cdots + (\sum_{j=1}^m u_n a_{nj} W^j_t)\right)\right) \\ &= \rho\left(-\left((\sum_{i=1}^n u_i a_{i1}) W^1_t + (\sum_{i=1}^n u_i a_{i2}) W^2_t + \cdots + (\sum_{i=1}^n u_i a_{in}) W^m_t\right)\right). \end{aligned}$$

If we let

$$d_j = \sum_{k=1}^n u_k a_{kj}, \tag{2.3.13}$$

we can write a more compact form

$$f^t_\rho(u_1,u_2,\ldots,u_n) = \rho\left(-(d_1 W^1_t + d_2 W^2_t + \cdots + d_m W^m_t)\right) . \tag{2.3.14}$$

By using the independence of principal factors W^i, we have, for $t\in[0,T]$

$$\begin{aligned} \ln\left(\mathbb{E}(e^{s(d_1 W^1_t + d_2 W^2_t + \cdots + d_m W^m_t)})\right) &= \ln\left(\Pi_{j=1}^m \mathbb{E}(e^{s d_j W^j_t})\right) \\ \sum_{j=1}^m \ln\left(\mathbb{E}(e^{s d_j W^j_t})\right) &= t\sum_{j=1}^m \phi_j(s d_j) , \end{aligned}$$

where the last equality comes from $\mathbb{E}(e^{sW^j_t}) = e^{t\phi_j(s)}$.

If we specialize the above equations to the case of EVaR, then equation (2.3.14) becomes, for $t\in[0,T]$

$$f^t_{EVaR_\beta}(u_1,u_2,\ldots,u_n) = EVaR_\beta\left(-(d_1 W^1_t + d_2 W^2_t + \cdots + d_m W^m_t)\right) = \inf_{s\geq 0} \frac{t\sum_{j=1}^m \phi_j(s d_j) - \ln\beta}{s} . \tag{2.3.15}$$

Now, consider the right-hand side of equation (2.3.15). By taking derivatives with respect to s we have, for $t\in[0,T]$,

$$\frac{\partial}{\partial s}\left(\frac{t\sum_{j=1}^m \phi_j(s d_j) - \ln\beta}{s}\right) = \frac{st\sum_{j=1}^m d_j \phi'_j(s d_j) - t\sum_{j=1}^m \phi_j(s d_j) + \ln\beta}{s^2}. \tag{2.3.16}$$

By setting equation (2.3.16) equal to zero, we can find the value $s^*(t,u^1,\ldots,u^n)$ that minimizes the right-hand side in (2.3.15). As indicated by the notation, this minimum value $s^*(t,u^1,\ldots,u^n)$ is a function of t and u_i for $1\leq i\leq n$ but in the following we use the more simple notation s^* for this value. Notice that the value s^* is in fact the infimum too. Based on convexity property of Laplace transform for one-sided Lévy processes and the condition $\phi'_j(0)\geq 0$, the infimum in (2.3.15) should be reached at some point we denote s^* (see [63]).

According to Theorem 2.3.3, the *Euler* allocation is the only possible allocation method for EVaR_β. So, in order to find the capital allocation, it is sufficient to find the derivative of the right-hand side of equation (2.3.15) with respect to the variable u_i and evaluate it at the point $u = (1, 1, \dots, 1)$. Straight-forward differentiation yields, for $i = 1, \dots, n$ and $t \in [0, T]$

$$\frac{\partial}{\partial u_i} f^t_{EVaR_\beta}(u_1, u_2, \dots, u_n) = \frac{s^* t \sum_{j=1}^m (s_i^* d_j + a_{ij} s^*)\phi_j'(s^* d_j) - t s_i^* \sum_{j=1}^m \phi_j(s^* d_j) + s_i^* \ln \beta}{s^{*2}} \ , \tag{2.3.17}$$

where we use the notation $s_i^* = \frac{\partial s^*}{\partial u_i}$.

Since s^* is the solution of setting equation (2.3.16) equal to zero, we can simplify (2.3.17) as follows, for $i = 1, \dots, n$.

$$\frac{\partial}{\partial u_i} f^t_{EVaR_\beta}(u_1, u_2, \dots, u_n) = t \sum_{j=1}^m a_{ij} \phi_j'(s^* d_j) \ . \tag{2.3.18}$$

Evaluating equation (2.3.18) at the point $u = (1, 1, \dots, 1)$ yields the allocated capital associated to the i^{th} department at time $t \in [0, T]$. Namely, for $i = 1, \dots, n$

$$K_t^i = \frac{\partial}{\partial u_i} f^t_{EVaR_\beta}(u_1, u_2, \dots, u_n) = t \sum_{j=1}^m a_{ij} \phi_j'(s^* \sum_{k=1}^n a_{kj}) \ . \tag{2.3.19}$$

Using Corollary 2.3.2 and integrating K_t^i in (2.3.19) yields the allocated capital satisfying Definition 2.3.4 with respect to the risk measure CEVaR_β. Thus, the allocated capital to i^{th} department over the period $[0, T]$ with respect to CEVaR_β is

$$L^i = \frac{1}{T} \int_0^T K_t^i dt - c^i \frac{T}{2} \ .$$

This completes the proof. □

2.4. Examples

In this section, we are interested in examining Theorem 2.3.4 for some examples in order to illustrate how this capital allocation can be computed. We present capital allocations for the examples already discussed in Subsection 2.2.1.

As we will see, there are some cases for which we can obtain an explicit expression for the capital allocation. In others, such an explicit form is not available but a solution can still be obtained by standard numerical methods. The difficulty lies in solving equation (2.3.16) when is set to be equal to zero.

2.4.1. Brownian Motion with Scale Parameter

Consider the general set-up defined through equation (2.3.10). Let the principal factors $W_t^1, \dots, W_t^m$ to be m independent Brownian motions with different scale parameters $\sigma_i > 0$ and m.g.f $\mathbb{E}(e^{sW_t^i}) = e^{\frac{1}{2}\sigma_i^2 s^2 t}$. We now only need to apply

Theorem 2.3.4. By solving equation (2.3.16) equal to zero we get, for $t \in [0, T]$

$$s^2 t \sum_{j=1}^{m} d_j^2 \sigma_j^2 - \frac{1}{2} s^2 t \sum_{j=1}^{m} d_j^2 \sigma_j^2 + \ln \beta = 0, \tag{2.4.1}$$

where d_j is given in (2.3.13). Or equivalently,

$$s^* = \left(\frac{-2 \ln \beta}{t \sum_{j=1}^{m} d_j^2 \sigma_j^2} \right)^{\frac{1}{2}} . \tag{2.4.2}$$

Substituting (2.4.2) into equation (2.3.19) at the point $u = (1, 1, \ldots, 1)$ we can compute the value K_t^i for $i = 1, \ldots, n$. That is,

$$K_t^i = t^{\frac{1}{2}} \left(\frac{-2 \ln \beta}{\sum_{j=1}^{m} \sigma_j^2 (\sum_{k=1}^{n} a_{kj})^2} \right)^{\frac{1}{2}} \sum_{j=1}^{m} \sigma_j^2 a_{ij} \sum_{k=1}^{n} a_{kj} \ , \tag{2.4.3}$$

for $t \in [0, T]$. Thus, the allocated capital to the i^{th} department with respect to CEVaR_β can be computed to be

$$L^i = \frac{2}{3} T^{\frac{1}{2}} \left(\frac{-2 \ln \beta}{\sum_{j=1}^{m} \sigma_j^2 (\sum_{k=1}^{n} a_{kj})^2} \right)^{\frac{1}{2}} \sum_{j=1}^{m} \sigma_j^2 a_{ij} \sum_{k=1}^{n} a_{kj} - c^i \frac{T}{2}. \tag{2.4.4}$$

Now as a special case, let the principal factors $W_t^1, \ldots, W_t^m$ to be m independent Brownian motions with common scale parameter $\sigma > 0$ and common Laplace transform $\mathbb{E}(e^{sW_t^i}) = e^{\frac{1}{2}\sigma^2 s^2 t}$. So, (2.4.2) will be reduced to

$$s^* = \left(\frac{-2 \ln \beta}{\sigma^2 t \sum_{j=1}^{m} d_j^2} \right)^{\frac{1}{2}} , \tag{2.4.5}$$

and the value K_t^i is then, for $t \in [0, T]$

$$K_t^i = \sigma t^{\frac{1}{2}} \left(\frac{-2 \ln \beta}{\sum_{j=1}^{m} (\sum_{k=1}^{n} a_{kj})^2} \right)^{\frac{1}{2}} \sum_{j=1}^{m} a_{ij} \sum_{k=1}^{n} a_{kj} \ . \tag{2.4.6}$$

Thus, the allocated capital, L^i, to the i^{th} department satisfying Definition 2.3.4 with respect to CEVaR_β for this special case can be written as

$$L^i = \frac{2}{3} T^{\frac{1}{2}} \sigma \left(\frac{-2 \ln \beta}{\sum_{j=1}^{m} (\sum_{k=1}^{n} a_{kj})^2} \right)^{\frac{1}{2}} \sum_{j=1}^{m} a_{ij} \sum_{k=1}^{n} a_{kj} - c^i \frac{T}{2}. \tag{2.4.7}$$

2.4.2. Cramér- Lundberg Process

Consider again the general set-up defined through equation (2.3.10). We let the principal factors $W_t^1, \ldots, W_t^m$ to be m independent compound Poisson processes with different jump means $\frac{1}{\mu_i} > 0$. i.e.,

$$W_t^i = \sum_{k=1}^{N_t^i} Z_k^i, \tag{2.4.8}$$

where the number of claims is assumed to follow a Poisson process $(N_t^i)_{0 \le t \le T}$ with intensity λ_i which is independent of the positive and i.i.d random variables $(Z_n^i)_{n \ge 1}$ representing claim sizes. The m.g.f for (2.4.8) is

$$\mathbb{E}(e^{sW_t^i}) = \exp\left[\lambda_i(\psi_i(s) - 1)\right], \tag{2.4.9}$$

where $\psi_i(s)$ is the m.g.f for claim process for $i = 1, \ldots, m$. By solving equation (2.3.16) equal to zero we get, for $t \in [0, T]$

$$t \sum_{j=1}^{m} \lambda_j \left(sd_j \psi_j'(sd_j) - \psi_j(sd_j) + 1 \right) + \ln \beta = 0, \tag{2.4.10}$$

where d_j is given in (2.3.13). Let s^* be a solution of (2.4.10) satisfying (2.3.15) which also is a function of t. Evaluating s^* at the point $u = (1, 1, \ldots, 1)$ and substituting into (2.3.19) yields the capital allocation value K_t^i for $i = 1, \ldots, n$. That is,

$$K_t^i = t \sum_{j=1}^{m} a_{ij} \lambda_j \psi_j'(s^* \sum_{k=1}^{n} a_{kj}), \tag{2.4.11}$$

for $t \in [0, T]$. Thus, the allocated capital to the i^{th} department satisfying Definition 2.3.4 with respect to CEVaR_β is given by

$$L^i = \frac{1}{T} \int_0^T K_t^i dt - c^i \frac{T}{2}, \tag{2.4.12}$$

for $i = 1, \ldots, n$. As a special case assume that claim sizes follow exponential distributions with mean $\frac{1}{\mu_i}$ for $i = 1, \ldots, m$. In this case we have $\psi^i(s) = \frac{\mu_i}{\mu_i - s}$ for $s < \mu_i$. Therefore, equation (2.4.11) will be reduced to the following equation for $t \in [0, T]$.

$$st \sum_{j=1}^{m} d_j \lambda_j \left(\frac{\mu_j}{(\mu_j - sdj)^2} - \frac{1}{\mu_j - sdj} \right) + \ln \beta = 0\,, \quad s < \min \left(\frac{\mu_j}{d_j} \right)_{1 \le j \le m}, \tag{2.4.13}$$

where d_j is given in (2.3.13). This is not as straight-forward as the equivalent equation for the previous example. Nonetheless, the value s^* satisfying (2.4.13) and (2.3.15) can be obtained numerically. Evaluating at the point $u =$

$(1, 1, \ldots, 1)$ and substituting into (2.3.19) yields the capital allocation value K_t^i for $i = 1, \ldots, n$. That is,

$$K_t^i = t \sum_{j=1}^{m} a_{ij} \left(\frac{\lambda_j \mu_j}{(\mu_j - s^* \sum_{k=1}^{n} a_{kj})^2} \right), \tag{2.4.14}$$

for $t \in [0, T]$ and where s^* is the solution of equation (2.4.13). Thus, the allocated capital to the i^{th} department satisfying Definition 2.3.4 with respect to CEVaR_β is given by

$$L^i = \frac{1}{T} \int_0^T K_t^i dt - c^i \frac{T}{2}, \tag{2.4.15}$$

for $1 \leq i \leq n$.

2.4.3. Gamma Subordinator

Assume the general set-up defined through equation (2.3.10). We let the principal factors $W_t^1, \ldots, W_t^m$ to be m independent gamma processes with different parameters $\alpha_i, b_i > 0$ and m.g.f

$$\mathbb{E}(e^{sW_t^i}) = \left(1 - \frac{s}{b_i}\right)^{-\alpha_i t} = \exp\left[-t\alpha_i \ln\left(1 - \frac{s}{b_i}\right)\right], \qquad s < b_i . \tag{2.4.16}$$

We now only need to apply Theorem 2.3.4. By solving equation (2.3.16) equal to zero we get, for $t \in [0, T]$

$$t \sum_{j=1}^{m} \alpha_j \left(\ln(1 - \frac{sd_j}{b_j}) + s\frac{d_j}{b_j - sd_j} \right) + \ln\beta = 0 , \quad s < \min\left(\frac{b_j}{d_j}\right)_{1 \leq j \leq m} , \tag{2.4.17}$$

where d_j is given in (2.3.13). This equation like equation (2.4.13) is not as straight-forward as the equivalent equation for the example with Brownian motion. Nonetheless, the value s^* satisfying (2.4.17) and (2.3.15) can be obtained numerically. Evaluating at the point $u = (1, 1, \ldots, 1)$ and substituting into (2.3.19) yields the capital allocation value K_t^i for $i = 1, \ldots, n$. That is,

$$K_t^i = t \sum_{j=1}^{m} a_{ij} \left(\frac{\alpha_j}{b_j - s^* \sum_{k=1}^{n} a_{kj}} \right) , \tag{2.4.18}$$

for $t \in [0, T]$ and where s^* is the solution of equation (2.4.17). Thus, the allocated capital to the i^{th} department satisfying Definition 2.3.4 with respect to CEVaR_β is given by

$$L^i = \frac{1}{T} \int_0^T K_t^i dt - c^i \frac{T}{2} , \tag{2.4.19}$$

for $1 \leq i \leq n$.

Chapter 3

OPTIMAL PORTFOLIO PROBLEM USING ENTROPIC VALUE AT RISK: WHEN THE UNDERLYING DISTRIBUTION IS NON-ELLIPTICAL

Abstract

This chapter is devoted to study the optimal portfolio problem. Harry Markowitz's Ph.D. thesis prepared the ground for the mathematical theory of finance [70]. In modern portfolio theory, we typically find asset returns that are modeled by a random variable with an elliptical distribution and the notion of portfolio risk is described by an appropriate risk measure. In this chapter, we propose new stochastic models for the asset returns that are based on *Jumps-Diffusion* **(J-D)** distributions [76, 80]. This family of distributions is more compatible with stylized features of asset returns. On the other hand, in the past decades, we find attempts in the literature to use well-known risk measures, such as *Value at Risk* and *Expected Shortfall*, in this context. Unfortunately, one drawback with these previous approaches is that no explicit formulas are available and numerical approximations are used to solve the optimization problem. In this chapter, we propose to use a new coherent risk measure, so-called, *Entropic Value at Risk(EVaR)* [2], in the optimization problem. For certain models, including a jump-diffusion distribution, this risk measure yields an explicit formula for the objective function so that the optimization problem can be solved without resorting to numerical approximations.

This chapter is a joint research work with Andrew Luong; see [74] [1].

[1]The contribution I have made to this project was in different ways. My first contribution was to define the main problem of the project and how to solve it. After discussing with Prof. Luong, we jointly identify the framework and body of the project, then I proposed to apply Entropic Value at Risk (EVaR) as a risk measure in modern portfolio problem. I also proposed to use different Jump-Diffusion processes to model returns in modern portfolio problem. I prepared the mathematical ground for the project by recalling required theorems and definitions in the

3.1. Introduction

The problem of optimal portfolio which is nowadays introduced in a new framework, called *Modern Portfolio Theory (MPT)*, has been extensively studied in the past decades. The MPT is one of the most important problems in financial mathematics. Harry Markowitz [70] introduced a new approach to the problem of optimal portfolio so called *Mean-Variance* analysis. He chose a preferred portfolio by taking into account the following two criteria. The expected portfolio return and the variance of the portfolio return. In fact, Markowitz preferred one portfolio to another one if it has higher expected return and lower variance.

Later, we find attempts in the literature to replace variance with well-known risk measures, such as *Value at Risk* and *Expected Shortfall*. For instance, Embrechts et al.[39] have shown that replacing mean-variance with any other risk measure having the translation invariant and positively homogeneous properties under elliptical distributions yields to the same optimal solution. Basak and Shapiro [13] studied an alternative version of Markowitz problem by applying VaR for controlling the incurred risk in an expected utility maximization framework which allows to maximize the profit of the risk takers. Studying the Markowitz model has been done in the same framework by considering the CVaR as risk measure [78]. Later, Acerbi and Simonetti [1] studied the same problem as the one studied in [13] with spectral risk measures. Recently, Cahuich and Hernandez [21] solved the same problem within the framework of utility maximization using the class of distortion risk measures [87].

There are both practical and theoretical weaknesses that can be made about the relevant framework of optimal portfolio problem in the literature. One of such criticisms relates to the asset returns model itself. In fact, elliptical distribution is the most and relevant distribution which is used to model asset returns in MPT. One of the reason for choosing this distribution ties with the tractability of this class of distribution. But, in practice financial returns do not follow an elliptical distribution (see [29]). A second objection focuses in the choice of a measure of risk for the portfolio. Unfortunately, one drawback with the previous works, for instance [13, 78], is that no explicit formulas are available and numerical approximations are used to solve the optimization problem. The stochastic models which we are proposing for the asset returns in this chapter are based on *Jumps-Diffusion* **(J-D)** distributions [76, 80]. This family of distributions is more compatible with stylized features of asset returns and also allows for a straight-forward statistical inference from readily available data. We also tackle

context. I also contributed in this project by writing up the first draft of it, then polishing it with Prof. Luong to get the final version.

the second issue by choosing a suitable (coherent) risk measure as our objective function. In this chapter, we propose to use a new coherent risk measure, so-called, *Entropic Value at Risk(EVaR)* [2, 46], in the optimization problem. As this risk measure is based on Laplace transform of asset returns, applying it to the jump-diffusion models yields an explicit formula for the objective function so that the optimization problem can be solved without using numerical approximations.

The organization of this chapter is as follows. In Section 2, we provide a summary of properties about coherent risk measures and *Entropic Value at Risk* measure. We also continue this section by presenting a typical representation of optimal portfolio problem where we minimize the risk of the portfolio for a given level of portfolio return. In Section 3, we introduce our two models to fit as asset returns and we apply them into the optimization problem. We also derive some distributional properties for these models and finish Section 3 by discussing about the KKT conditions and optimal solutions. In Section 4, we discuss about parameters estimation method which we have used in this chapter. We also provide a numerical example for three different stocks and analyze the efficient frontiers for EVaR, mean-variance and VaR for these three stocks. In this chapter we use optimization package in MATLAB to do the computations.

3.2. Preliminaries

3.2.1. Coherent Risk Measures

We are considering $L^\infty(\Omega, \mathcal{F}, \mathbb{P})$ as the set of all bounded random variables representing profit/loss for financial positions. The following definition is taken from [45].

Definition 3.2.1. . *A function $\rho : L^\infty(\Omega, \mathcal{F}, \mathbb{P}) \to \mathbb{R}$ is a Coherent Risk measure if*

1- $\rho(\lambda X + (1-\lambda)Y) \leq \lambda\rho(X) + (1-\lambda)\rho(Y)$ *for any* $X, Y \in L^\infty$ *and* $\lambda \in [0,1]$.*(Convexity)*

2- $\rho(\lambda X) = \lambda\rho(X)$ *for any* $X \in L^\infty$ *and* $\lambda > 0$.*(Positive Homogeneity)*

3- $\rho(X + m) = \rho(X) - m$ *for any* $X \in L^\infty$ *and* $m \in \mathbb{R}$.*(Translation Invariant)*

4- $\rho(Y) \leq \rho(X)$ *if* $X, Y \in L^\infty$ *and* $X \leq Y$.*(Decreasing)*

In this chapter, we propose to use the *Entropic Value at Risk* measure (EVaR_α) which is a coherent risk measure. Following [2] we now give a first definition.

Definition 3.2.2. *Let X be a random variable in $L^\infty(\Omega, \mathcal{F}, \mathbb{P})$ such that*

$$\mathbb{E}[\exp(-s\,X)] < \infty\,, \qquad s > 0\,.$$

Then the **Entropic Value at Risk***, denoted by* $EVaR_\alpha$*, is given by*

$$EVaR_\alpha(X) := \inf_{s>0} \frac{\ln \mathbb{E}[\exp(-s\,X)] - \ln \alpha}{s}, \tag{3.2.1}$$

For a given level $\alpha \in (0,1)$*.*

The following key result for EVaR_α can be found in [2, 46].

Theorem 3.2.1. *The risk measure* $EVaR_\alpha$ *from Definition 3.2.2 is a coherent risk measure. Moreover, for any* $X \in L^\infty(\Omega, \mathcal{F}, \mathbb{P})$ *having Laplace transform, its dual representation has the form*

$$EVaR_\alpha(X) = \sup_{f\in\mathcal{D}} \mathbb{E}(-fX)\ ,$$

where $\mathcal{D} = \{f \in L^1_+(\Omega, \mathcal{F}, \mathbb{P}) \mid \mathbb{E}[f\ln(f)] \le -\ln\alpha\}$ *and* $L^1_+(\Omega, \mathcal{F}, \mathbb{P})$ *is defined in* (1.3.7)*.*

For a comprehensive study on this risk measure we may refer to [2, 46].

3.2.2. Optimal Portfolio Problem

Consider a portfolio in a financial market with n different assets. Denote the assets returns by the vector $R = (R_1, \ldots, R_n)$ in which R_i shows the return of the i-th asset. The returns are random variables and their mean is denoted by $\mu = (\mu_1, \ldots, \mu_n)$ where μ_i is the the expected return of the i-th asset, $\mu_i = \mathbb{E}(R_i)$. Moreover, assume ρ as a risk measure. Then following [78], we have this definition.

Definition 3.2.3. *the optimal portfolio problem can be written mathematically as follows.*

$$\begin{aligned}
\min_{\omega} \quad & \rho(\sum_{i=1}^{n} \omega_i R_i) \\
\textit{subject to} \quad & \sum_{i=1}^{n} \omega_i \mu_i = \mu^*, \\
& \sum_{i=1}^{n} \omega_i = 1, \\
& \omega_i \ge 0,
\end{aligned} \tag{3.2.2}$$

where μ^* *is a given level of return.*

Applying various risk measures along with different models for random returns yields to interesting problems in both theoretical and practical point of views. For instance, the classical mean-variance model introduced by Markowitz [70] is a special case of the model introduced in Definition 3.2.3. In fact, Markowitz used variance as a risk measure and apply it into the objective function given in (3.2.2) and he also considered returns from the portfolio are normally distributed.

Remark 3.2.1. *It has been shown in [39] that if we assume the return variables follow elliptical distributions(like multivariate normal distribution), then the solution for the Markowitz mean-variance problem will be the same as the optimal solution for optimal portfolio problem* (3.2.2) *by minimizing any other risk measure having the translation invariant and positively homogeneous properties for a given level of return.*

[54] has shown in his PhD thesis that for two different examples of elliptical distributions(normal and Student t) the portfolio decomposition for Expected Shortfall and Value at Risk are the same as the one for standard deviation.

Referring to Remark 3.2.1 we see that if the underlying distribution is elliptical, then for any coherent risk measure the optimal solution for the problem in (3.2.2) is the same as the optimal solution for the classical model by Markowitz.

3.3. Set up of the Models

In this section, we propose two multivariate models which do not follow elliptical distributions. These models which are based on jump-diffusion distributions can be fitted as the underlying models for returns. Distributional properties of these models will be also studied.

3.3.1. Non-Elliptical Multivariate Models 1,2

Multivariate Model 1. Consider the following multivariate model:

$$R = X + H + \sum_{k=1}^{M} W_k, \tag{3.3.1}$$

where R, X, H, W_k are n-variate vectors such that

$$\begin{aligned} R &= (R_1, \ldots, R_n), \\ X &= (X_1, \ldots, X_n), \\ W_k &= (W_{k1}, \ldots, W_{kn}), \\ H &= \left(\sum_{k=1}^{N_1} Y_{k1}, \ldots, \sum_{k=1}^{N_n} Y_{kn} \right). \end{aligned}$$

Here, X_i follows the normal distribution with $X_i \sim N(\tilde{\mu}_i, \sigma^2)$ and X_i's are mutual independent for $i = 1, \ldots, n$. $W_k = (W_{k1}, \ldots, W_{kn})$ is assumed to follow the multivariate normal distribution with $W_k \sim N(\mu, A)$ for each k where

$\mu = (\mu_1, \ldots, \mu_n)$ is mean and A is $n \times n$ covariance matrix. Moreover, W_k's are assumed to be mutually independent. The random variable M follows the Poisson distribution with intensity γ and is independent of W_k for each k. N_k are assumed to have Poisson distribution with intensity λ_k and mutually independent for $k = 1, \ldots, n$. The Y_{ki} are assumed to be mutually independent for all k and all $i = 1, \ldots, n$ and Y_{ki} is normal distributed with $Y_{ki} \sim N(\theta_i, \sigma_i^2)$. Finally, N_k and Y_{kn} are mutually independent as well as $X, H, \sum_{k=1}^{M} W_k$.

This model can be driven from a jump-diffusion model which is the solution for a stochastic differential equation [76]. We can rewrite this multivariate model as follows.

$$\begin{aligned}
R_1 &= X_1 + \sum_{k=1}^{N_1} Y_{k1} + \sum_{k=1}^{M} W_{k1}, \\
R_2 &= X_2 + \sum_{k=1}^{N_2} Y_{k2} + \sum_{k=1}^{M} W_{k2}, \\
&\vdots \\
R_n &= X_n + \sum_{k=1}^{N_n} Y_{kn} + \sum_{k=1}^{M} W_{kn}.
\end{aligned}$$

Multivariate Model 2. The model (6.5) in [80] prepared the ground to introduce another non-elliptical multivariate model which can be fitted for portfolio returns. This proposed model is given as follows.

$$R = X + \sum_{k=1}^{M} W_k. \tag{3.3.2}$$

Here, R, X, W_k are n-variate vectors such that

$$\begin{aligned}
R &= (R_1, \ldots, R_n), \\
X &= (X_1, \ldots, X_n), \\
W_k &= (W_{k1}, \ldots, W_{kn}),
\end{aligned}$$

where $X = (X_1, \ldots, X_n)$ follows the multivariate normal distribution with $X \sim N(\tilde{\mu}, Q)$ where $\tilde{\mu}$ is mean and Q is $n \times n$ covariance matrix. $W_k = (W_{k1}, \ldots, W_{kn})$ is assumed to follow the multivariate normal distribution with $W_k \sim N(\mu, A)$ for each k where $\mu = (\mu_1, \ldots, \mu_n)$ is mean and A is $n \times n$ covariance matrix. Moreover, W_k's are assumed to be mutually independent. The random variable M follows the Poisson distribution with intensity λ and is independent of W_k for each k. Also, $X, \sum_{k=1}^{M} W_k$ are mutually independent.

Like the model (3.3.1) introduced in Subsection 3.1 we can rewrite the multivariate model (3.3.2) as

$$\begin{aligned}
R_1 &= X_1 + \sum_{k=1}^{M} W_{k1},\\
R_2 &= X_2 + \sum_{k=1}^{M} W_{k2},\\
&\vdots\\
R_n &= X_n + \sum_{k=1}^{M} W_{kn}.
\end{aligned}$$

3.3.2. Distributional Properties of the Multivariate Models 1, 2

Consider the multivariate models (3.3.1) and (3.3.2). As these models are given in terms of summation of multivariate normal and compound Poisson distributions we can provide the joint density functions for each of these models. [77] gives the following presentation for the density function of model (3.3.1) and also provides a proof but we give a proof here for the sake of completeness.

Proposition 3.3.1. *Consider the model* (3.3.1)*. Then the joint density functions of the vector R is given by*

$$f_R(r) = \sum_{k_1=0}^{\infty} \cdots \sum_{k_n=0}^{\infty} \sum_{m=0}^{\infty} \left(\frac{e^{-\lambda_1}\lambda_1^{k_1}}{k_1!}\right) \cdots \left(\frac{e^{-\lambda_n}\lambda_n^{k_n}}{k_n!}\right) \left(\frac{e^{-\gamma}\gamma^{m}}{m!}\right) \frac{e^{-\frac{1}{2}(r-u)T^{-1}(r-u)'}}{(2\pi)^{\frac{n}{2}}|T|^{\frac{1}{2}}}, \tag{3.3.3}$$

where $r = (r_1, \ldots, r_n) \in \mathbb{R}^n$, $u = (\tilde{\mu_1} + k_1\theta_1, \ldots, \tilde{\mu_n} + k_n\theta_n) + m\mu$ *and* $T = mA + diag(\sigma^2 + k_1\sigma_1^2, \ldots, \sigma^2 + k_n\sigma_n^2)$.

Proof. The idea we put forward to prove this proposition is using conditional density function. Since the X_i are mutual independent with normal distribution so the vector $X = (X_1, \ldots, X_n)$ follows a multivariate normal distribution with mean $(\tilde{\mu_1}, \ldots, \tilde{\mu_n})$ and covariance matrix $\sigma^2 I_n$ where I_n is the identity matrix of order n. Moreover by conditioning on each of N_i and using independency between

Y_{ji} we obtain

$$\mathcal{L}\left(\sum_{j=1}^{N_i} Y_{ji}|N_i = k_i\right) = N(k_i\theta_i, k_i\sigma_i^2), \tag{3.3.4}$$

for each $1 \le i \le n$. Thus, independency between N_i and Y_{ji} for all i and j yields

$$\mathcal{L}\left(H|N_1 = k_1, \dots, N_n = k_n\right) = N\left((k_1\theta_1 + \cdots + k_n\theta_n), diag(k_1\sigma_1^2, \dots, k_n\sigma_n^2)\right). \tag{3.3.5}$$

Conditioning on the random variable M and using the independency between W_i and M gives the following conditional distribution.

$$\mathcal{L}\left(\sum_{i=1}^{M} W_i|M = m\right) = N(m\mu, mA). \tag{3.3.6}$$

Putting (3.3.5) and (3.3.6) together and using independency between X, H and $\sum_{i=1}^{M} W_i$ provide the conditional distribution of R given $N_1 = k_1, \dots, N_n = k_n, M = m$. i.e.,

$$\mathcal{L}\left(R|N_1 = k_1, \dots, N_n = k_n, M = m\right) = N(u, T). \tag{3.3.7}$$

(3.3.7) gives the conditional density of R given $N_1 = k_1, \dots, N_n = k_n, M = m$. To get the density function of R we need to multiply the conditional density by the probability functions associated to each N_i and M and add them up. This completes the proof. □

If we follow the same procedure done for Proposition 3.3.1 and apply it for the model (3.3.2) we can obtain the density function for the vector R.

Remark 3.3.1. *The density function for the model* (3.3.2) *is*

$$f_R(r) = \sum_{m=0}^{\infty}\left(\frac{e^{-\lambda}\lambda^m}{m!}\right)\frac{e^{-\frac{1}{2}(r-u)T^{-1}(r-u)'}}{(2\pi)^{\frac{n}{2}}|T|^{\frac{1}{2}}}, \tag{3.3.8}$$

where $r = (r_1, \dots, r_n) \in \mathbb{R}^n$, $u = \tilde{\mu} + m\mu$ *and* $T = Q + mA$.

In the sequel of this part we provide the Laplace exponents for both models (3.3.1) and (3.3.2).

Lemma 3.3.1. *Consider the multivariate model* (3.3.1)*. Then the Laplace exponent for the vector* $R = (R_1, \dots, R_n)$ *at* $u = (u_1, \dots, u_n)$ *is*

$$\log \mathbb{E}(e^{-uR}) = -u\tilde{\mu}^t + \frac{\sigma^2}{2}uu^t + \gamma(e^{-u\mu^t + uAu^t} - 1) + \sum_{k=1}^{n}\lambda_k(e^{-\theta_k u_k + \frac{\sigma_k^2 u_k^2}{2}} - 1), \tag{3.3.9}$$

where u^t, μ^t *and* $\tilde{\mu}^t$ *are the column vectors associated to the row vectors* u, μ *and* $\tilde{\mu} = (\tilde{\mu_1}, \dots, \tilde{\mu_n})$ *respectively.*

Proof. The independency between X, H and $\sum_{k=1}^{M} W_k$ and using this point that Laplace transform for normal and compound Poisson distributions exists, yield the result. $\square$

Lemma 3.3.2. *Consider the multivariate model (3.3.2). Then the Laplace exponent for the vector $R = (R_1, \ldots, R_n)$ at $u = (u_1, \ldots, u_n)$ is*

$$\log \mathbb{E}(e^{-uR}) = -u\tilde{\mu}^t + uQu^t + \lambda(e^{-u\mu^t + uAu^t} - 1). \tag{3.3.10}$$

Proof. The Laplace transform for Gaussian distributions and compound Poisson distributions exists. So, (3.3.10) can be driven by using the independency between X and $\sum_{k=1}^{M} W_k$. $\square$

Now, we apply EVaR_α along with the model proposed in (3.3.1) to the optimal portfolio problem (3.2.2). Thus (3.2.2) is written as follows.

$$\begin{aligned}
\min_{\omega,s} \quad & \Bigg\{ \sum_{k=1}^{n} (-\tilde{\mu}_k \omega_k + \frac{\sigma^2}{2} s\omega_k^2) + \frac{\gamma(e^{-s\sum_{k=1}^{n} \mu_k\omega_k + s^2\omega A\omega^t} - 1)}{s} \\
& \quad + \frac{\sum_{k=1}^{n} \lambda_k (e^{-\theta_k\omega_k s + \frac{s^2\sigma_k^2\omega_k^2}{2}} - 1) - \ln\alpha}{s} \Bigg\} \\
\text{subject to} \quad & \sum_{i=1}^{n} (\tilde{\mu}_i + \lambda_i\theta_i + \mu_i\gamma)\omega_i = \mu^*, \\
& \sum_{i=1}^{n} \omega_i = 1, \\
& \omega_i \geq 0, s \geq 0.
\end{aligned} \tag{3.3.11}$$

Applying EVaR_α and the model (3.3.2) into the optimal portfolio problem (3.2.2) yield

$$\begin{aligned}
\min_{\omega,s} \quad & \sum_{k=1}^{n} -\tilde{\mu}_k\omega_k + s\omega Q\omega^t + \frac{\lambda(e^{-s\sum_{k=1}^{n} \mu_k\omega_k + s^2\omega A\omega^t} - 1) - \ln\alpha}{s} \\
\text{subject to} \quad & \sum_{i=1}^{n} (\tilde{\mu}_i + \mu_i\lambda)\omega_i = \mu^*, \\
& \sum_{i=1}^{n} \omega_i = 1, \\
& \omega_i \geq 0, s \geq 0.
\end{aligned} \tag{3.3.12}$$

3.3.3. Necessary and Sufficient Conditions for Optimal Problems, KKT Conditions

In this section we would like to identify the necessary and sufficient conditions for optimality of problems (3.3.11) and (3.3.12). In fact, we want to examine the Karush-Kuhn-Tucker (KKT) conditions for these problems and study whether the constrained problems in the last two sections have optimal solutions. Being the objective functions for both problems (3.3.11) and (3.3.12) smooth enough (they are continuously differentiable functions), will help us to verify the KKT conditions much easier.

3.3.3.1. *KKT Conditions for Optimal Problem with the multivariate model 1*

The KKT conditions provide necessary conditions for a point to be optimal point for a constrained nonlinear optimal problem. We refer to Chapter 5 page 241 [14] for a comprehensive study of KKT conditions for nonlinear optimal problems. Here, we study these conditions for the model (3.3.11) by using the same notation used in page 200 [14]. We rewrite problem (3.3.11) as follows.

$$
\begin{aligned}
\min_{\omega,s} \quad & f(\omega,s) = \Bigg\{ \sum_{k=1}^{n}(-\tilde{\mu}_k\omega_k + \frac{\sigma^2}{2}s\omega_k^2) + \frac{\gamma(e^{-s\sum_{k=1}^{n}\mu_k\omega_k + s^2\omega A\omega^t} - 1)}{s} \\
& \qquad + \frac{\sum_{k=1}^{n}\lambda_k(e^{-\theta_k\omega_k s + \frac{s^2\sigma_k^2\omega_k^2}{2}} - 1) - \ln\alpha}{s} \Bigg\} \\
\text{subject to} \quad & h_1(\omega,s) = \sum_{i=1}^{n}(\tilde{\mu}_i + \lambda_i\theta_i + \mu_i\gamma)\omega_i - \mu^* = 0, \\
& h_2(\omega,s) = \sum_{i=1}^{n}\omega_i - 1 = 0, \\
& g_i(\omega,s) = -\omega_i \le 0, \ \forall 1 \le i \le n, \\
& g_{n+1}(\omega,s) = -s \le 0. \qquad (3.3.13)
\end{aligned}
$$

Let $\omega_s = (\omega, s)$ be a *regular point*[2] for the problem (3.3.11). Then, the point ω_s is a local minimum of f subject to the constraints in (3.3.13) if there exists Lagrange multipliers $\nu_1, \ldots, \nu_{n+1}$ and η_1, η_2 for the Lagrangian function $L = f(\omega_s) + \sum_{k=1}^{n+1}\nu_k g_k(\omega_s) + \sum_{j=1}^{2}\eta_j h_j(\omega_s)$ such that the followings are true.

(1) $\frac{\partial L}{\partial \omega_i} = -\tilde{\mu}_i + s\sigma^2\omega_i + \gamma(-\mu_i + 2s^2(\omega A)_i)\frac{(e^{-s\sum_{k=1}^{n}\mu_k\omega_k + s^2\omega A\omega^t})}{s} + \lambda_i(-\theta_i + s^2\sigma^2\omega_i)\frac{e^{-\theta_k\omega_k s + \frac{s^2\sigma_k^2\omega_k^2}{2}}}{s} - \nu_i + (\tilde{\mu}_i + \lambda_i\theta_i + \mu_i\gamma)\eta_1 + \eta_2 = 0, \ i = 1, \ldots, n,$

[2]Let ω_s be a feasible point. Then, ω_s is said to be a regular point if the gradient vectors $\nabla g_i(\omega_s)$ for $i \in \{i : g_i(\omega_s) = 0, \ i = 1, \ldots, n+1\}$ are linearly independent.

(2) $\frac{\partial L}{\partial s} = \sum_{k=1}^{n} \frac{\sigma^2}{2}\omega_k^2 + \frac{(-\gamma\sum_{k=1}^{n}\mu_k\omega_k+\gamma s^2\omega A\omega^t-\gamma)(e^{-s\sum_{k=1}^{n}\mu_k\omega_k+s^2\omega A\omega^t})+\gamma}{s^2} + \frac{\sum_{k=1}^{n}\left(\left(\lambda_k(-\theta_k\omega_k s+s^2\sigma_k^2\omega_k^2-1)\right)\left(e^{-\theta_k\omega_k s+\frac{s^2\sigma_k^2\omega_k^2}{2}}\right)\right)+\sum_{k=1}^{n}\lambda_k+\ln\alpha}{s^2} = 0,$

(3) $\nu_k \geq 0,\ k = 1, \dots, n+1,$

(4) $\nu_k g_k = 0,\ k = 1, \dots, n+1,$

(5) $g_k \leq 0,\ k = 1, \dots, n+1,$ and $h_j = 0,\ j = 1, 2,$

where $(\omega A)_i$ is the ith entry of the row vector (ωA).

Remark 3.3.2. *Since, the functions h_1 and h_2 in (3.3.13) are linear and the functions g_i for $i = 1, \dots, n$ are convex, then by referring to Section 5.7 of [14] we see that the feasible region $\Omega = \{\omega_s : g_k(\omega_s) \leq 0,\ k = 1, \dots, n+1,$ and $h_j(\omega_s) = 0,\ j = 1, 2\}$ is a convex set. On the other hand, the risk measure $\rho = EvaR_{1-\alpha}$ is a convex function subject to the variables ω_i and s for all $i = 1, \dots, n$. We refer to [2] for a proof. Thus, the objective function f in problem (3.3.13) is convex too. We see that any* **local minimum** *for problem (3.3.13) is a* **global minimum** *too and the KKT conditions are also* **sufficient**. *See [14] page 212.*

3.3.3.2. *KKT Conditions for Optimal Problem with the multivariate model 2*

In this section we will provide the KKT conditions for the optimal problem (3.3.12). We show that these conditions are also sufficient for a solution to be an optimal one. First, we rewrite the problem (3.3.12) in the following way.

$$\begin{aligned}
\min_{\omega,s} \quad & f(\omega, s) = \sum_{k=1}^{n} -\tilde{\mu_k}\omega_k + s\omega Q\omega^t + \frac{\lambda(e^{-s\sum_{k=1}^{n}\mu_k\omega_k+s^2\omega A\omega^t} - 1) - \ln\alpha}{s} \\
\text{subject to} \quad & h_1(\omega, s) = \sum_{i=1}^{n}(\tilde{\mu_i} + \mu_i\lambda)\omega_i - \mu^* = 0, \\
& h_2(\omega, s) = \sum_{i=1}^{n}\omega_i - 1 = 0, \\
& g_i(\omega, s) = -\omega_i \leq 0,\ \forall 1 \leq i \leq n, \\
& g_{n+1}(\omega, s) = -s \leq 0.
\end{aligned} \tag{3.3.14}$$

By applying the same definition and notation used in the previous section we can provide the KKT conditions as follows.

(1) $\frac{\partial L}{\partial \omega_i} = -\tilde{\mu_i} + 2s(\omega Q)_i + \lambda(-s\mu_i + 2s^2(\omega A)_i)\frac{(e^{-s\sum_{k=1}^{n}\mu_k\omega_k+s^2\omega A\omega^t})}{s} = 0,\ i = 1, \dots, n,$

(2) $\frac{\partial L}{\partial s} = \omega A\omega^t + \frac{\lambda(-\sum_{k=1}^{n}\mu_k\omega_k+2s^2\omega A\omega^t-1)(e^{-s\sum_{k=1}^{n}\mu_k\omega_k+s^2\omega A\omega^t})+\lambda+\ln\alpha}{s^2} = 0,$

(3) $\nu_k \geq 0,\ k = 1, \ldots, n+1,$
(4) $\nu_k g_k = 0,\ k = 1, \ldots, n+1,$
(5) $g_k \leq 0,\ k = 1, \ldots, n+1,$ and $h_j = 0,\ j = 1, 2,$
where $(\omega A)_i$ and $(\omega Q)_i$ are the ith entry of the row vectors (ωA) and (ωQ) respectively.

Remark 3.3.3. *Since the feasible region $\Omega = \{\omega_s : g_k(\omega_s) \leq 0,\ k = 1, \ldots, n+1,$ and $h_j(\omega_s) = 0,\ j = 1, 2\}$ and the objective function for the optimal problem (3.3.14) are convex, so again by referring to [14] we can see that the KKT conditions are also* **sufficient** *and any* **local minimum** *for problem (3.3.14) is a* **global minimum** *as well.*

3.4. Efficient Frontier Analysis

In this section we study the optimization problem (3.2.2) for multivariate model 1 given in (3.3.1). In fact, we analyze the efficient frontier for this problem when the risk measures are EVaR and standard deviation. Our analysis shows that we have different portfolio decomposition corresponding to EVaR and standard deviation as the underlined model for returns is followed by a non-elliptical distribution (model 1). Thanks to the closed form for EVaR we can use optimization packages in mathematical software to solve the optimization problem (3.3.11) without using simulation techniques like Monte Carlo simulation.

3.4.1. Parameters Estimation

Studying the optimization problems (3.3.11) and (3.3.12) requires knowing the parameters of the multivariate models (3.3.1) and (3.3.2). To estimate these parameters we use a method of estimation for joint parameters so called *Extended Least Square(ELS)*[89]. In fact, assume that we are given a sample of n individuals. Let $y_i = [y_{i1} \ldots, y_{ip_i}]$ denote the i^{th} subject's $1 \times p_i$ vector of repeated measurements where the y_i are assumed to be independently distributed with mean and covariance matrices given by

$$\begin{aligned} \mathbb{E}(y_i) &= \bar{\mu}_{\mathbf{i}}(\beta) \\ Cov(y_i) &= G_i(\beta, \theta), \end{aligned} \tag{3.4.1}$$

where β and θ are vectors of unknown parameters which should be estimated. Extended Least Square(ELS) estimates are obtained by minimizing the following objective function.

$$f(\beta,\theta)=\sum_{i=1}^{n}\{(y_i-\bar{\mu}_i(\beta))G_i^{-1}(\beta,\theta)(y_i-\bar{\mu}_i(\beta))'+\ln|G_i(\beta,\theta)|\}, \tag{3.4.2}$$

where $\bar{\mu}_i(\beta)$ and $G_i(\beta,\theta)$ are defined in (3.4.1) and $|G_i|$ is the determinant of the positive definite covariance matrix G_i. Following [89] it can be seen that ESL is joint normal theory maximum likelihood estimation. In fact, minimizing (3.4.2) is equivalent to maximizing the log-likelihood function of the y_i when the y_i are independent and normally distributed with mean anc covariance matrices given by (3.4.1).

3.4.2. Data Sets

We construct the portfolio by choosing 3 stocks which are Intel Corp. (`http://finance.yahoo.com/q/hp?s=INTC&a=08&b=20&c=2010&d=07&e=26&f=2013&g=w`), Apple Inc. (`http://finance.yahoo.com/q/hp?a=08&b=20&c=2010&d=07&e=26&f=2013&g=w&s=AAPL%2C+&ql=1`) and Pfizer Inc. (PFE) (`http://finance.yahoo.com/q/hp?a=08&b=20&c=2010&d=07&e=26&f=2013&g=w&s=PFE%2C+&ql=1`). We use the close data ranged from 20/09/2010 to 26/08/2013. The weekly close data are converted to log return. i.e., if we consider P_n as the close price for the week n^{th} then log return is $R_n = \ln P_n - \ln P_{n-1}$.

Now consider the model (3.3.1). We try to apply this model to these three stocks and determine the parameters in (3.4.1) in order to solve the optimization problem (3.4.2). In this case we have $n = 153$, the number of our sample and y_i is a 1×3 row vector associated to the mean of returns. Then the vector $\bar{\mu}_i$ is

$$\bar{\mu_i}=(\tilde{\mu_1}+\lambda_1\theta_1+\mu_1\gamma,\tilde{\mu_2}+\lambda_2\theta_2+\mu_2\gamma,\tilde{\mu_3}+\lambda_3\theta_3+\mu_3\gamma)\,, \tag{3.4.3}$$

for all $1\le i\le 153$. Let $A=(a_{ij})_{3\times 3}$ be the covariance matrix for the multivariate normal distribution W_k. Then, the covariance matrix G_i in (3.4.1) has the following representation.

$$G_i=\begin{pmatrix}\sigma^2+\lambda_1(\theta_1^2+\sigma_1^2)+\gamma(a_{11}+\mu_1^2) & \gamma(a_{12}+\mu_1\mu_2) & \gamma(a_{13}+\mu_1\mu_3)\\ \gamma(a_{12}+\mu_1\mu_2) & \sigma^2+\lambda_2(\theta_2^2+\sigma_2^2)+\gamma(a_{22}+\mu_2^2) & \gamma(a_{23}+\mu_2\mu_3)\\ \gamma(a_{13}+\mu_1\mu_3) & \gamma(a_{23}+\mu_2\mu_3) & \sigma^2+\lambda_3(\theta_3^2+\sigma_3^2)+\gamma(a_{33}+\mu_3^2)\end{pmatrix}, \tag{3.4.4}$$

for all $1\le i\le 153$. Therefore, by plugging (3.4.3) and (3.4.4) into (3.4.2) we get the objective function for the ELS method. Doing the same procedure for the model (3.3.2) we can find the parameters in (3.4.1). Let $Q=(q_{ij})_{3\times 3}$ and $A=(a_{ij})_{3\times 3}$ be the covariance matrices for the multivariate normal distribution X and W_k respectively. Then we have

$$\bar{\mu_i}=(\tilde{\mu_1}+\mu_1\lambda,\tilde{\mu_2}+\mu_2\lambda,\tilde{\mu_3}+\mu_3\lambda)\,, \tag{3.4.5}$$

and

$$G_i = \begin{pmatrix} q_{11} + \lambda(a_{11} + \mu_1^2) & q_{12} + \lambda(a_{12} + \mu_1\mu_2) & q_{13} + \lambda(a_{13} + \mu_1\mu_3) \\ q_{12} + \lambda(a_{12} + \mu_1\mu_2) & q_{22} + \lambda(a_{22} + \mu_2^2) & q_{23} + \lambda(a_{23} + \mu_2\mu_3) \\ q_{13} + \lambda(a_{13} + \mu_1\mu_3) & q_{23} + \lambda(a_{23} + \mu_2\mu_3) & q_{33} + \lambda(a_{33} + \mu_3^2) \end{pmatrix}, \tag{3.4.6}$$

for all $1 \leq i \leq 153$.

In the following we provide the results for the portfolio decomposition corresponding to the three stocks, $\mathrm{EVaR}_{95\%}$ and standard deviation. This results have been driven for the model 1 given in (3.3.1). In order to estimate our parameters for the model 1 we call *fminsearch* in MATLAB, where the function to be optimized is the objective function introduced in (3.4.2). To find the efficient frontiers of $\mathrm{EVaR}_{95\%}$ we also call *fmincon* in MATLAB, where the function to be optimized is the objective function in (3.3.11). Figure 3.1 and 3.2 show the efficient frontiers based on model 1 for $\mathrm{EVaR}_{95\%}$ and standard deviation respectively. Tables 3.1 and 3.2 show the portfolio compositions and the corresponding $\mathrm{EVaR}_{95\%}$ and standard deviation respectively.

Return	$\mathrm{EVaR}_{95\%}$	Apple	Intel	PFE
0.0400	0.0738	0.2743	0.4140	0.3117
0.0480	0.0604	0.3210	0.3482	0.3308
0.0560	0.0494	0.3682	0.2827	0.3491
0.0640	0.0410	0.4159	0.2175	0.3667
0.0720	0.0351	0.4638	0.1524	0.3838
0.0800	0.0316	0.5120	0.0875	0.4005
0.0880	0.0301	0.5602	0.0226	0.4172
0.0960	0.0334	0.6772	0.0000	0.3228
0.1040	0.0493	0.8308	0.0000	0.1692
0.1120	0.0740	0.9844	0.0000	0.0156

TAB. 3.1. Portfolio composition and corresponding $\mathrm{EvaR}_{95\%}$ under model 1

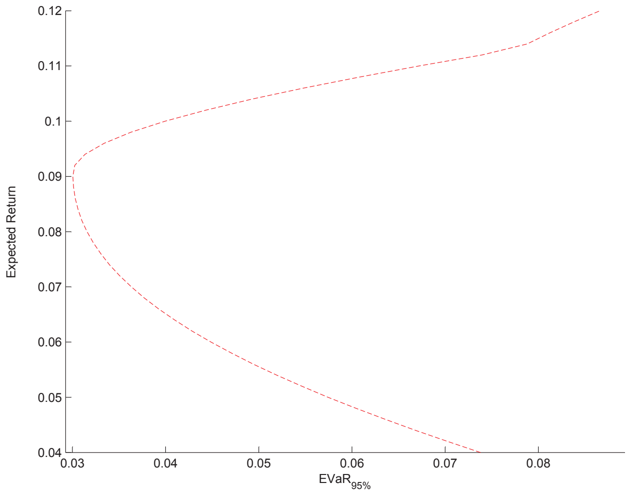


Fig. 3.1. Efficient frontier for $\mathrm{EVaR}_{95\%}$ under non-elliptical model 1

Return	Deviation	Apple	Intel	PFE
0.0400	0.1219	0.2674	0.4098	0.3228
0.0480	0.1143	0.3194	0.3472	0.3333
0.0560	0.1119	0.3715	0.2847	0.3438
0.0640	0.1147	0.4235	0.2222	0.3543
0.0720	0.1227	0.4755	0.1596	0.3648
0.0800	0.1359	0.5276	0.0971	0.3754
0.0880	0.1543	0.5797	0.0346	0.3857
0.0960	0.1811	0.6772	0.0000	0.3228
0.1040	0.2409	0.8308	0.0000	0.1692
0.1120	0.3386	0.9844	0.0000	0.0156

Tab. 3.2. Portfolio composition and corresponding standard deviation under model 1

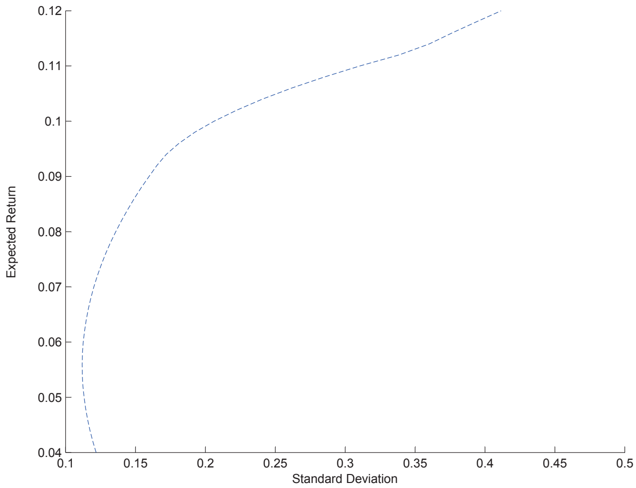


FIG. 3.2. Efficient frontier for standard deviation under non-elliptical model 1

Chapter 4

MULTIVARIATE DATA-BASED NATURAL RISK STATISTICS

Abstract

In [53], the concept of *natural risk statistics* is introduced as a data-based risk measure, i.e. as an axiomatic risk measure defined in the space $\mathbb{R}^n$. In this chapter, we set to generalize this notion to bivariate data sets (more generally, multivariate data sets) by defining an axiomatic risk measure on the space $\mathbb{R}^n \times \mathbb{R}^n$. This construction requires the definition of a suitable order in $\mathbb{R}^2$. This allows us to give a coherent characterization of these bivariate natural risk statistics. Some examples and applications of these measures are provided.

This chapter is a joint research work with Mélina Mailhot and Manuel Morales; see [69] [1].

4.1. Introduction

Designing risk measures with the right properties is an important problem from a practical point of view and, at the same time, it leads to interesting mathematical constructions. The usual approach is to postulate some reasonable axioms and then characterize the set of risk measures that satisfy these axioms. Coherent risk measures presented in [6] and insurance risk measures provided in [90] are examples of such constructions.

Along these lines, the concept of *natural risk statistics* has been introduced and studied in [8] and [53]. An interesting feature of this risk measure is that is defined on $\mathbb{R}^n$, i.e. the new risk measure assigns a value to a finite sample

[1] In this project, my contribution was in different ways. The main problem in this project was mostly joint identified by Prof. Morales, Prof. Mailhot and I. Then, I came up with the proof of the main theorem. Applying main results of the project in actuarial and financial context, providing different examples and interpretation of results obtained from examples were other contributions I have made to this project. I also contributed in this project by writing up the first draft of it, then polishing it with Prof. Morales and Prof. Mailhot to get the final version.

$(x_1, \ldots, x_n)$. This function measures the risk associated with a data sample from a financial (or insurance) position without further assumption on the underlying distribution. One can argue that this is often the kind of information upon which a risk manager relies to perform any risk analysis.

A *natural risk statistics* is a risk measure ρ_n that assigns a numerical value to a finite collection of data $(x_1, \ldots, x_n)$. Any collection of data can be seen as an element of $\mathbb{R}^n$ where n is the number of data available at the time. Under a set of axioms, it can be shown that *natural risk statistics* are characterized by the existence of a weight set $\mathcal{W}_n$ such that

$$\rho_n(x) = \sup_{w \in \mathcal{W}_n} \sum_{i=1}^{n} w_i x_{(i)},$$

where $x_{(i)}$ is the order statistics.

This characterization is what makes *natural risk statistics* consistent with industrial practices. These risk measures can be found as the supremum over a set of different scenarios defined by w_i. The main goal of this note is to extend this notion of *natural risk statistics* to the space $\mathbb{R}^n \times \mathbb{R}^n$ so that we can deal with bivariate data samples. We also provide illustrations showing the straightforward extension to higher dimensions.

In this setting, we suppose that we have a bivariate data set denoted by $\{(x_1, y_1), \ldots, (x_n, y_n)\}$. Naturally, each observation can be seen as element of $\mathbb{R}^2$ but, for our approach, we choose to see the whole data set as an element of $\mathbb{R}^n \times \mathbb{R}^n$, i.e. as $\{(x_1, \ldots, x_n), (y_1, \ldots, y_n)\}$. In this note, we discuss how an axiomatic risk measure ρ can be defined on the space $\mathbb{R}^n \times \mathbb{R}^n$. In particular, we give a characterization of such risk measures and some of their interesting properties.

Our motivation for studying functions on $\mathbb{R}^n \times \mathbb{R}^n$ is then two-fold, this space allows us to measure the risk associated with all finite collections of bivariate observations. On the other hand, extending the theory of coherent and convex risk measures to include risk measures on $\mathbb{R}^n \times \mathbb{R}^n$ is an interesting mathematical exercise on its own right.

Now, since we are interested in measuring the risk of bivariate portfolios, we felt compelled to define our risk measure through a vector-valued risk function. We find such an extension in [56] where they generalize the notion of coherent risk measures (as introduced in [6]) in order to allow for random portfolios taking values in $\mathbb{R}^d$. Their motivation to introduce such vector-valued risk measures is that investors are in general not able to aggregate their portfolio because of liquidity problems and/or transaction costs between the different security markets. In other words, a vector-valued risk measure is defined as a map from $\underbrace{L^\infty \times \cdots \times L^\infty}_{d-times}$

(the set of bounded random portfolios) into a subset of $\mathbb{R}^n$ for some $n \leq d$. This is, the risk of a multivariate portfolio composed of n positions is measured by a vector of d risk outcomes, i.e. $\rho : \underbrace{L^\infty \times \cdots \times L^\infty}_{d-times} \to \mathbb{R}^n$.

Here, we extend the notion of the natural risk statistics, introduced in [53], to the space $\mathbb{R}^n \times \mathbb{R}^n$. This generalization is carried out along the lines of the construction in [56] and as such, the end result is a set-valued risk measure which follows an axiomatic structure. This is carried out through the definition of a couple ordering on the space $\mathbb{R}^2$ which in turn allows us to order elements in $\mathbb{R}^n \times \mathbb{R}^n$.

The chapter is organized as follows. In Section 4.2 we provide a brief discussion of the concept of *natural risk statistics*, formal definitions and related theorems. In Section 4.3, we detail the axiomatic construction of a vector-valued natural risk statistics on $\mathbb{R}^n \times \mathbb{R}^n$. In this section, we define a couple ordering for the space $\mathbb{R}^2$, an axiomatic definition of bivariate data-based risk measure as well as a robust representation. In Section 4.4, we discuss an alternative characterization of the bivariate data-based risk measure via acceptance sets. Finally, in Section 4.5, we provide some practical examples that illustrate the interesting features of the proposed risk measures.

4.2. Natural risk statistics

The concept of *natural risk statistics* was first introduced in [53]. This notion attempts to move away from a modeled-based risk measure towards a data-based risk measure. Indeed, more often than not, a risk manager would have a data sample of a risky position from which risk must be assessed. Unlike coherent and convex risk measures, as introduced in [6] and [45], the notion of natural risk statistics makes no further assumption on the model behind the observations. In fact, in the presence of data, coherent risk measures lack robustness features with respect to outliers in a given data sample $(x_1, ..., x_n)$ (see for instance [31] and [53]). It turns out that there is an incompatibility between robustness and coherence for natural risk statistics (see [31]). This fact is documented in [3] and is a consequence of the very characterization of natural risk statistics. This is yet another motivation behind this contribution.

In order to proceed with our discussion, we briefly present in this section some definitions and results regarding *natural risk statistics*. We start with the axiomatic definition of a *natural risk statistics* which is stated here for finite ($\mathbb{R}^n$) and infinite (l^∞ or c_l) data sets.

4.2.1. Definition and properties

In the following definition we consider a data sets representing "loss" data for an insurance company rather than profit.

Definition 4.2.1. *([8, 53]) Let $\mathbb{A}$ be either one of spaces $\mathbb{R}^n$, l^∞ or c_l. A function $\rho : \mathbb{A} \longrightarrow \mathbb{R}$ is a natural risk statistics if it is:*

(1) Component wise positive homogeneous, i.e.

$$\rho(\lambda X) = \lambda\rho(X), \; \forall\, X \in \mathbb{A}\,,$$

for any $\lambda \geq 0$.

(2) Component wise translation invariant, i.e.

$$\rho(X + c\mathbf{1}) = \rho(X) + c, \forall\, X \in \mathbb{A}, c \in \mathbb{R},$$

where $\mathbf{1} = (\underbrace{1, \ldots, 1}_{n-times})$ if $\mathbb{A} = \mathbb{R}^n$ and $\mathbf{1} = (1, 1, \ldots)$ if $\mathbb{A} = l^\infty$ or c_l.

(3) Component wise increasing, i.e.

$$\rho(X) \leq \rho(Y)\ ,$$

for all $X \leq Y$ in $\mathbb{A}$. Here, the inequality $X \leq Y$ must be understood in the component wise sense.

(4) Component wise comonotone subadditive, i.e. if $(x_i - x_j)(y_i - y_j) \geqslant 0$ for any $j \neq i$, then

$$\rho(x_1 + y_1, \ldots, x_n + y_n) \leq \rho(x_1, \ldots, x_n) + \rho(y_1, \ldots, y_n)\ ,$$

for all $X, Y \in \mathbb{R}^n$, and

$$\rho(x_1 + y_1, x_2 + y_2, \ldots) \leq \rho(x_1, x_2, \ldots) + \rho(y_1, y_2, \ldots)\ ,$$

for all $X, Y \in l^\infty$ or c_l.

(5) Symmetric, i.e.

$$\rho(X) = \rho(X^{ij})\ ,$$

for all $X \in \mathbb{A}$ and all $i, j > 0$. Here the sequence X^{ij} is the element in $\mathbb{A}$ which is equal component wise to X except for the i-th and j-th component which are interchanged.

Moreover, if ρ satisfies only 2., 3. and 5. we say it is a general symmetric risk measure.

We notice that if $\mathbb{A} = \mathbb{R}^n$, then Definition 4.2.1 is the one in [3] and [53]. If $\mathbb{A} = l^\infty$ or c_l, then Definition 4.2.1 is an extended definition of *natural risk statistics* for infinite data sets.

Natural risk statistics can be interpreted as a weight vector allowed to each components of a vector of data. In this section we recall two main representations presented for natural risk statistics and subadditive natural risk statistics.
We refer to [53] for a thorough discussion related to the proof of the following theorems.

Theorem 4.2.1. *Let* $D := \{X \in \mathbb{R}^n \mid x_1 \leq x_2 \leq ... \leq x_n\}$ *and let* X_{os} *represent the order statistics of* X*, i.e.* $X_{os} := (x_{(1)}, ..., x_{(n)}) := X_\pi$ *for some* $\pi \in S_n$*, such that* $X_\pi \in D$ *where* S_n *is the set of all permutations for* $\{1, \ldots, n\}$*. Moreover, assume* $\mathcal{P} = \{\tilde{\omega} = (\omega_1, \ldots, \omega_n) \in \mathbb{R}^n \mid \sum_{i=1}^n \omega_i = 1\}$.

(1) For an arbitrarily given set of weights $\mathcal{W} \subset \mathcal{P}$*, the function*

$$\hat{\rho}(X) := \sup_{W \in \mathcal{W}} \sum_{i=1}^{n} w_i x_{(i)}, \quad \forall X \in \mathbb{R}^n, \tag{4.2.1}$$

is a natural risk statistic.

(2) Conversely, if $\hat{\rho}$ *is a natural risk statistic, then there exists a closed convex set of weights* $\mathcal{W} \subset \mathcal{P}$ *such that (4.2.1).*

The following is a somewhat stronger representation characterizing subadditive natural risk statistics.

Theorem 4.2.2. *Let* D*,* $\mathcal{P}$ *and* X_{os} *be as in Theorem 4.2.1.*

(1) For an arbitrarily given set of weights $\mathcal{W} \subset \mathcal{P} \cap D$*, the function*

$$\hat{\rho}(X) := \sup_{W \in \mathcal{W}} \sum_{i=1}^{n} w_i x_{(i)}, \quad \forall X \in \mathbb{R}^n,$$

is a subadditive natural risk statistic, i.e. satisfies axiom 4.

(2) Conversely, suppose the natural risk statistic $\hat{\rho}$ *is subadditive. Then there exists a closed convex set of weights* $\mathcal{W} \subset \mathcal{P} \cap D$ *such that (4.2.1).*

A natural risk statistic represents a risk measure of the observed data $X = (x_1, ..., x_n)$. For a justification of the concept, and a thorough comprehensive study of natural risk statistics as well as a detailed comparison to other classes of risk measure, we refer to [53] and [90].

4.3. Multivariate Data-based Natural Risk Statistics

In the last decade, financial industry as well as actuarial researchers and quantitative risk managers have started to accord great importance to the risk assessment of dependence structures. It is indeed common to have situations, in both insurance and finance, where multivariate models are needed in order to capture the risk of dependence across potentially correlated positions or losses. Motivated by this need, recent research has concentrated on studying risk measures in the multivariate setting. We find, for instance, in [56], the authors introduce an axiomatic definition for coherent vector-valued risk measures which is defined based on a proposed portfolio ordering. In [16], the author introduces a quantile-based risk measure for multivariate financial positions "the vector-valued Tail-conditional-expectation (TCE)". This extension is compatible with the multivariate framework introduced in [56]. In [22], the authors describe a general framework for multivariate risk measures, where the risk measure takes values in an abstract cone. They apply depth-trimmed (or central) regions and show that there is a close relation between this concept and the axiomatic definition for multivariate risk measures they propose. In [40], the authors introduce the multivariate lower and upper orthant VaR. In [32], the authors study the bivariate lower and upper orthant VaR. In particular, they provide new characterizations of the bivariate lower and upper orthant VaR and also derive desired properties of these bivariate risk measures. The author in [52] also studies set-valued risk measures.

Just like in the univariate case, existing multivariate risk measures are model-based as opposed to data-based. We set to study data-based risk measures in the multivariate case. The natural framework for this would be the development of a comprehensive construction of the natural risk statistics. Thus, the main contribution of this chapter is to construct a multivariate data-based risk measure or a multivariate version of the concept of *natural risk statistics* as introduced in [53]. Such a construction will endow us with a novel way of assessing dependence risk within multivariate data.

In this section, we give a definition of *bivariate data-based risk measure* and a representation for it in the bivariate setting. Natural extensions to higher dimensions are straight-forward but notationally cumbersome. We suppose that we have a bivariate data set denoted by $\{(x_1, y_1), \ldots, (x_n, y_n)\}$. These should be understood as observations coming from a portfolio composed of two positions or lines of business. Naturally, each observation can be seen as element of $\mathbb{R}^2$ but, for our purposes, we choose to see the whole data set as an element of $\mathbb{R}^n \times \mathbb{R}^n$, i.e. as $\{(x_1, \ldots, x_n), (y_1, \ldots, y_n)\}$. In other words, a data set of n paired observations is now seen as coupled pair of vectors each in $\mathbb{R}^n$.

In order to carry out the axiomatic construction of bivariate data-based risk measures along the lines of the robust representation of data-based univariate natural risk statistics, we first need a suitable definition couple ordering in $\mathbb{R}^2$.

Definition 4.3.1. *Let (x_1, y_1) and (x_2, y_2) be two elements of $\mathbb{R}^2$. Then, we say that $(x_1, y_1) \leq_{co} (x_2, y_2)$ if and only if $x_1 + y_1 \leq x_2 + y_2$ (and $(x_1, y_1) <_{co} (x_2, y_2)$ if and only if $x_1 + y_1 < x_2 + y_2$). Here we refer "co" to couple ordering.*

Notice that this ordering makes sense from a financial or insurance point of view where each pair (x_i, y_i) represents one observation coming from a portfolio composed of two, different and possibly dependent, positions or lines of business. Then, the sum $x_i + y_i$ is modeling one observation of the aggregate value of the whole portfolio. Ordering two data sets of such portfolio according to this aggregate order gives a certain degree of importance to the total value of the position as opposed to the value of each component.

Interestingly, we can now define the *order statistics* for a given bivariate data set using this ordering. Given n observations of a bivariate portfolio, $\{(x_1, y_1), \ldots, (x_n, y_n)\}$, we can now write $(x, y)_{(i)}$ for the i^{th} smallest pair of observations $\{(x_j, y_j)\}_{1\le j\le n}$ according to the aggregate ordering given in Definition 4.3.1, i.e.

$$(x, y)_{(1)} \le_{co} (x, y)_{(2)} \le_{co} \cdots \le_{co} (x, y)_{(n-1)} \le_{co} (x, y)_{(n)} \, . \tag{4.3.1}$$

In the case of having more than one pair with the same aggregate value, we come up with different versions of the order statistics associated to the given bivariate data set.

Moreover, this couple ordering implies a convenient ordering on the space $\mathbb{R}^n \times \mathbb{R}^n$, which is the space we need for our application since we chose to see this data in the form of $((x_1, \ldots, x_n), (y_1, \ldots, y_n))$. Indeed, this definition allows us to order two bivariate data-sets as follows,

Definition 4.3.2. *Let* $((x_1^1, x_2^1, \ldots, x_n^1), (y_1^1, y_2^1, \ldots, y_n^1))$ *and* $((x_1^2, x_2^2, \ldots, x_n^2), (y_1^2, y_2^2, \ldots, y_n^2))$ *be two elements of* $\mathbb{R}^n \times \mathbb{R}^n$*. Then, we say that* $((x_1^1, x_2^1, \ldots, x_n^1), (y_1^1, y_2^1, \ldots, y_n^1)) \le_{co} ((x_1^2, x_2^2, \ldots, x_n^2), (y_1^2, y_2^2, \ldots, y_n^2))$ *if and only if* $(x_k^1, y_k^1) \le_{co} (x_k^2, y_k^2)$ *for all* $k = 1, \ldots, n$*.*

Definition 4.3.2 can be easily extended to the multivariate version. In fact, we use the same couple ordering introduced in Definition 4.3.1 to compare elements in $\mathbb{R}^n \times \cdots \times \mathbb{R}^n$.

This couple ordering induces an equivalence relation on $\mathbb{R}^n \times \mathbb{R}^n$.

Definition 4.3.3. *Let* $((x_1^1, x_2^1, \ldots, x_n^1), (y_1^1, y_2^1, \ldots, y_n^1))$ *and* $((x_1^2, x_2^2, \ldots, x_n^2), (y_1^2, y_2^2, \ldots, y_n^2))$ *be two elements of* $\mathbb{R}^n \times \mathbb{R}^n$*. Then, we say that*

$$\Big((x_1^1, x_2^1, \ldots, x_n^1), (y_1^1, y_2^1, \ldots, y_n^1)\Big) \sim \Big((x_1^2, x_2^2, \ldots, x_n^2), (y_1^2, y_2^2, \ldots, y_n^2)\Big) \ ,$$

if and only if

$$\Big((x_1^1, x_2^1, \ldots, x_n^1), (y_1^1, y_2^1, \ldots, y_n^1)\Big) \le_{co} \Big((x_1^2, x_2^2, \ldots, x_n^2), (y_1^2, y_2^2, \ldots, y_n^2)\Big) \ ,$$

and

$$\Big((x_1^1,x_2^1,\ldots,x_n^1),(y_1^1,y_2^1,\ldots,y_n^1)\Big) \geq_{co} \Big((x_1^2,x_2^2,\ldots,x_n^2),(y_1^2,y_2^2,\ldots,y_n^2)\Big) \ ,$$

This relation gives us equivalence classes for $\mathbb{R}^n \times \mathbb{R}^n$. Without loss of generality we consider all portfolios of a class to be equally risky in an aggregate value sense. In this context, such ordering seems reasonable when the overall value of the portfolio is the main feature of concern. Under this perspective, two portfolios can be compared if one of them consistently produces observations that have a larger aggregate value.

4.3.1. Definition and Representation Theorem

We can now give a formal definition of a bivariate data-based risk measure. In fact, any bivariate data-based risk measure $\rho : \mathbb{R}^n \times \mathbb{R}^n \to \mathcal{P}(\mathbb{R}^2)$ is a set-valued function which for any bivariate data set in $\mathbb{R}^n \times \mathbb{R}^n$ assigns a subset of $\mathbb{R}^2$ containing of vectors. Here, $\mathcal{P}(\mathbb{R}^2)$ is the power set of $\mathbb{R}^2$. These vectors are all equals in the sense of couple ordering. That is, they all have the same aggregate value under the couple ordering in Definition 4.3.1. In the following, we provide an axiomatic definition for data-based risk measures assigning a single vector in $\mathbb{R}^2$. Thanks to the couple ordering in Definition 4.3.1 which enables us to reduce studying the case of set-valued data-based risk measures to the single vector-valued data-based risk measures.

Definition 4.3.4. *A function $\rho : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^2$ is a bivariate data-based risk measure if:*

(1) It is positive homogeneous and translation invariant component wise, i.e.,

$$\rho\left[a\,(\tilde{x},\tilde{y}) + b\,(\mathbf{1},\mathbf{1})\right] = a\,\rho\left[(\tilde{x},\tilde{y})\right] + b\,(\mathbf{1},\mathbf{1}) \ , \quad \forall \quad (\tilde{x},\tilde{y}) \in \mathbb{R}^n \times \mathbb{R}^n \ , \quad a \geq 0 \ , \ b \in \mathbb{R} \ ,$$

where $\mathbf{1} = (1,1,\ldots,1) \in \mathbb{R}^n$.

(2) It is component-wise monotonic, i.e.,

$$\rho\left[(\tilde{x}^1,\tilde{y}^1)\right] \leq_{co} \rho\left[(\tilde{x}^2,\tilde{y}^2)\right] \ ,$$

for all $(\tilde{x}^1,\tilde{y}^1) \leq_{co} (\tilde{x}^2,\tilde{y}^2)$ in $\mathbb{R}^n \times \mathbb{R}^n$.

(3) It is component-wise comonotone subadditive, i.e.,

$$\rho\left[(\tilde{x}^1,\tilde{y}^1) + (\tilde{x}^2,\tilde{y}^2)\right] \leq_{co} \rho\left[(\tilde{x}^1,\tilde{y}^1)\right] + \rho\left[(\tilde{x}^2,\tilde{y}^2)\right] \ ,$$

for all $(\tilde{x}^1,\tilde{y}^1)$ and $(\tilde{x}^2,\tilde{y}^2)$ in $\mathbb{R}^n \times \mathbb{R}^n$ such that,

$$\left[(x_i^1+y_i^1) - (x_j^1+y_j^1)\right]\left[(x_i^2+y_i^2) - (x_j^2+y_j^2)\right] \geq 0 \ ,$$

for any $i \neq j$.

(4) It is permutation invariant, i.e.,

$$\rho\left[((x_1,\dots,x_n),(y_1,\dots,y_n))\right]=\rho\left[((x_{i_1},\dots,x_{i_n}),(y_{i_1},\dots,y_{i_n}))\right]\ ,$$

for any permutation $(i_1,\dots,i_n)$ *of* $(1,2,\dots,n)$.

Remark 4.3.1. *We can extend the axioms proposed in Definition 4.3.4 in the case of set-valued data-based risk measures too. Assume that* ρ *is a set-valued data-based risk measure. For instance, we say* ρ *satisfies in the component-wise monotonic property if the following hold.*

$$(u,v)\leq_{co}(w,z),$$

for all $(\tilde{x}^1,\tilde{y}^1)\leq_{co}(\tilde{x}^2,\tilde{y}^2)$ *in* $\mathbb{R}^n\times\mathbb{R}^n$ *and for all* $(u,v)\in\rho\left[(\tilde{x}^1,\tilde{y}^1)\right]$, $(w,z)\in\rho\left[(\tilde{x}^2,\tilde{y}^2)\right]$.

We can easily see that the first two axioms in Definition 4.3.4 are enough to guarantee the continuity of the function ρ. We state this in the following result.

Proposition 4.3.1. *A function* $\rho:\mathbb{R}^n\times\mathbb{R}^n\to\mathbb{R}^2$ *satisfying the axioms 1. and 2. in Definition 4.3.4 is a continuous function.*

Proof. For every $(\tilde{x}^1,\tilde{y}^1)\in\mathbb{R}^n\times\mathbb{R}^n$, $\epsilon>0$ and $(\tilde{x}^2,\tilde{y}^2)$ satisfying $|x_i^1-x_i^2|<\epsilon$ and $|y_i^1-y_i^2|<\epsilon$, $i=1,\dots,n$, we have $(\tilde{x}^1,\tilde{y}^1)-\epsilon(\mathbf{1},\mathbf{1})<_{co}(\tilde{x}^2,\tilde{y}^2)<_{co}(\tilde{x}^1,\tilde{y}^1)+\epsilon(\mathbf{1},\mathbf{1})$. By the monotonicity in Axiom 2., we have $\rho\left[(\tilde{x}^1,\tilde{y}^1)\right]-\epsilon(\mathbf{1},\mathbf{1}))<_{co}\rho\left[(\tilde{x}^2,\tilde{y}^2)\right]<_{co}\rho\left[(\tilde{x}^1,\tilde{y}^1)\right]+\epsilon(\mathbf{1},\mathbf{1}))$. Using Axiom 1., this last inequality becomes,

$$\rho\left[(\tilde{x}^1,\tilde{y}^1)\right]-\epsilon(1,1)<_{co}\rho\left[(\tilde{x}^2,\tilde{y}^2)\right]<_{co}\rho\left[(\tilde{x}^1,\tilde{y}^1)\right]+\epsilon(1,1)\ ,$$

which establishes the continuity of ρ. □

We are now in a position to state the main result of this chapter. The following theorem, gives a necessary and sufficient condition to characterize bivariate data-based risk measures. In fact, we show that every bivariate data-based risk measure has a representation as follows. In the following theorem, we assume that we have just a unique order statistics for a given bivariate data set meaning that for a given bivariate data set we do not have more than one pair with the same aggregate value under the couple ordering given in Definition 4.3.1. We extend the following theorem for the case of having more than one version of order statistics for a given bivariate data set after the proof of the following theorem.

Theorem 4.3.1. *Let* $(\tilde{x},\tilde{y})=((x_1,x_2,\dots,x_n),(y_1,y_2,\dots,y_n))$ *be an element of* $\mathbb{R}^n\times\mathbb{R}^n$ *representing a bivariate data set of observations* $\{(x_1,y_1),\dots,(x_n,y_n)\}$. *And let* $(x,y)_{(1)},\dots,(x,y)_{(n)}$ *be the associated order statistics with respect to the couple ordering as given in (4.3.1).*

Then a function $\rho : \mathbb{R}^n \times \mathbb{R}^n \to \mathcal{P}(\mathbb{R}^2)$ is a bivariate data-based risk measure if and only if there exists a set of weights $\mathcal{W} = \{\tilde{w} = (w_1, \ldots, w_n) \mid \sum_{i=1}^n w_i = 1 \text{ and } w_i \geq 0 \,\forall\, 1 \leq i \leq n\} \subset \mathbb{R}^n$ such that

$$\rho\left[(\tilde{x}, \tilde{y})\right] = \sup_{\tilde{w} \in \mathcal{W}} \sum_{i=1}^{n} w_i (x, y)_{(i)}, \;\; \forall (\tilde{x}, \tilde{y}) \in \mathbb{R}^n \times \mathbb{R}^n. \tag{4.3.2}$$

Proof. First, we show the "if" part of the theorem. Suppose ρ accepts the representation in (4.3.2). Then it is obvious that ρ satisfies axioms 1. and 4. of Definition 4.3.4 in the sense we explained in Remark 4.3.1. In order to verify axiom 2., let us take two elements $(\tilde{x}^1, \tilde{y}^1) \leq_{co} (\tilde{x}^2, \tilde{y}^2)$ of $\mathbb{R}^n \times \mathbb{R}^n$. By definition, this means $(x_i^1, y_i^1) \leq_{co} (x_i^2, y_i^2),\ i = 1, \ldots, n$. In terms of the order statistics we can write,

$$(x^1, y^1)_{(i)} \leq_{co} (x^2, y^2)_{(i)} \;\forall i = 1, \ldots, n.$$

In turn, this implies that,

$$\rho\left[(\tilde{x}^2, \tilde{y}^2)\right] = \sup_{\tilde{w} \in \mathcal{W}} \sum_{i=1}^{n} w_i (x^2, y^2)_{(i)} \geq_{co} \sup_{\tilde{w} \in \mathcal{W}} \sum_{i=1}^{n} w_i (x^1, y^1)_{(i)} = \rho\left[(\tilde{x}^1, \tilde{y}^1)\right],$$

which shows that ρ satisfies axiom 2.
In order to verify axiom 3., consider again two elements $(\tilde{x}^1, \tilde{y}^1), (\tilde{x}^2, \tilde{y}^2)$ of $\mathbb{R}^n \times \mathbb{R}^n$. We notice that, if $(\tilde{x}^1, \tilde{y}^1)$ and $(\tilde{x}^2, \tilde{y}^2)$ are comonotonic, then there exists a permutation $(i_1, \ldots, i_n)$ of $(1, 2, \ldots, n)$ such that $(x_{i_1}^1, y_{i_1}^1) \leq_{co} (x_{i_2}^1, y_{i_2}^1) \leq_{co} \cdots \leq_{co} (x_{i_n}^1, y_{i_n}^1)$ and $(x_{i_1}^2, y_{i_1}^2) \leq_{co} (x_{i_2}^2, y_{i_2}^2) \leq_{co} \cdots \leq_{co} (x_{i_n}^2, y_{i_n}^2)$.
Hence, we have $((\tilde{x}^1, \tilde{y}^1) + (\tilde{x}^2, \tilde{y}^2))_{(j)} = (x_{i_j}^1, y_{i_j}^1) + (x_{i_j}^2, y_{i_j}^2) = (x^1, y^1)_{(j)} + (x^2, y^2)_{(j)}$ for $j = 1, 2, \ldots, n$. Therefore,

$$\begin{aligned}
\rho\left[(\tilde{x}^1, \tilde{y}^1) + (\tilde{x}^2, \tilde{y}^2)\right] &= \rho\left[(\tilde{x}^1 + \tilde{x}^2), (\tilde{y}^1 + \tilde{y}^2)\right] \\
&= \sup_{\tilde{w} \in \mathcal{W}} \sum_{i=1}^{n} w_i \left((x^1 + x^2), (y^1 + y^2)\right)_{(i)} \\
&= \sup_{\tilde{w} \in \mathcal{W}} \sum_{i=1}^{n} w_i \left((x^1, y^1)_{(i)} + (x^2, y^2)_{(i)}\right) \\
&\leq_{co} \sup_{\tilde{w} \in \mathcal{W}} \sum_{i=1}^{n} w_i (x^1, y^1)_{(i)} \\
&+ \sup_{\tilde{w} \in \mathcal{W}} \sum_{i=1}^{n} w_i (x^2, y^2)_{(i)} \\
&= \rho\left[(\tilde{x}^1, \tilde{y}^1)\right] + \rho\left[(\tilde{x}^2, \tilde{y}^2)\right],
\end{aligned}$$

which implies that ρ satisfies axiom 3.

Now, for the " only if" part, it turns out to be a more complex task. We need to exhibit the set of weights $\mathcal{W}$ described in Theorem 4.3.1 and show that any function ρ satisfying the axioms accepts such a representation. Inspired by the proof of the main theorem in [53], we start by defining the set $\mathcal{B} = \{(\tilde{x}, \tilde{y}) \in \mathbb{R}^n \times \mathbb{R}^n \mid (x_1, y_1) \leq_{co} (x_2, y_2) \leq_{co} \cdots \leq_{co} (x_n, y_n)\}$. In fact, we only need to prove the "only part" for this set $\mathcal{B}$ as by using axiom 4 we can always have a new rearrangement for any element $(\tilde{x}, \tilde{y}) \in \mathbb{R}^n \times \mathbb{R}^n$ such that this new element will be in $\mathcal{B}$. The following two lemmas about this set $\mathcal{B}$ are needed in order to continue with the proof of the "only if" part. In these lemmas, we are going to find a set of weights $\mathcal{W}$ which used to reach the representation (4.3.2) for $\rho\left[(\cdot,\cdot)\right]$. Finding this set of weights needs to apply an important theorem in convex analysis so-called supporting hyperplane theorem. More technical detail are provided in the following lemmas. Thanks to the continuity of $\rho\left[(\cdot,\cdot)\right]$, we just need to prove these lemmas for the interior points of $\mathcal{B}$. The results for boundary points can be obtained by approximating the boundary points by the interior points, and by using continuity of the function $\rho\left[(\cdot,\cdot)\right]$. It is worth to point out that in the following two lemmas we assume that $\rho\left[(\cdot,\cdot)\right]$ is a vector-valued risk measure which takes its value in $\mathbb{R}^2$. We then finish the proof of Theorem 4.3.1 for the case $\rho : \mathbb{R}^n \times \mathbb{R}^n \to \mathcal{P}(\mathbb{R}^2)$.

Lemma 4.3.1. *Let $\mathcal{B} = \{(\tilde{x}, \tilde{y}) \in \mathbb{R}^n \times \mathbb{R}^n \mid (x_1, y_1) \leq_{co} (x_2, y_2) \leq_{co} \cdots \leq_{co} (x_n, y_n)\}$ and denote $\mathcal{B}^o$ to be the interior of $\mathcal{B}$. $(\tilde{z}, \tilde{z}')$ Moreover, let us consider a fixed $(\tilde{z}, \tilde{z}') \in \mathcal{B}^o$ and any $\rho : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^2$ satisfying axioms $1-4$ in Definition 4.3.4. If $\rho\left[(\tilde{z}, \tilde{z}')\right] = (1,1)$ then there exists a weight $\tilde{w} = (w_1, \ldots, w_n) \subset \mathbb{R}^n$ such that the homogeneous function $\lambda\left[(\tilde{x}, \tilde{y})\right] := \sum_{i=1}^n w_i(x_i, y_i)$, satisfies*

$$
\begin{aligned}
\lambda\left[(\tilde{z}, \tilde{z}')\right] &= (1,1), && (4.3.3)\\
\lambda\left[(\tilde{x}, \tilde{y})\right] &<_{co} (1,1)\ \forall (\tilde{x}, \tilde{y}) \in \mathcal{B} \text{ and } \rho\left[(\tilde{x}, \tilde{y})\right] <_{co} (1,1). && (4.3.4)
\end{aligned}
$$

Proof. Consider the set $D = \{(\tilde{x}, \tilde{y}) \mid \rho\left[(\tilde{x}, \tilde{y})\right] <_{co} (1,1)\} \cap \mathcal{B}$. Now, consider a function $\rho : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^2$ satisfying axioms 1.-4. in Theorem 4.3.1. Take two arbitrary elements, $(\tilde{x}^1, \tilde{y}^1)$ and $(\tilde{x}^2, \tilde{y}^2)$ of $\mathbb{R}^n \times \mathbb{R}^n$. By construction, we know that $(\tilde{x}^1, \tilde{y}^1)$ and $(\tilde{x}^2, \tilde{y}^2)$ are comonotonic. It now follows from axioms 1. and 3. imply that the set D is convex, and, therefore, the closure $\bar{D}$ of D is also convex. For any $\epsilon > 0$, since, $\rho\left[(\tilde{z}, \tilde{z}') - \epsilon(\mathbf{1}, \mathbf{1})\right] = \rho\left[(\tilde{z}, \tilde{z}')\right] - \epsilon(1,1) = (1,1) - \epsilon(1,1) <_{co} (1,1)$, it follows that $(\tilde{z}, \tilde{z}') - \epsilon(\mathbf{1}, \mathbf{1}) \in D$. Since, $(\tilde{z}, \tilde{z}') - \epsilon(\mathbf{1}, \mathbf{1})$ tends $(\tilde{z}, \tilde{z}')$ as $\epsilon \downarrow 0$, we know that $(\tilde{z}, \tilde{z}')$ is a boundary point of D because $\rho\left[(\tilde{z}, \tilde{z}')\right] = (1,1)$.

Now, define $U = \{\tilde{u} = (x_i + y_i)_{1\le i\le n} \mid (\tilde{x},\tilde{y}) \in D\}$. Therefore, by one application of the supporting hyperplane theorem [2] there exists a linear functional for $\bar{U}$ at the point $\tilde{u}' = (z_i + z_i')_{1\le i\le n}$, i.e., there exists a nonzero vector $\tilde{w} = (w_1, \ldots, w_n) \in \mathbb{R}^n$ such that $F(\tilde{u}) := \sum_{i=1}^n w_i(x_i + y_i)$ satisfies $F(\tilde{u}) \le F(\tilde{u}')$ for all $\tilde{u} \in \bar{U}$.
Now we define the new function $\lambda[(\tilde{x},\tilde{y})] := \sum_{i=1}^n w_i(x_i, y_i)$. The function $\lambda[(\cdot,\cdot)]$ satisfies in $\lambda[(\tilde{x},\tilde{y})] \le_{co} \lambda[(\tilde{z},\tilde{z}')]$ for all $(\tilde{x},\tilde{y}) \in \bar{D}$. In particular, we have

$$\lambda[(\tilde{x},\tilde{y})] \le_{co} \lambda[(\tilde{z},\tilde{z}')], \; \forall(\tilde{x},\tilde{y}) \in D. \tag{4.3.5}$$

By Axiom 1 we have $\rho[(\mathbf{0},\mathbf{0})] = (0,0)$ therefore, $(\mathbf{0},\mathbf{0}) \in D$. Letting $(\tilde{x},\tilde{y}) = (\mathbf{0},\mathbf{0})$ in (4.3.5) yields $\lambda[(\tilde{z},\tilde{z}')] >_{co} (0,0)$. So, we can re-scale λ such that

$$\lambda[(\tilde{z},\tilde{z}')] = (1,1) = \rho[(\tilde{z},\tilde{z}')].$$

Now, the proof is finished if we show that the inequality appeared in (4.3.5) is strict. i.e. we need to show that

$$\lambda[(\tilde{x},\tilde{y})] <_{co} \lambda[(\tilde{z},\tilde{z}')], \; \forall(\tilde{x},\tilde{y}) \in D.$$

This can be shown by using absurdity method and applying axioms 1-4. The rest of the proof will be the same as one for univariate case. We refer to [53] to follow the procedure. □

In the following lemma, we try to get one more step closer to complete the plot which helps us to finish the proof of Theorem 4.3.1. This lemma which is mainly based on Lemma 4.3.1, will enable us to find a set of non-negative weights which are applied in the presentation (4.3.2) in Theorem 4.3.1.
Lemma 4.3.2. *Let $\mathcal{B} = \{(\tilde{x},\tilde{y}) \in \mathbb{R}^n \times \mathbb{R}^n \mid (x_1,y_1) \le_{co} (x_2,y_2) \le_{co} \cdots \le_{co} (x_n,y_n)\}$, and denote $\mathcal{B}^o$ to be the interior of $\mathcal{B}$. For any fixed $(\tilde{z},\tilde{z}') \in \mathcal{B}^o$ and any ρ satisfying Axioms $1-4$ in Definition 4.3.4, there exists a weight $\tilde{w} = (w_1,\ldots,w_n)$ such that*

$$\sum_{i=1}^n w_i = 1, \tag{4.3.6}$$

$$w_i \ge 0 \quad , \quad i = 1,2,\ldots,n, \tag{4.3.7}$$

$$\rho[(\tilde{x},\tilde{y})] \ge_{co} \sum_{i=1}^n w_i(x_i,y_i), \; \forall(\tilde{x},\tilde{y}) \in \mathcal{B}, \text{ and } \rho[(\tilde{z},\tilde{z}')] = \sum_{i=1}^n w_i(z_i,z_i'). \tag{4.3.8}$$

[2]This theorem states that if S is a closed convex set in a topological vector space X, and x_0 is a point on the boundary of S, then there exists a supporting hyperplane containingx_0. i.e., there exists a continuous linear functional F on X such that $F(x) \le F(x_0)$ for all $x \in S$.

Proof. We give a sketch of the proof since the lemma is the multivariate version of the univaraite one discussed in [53]. To prove this lemma we can consider three following cases on $\rho\,[(\tilde{z},\tilde{z}')]$.

(1) $\rho(\tilde{z},\tilde{z}') = (1,1)$.

(2) $\rho(\tilde{z},\tilde{z}') \neq (1,1)$, and $\rho(\tilde{z},\tilde{z}') >_{co} (0,0)$.

(3) $\rho(\tilde{z},\tilde{z}') \neq (1,1)$, and $\rho(\tilde{z},\tilde{z}') \leq_{co} (0,0)$.

Cases 2 and 3 can be studied using case 1 by applying axiom 1 where we can use positive homogeneity for the case 2 and translation invariant for case 3. so we just need to take care of case 1. In case 1 we use the weight $\tilde{w} = (w_1,\dots,w_n)$ and the function $\lambda\,[(\cdot,\cdot)]$ in Lemma 4.3.1 and show that all components of $\tilde{w}$ are non-negative and their sum is equal to 1. These along with Property (4.3.8) can be verified by doing some algebraic operations and an application of continuity of $\lambda\,[(\cdot,\cdot)]$. For technical details we refer to [53] where the authors prove this lemma in univariate case.

$\square$

Now, By applying the last two lemmas and one application of the continuity of the function $\rho\,[(\cdot,\cdot)]$, the proof of Theorem 4.3.1 will be completed.

By Axiom 4, we only need to show that there exists a set of weights $\mathcal{W} = \{\tilde{w} = (w_1,\dots,w_n)\} \subset \mathbb{R}^n$ with each $\tilde{w} \in \mathcal{W}$ satisfying $\sum_{i=1}^{n} w_i = 1$ and $w_i \geq 0,\ \forall 1 \leq i \leq n$, such that

$$\rho\,[(\tilde{x},\tilde{y})] = \sup_{\tilde{w}\in\mathcal{W}} \sum_{i=1}^{n} w_i\,(x_i,y_i)\,,\ \forall(\tilde{x},\tilde{y}) \in \mathcal{B}, \tag{4.3.9}$$

where recall that $\mathcal{B} = \{(\tilde{x},\tilde{y}) \in \mathbb{R}^n \times \mathbb{R}^n\ /\ (x_1,y_1) \leq_{co} (x_2,y_2) \leq co \cdots \leq_{co} (x_n,y_n)\}$.

By Lemma 4.3.2, for any point $(\tilde{x},\tilde{y}) \in \mathcal{B}^o$ there exists a weight $\tilde{w}^{(\tilde{x},\tilde{y})}$ which is non-negative, summation of its components is 1 and satisfies in (4.3.8). Therefore, we can take the set of such weights as

$$\mathcal{W} = \{\tilde{w}^{(\tilde{x},\tilde{y})} \mid (\tilde{x},\tilde{y}) \in \mathcal{B}^o\}. \tag{4.3.10}$$

We claim that the weight set $\mathcal{W}$ proposed in (4.3.10) is a set satisfying in Theorem 4.3.1 which leads to the representation (4.3.2) for $\rho\,[(\cdot,\cdot)]$. From (4.3.8), for any fixed $(\tilde{x}^1,\tilde{y}^1) \in \mathcal{B}^o$ we have

$$\rho\left[(\tilde{x}^1,\tilde{y}^1)\right] \geq_{co} \sum_{i=1}^{n} w_i^{(\tilde{x},\tilde{y})}(x_i^1,y_i^1),\ \forall(\tilde{x},\tilde{y}) \in \mathcal{B}^o, \tag{4.3.11}$$

$$\rho\left[(\tilde{x}^1,\tilde{y}^1)\right] = \sum_{i=1}^{n} w_i^{(\tilde{x}^1,\tilde{y}^1)}(x_i^1,y_i^1). \tag{4.3.12}$$

As $\tilde{w}^{(\tilde{x}^1,\tilde{y}^1)}$ may not be unique for $(\tilde{x}^1, \tilde{y}^1)$, then we can come up with a set of vectors evaluating the risk of $(\tilde{x}^1, \tilde{y}^1)$. In the other words, $\rho\left[(\tilde{x}^1, \tilde{y}^1)\right]$ is a set of vectors with the same aggregate value under the couple ordering in Definition 4.3.1. Therefore,

$$\rho\left[(\tilde{x}^1, \tilde{y}^1)\right] = \sup_{(\tilde{x},\tilde{y})\in\mathcal{B}^o} \sum_{i=1}^{n} w_i^{(\tilde{x},\tilde{y})}(x_i^1, y_i^1), \ \forall(\tilde{x}^1, \tilde{y}^1) \in \mathcal{B}^o, \tag{4.3.13}$$

where each $\tilde{w} \in \mathcal{W}$ given in (4.3.10). The proof is complete by using the continuity of $\rho\left[(\cdot,\cdot)\right]$ to show that equality (4.3.13) is also true for any point in the boundary of B, ∂B. The proof would be the same as the one for univariate case but in a multivariate framework. We refer to [53] for a thorough discussion on this issue for univariate case.

□

Remark 4.3.2. *We can extend Theorem 4.3.1 for the case of having more than one order statistics for a given bivariate data set under the couple ordering in Definition 4.3.1. In fact, a function $\rho : \mathbb{R}^n \times \mathbb{R}^n \to \mathcal{P}(\mathbb{R}^2)$ is a bivariate data-based risk measure if and only if there exists a set of weights $\mathcal{W} = \{\tilde{w} = (w_1, \ldots, w_n) \mid \sum_{i=1}^n w_i = 1 \ \ and \ \ w_i \geq 0 \ \forall\ 1 \leq i \leq n\} \subset \mathbb{R}^n$ such that*

$$\rho\left[(\tilde{x}, \tilde{y})\right] = \sup_{\tilde{w}\in\mathcal{W}, 1\leq j\leq k} \sum_{i=1}^{n} w_i(x, y)^j_{(i)}, \ \forall(\tilde{x}, \tilde{y}) \in \mathbb{R}^n \times \mathbb{R}^n, \tag{4.3.14}$$

where k is the number of different versions of the order statistics associated to a given bivariate data set $(\tilde{x}, \tilde{y}) \in \mathbb{R}^n \times \mathbb{R}^n$.

The way we extended Theorem 4.3.1 in Remark 4.3.2 can yield to have a set of vectors evaluating the bivariate risk for a given bivariate data set. In the following subsection we provide a method to choose one vector from this set of risk vectors.

4.3.2. Minimum Distance Data-based Risk Measures

Consider a bivariate data-based risk measure ρ with representation given in (4.3.2). It is possible to have more than one vector as a risk of observation $(\tilde{x}, \tilde{y})$. i.e., $\rho\left[(\tilde{x}, \tilde{y})\right]$ can be a set of vectors which are obtained by taking supremum over a set of weights $\mathcal{W}$ in the representation (4.3.2). Because of the nature of the couple ordering we proposed to use ordering the observations, we can have several vectors in $\mathbb{R}^2$ belonging to a same equivalence class. For instance, consider a typical observation $(\tilde{x}, \tilde{y})$. Then, it is possible to have $\rho\left[(\tilde{x}, \tilde{y})\right] = (1, 3) = (2, 2)$ because $(1, 3) \sim (2, 2)$which means $(1 + 3 = 2 + 2 = 4)$.

A question arising in this context is to choose a right pair of the equivalence class which contains $\rho\left[(\tilde{x}, \tilde{y})\right]$. There are several ways in which a risk measure can be chosen within an equivalent class. In this chapter we chose to use the Euclidean distance in $\mathbb{R}^2$. For this let us consider the loss observation $(\tilde{x}, \tilde{y})$ and $\rho\left[(\tilde{x}, \tilde{y})\right] \subset [(a,b)]$ where $[(a,b)]$ is an equivalence class containing the set of $\rho\left[(\tilde{x}, \tilde{y})\right]$. On the other hand, assume that $\rho\left[(\tilde{x}, \mathbf{0})\right]$ and $\rho\left[(\mathbf{0}, \tilde{y})\right]$ are the univariate risk statistics associated to the observations $\tilde{x}$ and $\tilde{y}$ driven from $\rho\left[(\cdot,\cdot)\right]$ by projecting $(\tilde{x}, \tilde{y})$ on the first and second components respectively. Now, we choose the unique risk associated to the observation $(\tilde{x}, \tilde{y})$ by the following way.

$$\rho_{min}\left[(\tilde{x}, \tilde{y})\right] = \min_{(c,d)\in\rho[(\tilde{x},\tilde{y})]} \sqrt{(c - \rho\left[(\tilde{x}, \mathbf{0})\right])^2 + (d - \rho\left[(\mathbf{0}, \tilde{y})\right])^2}.$$

Therefor, $\rho_{min}\left[(\tilde{x}, \tilde{y})\right]$ is the closest pair in $\mathbb{R}^2$ to $\rho\left[(\tilde{x}, \mathbf{0})\right]$ and $\rho\left[(\mathbf{0}, \tilde{y})\right]$ with respect to the Euclidean distance. So, this is our desired bivariate data-based risk measure.

4.3.3. Multivariate Extension

The concept of bivariate data-based risk measure and its characterization in Theorem 4.3.1, can be both easily extended to the multivariate case. We chose to thread our discussion around the bivariate case since it allows for a less cumbersome notation, nonetheless it is a straight-forward exercise to derive expressions in the multivariate case. In this small section, we present this generalization in the form of a proposition that will be used later in the examples section.

Proposition 4.3.2. *Let* $(\tilde{x}^{(1)}, ..., \tilde{x}^{(d)}) \in \mathbb{R}^n \times \cdots \times \mathbb{R}^n$ *and* $\{(x^{(1)}, \ldots, x^{(d)})_{(1)}, \ldots, (x^{(1)}, \ldots, x^{(d)})_{(n)}\}$ *be the order statistics of the observation* $(\tilde{x}^{(1)}, ..., \tilde{x}^{(d)})$ *with respect to the couple ordering introduced in Section 4.3. Then, a function* $\rho : \mathbb{R}^n \cdots \times \mathbb{R}^n \to \mathcal{P}(\mathbb{R}^d)$ *is a multivariate data-based risk measure if and only if there exists a set of weights* $\mathcal{W} = \{\tilde{w} = (w_1, \ldots, w_n) \mid \sum_{i=1}^n w_i = 1 \text{ and } w_i \geq 0 \, \forall \, 1 \leq i \leq n\} \subset \mathbb{R}^n$ *such that*

$$\rho\left[(\tilde{x}^{(1)}, ..., \tilde{x}^{(d)})\right] = \sup_{\tilde{w}\in\mathcal{W}} \sum_{i=1}^{n} w_i (x^{(1)}, \ldots, x^{(d)})_{(i)}, \ \forall (\tilde{x}^{(1)}, ..., \tilde{x}^{(d)}) \in \mathbb{R}^n \times \cdots \times \mathbb{R}^n. \tag{4.3.15}$$

The proof of Proposition 4.3.2 can easily be obtained as a trivial extension of the proof of Theorem 4.3.1 to higher dimensions.

In Section 5.5, we will provide examples showing how we can work with multivariate data-based risk measures using Proposition 4.3.2.

4.4. An Alternative Characterization of the Multivariate Date-Based Risk Measures via Acceptance Sets

Coherent risk measures on the space of the random variables can also be characterized via so-called acceptance sets. Here, we explore a similar characterization like the one proposed for natural risk statistics in [53]. For the sake simplicity, and just like in the previous sections, we focus on the bivariate case only. The axioms and characterization theorem provided in this section can be easily extended to the multivariate cases.

Definition 4.4.1. *Let $\mathcal{A} \subset \mathbb{R}^n \times \mathbb{R}^n$. We define the risk statistics $\rho_{\mathcal{A}}$ associated with $\mathcal{A}$ on $\mathbb{R}^n \times \mathbb{R}^n$ to be the function $\rho_{\mathcal{A}} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^2$,*

$$\rho_{\mathcal{A}}\left[(\tilde{x}, \tilde{y})\right] = \inf\{(m, m') \mid (\tilde{x} - m\mathbf{1}, \tilde{y} - m'\mathbf{1}) \in \mathcal{A}\}, \; \forall (\tilde{x}, \tilde{y}) \in \mathbb{R}^n \times \mathbb{R}^n \; . \quad (4.4.1)$$

Notice that this infimum is well-defined via couple ordering defined in Section 4.3 although it is also possible that $\rho_{\mathcal{A}}\left[(\tilde{x}, \tilde{y})\right]$ is a set of vectors in $\mathbb{R}^2$. We use the convention discussed in Subsection 4.3.2 to deal with such equivalence classes, i.e., we try to choose a vector in $\rho_{\mathcal{A}}\left[(\tilde{x}, \tilde{y})\right]$ in a way that it minimizes the Euclidean distance between elements of $\rho_{\mathcal{A}}\left[(\tilde{x}, \tilde{y})\right]$ with $\rho_{\mathcal{A}_1}\left[(\tilde{x}, \tilde{0})\right]$ and $\rho_{\mathcal{A}_2}\left[(\tilde{0}, \tilde{y})\right]$ where A_1 and A_2 are sets of the first of second coordinates of A respectively. Such a set will be referred as a *statistical acceptance set* and it can be understood as a set of observations whose risk is acceptable. Then, the associated function $\rho_{\mathcal{A}}$ would be the amount of capital that needs to be injected to a given set of observations to make it be an element of the acceptable set $\mathcal{A}$.

On the other hand, notice that a given bivariate data-based risk measure ρ naturally defines a subset in $\mathcal{A} \subset \mathbb{R}^n \times \mathbb{R}^n$, denoted by $\mathcal{A}_\rho$, in the following way,

$$\mathcal{A}_\rho := \{(\tilde{x}, \tilde{y}) \in \mathbb{R}^n \times \mathbb{R}^n \mid \rho\left[(\tilde{x}, \tilde{y})\right] \leq_{co} (0, 0)\}. \quad (4.4.2)$$

We can now define the notion of acceptance set.

Definition 4.4.2. *Let $\mathcal{A}$ be a subset of $\mathbb{R}^n \times \mathbb{R}^n$. Then, $\mathcal{A}$ is called a bivariate data-based acceptance set if it satisfies in the following axioms,*

- **C1.** *The set $\mathcal{A}$ contains $\mathbb{R}^n_- \times \mathbb{R}^n_-$ where $\mathbb{R}^n_- \times \mathbb{R}^n_- = \{(\tilde{x}, \tilde{y}) \in \mathbb{R}^n \times \mathbb{R}^n \mid x_i, y_i \leq 0, i = 1, \dots, n\}$,*
- **C2.** *The set $\mathcal{A}$ does not intersect the set $\mathbb{R}^n_{++} \times \mathbb{R}^n_{++}$, where*

$$\mathbb{R}^n_{++} \times \mathbb{R}^n_{++} = \{(\tilde{x}, \tilde{y}) \in \mathbb{R}^n \times \mathbb{R}^n \mid x_i, y_i > 0, i = 1, \dots, n\} \; ,$$

- **C3.** *If $(\tilde{x}^1, \tilde{y}^1)$ and $(\tilde{x}^2, \tilde{y}^2)$ are comonotonic and such that $(\tilde{x}^1, \tilde{y}^1) \in \mathcal{A}$, $(\tilde{x}^2, \tilde{y}^2) \in \mathcal{A}$, then $\alpha(\tilde{x}^1, \tilde{y}^1) + (1 - \alpha)(\tilde{x}^2, \tilde{y}^2) \in \mathcal{A}$, for all $\alpha \in [0, 1]$,*

C4. *The set $\mathcal{A}$ is positively homogeneous, i.e., if $(\tilde{x},\tilde{y}) \in \mathcal{A}$, then $\alpha(\tilde{x},\tilde{y}) \in \mathcal{A}$ for all $\alpha \geq 0$,*

C5. *If $(\tilde{x}^1,\tilde{y}^1) \leq_{co} (\tilde{x}^2,\tilde{y}^2)$ and $(\tilde{x}^2,\tilde{y}^2) \in \mathcal{A}$, then $(\tilde{x}^1,\tilde{y}^1) \in \mathcal{A}$,*

C6. *If $(\tilde{x},\tilde{y}) \in \mathcal{A}$, then $(x_{i_1}, x_{i_2}, \ldots, x_{i_n}, y_{i_1}, y_{i_2}, \ldots, y_{i_n}) \in \mathcal{A}$ for any permutation $(i_1, \ldots, i_n)$.*

Interestingly enough, a bivariate data-based risk statistics can be characterized through bivariate acceptance sets. This is ensured by the following result.

Theorem 4.4.1. *Let ρ be a function $\rho : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^2$ and $\mathcal{A}$ a subset of $\mathbb{R}^n \times \mathbb{R}^n$. Then,*

(1) if ρ is a bivariate data-based risk measure in the sense of Definition 4.3.4, then the set $\mathcal{A}_\rho$ associated with ρ is a bivariate acceptance set, i.e. it satisfies axioms $C1-C6$ in Definiton 4.4.2. Moreover, $\mathcal{A}_\rho$ is a closed set,

(2) if $\mathcal{A}$ is a bivariate data-based acceptance set, in the sense of Definiton 4.4.2, then its associated risk function $\rho_{\mathcal{A}}$, as defined in (4.4.1), is a bivariate data-based risk measure.

Proof. We get inspiration from the method of proof given in Hayde et al. [53]. for natural risk statistics to prove the theorem. We refer to [53] for the first part and we give the proof of second part which is slightly different than one in [53].

For $(\tilde{x},\tilde{y}) \in \mathbb{R}^n \times \mathbb{R}^n$ and $b \in \mathbb{R}$, we have

$$
\begin{aligned}
\rho_{\mathcal{A}}\left[(\tilde{x},\tilde{y}) + b(\mathbf{1},\mathbf{1})\right] &= \inf\{(m,m') \mid (\tilde{x} + (b-m)\mathbf{1}, \tilde{y} + (b-m')\mathbf{1}) \in \mathcal{A}\} \\
&= \inf\{(b+u, b+u') \mid (\tilde{x} - u\mathbf{1}, \tilde{y} - u'\mathbf{1}) \in \mathcal{A}\} \\
&= b(1,1) + \inf\{(u,u') \mid (\tilde{x} - u\mathbf{1}, \tilde{y} - u'\mathbf{1}) \in \mathcal{A}\} \\
&= b(1,1) + \rho_{\mathcal{A}}\left[(\tilde{x},\tilde{y})\right],
\end{aligned}
$$

which we used the couple ordering defined in the last section. This shows the vectorial translation invariance property of $\rho_{\mathcal{A}}$. To prove the positively homogeneous property, assume that $(\tilde{x},\tilde{y}) \in \mathbb{R}^n \times \mathbb{R}^n$ and $a \geq 0$, if $a = 0$, then

$$
\rho_{\mathcal{A}}\left[a(\tilde{x},\tilde{y})\right] = \inf\{(m,m') \mid (\mathbf{0} - m\mathbf{1}, \mathbf{0} - m'\mathbf{1}) \in \mathcal{A}\} = (0,0) = a\rho_{\mathcal{A}}\left[(\tilde{x},\tilde{y})\right],
$$

where the second equality follows from $C1$ and $C2$. If $a > 0$, then

$$
\begin{aligned}
\rho_{\mathcal{A}}\left[a(\tilde{x},\tilde{y})\right] &= \inf\{(m,m') \mid (a\tilde{x} - m\mathbf{1}, a\tilde{y} - m'\mathbf{1}) \in \mathcal{A}\} \\
&= \inf\{(au, au') \mid a(\tilde{x} - u\mathbf{1}, \tilde{y} - u'\mathbf{1}) \in \mathcal{A}\} \\
&= a\inf\{(u,u') \mid a(\tilde{x} - u\mathbf{1}, \tilde{y} - u'\mathbf{1}) \in \mathcal{A}\} \\
&= a\inf\{(u,u') \mid (\tilde{x} - u\mathbf{1}, \tilde{y} - u'\mathbf{1}) \in \mathcal{A}\}
\end{aligned}
$$

$$= \quad a\rho_{\mathcal{A}}\left[(\tilde{x},\tilde{y})\right],$$

by using Axiom $C4$. This shows the positively homogeneous property of $\rho_{\mathcal{A}}$. To show the monotonicity property of $\rho_{\mathcal{A}}$, suppose $(\tilde{x}^1,\tilde{y}^1) \leq_{co} (\tilde{x}^2,\tilde{y}^2)$. For any $m \in \mathbb{R}$, if $(\tilde{x}^2 - m\mathbf{1}, \tilde{y}^2 - m'\mathbf{1}) \in \mathcal{A}$, then $(\tilde{x}^1 - m\mathbf{1}, \tilde{y}^1 - m'\mathbf{1}) \leq_{co} (\tilde{x}^2 - m\mathbf{1}, \tilde{y}^2 - m'\mathbf{1})$ along with Axiom $C5$ imply that $(\tilde{x}^1 - m\mathbf{1}, \tilde{y}^1 - m'\mathbf{1}) \in \mathcal{A}$. Therefore, $\{(m,m') \mid (\tilde{x}^2 - m\mathbf{1}, \tilde{y}^2 - m'\mathbf{1}) \in \mathcal{A}\} \subset \{(m,m') \mid (\tilde{x}^1 - m\mathbf{1}, \tilde{y}^1 - m'\mathbf{1}) \in \mathcal{A}\}$. Now, by taking infimum on both sides of the above relation, we obtain $\rho_{\mathcal{A}}\left[(\tilde{x}^1,\tilde{y}^1)\right] \leq_{co} \rho_{\mathcal{A}}\left[(\tilde{x}^2,\tilde{y}^2)\right]$ which proves the monotonicity property of ρ. To verify comonotonic subadditivity property, let $(\tilde{x}^1,\tilde{y}^1)$ and $(\tilde{x}^2,\tilde{y}^2)$ be comonotonic. For any couples (m,m') and (n,n') such that $(\tilde{x}^1 - m\mathbf{1}, \tilde{y}^1 - m'\mathbf{1}) \in \mathcal{A}$, $(\tilde{x}^2 - n\mathbf{1}, \tilde{y}^2 - n'\mathbf{1}) \in \mathcal{A}$, since $(\tilde{x}^1 - m\mathbf{1}, \tilde{y}^1 - m'\mathbf{1})$ and $(\tilde{x}^2 - n\mathbf{1}, \tilde{y}^2 - n'\mathbf{1})$ are comonotonic, it follows from Axiom $C3$ that $\frac{1}{2}(\tilde{x}^1 - m\mathbf{1}, \tilde{y}^1 - m'\mathbf{1}) + \frac{1}{2}(\tilde{x}^2 - n\mathbf{1}, \tilde{y}^2 - n'\mathbf{1}) \in \mathcal{A}$. By positive homogeneity property of $\mathcal{A}$ and the previous formula we have $(\tilde{x}^1 + \tilde{x}^2 - (m+n)\mathbf{1}, \tilde{y}^1 + \tilde{y}^2 - (m'+n')\mathbf{1}) \in \mathcal{A}$. Therefore,

$$\rho_{\mathcal{A}}\left[(\tilde{x}^1,\tilde{y}^1) + (\tilde{x}^2,\tilde{y}^2)\right] \leq_{co} (m+n, m'+n') = (m,m') + (n,n').$$

Now, taking infimum of all couples (m,m') and (n,n') satisfying $(\tilde{x}^1 - m\mathbf{1}, \tilde{y}^1 - m'\mathbf{1}) \in \mathcal{A}$, $(\tilde{x}^2 - n\mathbf{1}, \tilde{y}^2 - n'\mathbf{1}) \in \mathcal{A}$, on both sides of above inequality yields

$$\rho_{\mathcal{A}}\left[(\tilde{x}^1,\tilde{y}^1) + (\tilde{x}^2,\tilde{y}^2)\right] \leq_{co} \rho_{\mathcal{A}}\left[(\tilde{x}^1,\tilde{y}^1)\right] + \rho_{\mathcal{A}}\left[(\tilde{x}^2,\tilde{y}^2)\right].$$

So, comonotonic subadditivity property of $\rho_{\mathcal{A}}$ was proved.
To show the permutation invariance property of $\rho_{\mathcal{A}}$, fix any $(\tilde{x},\tilde{y}) \in \mathbb{R}^n \times \mathbb{R}^n$ and any permutation $(i_1,\ldots,i_n)$. Then for any $m \in \mathbb{R}$, Axiom $C6$ implies that $(\tilde{x} - m\mathbf{1}, \tilde{y} - m'\mathbf{1}) \in \mathcal{A}$ if and only if $(x_{i_1} - m, \ldots, x_{i_n} - m, y_{i_1} - m', \ldots, y_{i_n} - m') \in \mathcal{A}$. Hence,

$$\{m \mid (\tilde{x} - m\mathbf{1}, \tilde{y} - m'\mathbf{1})\} = \{m \mid (x_{i_1} - m, \ldots, x_{i_n} - m, y_{i_1} - m', \ldots, y_{i_n} - m') \in \mathcal{A}\}.$$

Now, if we take infimum on both sides then, we have

$$\rho_{\mathcal{A}}\left[(\tilde{x},\tilde{y})\right] = \rho_{\mathcal{A}}\left[(x_{i_1}, x_{i_2}, \ldots, x_{i_n}, y_{i_1}, y_{i_2}, \ldots, y_{i_n})\right].$$

This proves the permutation invariance property of $\rho_{\mathcal{A}}$.

□

4.5. Examples

In this section, we present examples of different multivariate data-based risk measures. Reducing our statistics to the univariate framework allows to retrieve definitions found in [53]. We define Multivariate Value-at-Risk (MVaR), Multivariate

Tail Conditional Median (MTCM) and Multivariate Tail Conditional Expectation (MTCE). In this section we assume that for a given multivariate data-set, there is only one associated order statistics. For the general case, we follow the procedure explained in Remark 4.3.2. Consider $(\tilde{x}^{(1)}, ..., \tilde{x}^{(d)}) \in \mathbb{R}^n \times \cdots \times \mathbb{R}^n$ and let $\{(x^{(1)}, \ldots, x^{(d)})_{(1)}, \ldots, (x^{(1)}, \ldots, x^{(d)})_{(n)}\}$ be the order statistics of the observation $(\tilde{x}^{(1)}, ..., \tilde{x}^{(d)})$ with respect to the couple ordering introduced in Section 4.3. Here, we give the definitions for MVaR, MTCM and MTCE when data represent losses of a company with different dependent departments/portfolios. Traditionally, when using returns, we first convert the data into losses. Then, we apply the desired multivariate data-based risk measures. Suppose $(\tilde{x}^{(1)}, ..., \tilde{x}^{(d)})$ represents returns of a company and $\rho_\alpha\left[(\cdot, \ldots, \cdot)\right]$ is a multivariate data-based risk measure defined for loss data. Then, the risk measure $\tilde{\rho}_\alpha\left[(\cdot, \ldots, \cdot)\right]$ defined for return data is defined by

$$\tilde{\rho}_\alpha\left[(\tilde{x}^{(1)}, ..., \tilde{x}^{(d)})\right] = \rho_{(1-\alpha)}\left[(-\tilde{x}^{(1)}, ..., -\tilde{x}^{(d)})\right], \tag{4.5.1}$$

for $0 \leq \alpha \leq 1$. This definition is compatible with the representation of univariate risk measures defined on a set of random variables representing returns, based on a risk measure defined on a set of random variables representing losses, as detailed in [45].

4.5.1. MVaR, MTCM and MTCE

Multivariate Value-at-Risk: We define $\mathrm{MVaR}_\alpha\left[(\tilde{x}^{(1)}, ..., \tilde{x}^{(d)})\right]$ by

$$\mathrm{MVaR}_\alpha\left[(\tilde{x}^{(1)}, ..., \tilde{x}^{(d)})\right] = \inf_{1\leq i\leq n,\ \alpha\leq \frac{i}{n}} (x^{(1)}, \ldots, x^{(d)})_{(i)}, \tag{4.5.2}$$

where "inf" is taken under couple ordering. This definition is compatible with representation (4.3.15) for multivariate data-based risk measures. In this case, the set of weights $\mathcal{W}$ is a singleton and its element is chosen by setting $w_i = 1$ and $w_j = 0$ for $1 \leq i \leq n$ and $j \neq i$.

Multivariate Tail Conditional Median: We define $\mathrm{MTCM}_\alpha\left[(\tilde{x}^{(1)}, ..., \tilde{x}^{(d)})\right]$ as the MVaR at level $\frac{\alpha+1}{2}$, that is

$$\mathrm{MTCM}_\alpha\left[(\tilde{x}^{(1)}, ..., \tilde{x}^{(d)})\right] = \mathrm{MVaR}_{\frac{\alpha+1}{2}}\left[(\tilde{x}^{(1)}, ..., \tilde{x}^{(d)})\right] = \inf_{1\leq i\leq n,\ \frac{\alpha+1}{2}\leq \frac{i}{n}} (x^{(1)}, \ldots, x^{(d)})_{(i)}. \tag{4.5.3}$$

This definition is compatible with representation (4.3.15) for multivariate data-based risk measures. The set of weights $\mathcal{W}$ is a singleton and its element is chosen by setting $w_i = 1$ and $w_j = 0$ for $1 \leq i \leq n$ and $j \neq i$.

Multivariate Tail Conditional Expectation: We define $\mathrm{MTCE}_\alpha\left[(\tilde{x}^{(1)},...,\tilde{x}^{(d)})\right]$ by

$$\mathrm{MTCE}_\alpha\left[(\tilde{x}^{(1)},...,\tilde{x}^{(d)})\right] = \tag{4.5.4}$$
$$\inf_{1\leq i\leq n,\ \alpha\leq\frac{i}{n}} \frac{1}{1-\alpha}\left[\left(\frac{i}{n}-\alpha\right)\mathrm{MVaR}_\alpha\left[(\tilde{x}^{(1)},...,\tilde{x}^{(d)})\right]+\frac{1}{n}\sum_{j=i+1}^{n}\mathrm{MVaR}_{\frac{j}{n}}\left[(\tilde{x}^{(1)},...,\tilde{x}^{(d)})\right]\right].$$

This definition is also compatible with representation (4.3.15) for multivariate data-based risk measures. We obtain this element by choosing $w_i = 0$ and $w_j = \frac{1}{n(1-\alpha)}$ for $i+1 \leq j \leq n$ for $1 \leq j \leq i-1$.

Remark 4.5.1. *For the case of having more than one version for the order statistics associated to* $(\tilde{x}^{(1)},...,\tilde{x}^{(d)})$, *we take the* inf *over different order statistics in equations* (4.5.2), (4.5.3) *and* (4.5.4). *Then we follow the criterion proposed in Subsection 4.3.2 to choose the right risk vector for MVaR, MTCM and MTCE respectively.*

4.5.2. Numerical illustrations

In this section, we consider portfolios of dependent stocks. Illustration 1 presents a portfolio with 2 stocks: Intel Corp. (`http://finance.yahoo.com/q/hp?s=INTC&a=08&b=20&c=2010&d=07&e=26&f=2013&g=w`) and Apple Inc. (`http://finance.yahoo.com/q/hp?a=08&b=20&c=2010&d=07&e=26&f=2013&g=w&s=AAPL%2C+&ql=1`), respectively. Illustration 2 presents a portfolio composed of 3 stocks: Apple Inc., Intel Corp. and Pfizer Inc. (PFE) (`http://finance.yahoo.com/q/hp?a=08&b=20&c=2010&d=07&e=26&f=2013&g=w&s=PFE%2C+&ql=1`), respectively. We use close data ranged from 20/09/2010 to 26/08/2013 in both examples. Weekly close data are converted into log returns, i.e. if we consider P_n as the close price for the n^{th} week then the log return is $R_n = \ln P_n - \ln P_{n-1}$. We use negative returns to calculate MVaR, MCTM and MTCE for weekly losses. Finally, in Illustration 3, we present a portfolio of two dependent risks, representing losses. We use a bivariate Gumbel copula model, with varying dependence parameter.

4.5.2.1. *Illustration 1*

In Table 4.1, we provide numerical results for three different bivariate data-based risk measures for weekly losses (negative returns) data of Apple Inc. and Intel Corp.

α	$\mathrm{MVaR}_\alpha\left[(\tilde{x}^{(1)},\tilde{x}^{(2)})\right]$	$\mathrm{MTCM}_\alpha\left[(\tilde{x}^{(1)},\tilde{x}^{(2)})\right]$	$\mathrm{MTCE}_\alpha\left[(\tilde{x}^{(1)},\tilde{x}^{(2)})\right]$	$(\mathrm{MTCE}_\alpha-\mathrm{MTCM}_\alpha)/\mathrm{MTCE}_\alpha$
99.5%	(0.1281,0.0137)	(0.1281,0.0137)	(0.1281,0.0137)	(0.0000,0.0000)
99.0%	(0.0572,0.0733)	(0.1281,0.0137)	(0.1035,0.0344)	(-0.2377,0.6017)
95.0%	(0.0308,0.0664)	(0.0441,0.0715)	(0.0606,0.0573)	(0.2723,-0.2478)
50.0%	(0.0234,-0.0266)	(-0.0065, 0.0445)	(0.0266,0.0169)	(1.2444,-1.6331)

Tab. 4.1. MVaR, MTCM and MTCE for Apple Inc. and Intel Corp. weekly losses (negative returns) from 20/09/2010 to 26/08/2013.

Remark 4.5.2. *We can also directly compute bivariate data-based risk measures $MVaR_\alpha\left[(\cdot,\cdot)\right]$, $MTCM_\alpha\left(\cdot,\cdot\right)]$ and $MTCE_\alpha\left[(\cdot,\cdot)\right]$ at arbitrary levels $\alpha\in[0,1]$ for data sets representing returns, from (4.5.1). Consider $(\tilde{u},\tilde{w})$ as the returns for Apple Inc. and Intel Corp., respectively. For example, suppose $\alpha=0.95$. Then,*

(1) $\widetilde{MVaR}_{0.05}\left[(\tilde{u},\tilde{w})\right]=MVaR_{0.95}\left[(\tilde{x}^{(1)},\tilde{x}^{(2)})\right]=(0.0308,0.0664)$,

(2) $\widetilde{MTCM}_{0.05}\left[(\tilde{u},\tilde{w})\right]=MTCM_{0.95}\left[(\tilde{x}^{(1)},\tilde{x}^{(2)})\right]=(0.0441,0.0715)$, *and*

(3) $\widetilde{MTCE}_{0.05}\left[(\tilde{u},\tilde{w})\right]=MTCE_{0.95}\left[(\tilde{x}^{(1)},\tilde{x}^{(2)})\right]=(0.0606,0.0573)$.

Remark 4.5.3. *Based on numerical results of Table 4.1, we illustrate the positive homogeneous and component wise translation invariant property for bivariate data-based risk measures. Let 0.1281 be the larger loss (-0.1281 being the lower return) data for the two stocks combined, and suppose $\alpha=.95$. Then, we have*

(1) $MVaR_{0.95}\left[(\tilde{x}^{(1)},\tilde{x}^{(2)})+0.1281\,(\mathbf{1},\mathbf{1})\right]=MVaR_{0.95}\left[(\tilde{x}^{(1)},\tilde{x}^{(2)})\right]+0.1281(1,1)=(0.1589,0.1945)$,

(2) $MTCM_{0.95}\left[(\tilde{x}^{(1)},\tilde{x}^{(2)})+0.1281\,(\mathbf{1},\mathbf{1})\right]=MTCM_{0.95}\left[(\tilde{x}^{(1)},\tilde{x}^{(2)})\right]+0.1281(1,1)=(0.1722,0.1996)$, *and*

(3) $MTCE_{0.95}\left[(\tilde{x}^{(1)},\tilde{x}^{(2)})+0.1281\,(\mathbf{1},\mathbf{1})\right]=MTCE_{0.95}\left[(\tilde{x}^{(1)},\tilde{x}^{(2)})\right]+0.1281(1,1)=(0.1887,0.1854)$.

4.5.2.2. *Illustration 2*

Table 4.2 provides numerical results for multivariate data-based risk measures, introduced at the beginning of this section, for weekly losses of three different stocks Apple Inc., Intel Corp. and Pfizer Inc. Multivariate data-based risk measures are easily extendable to higher dimensions. MVaR, MTCM and MTCE are fast to compute.

α	$\mathrm{MVaR}_\alpha\left[(\tilde{x}^{(1)},\tilde{x}^{(2)},\tilde{x}^{(3)})\right]$	$\mathrm{MTCM}_\alpha\left[(\tilde{x}^{(1)},\tilde{x}^{(2)},\tilde{x}^{(3)})\right]$	$\mathrm{MTCE}_\alpha\left[(\tilde{x}^{(1)},\tilde{x}^{(2)},\tilde{x}^{(3)})\right]$	$(\mathrm{MTCE}_\alpha-\mathrm{MTCM}_\alpha)/\mathrm{MTCE}_\alpha$
99.5%	(-0.0026,0.0097,0.3956)	(-0.0026,0.0097,0.3956)	(-0.0026,0.0097,0.3956)	(0.0000,0.0000,0.0000)
99.0%	(0.00001,-0.0056,0.3497)	(-0.0026,0.0097,0.3956)	(-0.0017,0.0044,0.3797)	(-0.5294,-1.2045,-0.0419)
95.0%	(0.0646,0.0171,0.0333)	(-0.0177,-0.0054,0.3575)	(0.0129,0.0159,0.2441)	(2.3721,1.3396,-0.4646)
50.0%	(-0.0188,-0.0005,0.0124)	(0.0958,-0.0345,-0.0258)	(0.0220,0.0158,0.0232)	(-3.3545,3.1835,2.1121)

Tab. 4.2. MVaR, MTCM and MTCE for Apple Inc., Intel Corp. and Pfizer Inc. weekly losses (negative returns) from 20/09/2010 to 26/08/2013.

Table 4.1 and Table 4.2 show that, for fixed level α and different number of stocks in two similar portfolios, multivariate data-based risk statistics are varying. This can be interpreted in terms of risk diversification. In fact, aggregated risks are diversified and global risk is distributed between each component of the portfolio. When the dimension of data changes, the risk associated to each element of the portfolio changes as well.

4.5.2.3. *Illustration 3*

Here, we consider a bivariate Gumbel copula as benchmark model. The bivariate Gumbel copula is defined by

$$C_\theta^{GC}(u_1, u_2) = \exp\left[-\left((-\ln u_1)^\theta + (-\ln u_2)^\theta\right)^{\frac{1}{\theta}}\right],$$

where $\theta > 1$ and $0 \leq u_1, u_2 \leq 1$, as presented in [28]. It can be shown that for the Gumbel copula the coefficient of upper tail dependence is $\lambda_u = 2 - 2^{\frac{1}{\theta}}$. This shows that the Gumbel copula has upper tail dependence when $\theta > 1$.
We simulate using MATLAB function "copularnd('Gumbel',θ, N)" where θ is the dependence parameter for the Gumbel copula and N is the number of returns. We simulate 30 times, with $N = 100$, $\theta = 1.6$ (Table 4.3) and $\theta = 20$ (Table 4.4). We obtain 3000 returns. We calculate $\mathrm{MVaR}_\alpha\left[(\tilde{x}^{(1)}, \tilde{x}^{(2)})\right]$, $\mathrm{MTCM}_\alpha\left[(\tilde{x}^{(1)}, \tilde{x}^{(2)})\right]$ and $\mathrm{MTCE}_\alpha\left[(\tilde{x}^{(1)}, \tilde{x}^{(2)})\right]$ by averaging the statistics obtained for each simulation.

α	$\mathrm{MVaR}_\alpha\left[(\tilde{x}^{(1)}, \tilde{x}^{(2)})\right]$	$\mathrm{MTCM}_\alpha\left[(\tilde{x}^{(1)}, \tilde{x}^{(2)})\right]$	$\mathrm{MTCE}_\alpha\left[(\tilde{x}^{(1)}, \tilde{x}^{(2)})\right]$	$(\mathrm{MTCE}_\alpha - \mathrm{MTCM}_\alpha)/\mathrm{MTCE}_\alpha$
99.5%	(0.9730,0.9840)	(0.9730,0.9840)	(0.9730,0.9840)	(0.0000,0.0000)
99.0%	(0.9475,0.9560)	(0.9730,0.9840)	(0.9730,0.9840)	(0,0)
95.0%	(0.8985,0.8641)	(0.9094,0.9475)	(0.9210,0.9512)	(0.0126,0.0039)
50.0%	(0.5893,0.3333)	(0.7359,0.5812)	(0.6910,0.6932)	(-0.0650,0.1616)

Tab. 4.3. MVaR, MTCM and MTCE for the bivariate Gumbel copula with dependence parameter $\theta = 1.6$.

α	$\mathrm{MVaR}_\alpha\left[(\tilde{x}^{(1)}, \tilde{x}^{(2)})\right]$	$\mathrm{MTCM}_\alpha\left[(\tilde{x}^{(1)}, \tilde{x}^{(2)})\right]$	$\mathrm{MTCE}_\alpha\left[(\tilde{x}^{(1)}, \tilde{x}^{(2)})\right]$	$(\mathrm{MTCE}_\alpha - \mathrm{MTCM}_\alpha)/\mathrm{MTCE}_\alpha$
99.5%	(0.9931,0.9932)	(0.9931,0.9932)	(0.9931,0.9932)	(0,0)
99.0%	(0.9845,0.9836)	(0.9931,0.9932)	(0.9931,0.9932)	(0,0)
95.0%	(0.9277,0.9216)	(0.9656,0.9664)	(0.9667 0.9664)	(0.0011,0)
50.0%	(0.4872, 0.4845)	(0.7521 0.7511)	(0.7393,0.7401)	(-0.0076, -0.0027)

Tab. 4.4. MVaR, MTCM and MTCE for the bivariate Gumbel copula with dependence parameter $\theta = 20$.

Remark 4.5.4. *In Tables 4.3 and 4.4, we clearly see that for a fixed level of risk α, each one of MVaR, MTCM and MTCE are increasing functions of the dependence parameter θ. Hence,*

$$\begin{aligned} MVaR_{\alpha}^{\theta_1}\left[(\tilde{x}^{(1)},\tilde{x}^{(2)})\right] &\leq_{co} MVaR_{\alpha}^{\theta_2}\left[(\tilde{u},\tilde{v})\right], \\ MTCM_{\alpha}^{\theta_1}\left[(\tilde{x}^{(1)},\tilde{x}^{(2)})\right] &\leq_{co} MTCM_{\alpha}^{\theta_2}\left[(\tilde{u},\tilde{v})\right], \\ MTCE_{\alpha}^{\theta_1}\left[(\tilde{x}^{(1)},\tilde{x}^{(2)})\right] &\leq_{co} MTCE_{\alpha}^{\theta_2}\left[(\tilde{u},\tilde{v})\right], \end{aligned} \tag{4.5.5}$$

where $(\tilde{x}^{(1)},\tilde{x}^{(2)})$ and $(\tilde{u},\tilde{v})$ are two bivariate data sets simulated from Gumbel copula with the same marginal, for different θ_1 and θ_2 respectively.
Since Gumbel copula has upper tail dependence when $\theta > 1$, we see from the last column of Table 4.3 and Table 4.4 that when the dependence parameter $\theta \to \infty$, the components of a bivariate data-based risk measure at a fixed level of risk α get closer to the same value. This is due to the upper tail dependence of Gumbel copula.

4.5.2.4. *General Remark*

Remark 4.5.5. *Let $(\tilde{x}^{(1)},...,\tilde{x}^{(d)})$ be a multivariate data set. Then, the multivariate data-based risk measures introduced in this section are increasing functions w.r.t the risk level α. That is,*

$$\begin{aligned} MVaR_{\alpha_1}\left[(\tilde{x}^{(1)},...,\tilde{x}^{(d)})\right] &\leq_{co} MVaR_{\alpha_2}\left[(\tilde{x}^{(1)},...,\tilde{x}^{(d)})\right], \\ MTCM_{\alpha_1}\left[(\tilde{x}^{(1)},...,\tilde{x}^{(d)})\right] &\leq_{co} MVaR_{\alpha_2}\left[(\tilde{x}^{(1)},...,\tilde{x}^{(d)})\right], \\ MTCE_{\alpha_1}\left[(\tilde{x}^{(1)},...,\tilde{x}^{(d)})\right] &\leq_{co} MVaR_{\alpha_2}\left[(\tilde{x}^{(1)},...,\tilde{x}^{(d)})\right], \end{aligned}$$

for $0 \leq \alpha_1 \leq \alpha_2 \leq 1$. Moreover, for a fixed level of risk α, we have

$$\begin{aligned} MVaR_{\alpha}\left[(\tilde{x}^{(1)},...,\tilde{x}^{(d)})\right] &\leq_{co} MTCM_{\alpha}\left[(\tilde{x}^{(1)},...,\tilde{x}^{(d)})\right], \\ MVaR_{\alpha}\left[(\tilde{x}^{(1)},...,\tilde{x}^{(d)})\right] &\leq_{co} MTCE_{\alpha}\left[(\tilde{x}^{(1)},...,\tilde{x}^{(d)})\right], \end{aligned} \tag{4.5.6}$$

where $\leq_{co}$ is the ordering introduced in Section 4.3. The first inequality in (4.5.6) can be driven using monotonicity property of $MVaR_{\beta}\left[(\cdot,...,\cdot)\right]$. Replacing $\beta = \alpha$ and $\beta = \frac{\alpha+1}{2}$ in $MVaR_{\beta}\left[(\cdot,...,\cdot)\right]$ yields the inequality. The second inequality in (4.5.6) can be also driven from monotonicity property of $MVaR_{\beta}\left[(\cdot,...,\cdot)\right]$ and definition of $MTCE_{\beta}\left[(\cdot,...,\cdot)\right]$. We refer to Tables 4.1, 4.3 and Table 4.4 for a numerical verification of (4.5.6) for bivariate data-based risk measures.

Remark 4.5.6. *In each illustration, we observe that $MVaR_{.995} = MTCM_{.99}$, which is consistent with our definition of $MTCM_{\frac{\alpha+1}{2}} = MVaR_{\alpha}$.*

Chapter 5

ON THE DEPLETION PROBLEM FOR AN INSURANCE RISK PROCESS: NEW NON-RUIN QUANTITIES IN COLLECTIVE RISK THEORY

Abstract

The field of risk theory has traditionally focused on ruin-related quantities. In particular, the so-called Expected Discounted Penalty Function [50] has been the object of a thorough study over the years. Although interesting in their own right, ruin related quantities do not seem to capture path-dependent properties of the reserve. In this article we aim at presenting the probabilistic properties of drawdowns and the speed at which an insurance reserve depletes as a consequence of the risk exposure of the company. These new quantities are not ruin related yet they capture important features of an insurance position and we believe it can lead to the design of a meaningful risk measures. Studying drawdowns and speed of depletion for Lévy insurance risk processes represent a novel and challenging concept in insurance mathematics. In this chapter, all these concepts are formally introduced in an insurance setting. Moreover, using recent results in fluctuation theory for Lévy processes [71], we derive expressions for the distribution of several quantities related to the depletion problem. Of particular interest are the distribution of drawdowns and the Laplace transform for the speed of depletion. These expressions are given for some examples of Lévy insurance risk processes for which they can be calculated, in particular for the classical Cramer-Lundberg model.

This chapter is a joint research work with Zied Ben-Salah, Hélène Guérin and Manuel Morales; see [15] [1].

[1] I contributed in this project in different ways. Prof. Morales came up with the general idea of the project, then I worked on it to derive different solutions for quantities of interest in the project. Applying main results of the project to different examples and interpreting the obtained solutions were other contributions I have made to this project. I also contributed in this project by writing up the first draft of it, then polishing it with Prof. Morales, Prof. Guérin and Dr. Ben-Salah to get the final version.

5.1. Introduction

Traditionally, collective risk theory is mainly concerned with the ruin problem which is nicely encapsulated in the concept of Expected Discounted Penalty Function (EDPF) introduced in [50]. This so-called Gerber-Shiu function is a functional of the ruin time (i.e., the first time the reserve level of a firm becomes negative), the surplus prior to ruin, and the deficit at ruin. The EDPF has been extensively studied and generalized to various scenarios and there is now a wide range of models for which expressions of the EDPF are available. All of these models incorporate different levels of complexity into the picture.

In particular, the so-called Lévy insurance risk processes have been the object of much attention in the last decade, mainly because they nicely generalize the Cramer-Lundberg model while allowing to bring new insight into the field of ruin theory through the well-developed theory of fluctuations for such processes. Several families of Lévy processes have been put forward as risk models and we now have a well-established literature on the subject. For a thorough discussion on the suitability of these processes as risk models we refer the reader to [49, 72] and references therein.

As it turns out, the first-passage problem for Lévy processes is well understood and recent results in this area have been applied to the ruin problem in order to gain interesting insight (see for instance [19, 20, 60]). In this chapter, we focus yet again on Lévy insurance risk processes because of the extensive set of tools available for this family of stochastic processes. Through concepts originally developed for the study of the first-passage time problem, we can now study questions that go beyond the ruin problem and that are connected to path-properties of the process that give a telling picture of how depletion occurs.

Quantities such as the speed of depletion and drawdowns have been studied in finance in connection to the concept of market crash [93]. Indeed, in finance one would be interested in knowing how fast and how frequent drawdowns of a certain size occur. In insurance, these questions have not been studied yet, despite the fact that these concepts are meaningful from an insurance risk management point of view. Clearly knowing how your insurance reserve is affected by drawdowns and how fast and frequent these are could be useful to devise risk management tools. These quantities provide a measure of riskiness that is not linked to the ruin event but rather to the depletion features of the reserve. However, this problem is technically challenging due to the jump nature of insurance models.

The aim of this chapter is two-fold. On one hand, we aim at introducing the problem of depletion into the theory of collective risk theory as a meaningful question from a risk management point of view. We formally define new non-ruin quantities within the classical risk theory framework and we discuss their main features and advantages over traditional ruin-related quantities. Indeed, it is interesting to notice that all of the available research focuses on ruin-related quantities which, by their very nature, fail to explain how an insurance reserve depletes over time. Thus, although ruin theory provides a good probabilistic picture of the problem of insolvency of an insurance reserve, it cannot explain other features that are equally representative of the riskiness of an insurance reserve such as its speed of depletion and the frequency of drawdowns. The ruin event is an object of concern over the long-run but a risk manager might also keep an eye on any series of particularly large drawdowns especially if they happen particularly fast. So concerning oneself with the ruin event overlooks other risky events that also have an impact on the solvency and financial planning of an insurance company.

A second objective is to actually derive expressions for the distribution of several depletion-related random variables. As it turns out, recent results in the theory of fluctuations for one-sided Lévy processes [71] can be used to derive expressions for these depletion-related quantities. Key to the derivation of such expressions is the scale function of the process driving the insurance risk model. As we discuss, general and non-explicit expressions for the distribution of random variables in the depletion problem can only be simplified if a simple form for the scale function is available. Hence, we derive explicit expressions for the case of a classical Cramer-Lundberg model driven by a compound Poisson process with exponential jumps. Not surprisingly, this simple case has always yielded text-book examples of closed-form solutions in the risk theory literature. The problem of depletion is no exception and we present explicit expressions for the distribution of several depletion-related random variables in this case. We also provide a similar analysis for other examples of insurance risk models possessing simple scale functions, namely the gamma subordinator and the stable family of processes.

This chapter is organized as follows. In Section 5.2 we introduce a general model based on a Lévy risk process for which we define the depletion problem and the notions of *drawdowns* and *speed of depletion* as well as related variables. Some preliminary results from the theory of fluctuations for Lévy processes are given in Section 5.3. In Section 5.4 we study the problem of depletion for an insurance Lévy risk process and we give general expressions for the distribution of depletion random variables of interest. In Section 5.5, we derive explicit expressions for all depletion-related quantities for three examples of Lévy insurance risk processes.

Finally, in Section 5.6, we simulate some of risk models which discussed in Section 5.5, and study the empirical results for these risk models.

5.2. The Depletion Problem for an Insurance Risk Model

We consider a very general setup that generalizes the standard Cramer-Lundberg model. We consider in this chapter an insurance risk process $X = (X_t)_{t \geqslant 0}$ with X a spectrally negative Lévy process defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geqslant 0}, \mathbb{P})$ and $X_0 = 0$ $\mathbb{P} - a.s.$ For technical reasons that will become clear later in the chapter, we restrict ourselves to those processes having paths of unbounded variation or paths of bounded variation as well as a Lévy measure which is absolutely continuous with respect to the Lebesgue measure. In order to avoid the case of trivial reflected processes we exclude processes with monotone paths. As is customary, the symbols $\mathbb{E}_x$ and $\mathbb{P}_x$ will denote the expectation and the probability measure related to the process started at x, i.e. the expectation and law of the process $x + X$ under $\mathbb{P}$ and if the process is started from zero we will use simple notations $\mathbb{E}$ and $\mathbb{P}$.

Notice that a such a model contains all elements of a traditional risk model and encompasses, among others, the risk models studied in [19, 48, 55, 72]. Indeed, the constant rate premium is included as the drift of X, the so-called perturbation comes in as the Brownian component of X and the pure aggregate claims is present as the jump part of X, which could be set as a compound Poisson or an infinite activity process. With this in mind, we assume the process X to have a positive drift such that $\mathbb{E}[X_1] > 0$. Notice that in traditional ruin theory, this assumption responds to both, technical and practical reasons. Technically, it is needed in order to avoid the possibility that X becomes negative almost surely whereas from a practical point of view it makes sense since it is common practice in insurance to work with loaded premiums. Indeed, it is standard to write the drift component within X in terms of a safety loading. For instance, notice that we can recuperate the classical Cramer-Lundberg model if $X_t = ct - S_t$ where $c := (1+\theta)\mathbb{E}[S_1]$ and S is a compound Poisson process modeling aggregate claims. The drift c, with a positive safety loading $\theta > 0$, is the collected premium rate. In the context of the depletion problem we do not need this condition. We keep it here for purely practical reasons as it is common practice to have insurance loaded premiums.

One of the advantages of considering a general Lévy risk model is that we can use the tools and methods of the fluctuation theory of Lévy processes, allowing for a somewhat deeper understanding of the ruin problem but also of the depletion problem which can prove to yield just as interesting information about the riskiness of the reserve.

For a more extensive discussion on Lévy risk models we refer to [49]. In this chapter we will specialize this setting to three examples of Lévy processes that have been studied in the literature in the context of the ruin problem.

One of the main objects of interest in ruin theory is the *ruin time*, τ, representing the first passage time of an insurance Lévy risk process X below zero when $X_0 = x$, i.e.

$$\tau := \inf\{t > 0 \ : \ X_t < 0\}, \tag{5.2.1}$$

where we set $\tau = +\infty$ if $X_t \geq 0$ for all $t \geqslant 0$.

In this chapter, our main object of concern is the depletion problem that has two different random times as its main building blocks. In order to give a thorough definition of these concepts we need to introduce some notation.

We define the running infimum and the running supremum of a given Lévy process X by

$$\underline{X}_t := \inf_{0 \leqslant s \leqslant t} X_s \qquad \text{and} \qquad \overline{X}_t := \sup_{0 \leqslant s \leqslant t} X_s \ .$$

Now we characterize the depletion problem for X. We first define the *drawdown* process $Y = (Y_t)_{t \geqslant 0}$, associated with a given risk process X, to be

$$Y_t := \overline{X}_t - X_t \ , \qquad t \geqslant 0 \ . \tag{5.2.2}$$

The first-passage time over a level $a > 0$ of the drawdown process Y is then defined to be

$$\tau_a := \inf\{t \geqslant 0 : Y_t > a\} \ . \tag{5.2.3}$$

It is well-known that $\tau_a < \infty$ $\mathbb{P}$-almost surely (see [12], Theorem 1). Just like the ruin time in (5.2.1), this new random time in (5.2.3) contains relevant information on potentially risky behavior of the reserve. Their distribution can be used to measure the likeliness of path-related events that might have a negative impact on the financial health of the reserve. The random time τ_a records the time at which a drawdown in the reserve is larger than a, previously agreed upon, critical level a. An interesting set of associated tale-telling random variables can be built upon the random time (5.2.3). First, we need to define a process that will be useful in constructing meaningful non-ruin quantities. The last time before t that X reaches its running supremum, denoted by $\overline{G}_t$, is defined as

$$\overline{G}_t := \sup\{s \leq t : X_s \text{ or } X_{s-} = \overline{X}_s\} \ . \tag{5.2.4}$$

Thus the time τ_a of the critical drawdown of size a along with the following quantities characterize the depletion problem for X:

- the last time the reserve was at its maximum level prior to critical drawdown, $\overline{G}_{\tau_a}$;
- the *speed of depletion*, $\tau_a - \overline{G}_{\tau_a}$;
- the maximum reserve level attained before critical drawdown is observed, $\overline{X}_{\tau_a}$;
- the minimum reserve level prior to critical drawdown, $\underline{X}_{\tau_a}$;
- the largest drawdown observed before critical drawdown of size a, $Y_{\tau_a -}$;
- the overshoot of the critical drawdown over level a, $Y_{\tau_a} - a$.

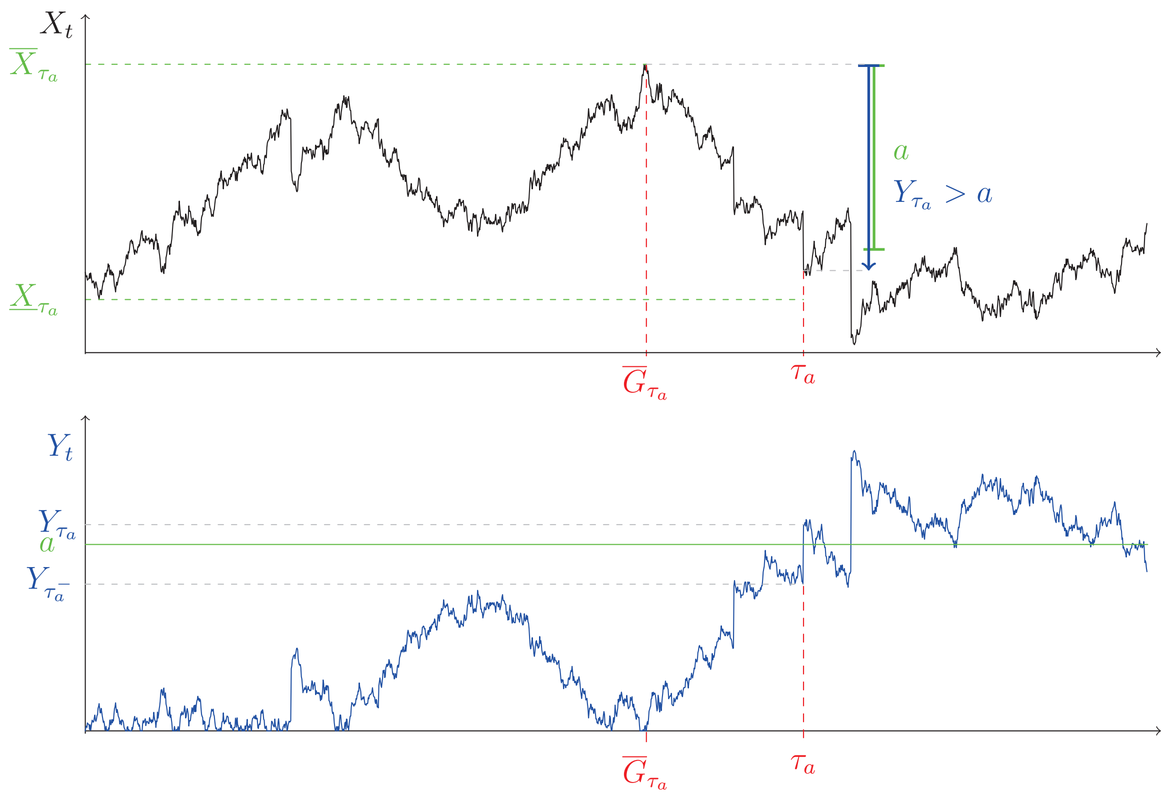


FIG. 5.1. A path of $X_t = 10 + t + 2B_t - S_t$, the corresponding drawdown process Y, and their related depletion quantities, where $(B_t)_{t\geq 0}$ is a standard Brownian motion and S is an independent compound Poisson process with Lévy measure $\nu(dx) = e^{-2y}dx$.

Clearly, these variables contain information on the how the insurance deserve depletes over time. All of these quantities encapsulate relevant knowledge about the critical drawdown event. A risk manager would be potentially interested in gaining information regarding the distribution of the time of the critical drawdown of size a, i.e. $\mathbb{P}_x(\tau_a \leqslant t)$. This gives information on how likely the reserve is to face a critical drawdown within a given time interval. Even more valuable information can be found in the distribution of the *speed of depletion*, this random variable indicates how fast critical drawdowns tend to occur. A drawdown is not as alarming if it happens over a long period than if it happens suddenly. Information about the distributions of the maximum and minimum reserve levels prior to critical drawdown and of the largest drawdown on record before critical drawdown sheds light on the structure of the depletion event. It is also interesting to know how large (or not) critical drawdowns tend to be, that is, by how much they overshoot the critical level a when they occur. In fact, the level a itself can be set by using the distribution of the overshoot. Since this distribution is a function of a we can decide what a critical drawdown size is depending on how likely certain levels are.

It is interesting to notice that there is a connection between ruin and depletion through the distribution of the minimum reserve level prior to critical drawdown. We will see that, if expressions are available, we can calculate the probability that ruin occurs before a critical drawdown of size a.

In general, just like in the ruin problem, knowledge on the probabilistic properties of such quantities could be relevant in risk management applications. Although critical drawdowns do not spell immediate doom for the company as ruin does, a large enough drawdown might be a warning sign that a risk manager might want to take into account. This information coupled with knowledge on how fast these critical drawdowns happen could be used to design risk measures and or management policies that will ensure the solvency of the reserve.

Once that these non-ruin quantities have been introduced, the aim of the chapter is to derive expressions for the probability measure of these random variables associated with the depletion problem. This will be done in detail for three examples of Lévy insurance risk processes. But before we need to introduce some preliminary results that are key to our analysis.

5.3. Drawdowns for Spectrally Negative Lévy Processes

In this section, we introduce some notions and results that are needed in the rest of the chapter. Let $X = (X_t)_{t\geqslant 0}$ be a spectrally negative Lévy process defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geqslant 0}, \mathbb{P})$. We impose the same restrictions as in [12], i.e. X has either paths of unbounded variation or paths of bounded

variation as well as a Lévy measure which is absolutely continuous with respect to the Lebesgue measure. Further, we exclude processes with monotone paths.
Since X has no positive jumps, the expectation $\mathbb{E}\Big[e^{sX_t}\Big]$ exists for all $s \geqslant 0$ and it is given by $\mathbb{E}\Big[e^{sX_t}\Big] = e^{t\psi(s)}$ where $\psi(s)$ is of the form

$$\psi(s) = d\,s + \frac{1}{2}\sigma^2 s^2 + \int_0^\infty (e^{-x\,s} - 1 + s\,x\mathbb{1}_{\{x<1\}})\,\nu(dx)\ , \tag{5.3.1}$$

where $d \in \mathbb{R}$, $\sigma > 0$ and ν is the Lévy measure associated with the process $-X$ (for a thorough account on Lévy process see [17, 63]).
For the right inverse of ψ, we shall write Φ on $[0, \infty)$. Formally, for each $q \geqslant 0$,

$$\Phi(q) := \sup\{s \geqslant 0 : \psi(s) = q\}\ . \tag{5.3.2}$$

Notice that since X is a spectrally negative Lévy process X, we have that $\Phi(q) > 0$ for $q > 0$ (see [63]).
It is well-known that, for every $q \geqslant 0$, there exists a function $W^{(q)} : \mathbb{R} \longrightarrow [0, \infty)$ such that $W^{(q)}(y) = 0$ for all $y < 0$ satisfying

$$\int_0^\infty e^{-\lambda y} W^{(q)}(y)dy = \frac{1}{\psi(\lambda) - q}, \qquad \lambda > \Phi(q). \tag{5.3.3}$$

This is the so-called *q-scale functions* $\{W^{(q)},\ q \geqslant 0\}$ of the process X (see [63]) and it is a key notion in the analysis of drawdowns for spectrally negative Lévy processes. Notice that for $q = 0$, equation (5.3.3) defines the so-called scale function and we simply write W.
Before discussing drawdowns, we need to introduce additional functions related to the q-scale function. Let $W_+^{\prime(q)}$ be the right derivative function of the q-scale. Following the notation in [71], we denote the ratio of the right derivative of the q-scale function and the q-scale function at $a > 0$ by

$$\lambda(a, q) := \frac{W_+^{\prime(q)}(a)}{W^{(q)}(a)}\ . \tag{5.3.4}$$

We can now define, for any $a > 0$ and $p, q > 0$, the mapping $F_{p,q,a} : \mathbb{R}_+ \to \mathbb{R}_+$

$$F_{p,q,a}(y) := \lambda(a, q)e^{-y\lambda(a,p)}\ . \tag{5.3.5}$$

Moreover, consider the q-resolvent measure $R_a^{(q)}(dy) = \mathbb{E}[\int_0^{\tau_a} e^{-qt}\mathbb{1}_{Y_t \in dy}dt]$ of Y killed upon first exit from $[0, a]$ which can be expressed in the following way (see [75], Theorem 1),

$$R_a^{(q)}(dy) := \Big[\lambda(a, q)^{-1} W^{(q)}(dy) - W^{(q)}(y)dy\Big]\ , \qquad y \in [0, a]\ , \tag{5.3.6}$$

and the function

$$\Delta^{(q)}(a) = \frac{\sigma^2}{2}\Big[W'^{(q)}(a) - \lambda(a,q)^{-1}W''^{(q)}(a)\Big] \tag{5.3.7}$$

with $\Delta^{(q)}(a) = 0$ when $\sigma = 0$.

The functions in (5.3.4), (5.3.5), (5.3.6) and (5.3.7) will frequently appear throughout the chapter. The following theorem will play a key role in our contribution. For a thorough discussion and a proof, we refer to [71].

Theorem 5.3.1. *Consider a spectrally negative Lévy process X such that $X_0 = x \in \mathbb{R}$. Moreover, X has paths of unbounded variation or has a Lévy measure which is absolutely continuous with respect to the Lebesgue measure. Let further Y be its associated drawdown process defined in (5.2.2). Let τ_a be the stopping time in (5.2.3) so we can define the following events, for a given $a > 0$,*

$$A_0 = \{\underline{X}_{\tau_a} \geqslant u, \overline{X}_{\tau_a} \in dv, Y_{\tau_a-} \in dy, Y_{\tau_a} - a \in dh\} \quad \text{and} \quad A_c = \Big\{\underline{X}_{\tau_a} \geqslant u, \overline{X}_{\tau_a} \in dv, Y_{\tau_a} = a\Big\}, \tag{5.3.8}$$

where u, v, y and h satisfy

$$u \leq x, \;\; y \in [0,a], \;\; v \geqslant x \vee (u+a) \; \text{and} \; h \in (0, v-u-a].$$

Then, for any $q, r \geqslant 0$ the following identities hold true:

$$\mathbb{E}_x\left[e^{-q\tau_a - r\overline{G}_{\tau_a}}\mathbb{1}_{A_0}\right] = \frac{W^{(q+r)}((x-u)\wedge a)}{W^{(q+r)}(a)}F_{q+r,q,a}(v-(x\vee(u+a)))R_a^{(q)}(dy)\nu(a-y+dh)dv, \tag{5.3.9}$$

$$\mathbb{E}_x\left[e^{-q\tau_a - r\overline{G}_{\tau_a}}\mathbb{1}_{A_c}\right] = \frac{W^{(q+r)}((x-u)\wedge a)}{W^{(q+r)}(a)}F_{q+r,q,a}(v-(x\vee(u+a)))\Delta^{(q)}(a)dv, \tag{5.3.10}$$

where $\mathbb{1}$ is the standard indicator function, ν is the Lévy measure of X that appears in (5.3.1), $x \vee y = \max(x,y)$, $x \wedge y = \min(x,y)$ and $\overline{G}_t$ is the process defined in (5.2.4).

We remark that the above theorem holds for spectrally negative Lévy process having paths of unbounded variation or having a Lévy measure which is absolutely continuous with respect to the Lebesgue measure. That is why we restrict ourselves to this type of processes. This is in no way restrictive since most of the risk insurance processes in the literature fall within this class, i.e. they are defined through a Lévy density.

We also remark that on the event A_0 defined in (5.3.8), the critical drawdown is performed by a jump of the Lévy process X while it is performed continuously on the event A_c. These two events and the expectations in (5.3.9) and (5.3.10) in Theorem 5.3.1 contain all information regarding the depletion problem. The aim of this chapter, to provide explicit expressions for the distribution of these depletion-related random variables under relevant insurance models.

5.4. Analysis of the Depletion Problem

In this section, we use the general setting described in Section 5.2 where X is a spectrally negative Lévy process either with paths of unbounded variation or paths of bounded variation with a Lévy measure absolutely continuous with respect to the Lebesgue measure. The main goal of this chapter is then to study the depletion event as told by the quantities in Theorem 5.3.1. In principle, we can study through the expectations in (5.3.9) and (5.3.10) the probability measure of all quantities involved as well as the Laplace transform of the speed of depletion. This can be accomplished by setting $q = r = 0$ and/or integrating over a suitable set those expressions in (5.3.9) and (5.3.10). How explicit these expressions are will depend on the form of the q-scale function and the Lévy measure of the model. Nonetheless, in this section we give some general results that bring insight into the problem.

It turns out that there is a link between the running infimum at time τ_a given by (5.2.3) and the ruin time τ given by (5.2.1). We can easily deduce that $\tau_a \leq \tau$ a.s. on the event $\{\underline{X}_{\tau_a} \geqslant 0\}$, while $\{\tau_a < \tau\}$ implies $\{\underline{X}_{\tau_a} \geqslant 0\}$.

Furthermore, from the definition of these quantities we can see that when the initial surplus x is strictly greater than a, then $\underline{X}_{\tau_a-} > 0$ and the hitting time of the critical drawdown is smaller than the ruin time, i.e. $\tau_a \leq \tau$ a.s. On the other hand, when $x < a$, ruin can occur before the critical drawdown.

We can now state a a result which makes a link between the ruin event and the depletion problem.

Theorem 5.4.1. *Consider an insurance risk process $(X_t)_{t\geqslant 0}$ with initial surplus $x \geqslant 0$ satisfying assumptions of Section 5.2 and let $a > 0$ be a fixed critical drawdown size. Then*

$$\mathbb{P}_x(\underline{X}_{\tau_a} < 0) = 1 - \frac{W(x\wedge a)}{W(a)} + \frac{W(x\wedge a)}{W(a)} \int_{y\in[0,a]} \int_{h>0} \left(1 - e^{-\lambda(a,0)h}\right) \nu(x \vee a - y + dh) R_a^{(0)}(dy), \tag{5.4.1}$$

where W is the scale function, ν the Lévy measure of X and $R_a^{(0)}$ is defined by (5.3.6) with $q = 0$.

Proof. We notice that $\mathbb{P}_x(\underline{X}_{\tau_a} < 0) = 1 - \mathbb{P}_x(\underline{X}_{\tau_a} \geqslant 0)$ and

$$\mathbb{P}_x(\underline{X}_{\tau_a} \geqslant 0) = \mathbb{P}_x(\underline{X}_{\tau_a} \geqslant 0,\, Y_{\tau_a} > a) + \mathbb{P}_x(\underline{X}_{\tau_a} \geqslant 0,\, Y_{\tau_a} = a).$$

Then putting $q = r = u = 0$ and integrating (5.3.9) and (5.3.10) with respect to $v \in [x \vee a, \infty)$, $y \in [0, a]$ and $h \in (0, v - a]$, we have

$$\mathbb{P}_x(\underline{X}_{\tau_a} \geqslant 0) = \frac{W(x\wedge a)}{W(a)} \left[\int_{y\in[0,a]} \left(\int_{v\geqslant x\vee a} F_{0,0,a}(v - x \vee a) \int_{h\in(0,v-a]} \nu(a - y + dh) dv \right) R_a^{(0)}(dy) + \int_{v\geqslant x\vee a} F_{0,0,a}(v - x\vee a) \Delta^{(0)}(a) dv \right].$$

Since $F_{0,0,a}$ is defined by (5.3.5), using Fubini's theorem, the previous expression gives

$$\mathbb{P}_x(\underline{X}_{\tau_a} \geqslant 0) = \frac{W(x \wedge a)}{W(a)} \left[e^{\lambda(a,0)x \vee a} \int_{y \in [0,a]} \int_{h>0} e^{-\lambda(a,0)(x \vee (h+a))} \nu(a-y+dh) R_a^{(0)}(dy) + \Delta^{(0)}(a) \right]. \tag{5.4.2}$$

We notice that taking $u \to -\infty$ and integrating (5.3.9) and (5.3.10) with respect to $v \in [x \vee a, \infty)$, $h \in (0, v-a]$ and $y \in [0,a]$ with $r = q = 0$,

$$\int_0^a R_a^{(0)}(dy) \int_0^\infty \nu(a-y+dh) + \Delta^{(0)}(a) = 1. \tag{5.4.3}$$

Using this remark, we deduce that (5.4.2) gives for $x \leq a$

$$\begin{aligned} \mathbb{P}_x(\underline{X}_{\tau_a} \geqslant 0) &= \frac{W(x)}{W(a)} \left[\int_{y \in [0,a]} \int_{h>0} e^{-\lambda(a,0)h} \nu(a-y+dh) R_a^{(0)}(dy) + \Delta^{(0)}(a) \right] \\ &= \frac{W(x)}{W(a)} \left[1 - \int_{y \in [0,a]} \int_{h>0} \left(1 - e^{-\lambda(a,0)h}\right) \nu(a-y+dh) R_a^{(0)}(dy)) \right], \end{aligned}$$

and for $x > a$,

$$\begin{aligned} &\mathbb{P}_x(\underline{X}_{\tau_a} \geqslant 0) \\ &= \int_{y \in [0,a]} \int_0^{x-a} \nu(a-y+dh) R_a^{(0)}(dy) + \int_{y \in [0,a]} \int_{x-a}^\infty e^{-\lambda(a,0)(h+a-x)} \nu(a-y+dh) R_a^{(0)}(dy) + \Delta^{(0)}(a) \\ &= 1 - \int_{y \in [0,a]} \int_{x-a}^\infty \left(1 - e^{-\lambda(a,0)(h+a-x)}\right) \nu(a-y+dh) R_a^{(0)}(dy) \\ &= 1 - \int_{y \in [0,a]} \int_{h>0} \left(1 - e^{-\lambda(a,0)h}\right) \nu(x-y+dh) R_a^{(0)}(dy). \end{aligned}$$

The theorem is proved.

□

Theorem 5.4.1 is of interest because $\mathbb{P}_x[\underline{X}_{\tau_a} < 0]$ is in fact the probability of ruin occurring before a critical drawdown of size a, i.e. it is the probability that the reserve falls below the level zero during the interval $[0, \tau_a]$.
In Section 5.4.1, we first give general expressions for the probability measures of depletion-related quantities. As it might be of interest, from a risk management point of view, to study the depletion problem when ruin does not occur before the critical drawdown time, we also compute the distribution of depletion-related quantities on the event $\{\tau_a < \tau\}$. This is carried out in Section 5.4.2.

5.4.1. Distributions of depletion quantities

Theorem 5.4.2. *Consider an insurance risk process $(X_t)_{t \geqslant 0}$ with initial surplus $x \geqslant 0$ satisfying assumptions of Section 5.2 and let $a > 0$ be a fixed critical drawdown size. Then,* *(1) the probability distribution of the drawdown observed just before critical drawdown is the following,*

$$\mathbb{P}_x(Y_{\tau_a-} \in dy) = \left(R_a^{(0)}(dy) \int_0^\infty \nu(a-y+dh) \right) \mathbb{1}_{y\in(0,a]} + \Delta^{(0)}(a)\delta_a(dy) \ , \quad (5.4.4)$$

(2) the probability distribution of the overshoot over the critical drawdown $Y_{\tau_a} - a$ is the following,

$$\mathbb{P}_x(Y_{\tau_a} - a \in dh) = \int_0^a \nu(a-y+dh) R_a^{(0)}(dy) \mathbb{1}_{h>0} + \Delta^{(0)}(a)\delta_0(dh) \ , \quad (5.4.5)$$

(3) the maximum reserve level attained before critical a drawdown, $\overline{X}_{\tau_a}$, follows a translated exponential distribution, i.e.

$$\mathbb{P}_x(\overline{X}_{\tau_a} \in dv) = \lambda(a,0) e^{-\lambda(a,0)(v-x)} \mathbb{1}_{v\geqslant x} dv \ .$$

Proof. In order to prove this theorem, we use Theorem 5.3.1 when $u \to -\infty$ with $r = q = 0$. By integrating,

(1) For $y \in [0,a)$, we have

$$\mathbb{E}_x[I_{\{Y_{\tau_a-}\in dy\}}] = R_a^{(0)}(dy) \int_x^\infty F_{0,0,a}(v-x)dv \int_0^\infty \nu(a-y+dh),$$

and for $y = a$, $\mathbb{P}(Y_{\tau_a-} = a) = \Delta^{(0)}(a) \int_x^\infty F_{0,0,a}(v-x)dv$.

(2) For $h > 0$, we have

$$\mathbb{E}_x[I_{\{Y_{\tau_a}-a\in dh\}}] = \int_x^\infty F_{0,0,a}(v-x)dv \int_0^a \nu(a-y+dh) R_a^{(0)}(dy),$$

and for $h = 0$, $\mathbb{P}(Y_{\tau_a} = a) = \int_x^\infty F_{0,0,a}(v-x)dv\Delta^{(0)}(a)$.

(3) Finally,

$$\mathbb{P}_x(\overline{X}_{\tau_a} \in dv) = \mathbb{E}_x\left[I_{\{\overline{X}_{\tau_a}\in dv\}}\right] = F_{0,0,a}(v-x)dv\left(\int_0^a R_a^{(0)}(dy)\int_0^\infty \nu(a-y+dh) + \Delta^{(0)}(a)\right).$$

Using (5.4.3) and (5.3.5) yields the result.

□

We notice from Theorem 5.4.2 that the distributions of $Y_{\tau_a^-}$ and of $Y_{\tau_a} - a$ do not depend on the initial surplus x and whatever are the characteristics of the Lévy process X, the distribution of the maximum reserve level $\overline{X}_{\tau_a}$ attained before critical drawdown is always an exponential distribution shifted by the initial surplus x. This result is a typical extension of the same result where we study the distribution of the maximum reserve level $\overline{X}_{\mathtt{e}_q}$ attained before an exponentially distributed random time $\mathtt{e}_q$ with parameter q.
Now we turn our attention to the random times τ_a and $\overline{G}_{\tau_a}$. We start by giving an interesting result concerning the speed of depletion $\tau_a - \overline{G}_{\tau_a}$. This is an

immediate consequence from to Theorem 5.3.1 that was not pointed out in [71], yet it is crucial in evaluating all components in the expression (5.3.9).

Proposition 5.4.1. *Under the same assumptions and definitions of Theorem 5.3.1, the random variables* $\overline{G}_{\tau_a}$ *and* $\tau_a - \overline{G}_{\tau_a}$ *are independent.*

PROOF. It can be easily verified that the statement in Theorem 5.3.1 still holds under weaker conditions on q and r. In fact, the result in (5.3.9) holds true for $q \geqslant 0$ and $q+r \geqslant 0$ and not only for $q, r \geqslant 0$ as indicated in the original statement. In fact, the conditions on q and r arise in the proof when we want to take the q and $q+r$ -scale functions into account in the expressions given in Theorem 5.3.1. As the scale functions $W^{(q)}, W^{(q+r)}$ are just well defined for $q, q+r \geqslant 0$.
By definition, $\overline{G}_{\tau_a}$ and $\tau_a - \overline{G}_{\tau_a}$ are positive $\mathbb{P}$-almost-surely finite random variables. It is well-known that, for $r, q \geqslant 0$, the bivariate Laplace transform $\mathbb{E}_x\left[e^{-r\overline{G}_{\tau_a} - q(\tau_a - \overline{G}_{\tau_a})}\right]$ characterizes the joint distribution of $\overline{G}_{\tau_a}$ and $\tau_a - \overline{G}_{\tau_a}$ (see for example [41]).
Clearly,

$$\mathbb{E}_x\left[e^{-r\overline{G}_{\tau_a} - q(\tau_a - \overline{G}_{\tau_a})}\right] = \mathbb{E}_x\left[e^{-q\tau_a - (r-q)\overline{G}_{\tau_a}}\right] .$$

An expression for the bivariate Laplace transform of $\overline{G}_{\tau_a}$ and $\tau_a - \overline{G}_{\tau_a}$ can be obtained through identities (5.3.9) and (5.3.10) in Theorem 5.3.1. Since $F_{r,q,a}(v)$ is the product of a function depending only on r and a function depending only on q, expressions (5.3.9) and (5.3.10) are also the product of a function depending only on r and a function depending only on q respectively, which concludes the proof.

□

Proposition 5.4.2. *Consider an insurance risk process* $(X_t)_{t\geqslant 0}$ *with initial surplus* $x \geqslant 0$ *satisfying the assumptions of Section 5.2 and let* $a > 0$ *be a fixed critical drawdown size. Then, for* $q \geqslant 0$, $q+r \geqslant 0$, *the bivariate Laplace transform of* τ_a *and* $\overline{G}_{\tau_a}$ *is given by*

$$\mathbb{E}_x\left[e^{-q\tau_a - r\overline{G}_{\tau_a}}\right] = \frac{\lambda(a,q)}{\lambda(a,q+r)}\left(\int_{y\in[0,a]}\int_{h>0} \nu(a-y+dh) R_a^{(q)}(dy) + \Delta^{(q)}(a)\right) . \tag{5.4.6}$$

PROOF. Now, just like in the proof of Proposition 5.4.1, we notice that the result of Theorem 5.3.1 is still valid with $q \geq 0$ and $r+q \geq 0$. Taking $u \to -\infty$ and integrating (5.3.9) and (5.3.10) with respect to v, h and y in Theorem 5.3.1, we obtain

$$\mathbb{E}_x\left[e^{-q\tau_a - r\overline{G}_{\tau_a}}\right] = \int_{v\geqslant x} F_{q+r,q,a}(v-x)dv\left(\int_{y\in[0,a]}\int_{h>0}\nu(a-y+dh)R_a^{(q)}(dy)+\Delta^{(q)}(a)\right).$$

From the definition (5.3.5) of $F_{q+r,q,a}(.)$, we have $\int_x^\infty F_{q+r,q,a}(v-x)dv = \frac{\lambda(a,q)}{\lambda(a,q+r)}$. Substituting this last equation yields,

$$\mathbb{E}_x\left[e^{-q\tau_a - r\overline{G}_{\tau_a}}\right] = \frac{\lambda(a,q)}{\lambda(a,q+r)}\left(\int_{y\in[0,a]}\int_{h>0}\nu(a-y+dh)R_a^{(q)}(dy)+\Delta^{(q)}(a)\right).$$

□

Remark 5.4.1. *Putting $r=0$ in (5.4.6), the Laplace transform of τ_a is given by*

$$\mathbb{E}_x\left[e^{-q\tau_a}\right] = \int_{y\in[0,a]}\int_{h>0}\nu(a-y+dh)R_a^{(q)}(dy)+\Delta^{(q)}(a).$$

Using (5.4.6) with $q=0$ and (5.4.3) the Laplace transform of $\overline{G}_{\tau_a}$ is given by

$$\mathbb{E}_x\left[e^{-r\overline{G}_{\tau_a}}\right] = \frac{\lambda(a,0)}{\lambda(a,r)}. \tag{5.4.7}$$

In the following, we are going to provide an expression for the Laplace transform of the depletion random variable, $\tau_a - \overline{G}_{\tau_a}$.

Theorem 5.4.3. *Consider an insurance risk process $(X_t)_{t\geqslant 0}$ with initial surplus $x\geqslant 0$ satisfying the assumptions of Section 5.2 and let $a>0$ be a fixed critical drawdown size. Then, the Laplace transform of the speed of depletion $\tau_a - \overline{G}_{\tau_a}$ is given by*

$$\mathbb{E}_x\left[e^{-q(\tau_a - \overline{G}_{\tau_a})}\right] = \frac{\lambda(a,q)}{\lambda(a,0)}\left(\int_{y\in[0,a]}\int_{h>0}\nu(a-y+dh)R_a^{(q)}(dy)+\Delta^{(q)}(a)\right).$$

Proof. By Proposition 5.4.1, we know that $\overline{G}_{\tau_a}$ and $\tau_a - \overline{G}_{\tau_a}$ are independent variables, then for $r,q\geqslant 0$

$$\mathbb{E}_x\left[e^{-r\overline{G}_{\tau_a}}\right]\mathbb{E}_x\left[e^{-q(\tau_a-\overline{G}_{\tau_a})}\right] = \mathbb{E}_x\left[e^{-r\overline{G}_{\tau_a}-q(\tau_a-\overline{G}_{\tau_a})}\right] = \mathbb{E}_x\left[e^{-q\tau_a-(r-q)\overline{G}_{\tau_a}}\right]. \tag{5.4.8}$$

We can now find an expression for the right-end of equation (5.4.8) by setting $q^*=q$ and $r^*=r-q$ and using (5.4.6). In other words, since $q^*\geqslant 0$ and $q^*+r^*\geqslant 0$, we can then write

$$\begin{aligned}\mathbb{E}_x\left[e^{-q\tau_a-(r-q)\overline{G}_{\tau_a}}\right] &= \mathbb{E}_x\left[e^{-q^*\tau_a-r^*\overline{G}_{\tau_a}}\right]\\ &= \frac{\lambda(a,q^*)}{\lambda(a,q^*+r^*)}\left(\int_{y\in[0,a]}\int_{h>0}\nu(a-y+dh)R_a^{(q^*)}(dy)+\Delta^{(q^*)}(a)\right).\end{aligned}$$

Substituting $q^*=q$ and $r^*=r-q$ into (5.4.10) and using equation (5.4.8) and equation (5.4.7) in Proposition 5.4.2 yield the result. □

5.4.2. Distributions of depletion quantities in risk management

In this subsection, we provide expressions for the conditional distribution of depletion quantities discussed in Section 5.4.1 given the event $\{\underline{X}_{\tau_a} \geqslant 0\}$. This set guarantees that ruin does not occur before the critical drawdown time. We use the notations $\mathbb{P}(.\,;A)$ and $\mathbb{E}[.\,;A]$ for $\mathbb{P}(.\cap A)$ and $\mathbb{E}[.\,\mathbb{1}_A]$ respectively.

Proposition 5.4.3. *Consider an insurance risk process $(X_t)_{t\geqslant 0}$ satisfying assumptions of Section 5.2 with initial surplus $x>0$ and let $a>0$ be a fixed critical drawdown size. Then,*

(1) *the conditional distribution of the drawdown observed just before critical drawdown given the event $\{\underline{X}_{\tau_a} \geqslant 0\}$ is*

$$\mathbb{P}_x(Y_{\tau_a-}\in dy|\ \underline{X}_{\tau_a}\geqslant 0)$$
$$=\frac{R_a^{(0)}(dy)e^{\lambda(a,0)x\vee a}\int_{h>0}e^{-\lambda(a,0)(x\vee(h+a))}\nu(a-y+dh)\mathbb{1}_{y\in[0,a]}+\Delta^{(0)}(a)\delta_a(dy)}{1-\int_{y\in[0,a]}\int_{h>0}(1-e^{-\lambda(a,0)h})\nu(x\vee a-y+dh)R_a^{(0)}(dy)},$$

(2) *the conditional distribution of the overshoot over the critical drawdown $Y_{\tau_a}-a$ given the event $\{\underline{X}_{\tau_a} \geqslant 0\}$ is*

$$\mathbb{P}_x(Y_{\tau_a}-a\in dh|\ \underline{X}_{\tau_a}\geqslant 0)$$
$$=\frac{e^{-\lambda(a,0)(x\vee(h+a)-x\vee a)}\int_0^a\nu(a-y+dh)R_a^{(0)}(dy)\mathbb{1}_{h>0}+\Delta^{(0)}(a)\delta_a(dh)}{1-\int_{y\in[0,a]}\int_{h>0}(1-e^{-\lambda(a,0)h})\nu(x\vee a-y+dh)R_a^{(0)}(dy)},$$

(3) *the conditional distribution of the maximum reserve level attained before critical drawdown of size a given the event $\{\underline{X}_{\tau_a} \geqslant 0\}$ is*

$$\mathbb{P}_x\Big(\overline{X}_{\tau_a}\in dv|\ \underline{X}_{\tau_a}\geqslant 0\Big)$$
$$=\frac{\lambda(a,0)e^{-\lambda(a,0)(v-x\vee a)}\Big(\int_0^{v-a}\int_0^a\nu(a-y+dh)R_a^{(0)}(dy)+\Delta^{(0)}(a)\Big)}{1-\int_{y\in[0,a]}\int_{h>0}(1-e^{-\lambda(a,0)h})\nu(x\vee a-y+dh)R_a^{(0)}(dy)}\mathbb{1}_{v\geqslant x\vee a}dv\,.$$

Proof. It is clear from Theorem 5.4.1 that

$$\mathbb{P}_x(\underline{X}_{\tau_a}\geqslant 0)=\frac{W(x\wedge a)}{W(a)}\left(1-\int_{y\in[0,a]}\int_{h>0}\left(1-e^{-\lambda(a,0)h}\right)\nu(x\vee a-y+dh)R_a^{(0)}(dy)\right).\tag{5.4.9}$$

Using Theorem 5.3.1 with $u=0$ and $r=q=0$ yields:

(1) For $y\in[0,a)$, we have

$$\mathbb{P}_x(Y_{\tau_a-}\in dy;\underline{X}_{\tau_a}\geqslant 0)=R_a^{(0)}(dy)\frac{W(x\wedge a)}{W(a)}\int_{x\vee a}^{\infty}F_{0,0,a}(v-(x\vee a))\int_{h\in(0,v-a]}\nu(a-y+dh)dv,\tag{5.4.10}$$

and for $y = a$, $\mathbb{P}(Y_{\tau_a -} = a; \underline{X}_{\tau_a} \geqslant 0) = \frac{W(x\wedge a)}{W(a)}\Delta^{(0)}(a)\int_{x\vee a}^{\infty} F_{0,0,a}(v - x \vee a)dv$. By applying Fubini's Theorem and equation (5.3.5) together into equation (5.4.10) we have

$$\mathbb{P}_x\left(Y_{\tau_a -} \in dy; \underline{X}_{\tau_a} \geqslant 0\right) = \frac{W(x\wedge a)}{W(a)}\left[R_a^{(0)}(dy)e^{\lambda(a,0)x\vee a}\int_{h>0} e^{-\lambda(a,0)(x\vee(h+a))}\nu(a-y+dh)\mathbb{1}_{y\in[0,a]} + \Delta^{(0)}(a)\delta_a(dy)\right]. \tag{5.4.11}$$

The first part of the Theorem is obtained by using (5.4.11) and (5.4.9).

(2) For $h > 0$, we have

$$\mathbb{E}_x\left[\mathbb{1}_{\{Y_{\tau_a}-a\in dh\}}; \underline{X}_{\tau_a} \geqslant 0\right] = \frac{W(x\wedge a)}{W(a)}\int_{x\vee(h+a)}^{\infty} F_{0,0,a}(v - x\vee a)dv\int_0^a \nu(a-y+dh)R_a^{(0)}(dy), \tag{5.4.12}$$

and for $h = 0$, $\mathbb{P}(Y_{\tau_a} = a; \underline{X}_{\tau_a} \geqslant 0) = \frac{W(x\wedge a)}{W(a)}\int_{x\vee a}^{\infty} F_{0,0,a}(v - x\vee a)dv\Delta^{(0)}(a)$. The proof of the second part of the Theorem is done by applying the last equation, (5.4.12) and (5.4.9) into the definition of conditional distribution.

(3) At the end, for $v \geqslant x \vee a$

$$\mathbb{P}_x(\overline{X}_{\tau_a} \in dv; \underline{X}_{\tau_a} \geqslant 0) = \frac{W(x\wedge a)}{W(a)}F_{0,0,a}(v - x\vee a)\left(\int_0^{v-a}\int_0^a \nu(a-y+dh)R_a^{(0)}(dy) + \Delta^{(0)}(a)\right)dv. \tag{5.4.13}$$

The proof is complete if we use (5.4.13) and (5.4.9).

□

Proposition 5.4.4. *Consider an insurance risk process $(X_t)_{t\geqslant 0}$ satisfying assumptions of Section 5.2 with initial surplus $x > 0$ and let $a > 0$ be a fixed critical drawdown size. Then, for $q, r \geqslant 0$, we have*

$$\begin{aligned}&\mathbb{E}_x\left[e^{-q\tau_a - r\overline{G}_{\tau_a}}; \underline{X}_{\tau_a} \geqslant 0\right]\\ &= \frac{W^{(q+r)}(x\wedge a)}{W^{(q+r)}(a)}\frac{\lambda(a,q)}{\lambda(a,q+r)}\left(\int_{y\in[0,a]}\int_{h>0} e^{-\lambda(a,q+r)(x\vee(h+a)-x\vee a)}\nu(a-y+dh)R_a^{(q)}(dy) + \Delta^{(q)}(a)\right).\end{aligned} \tag{5.4.14}$$

Proof. This result can be obtained, like in the proof of the previous proposition, by putting $u = 0$ in (15) and (16) in Theorem 5.3.1 and integrating with respect to dv, dh and dy. □

We notice that in general, variables $\tau_a - \overline{G}_{\tau_a}$ and $\overline{G}_{\tau_a}$ given the event $\{\underline{X}_{\tau_a} \geqslant 0\}$ are no more independent variables (especially when the diffusion coefficient σ is positive).

5.5. Examples of Lévy Insurance Risk Processes

We study in this section particular examples of risk process X satisfying the general setting described in Section 5.2, when the q-scale function of the model has a tractable form.

We start by studying three examples of risk process X, without Brownian motion part, i.e. $\sigma = 0$ in its Laplace exponent (5.3.1). Thus the results in Section 5.4 apply to our problem and they endow us with tools to fully study the depletion problem. Since $\sigma = 0$ in the studied models without Brownian perturbation, the set A_c in Theorem 5.3.1 is empty and the coefficient $\Delta^{(q)}(a)$ will not appear in the sequel. We also discuss two examples of risk process X starting at an initial surplus $x \geq 0$, with a Brownian motion part, i.e. $\sigma \neq 0$ in its Laplace exponent (5.3.1).

In this chapter, we aim at computing expressions for the distribution of the depletion-related random variables for relevant insurance risk processes. As it turns out, the results in Theorems 5.3.1, 5.4.1, 5.4.2 and 5.4.3 lead to explicit expressions for the distribution of depletion-related random variables when the q-scale function of the model has a tractable form. In fact, we will see that a tractable form for the q-scale function is inherited by the functions λ, $F_{p,q,a}$ and $R_a^{(q)}$ defined in (5.3.4), (5.3.5) and (5.3.6) respectively. These functions are the key ingredients in the general expressions of Theorems 5.3.1, 5.4.1, 5.4.2 and 5.4.3. In this section, we show how there are some interesting examples of insurance models with tractable q-scale functions leading to relatively simple expressions for the distributions of depletion random variables. In the following, we will analyze in more detail some models for which we can have an explicit understanding of the depletion problem:

- the Classical Cramer-Lundberg model with exponential claims,
- the Gamma risk process,
- the Spectrally Negative Stable risk process,
- the Brownian perturbed model without claims,
- the Meromorphic risk process (Beta process).

5.5.1. Classical Cramer-Lundberg Model with Exponential Claims

The so-called classical or Cramer-Lundberg model was introduced in [68]. The risk process R is a compound Poisson process starting at $x \geqslant 0$, i.e.,

$$R_t = x + X_t \, , \tag{5.5.1}$$

where $X_t = ct - \sum_{i=1}^{N_t} Z_i$ and the number of claims is assumed to follow a Poisson process $(N_t)_{t\geqslant 0}$ with intensity λ which is independent of the positive and *iid* random variables $(Z_n)_{n\geqslant 1}$ representing claim sizes. The loaded premium c is of the form $c = (1+\theta)\lambda\mathbb{E}[Z_1]$ for some safety loading factor $\theta > 0$. The form of the q-scale function in this model is relatively simple when the claim sizes are exponentially distributed with mean $1/\mu$. In this case, the Lévy measure takes the simple form $\nu(dx) = \lambda K(dx)$ where K is the exponential probability measure associated to the claim sizes. In turn, the Laplace exponent in (5.3.1) becomes

$$\psi(s) = c\,s - \lambda[\phi_K(s) - 1]\,, \qquad s > 0\,, \tag{5.5.2}$$

where $\phi_K(s) = \frac{\mu}{\mu+s}$ is the Laplace transform of an exponential distribution (see for instance [63]). In this case, the premium rate is $c = \frac{\lambda(1+\theta)}{\mu}$ where $\theta > 0$ is a positive security loading.

A path of a such process is linear by parts, so its corresponding drawdown process is quite simple to draw.

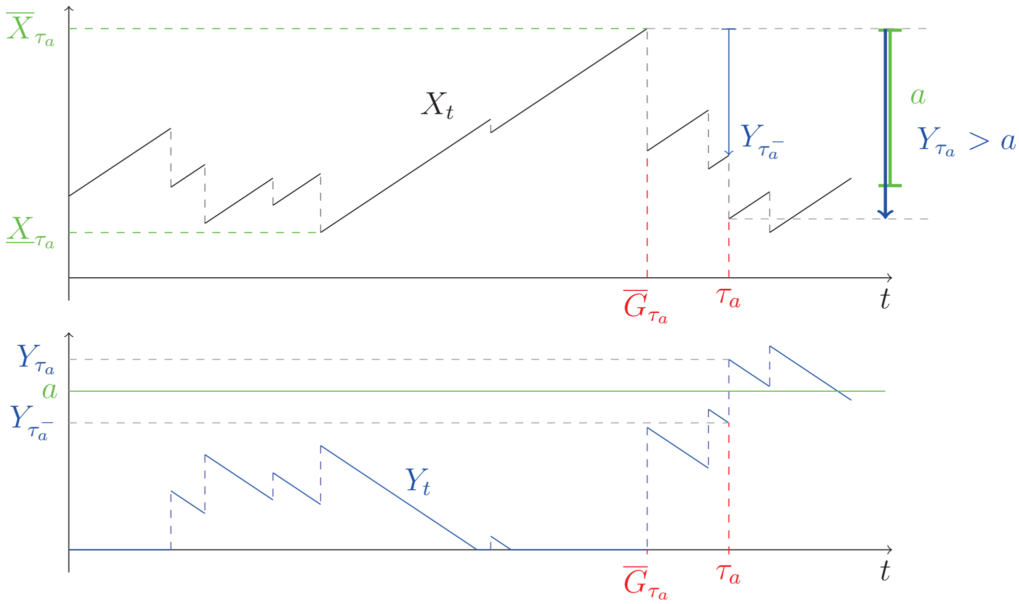


Fig. 5.2. A path of a compound Poisson process R, the corresponding drawdown process Y and their related depletion quantities.

This model has been for long a textbook example for which the distribution of ruin-related quantities can be explicitly computed. This is in fact possible thanks to the tractable form of the q-scale function although maybe this was not immediately recognized. It turns out that just like for the ruin problem, this tractable form of the q-scale function also allows for an explicit study of the depletion

problem. Here, we give the form of the q-scale function and study in detail the depletion problem for this particular example. Moreover, we derive explicit expressions for the distributions of depletion random variables of Theorems 5.3.1, 5.4.1, 5.4.2 and 5.4.3.
The expression for the q-scale function in this case is known and is given by

$$W^{(q)}(x) = \frac{1}{c^2(\Phi(q)+\mu)^2 - c\lambda\mu}\Big[c(\Phi(q)+\mu)^2\, e^{\Phi(q)x} - \lambda\mu e^{-\left(\mu - \frac{\lambda\mu}{c(\Phi(q)+\mu)}\right)x}\Big], \tag{5.5.3}$$

with $\Phi(q) = \frac{1}{2c}\Big(q+\lambda-c\mu+\sqrt{(q+\lambda-c\mu)^2+4q\mu c}\Big)$. For details see [62].
Now we also need expressions for the functions λ and $R_a^{(q)}$. These are given in the following result.

Proposition 5.5.1. *Consider the process $(R_t)_{t\geq 0}$ in (5.5.1) where the Z_i's are identically independent exponential random variables with mean $1/\mu$ and let $a>0$ be a critical drawdown size. Then*

$$\lambda(a,q) = \Phi(q) + \frac{\lambda\mu\Big[c(\Phi(q)+\mu)^2 - \lambda\mu\Big]}{c(\Phi(q)+\mu)\Big[(\Phi(q)+\mu)^2 c e^{\left(\Phi(q)+\mu-\frac{\lambda\mu}{(\Phi(q)+\mu)c}\right)a} - \lambda\mu\Big]}\ , \tag{5.5.4}$$

$$R_a^{(q)}(dy) = \frac{1}{c\lambda(a,q)}\delta_0(dy) + \left[\frac{1}{\lambda(a,q)}W'^{(q)}(y) - W^{(q)}(y)\right]\mathbb{1}_{y\in(0,a]}dy\ ,$$

with $W^{(q)}$ and $W'^{(q)}$ given by (5.5.3) and (5.5.5) respectively.

Proof. By definition

$$\lambda(a,q) = \frac{W_+'^{(q)}(a)}{W^{(q)}(a)}\ .$$

By referring to equation (5.5.3) we can see that $W^{(q)}$ is a derivable function on $(0,\infty)$ so by differentiating equation (5.5.3) we can directly obtain

$$\begin{aligned} W'^{(q)}(x) &= \Phi(q)W^{(q)}(x) + \frac{\lambda\mu}{c\Big[(\Phi(q)+\mu)^2 c - \lambda\mu\Big]}\left[\Phi(q) + \left(\mu - \frac{\lambda\mu}{(\Phi(q)+\mu)c}\right)\right]e^{-\left(\mu-\frac{\lambda\mu}{(\Phi(q)+\mu)c}\right)x} \\ &= \Phi(q)W^{(q)}(x) + \frac{\lambda\mu}{c^2(\Phi(q)+\mu)}e^{-\left(\mu-\frac{\lambda\mu}{(\Phi(q)+\mu)c}\right)x}. \end{aligned} \tag{5.5.5}$$

Combining (5.5.3) and (5.5.5) in the definition (5.3.4) yields the first result.
As $W^{(q)}$ is an increasing function and has a mass at $x=0$, recall that in this case (see [63]), $W^{(q)}(0^+) = 1/c$, we have $W^{(q)}(dx) = W'^{(q)}(x)dx + \frac{1}{c}\delta_0(dy)$. Direct substitution into the definition (5.3.6) of $R_a^{(q)}$ yields the second result. □

The main results regarding depletion-related quantities are given in terms of $W^{(q)}$ and $\lambda(a,q)$ with $q=0$. These expressions take on a more simple form in this case and are given in the following result.

Proposition 5.5.2. *Consider the process* $(R_t)_{t\geq 0}$ *in (5.5.1) where the* Z_i*'s are identically independent exponential random variables with mean* $1/\mu$ *and let* $a > 0$ *be a critical drawdown size. Then*

$$W(x) = \frac{\mu}{\lambda(1+\theta)\theta}\left(1+\theta - e^{\frac{-\mu\theta}{1+\theta}x}\right),$$

$$W'(x) = \frac{\mu^2}{\lambda(1+\theta)^2}\, e^{\frac{-\mu\theta}{1+\theta}x},$$

$$\lambda(a,0) = \frac{\mu\theta}{(1+\theta)\left((1+\theta)e^{\frac{\mu\theta}{1+\theta}a} - 1\right)},$$

$$F_{0,0,a}(y) = \frac{\mu\theta}{(1+\theta)\left((1+\theta)e^{\frac{\mu\theta}{1+\theta}a} - 1\right)} \exp\left(\frac{-\mu\theta y}{(1+\theta)\left((1+\theta)e^{\frac{\mu\theta}{1+\theta}a} - 1\right)}\right),$$

$$R_a^{(0)}(dy) = \frac{\mu}{\lambda\theta}\left(e^{\frac{\mu\theta}{1+\theta}(a-y)} - 1\right)\mathbb{1}_{y\in(0,a]}dy + \frac{1}{\lambda\theta}\left((1+\theta)e^{\frac{\mu\theta}{1+\theta}a} - 1\right)\delta_0(dy).$$

Proof. It is straight forward by setting $q = 0$ in Proposition 5.5.1 with $c = \frac{\lambda(1+\theta)}{\mu}$, and by using definitions (5.3.5) and (5.3.6). Recall that $\Phi(0) = 0$. □

We now give explicit representations for the distributions in Theorems 5.4.1, 5.4.2 and 5.4.3 as they are specialized to this case.

Proposition 5.5.3. *Consider an insurance risk process* $(R_t)_{t\geqslant 0}$ *of the form defined in (5.5.1) with an initial level* $x \geqslant 0$ *and let* $a > 0$ *be a fixed critical drawdown size. Then*

$$\mathbb{P}_x(\underline{X}_{\tau_a} < 0) = 1 - \frac{W(x\wedge a)}{W(a)}\left(1 - \frac{\lambda(a,0)}{\mu+\lambda(a,0)}e^{-\mu(x\vee a - a)}\right), \tag{5.5.6}$$

where $\lambda(a,0)$ *and the scale function* W *are given in Proposition 5.5.2.*

Proof. To show this proposition we use the expression of $\mathbb{P}_x(\underline{X}_{\tau_a} < 0)$ given in Theorem 5.4.1. But in this model $\sigma = 0$ so we have $\Delta^{(0)}(a) = 0$ and

$$\mathbb{P}_x(\underline{X}_{\tau_a} < 0) = 1 - \frac{W(x\wedge a)}{W(a)} + \frac{W(x\wedge a)}{W(a)}\int_{y\in[0,a]}\int_{h>0}\left(1 - e^{-\lambda(a,0)h}\right)\nu(x\vee a - y + dh)R_a^{(0)}(dy), \tag{5.5.7}$$

The Lévy measure for the process S_t is $\nu(dx) = \lambda\mu e^{-\mu x}dx$. So, by replacing this into the interior integral in (5.5.7) we have

$$\int_{h>0}\left(1 - e^{-\lambda(a,0)h}\right)\nu(x\vee a - y + dh) = \int_{h>0}\left(1 - e^{-\lambda(a,0)h}\right)\lambda\mu e^{-\mu(x\vee a - y + h)}dh$$

$$= \lambda e^{-\mu(x\vee a)}\left(e^{\mu y} - \frac{\mu}{\mu+\lambda(a,0)}e^{\mu y}\right). \tag{5.5.8}$$

On the other hand, we have

$$\lambda \int_0^a e^{\mu y} R_a^{(0)}(dy) = e^{\mu a}. \tag{5.5.9}$$

So by applying (5.5.9) and (5.5.8) into (5.5.7) we have

$$\mathbb{P}_x(\underline{X}_{\tau_a} < 0) = 1 - \frac{W(x\wedge a)}{W(a)} + \frac{W(x\wedge a)\lambda(a,0)e^{-\mu(x\vee a-a)}}{W(a)(\mu+\lambda(a,0))}$$

□

Proposition 5.5.4. *Consider an insurance risk process $(R_t)_{t\geqslant 0}$ of the form defined in (5.5.1) with an initial level $x \geqslant 0$ and let $a > 0$ be a fixed critical drawdown size. Then*

(1) *the largest drawdown observed before critical drawdown follows a mixture of a diffusive distribution on $(0,a]$ and the Dirac measure at 0. i.e.,*

$$\mathbb{P}_x(Y_{\tau_a-} \in dy) = \frac{\mu}{\theta}\left(e^{-\frac{\mu}{1+\theta}(a-y)} - e^{-\mu(a-y)}\right)\mathbb{1}_{y\in(0,a]}dy + \frac{1}{\theta}\left((1+\theta)e^{-\frac{\mu}{1+\theta}a} - e^{-\mu a}\right)\delta_0(dy)\ .$$

(2) *the overshoot over the critical drawdown $Y_{\tau_a} - a$ follows an exponential distribution with mean $1/\mu$. i.e.,*

$$\mathbb{P}_x(Y_{\tau_a} - a \in dh) = \mu e^{-\mu h}\mathbb{1}_{h>0}dh\ .$$

Proof. We prove this proposition by applying Theorem 5.4.2. Here $\sigma = 0$ and naturally $\Delta^{(0)}(a) = 0$.

(1) By plugging $\int_0^\infty \nu(a-y+dh) = \lambda e^{-\mu(a-y)}$ into (5.4.4) we have

$$\mathbb{P}_x(Y_{\tau_a-} \in dy) = \lambda e^{-\mu(a-y)} R_a^{(0)}(dy). \tag{5.5.10}$$

We conclude the first part of the theorem using the expression of $R_a^{(0)}(dy)$ given in Proposition 5.5.2.

(2) To prove the second part of the theorem we have

$$\nu(a-y+dh) = \lambda\mu e^{-(a-y+h)\mu}dh. \tag{5.5.11}$$

for $h > 0,\ \ y \in [0,a]$. Substituting (5.5.11) in (5.4.5) gives $\mathbb{P}_x(Y_{\tau_a} - a \in dh) = \mu e^{-\mu h}dh$.

□

We now provide explicit expressions for joint Laplace transform of $\overline{G}_{\tau_a}$ and τ_a as well as the Laplace transform of the speed of depletion. Notice that the Laplace transform of $\overline{G}_{\tau_a}$ is a simple function of $\lambda(a,q)$ and it can be computed in a straightforward way using Proposition 5.4.2 and equation (5.5.4).

Proposition 5.5.5. *Consider an insurance risk process $(R_t)_{t\geqslant 0}$ of the form defined in (5.5.1) with an initial level $x \geqslant 0$ and let $a > 0$ be a fixed critical drawdown size. Then for $q, r \geqslant 0$*

(1) the bivariate Laplace transform of τ_a and $\overline{G}_{\tau_a}$ is given by

$$\mathbb{E}_x\left[e^{-q\tau_a - r\overline{G}_{\tau_a}}\right] = \left(\frac{\Phi(q)+\mu}{\mu}\right)\left(\frac{\lambda(a,q)-\Phi(q)}{\lambda(a,q+r)}\right)e^{\Phi(q)a}\ , \tag{5.5.12}$$

with $\Phi(q) = (2c)^{-1}\Big(q+\lambda-c\mu+\sqrt{(q+\lambda-c\mu)^2+4q\mu c}\Big)$,

(2) the Laplace transform of the speed of depletion $\tau_a - \overline{G}_{\tau_a}$ is given by

$$\mathbb{E}_x\Big[e^{-q(\tau_a-\overline{G}_{\tau_a})}\Big] = \left(\frac{\Phi(q)+\mu}{\mu}\right)\left(\frac{\lambda(a,q)-\Phi(q)}{\lambda(a,0)}\right)e^{\Phi(q)a}\ .$$

Recall that $\lambda(a,.)$ is given in Proposition 5.5.2.

Proof. As $\sigma = 0$, then $\Delta^{(0)}(a) = 0$.

(1) The Lévy measure of the process X is $\nu(dx) = \lambda\mu e^{-\mu x}dx$ and so $\int_0^\infty \nu(a - y + dh) = \lambda e^{-\mu(a-y)}$. Now using Proposition 5.4.2 as it specializes to this case yields

$$\mathbb{E}_x\left[e^{-q\tau_a - r\overline{G}_{\tau_a}}\right] = \frac{\lambda(a,q)\,\lambda e^{-\mu a}}{\lambda(a,q+r)}\int_{y\in[0,a]} e^{\mu y}R_a^{(q)}(dy)\ . \tag{5.5.13}$$

Now, using Proposition 5.5.13 we can compute the following integral.

$$\lambda e^{-\mu a}\int_{y\in[0,a]} e^{\mu y}R_a^{(q)}(dy) = \frac{\lambda}{\lambda(a,q)}W^{(q)}(a) - \lambda e^{-\mu a}\Big(\frac{\mu}{\lambda(a,q)}+1\Big)\int_0^a e^{\mu y}W^{(q)}(y)dy. \tag{5.5.14}$$

In order to finish the proof it would be sufficient to find an expression for $\int_0^a e^{\mu y}W^{(q)}(dy)$. By using equation (5.5.3), we have

$$\int_0^a e^{\mu y}W^{(q)}(y)dy = \frac{\Phi(q)+\mu}{(\Phi(q)+\mu)^2c-\lambda\mu}\left(e^{(\Phi(q)+\mu)a} - e^{\frac{\lambda\mu}{(\Phi(q)+\mu)c}a}\right)\ . \tag{5.5.15}$$

Now, by combining equations (5.5.15) and (5.5.14) into (5.5.13) we obtain

$$\mathbb{E}_x\left[e^{-q\tau_a - r\overline{G}_{\tau_a}}\right] = \frac{\lambda}{\lambda(a,q+r)}\left(W^{(q)}(a) - (\mu+\lambda(a,q))\frac{\Phi(q)+\mu}{(\Phi(q)+\mu)^2c-\lambda\mu}\left(e^{\Phi(q)a} - e^{-\left(\mu-\frac{\lambda\mu}{(\Phi(q)+\mu)c}\right)a}\right)\right).$$

Using the expressions (5.5.3) and (5.5.4) for $W^{(q)}(a)$ and $\lambda(a,q)$ respectively, we deduce that,

$$\begin{aligned}\mathbb{E}_x\left[e^{-q\tau_a - r\overline{G}_{\tau_a}}\right] &= \frac{\lambda}{\lambda(a,q+r)}\frac{((\Phi(q)+\mu)^2c-\lambda\mu)e^{\Phi(q)a}}{c\Big((\Phi(q)+\mu)^2ce^{\left(\Phi(q)+\mu-\frac{\lambda\mu}{(\Phi(q)+\mu)c}\right)a}-\lambda\mu\Big)}\\ &= \frac{(\Phi(q)+\mu)(\lambda(a,q)-\Phi(q))e^{\Phi(q)a}}{\mu\lambda(a,q+r)}.\end{aligned}$$

This completes the proof.

(2) In a similar way, specializing Theorem 5.4.3 to this case using the expressions in Proposition 5.5.13 yields the result.

□

In the following proposition, we provide the conditional distribution for the depletion quantities given the event $\{\underline{X}_{\tau_a} \geqslant 0\}$. Expressing the joint Laplace transform for $\overline{G}_{\tau_a}$ and τ_a in the presence of the event $\{\underline{X}_{\tau_a} \geqslant 0\}$ is also of interest so that for which we discuss in the following proposition.

Proposition 5.5.6. *Consider an insurance risk process $(R_t)_{t\geqslant 0}$ of the form defined in (5.5.1) with an initial level $x \geqslant 0$ and let $a > 0$ be a fixed critical drawdown size. Then*

(1) $\mathbb{P}_x\left(Y_{\tau_a-}\in dy|\ \underline{X}_{\tau_a}\geqslant 0\right)=\lambda e^{-\mu(a-y)}R_a^{(0)}(dy)\mathbb{1}_{[0,a]}(y)$,

(2) $\mathbb{P}_x\left(Y_{\tau_a}-a\in dh|\ \underline{X}_{\tau_a}\geqslant 0\right)=\frac{\mu e^{-\lambda(a,0)(x\vee(h+a)-x\vee a)-\mu h}dh}{1-\frac{\lambda(a,0)}{\mu+\lambda(a,0)}e^{-\mu(x\vee a-a)}}\mathbb{1}_{h>0}$,

(3) $\mathbb{P}_x\left(\overline{X}_{\tau_a}\in dv|\ \underline{X}_{\tau_a}\geqslant 0\right)=\frac{\lambda(a,0)e^{-\lambda(a,0)x\vee a}\left(e^{-\lambda(a,0)v}-e^{-(\lambda(a,0)+\mu)v+\mu a}\right)}{1-\frac{\lambda(a,0)}{\mu+\lambda(a,0)}e^{-\mu(x\vee a-a)}}\mathbb{1}_{v\geqslant x\vee a}dv$,

(4) $\mathbb{E}_x\left[e^{-q\tau_a-r\overline{G}_{\tau_a}};\underline{X}_{\tau_a}\geqslant 0\right]=\frac{W^{(q+r)}(x\wedge a)}{W^{(q+r)}(a)}\frac{(\mu+\Phi(q))(\lambda(a,q)-\Phi(q))}{\mu\lambda(a,q+r)}\left(1-\frac{\lambda(a,q+r)}{\mu+\lambda(a,q+r)}e^{-\mu(x\vee a-a)}\right)$.

Proof. To prove this proposition we need to use Propositions 5.4.3 and 5.4.4.

(1) To show the first part we compute the expression given in (5.4.11). In fact by considering different cases between x and a we have

$$\mathbb{P}_x(Y_{\tau_a-}\in dy;\underline{X}_{\tau_a}\geqslant 0)=\frac{W(x\wedge a)}{W(a)}\left(\lambda e^{-\mu(a-y)}R_a^{(0)}(dy)\left(1-\frac{\lambda(a,0)}{\mu+\lambda(a,0)}e^{-\mu(x\vee a-a)}\right)\right). \tag{5.5.16}$$

On the other side, by (5.5.6) we have

$$\mathbb{P}_x(\underline{X}_{\tau_a}\geqslant 0)=\frac{W(x\wedge a)}{W(a)}\left(1-\frac{\lambda(a,0)}{\mu+\lambda(a,0)}e^{-\mu(x\vee a-a)}\right) \tag{5.5.17}$$

Applying (5.5.17) and (5.5.16) to (5.4.11) yields the result in the first part of the theorem.

(2) To show this part we also compute the expression given in (5.4.12). After simplifying the expression we have

$$\mathbb{P}_x(Y_{\tau_a} - a \in dh; \underline{X}_{\tau_a} \geqslant 0) = \frac{W(x \wedge a)}{W(a)} \Big(\mu e^{-\lambda(a,0)(x \vee (h+a) - x \vee a) - \mu h} \Big) dh. \quad (5.5.18)$$

Once again applying (5.5.18) and (5.5.17) to (5.4.12) proves the result in the second part of the theorem.

(3) This part can be proven by computing the expression given in (5.4.13) and using (5.5.17). In fact after some simplifications we get

$$\mathbb{P}_x\Big(\overline{X}_{\tau_a} \in dv; \underline{X}_{\tau_a} \geqslant 0\Big) = \frac{W(x \wedge a)}{W(a)} \Big(\lambda(a,0) e^{-\lambda(a,0) x \vee a} \Big(e^{-\lambda(a,0)v} - e^{-(\lambda(a,0)+\mu)v + \mu a} \Big)\Big) dv. \quad (5.5.19)$$

To end the proof we just need to replace (5.5.19) and (5.5.17) to (5.4.13).

(4) The last part of the theorem can be shown directly by computing the expression in (5.4.14). It is clear from (5.5.12) that

$$\lambda \int_{y \in [0,a]} e^{\mu y} R_a^{(q)}(dy) = \left(\frac{\Phi(q) + \mu}{\mu} \right) \left(\frac{\lambda(a,q) - \Phi(q)}{\lambda(a,q)} \right) e^{(\Phi(q)+\mu)a}. \quad (5.5.20)$$

So the proof is complete if we apply (5.5.20) to the expression in (5.4.14) and simplify the expression.

□

Remark 5.5.1. *Under the same assumptions of Proposition 5.5.6, it can be seen that*

(1) from the first part of Proposition 5.5.6, the event $\{Y_{\tau_a -} \in dy\}$ is independent of the event $\{\underline{X}_{\tau_a} \geqslant 0\}$. In fact by recalling (5.5.10) we have $\mathbb{P}_x(Y_{\tau_a -} \in dy | \ \underline{X}_{\tau_a} \geqslant 0) = \lambda e^{-\mu(a-y)} R_a^{(0)}(dy) = \mathbb{P}_x(Y_{\tau_a -} \in dy)$. Thus knowing $\underline{X}_{\tau_a}$ does not affect on the distribution of the largest drawdown before critical drawdown, $Y_{\tau_a -}$.

(2) from the second part of Proposition 5.5.6, if $x < a$, then the random variable $Y_{\tau_a} - a$ given $\{\underline{X}_{\tau_a} \geqslant 0\}$ follows an exponential distribution with parameter $\mu + \lambda(a,0)$. In other words,

$$\mathbb{P}_x(Y_{\tau_a} - a \in dh | \ \underline{X}_{\tau_a} \geqslant 0) = (\mu + \lambda(a,0)) e^{-(\mu + \lambda(a,0))h} \mathbb{1}_{h>0}.$$

(3) from the joint Laplace transform of $(\tau_a, \overline{G}_{\tau_a})$, we deduce that $\overline{G}_{\tau_a}$ and $\tau_a - \overline{G}_{\tau_a}$ are independent random variables on the event $\{\underline{X}_{\tau_a} \geqslant 0\}$. Thus, the Laplace transform of the depletion random variable , $\tau_a - \overline{G}_{\tau_a}$, given

the event $\{\underline{X}_{\tau_a} \geqslant 0\}$ *is*

$$\mathbb{E}_x\left[e^{-q\left(\tau_a-\overline{G}_{\tau_a}\right)} \mid \underline{X}_{\tau_a} \geqslant 0\right] = \frac{\mu+\Phi(q)}{\mu}\frac{\lambda(a,q)-\Phi(q)}{\lambda(a,0)}.$$

In fact, with the same proof as for Proposition 5.4.1 we can show that $\overline{G}_{\tau_a}$ *and* $\tau_a - \overline{G}_{\tau_a}$ *are independent random variables on the event* $\{\underline{X}_{\tau_a} \geqslant 0\}$*. Moreover, we have* $\mathbb{E}_x\left[e^{-q\left(\tau_a-\overline{G}_{\tau_a}\right)} \mid \underline{X}_{\tau_a} \geqslant 0\right] = \frac{\mathbb{E}_x\left[e^{-q\left(\tau_a-\overline{G}_{\tau_a}\right)}; \underline{X}_{\tau_a} \geqslant 0\right]}{\mathbb{P}_x(\underline{X}_{\tau_a} \geqslant 0)}$*, where* $\mathbb{E}_x\left[e^{-q\left(\tau_a-\overline{G}_{\tau_a}\right)}; \underline{X}_{\tau_a} \geqslant 0\right]$ *can be obtained from part 4 in Proposition 5.5.6 as* $\overline{G}_{\tau_a}$ *and* $\tau_a - \overline{G}_{\tau_a}$ *are independent random variables on the event* $\{\underline{X}_{\tau_a} \geqslant 0\}$*. Moreover,* $\mathbb{P}_x(\underline{X}_{\tau_a} \geqslant 0)$ *is given by equation (5.5.17).*

5.5.2. Gamma Risk Process

The gamma risk model was introduced in [38] and is defined by

$$R_t = x + X_t \,, \tag{5.5.21}$$

where $X_t = ct - S_t$ and the aggregate claims process $(S_t)_{t\geqslant 0}$ is assumed to follow a gamma subordinator with Lévy measure

$$\nu(dx) = \alpha x^{-1} e^{-\beta x} dx \,, \qquad x > 0 \,,$$

where $\alpha, \beta > 0$. The loaded premium c is of the form $c = (1+\theta)\mathbb{E}[S_1]$ for some safety loading factor $\theta > 0$. In turn, the Laplace exponent in (5.3.1) becomes

$$\psi_X(s) = c\, s - \alpha \ln(1 + \frac{s}{\beta}) \,. \qquad s > 0 \,,$$

We refer the reader to [49] for a discussion of subordinator models in risk theory. In this section we are going to provide expressions for $\underline{X}_{\tau_a}, \overline{X}_{\tau_a}, Y_{\tau_a -}$ and $Y_{\tau_a} - a$ associated to the process X given by $X_t = ct - S_t$.

To find these expressions we need first to provide $W(x)$ for X. Let the process X start at $x \geqslant 0$. Based on the result given in Chapter 8 of [63] for survival probability for a spectrally negative Lévy process we have

$$1 - \phi(x) = \begin{cases} \psi_X'(0^+) W(x) & \text{if } \psi_X'(0^+) > 0 \\ 0 & \text{Otherwise,} \end{cases} \tag{5.5.22}$$

where $\phi(x)$ is the probability of ruin and ψ_X is the Laplace exponent for X. On the other hand, [38] gives another expression for survival probability for $X_t = ct - S_t$

when S is a gamma subordiantor. That is

$$1-\phi(x)=\frac{\theta}{1+\theta}\sum_{n\geqslant 0}\frac{1}{(1+\theta)^n}M^{*n}(x), \tag{5.5.23}$$

where $M(x)=\beta\int_0^x\int_{\beta t}^{\infty}u^{-1}e^{-u}dudt=1+\beta xE_1(\beta x)$. Here $E_1(x)=\int_x^{\infty}u^{-1}e^{-u}du$ is the exponential integral function and $M^{*n}(x)=\int_0^x M^{*(n-1)}(x-y)M'(y)dy$ is the nth-fold convolution where $M'(y)=\beta E_1(\beta y)$.
As $\psi_X'(0^+)=c-\psi_S'(0^+)=c-\frac{\alpha}{\beta}$ and $c=\frac{\alpha(1+\theta)}{\beta}$, we have $\psi_X'(0^+)=\frac{\alpha\theta}{\beta}>0$. Now, by equalizing (5.5.22) and (5.5.23) we can get the expression for $W(x)$. More precisely, we have

$$W(x)=\frac{\beta}{(1+\theta)\alpha}\sum_{n\geqslant 0}\frac{1}{(1+\theta)^n}M^{*n}(x). \tag{5.5.24}$$

Taking derivative of $W(x)$ in (5.5.24) yields

$$W'(x)=\frac{\beta}{(1+\theta)\alpha}\sum_{n\geqslant 0}\frac{1}{(1+\theta)^n}(M^{*n})'(x)=\frac{\beta}{(1+\theta)\alpha}\sum_{n\geqslant 0}\frac{1}{(1+\theta)^n}\int_0^x\frac{\partial M^{*(n-1)}}{\partial x}(x-y)*M'(y)dy. \tag{5.5.25}$$

So we have the following expression for $\lambda(a,0)$.

$$\lambda(a,0)=\frac{W'(a)}{W(a)}=\frac{\sum_{n\geqslant 0}\frac{1}{(1+\theta)^n}(M^{*n})'(a)}{\sum_{n\geqslant 0}\frac{1}{(1+\theta)^n}M^{*n}(a)}. \tag{5.5.26}$$

Using (5.5.24), (5.5.25) and (5.5.26) can also provide expressions for $F_{0,0,a}(x)$ and $R_a^{(0)}(dy)$.

Proposition 5.5.7. *Consider an insurance risk process $(R_t)_{t\geqslant 0}$ of the form defined in (5.5.21) with an initial level $x\geqslant 0$ and let $a>0$ be a fixed critical drawdown size. Then*

$$\mathbb{P}_x[\underline{X}_{\tau_a}<0]=1-\frac{W(x\wedge a)}{W(a)}+\frac{W(x\wedge a)e^{-\beta(x\vee a)}}{W(a)}\left(G_{\beta,\beta}(x\vee a)-G_{\beta+\lambda(a,0),\beta}(x\vee a)\right),$$

where $G_{\gamma,\beta}(v)$, $\lambda(a,0)$ and $W(a)$ are given by (5.5.27), (5.5.26) and (5.5.24) respectively.

Proof. Once again like the procedure we have done in the previous subsection, to show this proposition we need to use Theorem 5.4.1. Here as $\sigma=0$ so naturally $\Delta^{(0)}(a)=0$. We just need to compute the integrals in the expression given for $\mathbb{P}_x[\underline{X}_{\tau_a}<0]$ in Theorem 5.4.1. The Lévy measure for the process S_t is $\nu(dx)=\alpha x^{-1}e^{-\beta x}dx$. So, by replacing this into $\int_{h>0}\left(1-e^{-\lambda(a,0)h}\right)\nu(x\vee a-y+dh)$ we have

$$\int_{h>0}\left(1-e^{-\lambda(a,0)h}\right)\nu(x\vee a-y+dh)=\alpha\int_0^{\infty}\left(1-e^{-\lambda(a,0)h}\right)(x\vee a-y+h)^{-1}e^{-\beta(x\vee a-y+h)}dh$$

$$= \alpha\Big(E_1(\beta(x \vee a - y)) - e^{\lambda(a,0)(x\vee a-y)}E_1((\beta+\lambda(a,0))(x\vee a-y))\Big),$$

where E_1 is the exponential integral function.
Now, define the following function.

$$G_{\gamma,\beta}(v) = \int_0^\infty e^{-\gamma h}\int_0^a (v+h-y)^{-1}e^{\beta y}R_a^{(0)}(dy)dh, \tag{5.5.27}$$

for $\gamma, v > 0$. It is clear that $G_{\beta,\beta}(a) = \frac{e^{\beta a}}{\alpha}$ because of

$$\alpha\int_0^a E_1(\beta(a-y))R_a^{(0)}(dy) = 1. \tag{5.5.28}$$

We use equation (5.4.3) to get equation (5.5.28). So by applying numerical methods we can compute the function $G_{\gamma,\beta}$.
To end the proof it is sufficient to apply (5.5.27) in (5.4.1). Thus we have

$$\mathbb{P}_x(\underline{X}_{\tau_a} < 0) = 1 - \frac{W(x\wedge a)}{W(a)} + \frac{W(x\wedge a)e^{-\beta x\vee a}}{W(a)}\Big(G_{\beta,\beta}(x\vee a) - G_{\beta+\lambda(a,0),\beta}(x\vee a)\Big).$$

□

Now, we are going to provide representations for distributions of each of the random variables Y_{τ_a-} and $Y_{\tau_a} - a$.

Proposition 5.5.8. *Consider an insurance risk process $(R_t)_{t\geqslant 0}$ of the form defined in (5.5.21) with an initial level $x \geqslant 0$ and let $a > 0$ be a fixed critical drawdown size. Then*

(1) the distribution of Y_{τ_a-}, the largest drawdown observed before the critical drawdown of size a, is

$$\mathbb{P}_x(Y_{\tau_a-} \in dy) = \alpha E_1(\beta(a-y))R_a^{(0)}(dy).$$

(2) the overshoot of critical drawdown over level a is:

$$\mathbb{P}_x(Y_{\tau_a} - a \in dh) = \alpha e^{-\beta(a+h)}\int_0^a (v+h-y)^{-1}e^{\beta y}R_a^{(0)}(dy)dh.$$

Proof. As $\sigma = 0$ thus $\Delta^{(0)}(a) = 0$.

(1) It is clear that $\int_0^\infty \nu(a-y+dh) = \alpha E_1(\beta(a-y))$. By plugging it into (5.4.4) we have

$$\mathbb{E}_x[\mathbb{1}_{\{Y_{\tau_a-}\in dy\}}] = \alpha E_1(\beta(a-y))R_a^{(0)}(dy).$$

(2) To prove the second part of the theorem we have

$$\nu(a-y+dh) = \alpha e^{-\beta(a-y+h)}(a-y+h)^{-1}dh. \tag{5.5.29}$$

for $h > 0$, $y \in [0, a]$. Substituting (5.5.29) in (5.4.5) gives

$$\mathbb{E}_x[\mathbb{1}_{\{Y_{\tau_a} - a \in dh\}}] = \alpha e^{-\beta(a+h)} \int_0^a (a + h - y)^{-1} e^{\beta y} R_a^{(0)}(dy) dh.$$

□

5.5.3. Spectrally Negative Stable process

In this subsection we are studying the running minimum at the first-passage time over a level a of the drawdown process Y. In fact, we study the probability which the running minimum at the first-passage time goes below 0 when X_t is a spectrally negative stable process with stability parameter $\alpha \in (1, 2)$. Furthermore, we give representations for distributions of each of the stochastic processes $\overline{X}_{\tau_a}, Y_{\tau_a -}$ and $Y_{\tau_a} - a$ associated to the main process X given above.

Let $(X_t)_{t \geq 0}$ be a spectrally negative stable process with stability parameter $\alpha \in (1, 2)$ and Laplace exponent $\psi(s) = s^\alpha$ for $s > 0$. Moreover, the Lévy measure in (5.3.1) is $\nu(dx) = \frac{\beta}{x^{1+\alpha}} dx$ for $x, \beta > 0$. It can be seen (see for example [64, 71]) that

$$W^{(q)}(x) = \alpha x^{\alpha - 1} E'_{\alpha,1}(q x^\alpha), \tag{5.5.30}$$

for $x, q \geqslant 0$ where $E_{\alpha,1}(z) = \sum_{k \geqslant 0} \frac{z^k}{\Gamma(1+\alpha k)}$ is the Mittag-Leffler function and Γ is the gamma function.

Proposition 5.5.9. *Let $(X_t)_{t \geq 0}$ be a stable process with stability parameter $\alpha \in (1, 2)$ and let $a > 0$ be a critical drawdown size. Then*

$$\lambda(a, q) = \frac{\alpha - 1}{a} + q \alpha a^{\alpha - 1} \frac{E''_{\alpha,1}(q a^\alpha)}{E'_{\alpha,1}(q a^\alpha)} , \tag{5.5.31}$$

$$R_a^{(q)}(dy) = \left[\alpha y^{\alpha - 2} E'_{\alpha,1}(q y^\alpha) \left(\frac{\alpha - 1}{\lambda(a, q)} - y \right) + \frac{q \alpha^2 y^{2\alpha - 2}}{\lambda(a, q)} E''_{\alpha,1}(q y^\alpha) \right] dy . \tag{5.5.32}$$

Proof. Taking derivative of (5.5.30) with respect to x and substitute it in $\lambda(a, q)$ given by (5.3.4) (in $R_a^{(q)}$ given by (5.3.6) respectively) yields (5.5.31)((5.5.32) respectively). □

As a particular case of Proposition 5.5.9, we have

$$\lambda(a, 0) = \frac{\alpha - 1}{a} \quad \text{and} \quad R_a^{(0)}(dy) = \frac{\alpha y^{\alpha - 2}}{\Gamma(1 + \alpha)} (a - y) dy. \tag{5.5.33}$$

Proposition 5.5.10. *Consider a spectrally negative stable process $(X_t)_{t \geqslant 0}$ with stability parameter $\alpha \in (1, 2)$ with an initial surplus $x \geqslant 0$ and let $a > 0$ be a* *fixed critical drawdown size. Then*

$$\mathbb{P}_x(\underline{X}_{\tau_a} < 0) = 1 - \left(\frac{x \wedge a}{a}\right)^{\alpha-1} + \beta \left(\frac{x \wedge a}{a}\right)^{\alpha-1} \left(\frac{H_{0,0,0}(x \vee a)}{\alpha} - H_{\lambda(a,0),\lambda(a,0),0}(x \vee a)\right),$$

where $\lambda(a, 0)$ *is given by* (5.5.31) *and* $g_{a,\alpha,0}(v)$ *is defined by* (5.5.34).

Proof. To show this proposition we need to use Theorem 5.4.1 one more time. The Lévy measure for the process $-X_t$ is $\nu(dh) = \frac{\beta}{h^{1+\alpha}} dh$ for $\beta > 0$. So, by replacing this into $\int_{h>0} \left(1 - e^{-\lambda(a,0)h}\right) \nu(x \vee a - y + dh)$ we have

$$\begin{aligned} \int_{h>0} \left(1 - e^{-\lambda(a,0)h}\right) \nu(x \vee a - y + dh) &= \int_{h > x \vee a - y} \left(1 - e^{\lambda(a,0)(x \vee a - y)} e^{-\lambda(a,0)h}\right) \frac{\beta}{h^{1+\alpha}} dh \\ &= \frac{\beta}{\alpha} \left[\frac{1}{(x \vee a - y)^\alpha} - e^{\lambda(a,0)(x \vee a - y)} \int_{h > x \vee a - y} \frac{\alpha e^{-\lambda(a,0)h}}{h^{1+\alpha}} dh\right]. \end{aligned}$$

Now define

$$H_{\gamma_1,\gamma_2,q}(v) = \int_0^\infty e^{-\gamma_1 h} \int_0^a \frac{e^{-\gamma_2 y} R_a^{(q)}(dy)}{(v - y + h)^{\alpha+1}} dh. \tag{5.5.34}$$

We can use the numerical methods to compute $H_{\gamma_1,\gamma_2,q}(v)$ for a given value v. For the special case $\gamma_1 = \gamma_2 = 0$, $v = a$ and $q = 0$ we have

$$H_{0,0,0}(a) = \int_0^a \frac{R_a^{(0)}(dy)}{(a - y)^\alpha} = \frac{\alpha}{\beta}, \tag{5.5.35}$$

because of

$$\int_0^a \int_0^\infty \frac{\beta}{(a - y + h)^{1+\alpha}} dh R_a^{(0)}(dy) = 1. \tag{5.5.36}$$

We use (5.4.3) to get equation (5.5.36). To end the proof it is sufficient to replace (5.5.34) and (5.5.35) into (5.4.1).
Therefore, after some simplifications we get

$$\mathbb{P}_x(\underline{X}_{\tau_a} < 0) = 1 - \left(\frac{x \wedge a}{a}\right)^{\alpha-1} + \beta \left(\frac{x \wedge a}{a}\right)^{\alpha-1} \left(\frac{H_{0,0,0}(x \vee a)}{\alpha} - H_{\lambda(a,0),\lambda(a,0),0}(x \vee a)\right).$$

□

Proposition 5.5.11. *Consider a spectrally negative stable process* $(X_t)_{t \geqslant 0}$ *with stability parameter* $\alpha \in (1, 2)$ *with an initial surplus* $x \geqslant 0$ *and let* $a > 0$ *be a fixed critical drawdown size. Then*

(1) the distribution of $Y_{\tau_a -}$*, the largest drawdown observed before the critical drawdown of size* a *is*

$$\mathbb{P}_x(Y_{\tau_a -} \in dy) = \frac{\beta y^{\alpha-2}}{\Gamma(1 + \alpha)(a - y)^{\alpha-1}} dy.$$

(2) the overshoot of critical drawdown over level a is

$$\mathbb{P}_x(Y_{\tau_a} - a \in dh) = \beta \int_0^a \frac{R_a^{(0)}(dy)}{(a-y+h)^{\alpha+1}} dh.$$

Proof. As $\sigma = 0$, thus $\Delta^{(0)}(a) = 0$.

(1) It can be easily shown that $\int_0^\infty \nu(a-y+dh) = \frac{\beta}{\alpha(a-y)^\alpha}$. By plugging it into (5.4.4) we get

$$\mathbb{E}_x[\mathbb{1}_{\{Y_{\tau_a-} \in dy\}}] = \frac{\beta}{\alpha(a-y)^\alpha} R_a^{(0)}(dy). \tag{5.5.37}$$

The proof of first part is done by replacing (5.5.33) into (5.5.37).

(2) To prove the second part of the theorem we have

$$\nu(a-y+dh) = \frac{\beta}{(a-y+h)^{\alpha+1}} dh. \tag{5.5.38}$$

Substituting (5.5.38) in (5.4.5) gives

$$\mathbb{E}_x[\mathbb{1}_{\{Y_{\tau_a} - a \in dh\}}] = \beta \int_0^a \frac{R_a^{(0)}(dy)}{(a-y+h)^{\alpha+1}} dh.$$

$\square$

In the sequel of this subsection, we are going to provide the joint Laplace transform of τ_a and $\overline{G}_{\tau_a}$. In fact, we are seeking the Laplace transform of the depletion random variable, $\tau_a - \overline{G}_{\tau_a}$.

Proposition 5.5.12. *Consider a spectrally negative stable process $(X_t)_{t \geqslant 0}$ with stability parameter $\alpha \in (1,2)$ with an initial surplus $x \geqslant 0$ and let $a > 0$ be a fixed critical drawdown size.*

(1) The bivariate Laplace transform of τ_a and $\overline{G}_{\tau_a}$ is given by

$$\mathbb{E}_x\left[e^{-q\tau_a - r\overline{G}_{\tau_a}}\right] = \frac{\beta\lambda(a,q)}{\alpha\lambda(a,q+r)} H_{0,0,q}(a),$$

where $H_{0,0,q}(a)$ is defined by (5.5.34) and $\lambda(a,q)$ is given by (5.5.31).

(2) The Laplace transform of the speed of depletion $\tau_a - \overline{G}_{\tau_a}$ is given by,

$$\mathbb{E}_x\left[e^{-q(\tau_a - \overline{G}_{\tau_a})}\right] = \frac{\beta\lambda(a,q)}{\alpha\lambda(a,0)} H_{0,0,q}(a)\ .$$

Proof. As $\sigma = 0$ thus $\Delta^{(0)}(a) = 0$ in Proposition 5.4.2 and Theorem 5.4.3.

(1) The Lévy measure of the process $-X$ is $\frac{\beta}{x^{\alpha+1}}dx$ and so $\int_0^\infty \nu(a-y+dh) = \frac{\beta}{\alpha(a-y)^\alpha}$. Now using Proposition 5.4.2 as it specializes to this case yields

$$\mathbb{E}_x\left[e^{-q\tau_a - r\overline{G}_{\tau_a}}\right] = \frac{\beta\lambda(a,q)}{\alpha\lambda(a,q+r)}\int_0^a \frac{R_a^{(q)}(dy)}{(a-y)^\alpha} = \frac{\beta\lambda(a,q)}{\alpha\lambda(a,q+r)}H_{0,0,q}(a).$$

This ends the proof of the first part.

(2) In a similar way, specializing Theorem 5.4.3 to this case using the expressions in (5.5.31) and (5.5.32) yields the result.

□

5.5.4. Brownian Perturbed Model without Claims

The Brownian Perturbed risk process R without claims is a Brownian motion with drift, starting at $x \geqslant 0$, i.e.

$$R_t = x + X_t\ , \tag{5.5.39}$$

where X is a Brownian motion with a drift given by

$$X_t = ct + \sigma B_t\ , \tag{5.5.40}$$

and $(B_t)_{t\geqslant 0}$ is assumed to be a standard Brownian motion. The form of the q-scale function in this model is relatively simple since the claim sizes disappear. In this case, the Laplace exponent in (5.3.1) takes a simple form and becomes

$$\psi(s) = c\,s + \frac{\sigma^2}{2}\,s^2\ , \qquad s > 0\ . \tag{5.5.41}$$

We obtain an explicit expression for the q-scale function (see [62] for more details) as

$$W^{(q)}(x) = \frac{1}{\sqrt{c^2+2q\sigma^2}}\left[e^{(\sqrt{c^2+2q\sigma^2}-c)\frac{x}{\sigma^2}} - e^{-(\sqrt{c^2+2q\sigma^2}+c)\frac{x}{\sigma^2}}\right] \tag{5.5.42}$$

Following the same order of ideas as in previous examples, we obtain explicit expressions for the functions λ, $R_a^{(q)}$ and $\Delta_a^{(q)}$. These are given in the following result.

Proposition 5.5.13. *Consider the process $(R_t)_{t\geq 0}$ in (5.5.39) and let $a > 0$ be a critical drawdown size. Then*

$$\lambda(a,q) = -\frac{c}{\sigma^2} + \frac{\sqrt{c^2+2q\sigma^2}}{\sigma^2}\left[\frac{1+e^{-2\frac{a}{\sigma^2}\sqrt{c^2+2q\sigma^2}}}{1-e^{-2\frac{a}{\sigma^2}\sqrt{c^2+2q\sigma^2}}}\right], \tag{5.5.43}$$

$$R_a^{(q)}(dy) = \left[-\left(\frac{c}{\sigma^2\,\lambda(a,q)}+1\right)W^{(q)}(y) + \frac{1}{\sigma^2\,\lambda(a,q)}\left(e^{(\sqrt{c^2+2q\sigma^2}-c)\frac{y}{\sigma^2}} + e^{-(\sqrt{c^2+2q\sigma^2}+c)\frac{y}{\sigma^2}}\right)\right]\mathbb{1}_{y\in(0,a]}dy\ ,$$

$$\Delta^{(q)}(a) = -\left(\frac{c}{2}+\frac{c^2+q\sigma^2}{\sigma^2\,\lambda(a,q)}\right)W^{(q)}(y) + \left(\frac{1}{2}-\frac{c}{\sigma^2\,\lambda(a,q)}\right)\left(e^{(\sqrt{c^2+2q\sigma^2}-c)\frac{a}{\sigma^2}} + e^{-(\sqrt{c^2+2q\sigma^2}+c)\frac{a}{\sigma^2}}\right),$$

where $W^{(q)}$ is given by (5.5.42).

Proof. We have

$$\lambda(a,q) = \frac{W_+'^{(q)}(a)}{W^{(q)}(a)} \ .$$

By referring to equation (5.5.42) we can see that $W^{(q)}$ is a differentiable function on $(0,\infty)$ so by differentiating equation (5.5.42) we can directly obtain

$$\begin{aligned} W'^{(q)}(x) &= -\frac{c}{\sigma^2}W^{(q)}(x) + \frac{1}{\sigma^2}\left[e^{(\sqrt{c^2+2q\sigma^2}-c)\frac{x}{\sigma^2}} + e^{-(\sqrt{c^2+2q\sigma^2}+c)\frac{x}{\sigma^2}}\right] \\ &= -\frac{c}{\sigma^2}W^{(q)}(x) + \frac{1}{\sigma^2}\left[e^{-c\frac{x}{\sigma^2}}\left(e^{\sqrt{c^2+2q\sigma^2}\frac{x}{\sigma^2}} + e^{-\sqrt{c^2+2q\sigma^2}\frac{x}{\sigma^2}}\right)\right]. \quad (5.5.44) \end{aligned}$$

Applying (5.5.42) and (5.5.44) in the definition (5.3.4) yields the first result.
As $W^{(q)}$ is an increasing function and does not have a mass at $x = 0$, we have $W^{(q)}(dx) = W'^{(q)}(x)dx$ (see [63]). Direct substitution into the definition (5.3.6) of $R_a^{(q)}$ yields the second result.
Differentiating equation (5.5.44) we can obtain

$$\begin{aligned} W''^{(q)}(x) &= -\frac{c}{\sigma^2}W'^{(q)}(x) - \frac{c}{\sigma^4}\left(e^{(\sqrt{c^2+2q\sigma^2}-c)\frac{x}{\sigma^2}} + e^{-(\sqrt{c^2+2q\sigma^2}+c)\frac{x}{\sigma^2}}\right) + \frac{c^2+2q\sigma^2}{\sigma^4}W^{(q)}(x) \\ &= \frac{2c^2+2q\sigma^2}{\sigma^4}W^{(q)}(x) - \frac{2c}{\sigma^4}\left(e^{(\sqrt{c^2+2q\sigma^2}-c)\frac{x}{\sigma^2}} + e^{-(\sqrt{c^2+2q\sigma^2}+c)\frac{x}{\sigma^2}}\right), \quad (5.5.45) \end{aligned}$$

where in the last equality we have substituted $W'^{(q)}(x)$ by its expression provided in (5.5.44). Combining (5.5.44) and (5.5.45) in equation (5.3.7) for $\Delta^{(q)}(a)$ yields the third result. □

In the following propositions, we will give explicit representations for the distributions in Theorem 5.4.1, 5.4.2 and 5.4.3 for the risk model (5.5.39).

Proposition 5.5.14. *Consider a process $(R_t)_{t\geqslant 0}$ of the form defined in (5.5.39) with an initial level $x \geqslant 0$ and let $a > 0$ be a fixed critical drawdown size. Then*

(1)

$$\mathbb{P}_x(\underline{X}_{\tau_a} < 0) = \frac{e^{-\frac{2c(x\wedge a)}{\sigma^2}} - e^{-\frac{2ca}{\sigma^2}}}{1 - e^{-\frac{2ca}{\sigma^2}}}. \quad (5.5.46)$$

(2) the largest drawdown observed before critical drawdown follows a Dirac measure at a. That is

$$\mathbb{P}_x(Y_{\tau_a-} \in dy) = \delta_a(dy) \ .$$

(3) the overshoot over the critical drawdown $Y_{\tau_a} - a$ follows a Dirac measure at 0. That is

$$\mathbb{P}_x(Y_{\tau_a} - a \in dh) = \delta_0(dh) \ .$$

Proof. (1) The expression given for $\mathbb{P}_x(\underline{X}_{\tau_a} < 0)$ in Theorem 5.4.1 will be reduced to

$$\mathbb{P}_x(\underline{X}_{\tau_a} < 0) = 1 - \frac{W(x \wedge a)}{W(a)}, \tag{5.5.47}$$

since the Lévy measure vanishes. Moreover, $W(x)$ is given by

$$W(x) = \frac{1}{c}\Big(1 - e^{-\frac{2cx}{\sigma^2}}\Big). \tag{5.5.48}$$

Thus, (5.5.46) holds.

(2) Note that for $q = 0$, (5.5.43) will be reduced to

$$\lambda(a, 0) \quad = \quad \frac{2c}{\sigma^2}\Bigg(\frac{e^{\frac{-2ac}{\sigma^2}}}{1 - e^{\frac{-2ac}{\sigma^2}}}\Bigg) \, . \tag{5.5.49}$$

On the other hand, by differentiating (5.5.48) we have

$$\begin{aligned} \Delta^{(0)}(a) \quad &= \quad \frac{\sigma^2}{2}\Big[W'(a) - \lambda(a,0)^{-1}W''(a)\Big] \\ &= \quad 1 \, . \end{aligned} \tag{5.5.50}$$

Therefore, $\mathbb{P}_x(Y_{\tau_a -} \in dy) = \delta_a(dy)$.

(3) The third part of this proposition can also be obtained in the same way.

□

In the following, we provide explicit expressions for joint Laplace transform of $\overline{G}_{\tau_a}$ and τ_a as well as the Laplace transform of the speed of depletion. Notice that the corresponding integral with respect to $R_a^{(q)}(\cdot)$ in Proposition 5.4.2 and Theorem 5.4.3 disappears since X_t has no jumps. In this case, these Laplace transforms are a simple function of $W^{(q)}(a)$ given in (5.5.42), and they can be computed in a straightforward way using Proposition 5.4.2 and Theorem 5.4.3.

Proposition 5.5.15. *Consider a process $(R_t)_{t \geqslant 0}$ of the form defined in (5.5.39) with an initial level $x \geqslant 0$ and let $a > 0$ be a fixed critical drawdown size. Then for $q, r \geqslant 0$*

(1) the bivariate Laplace transform of τ_a and $\overline{G}_{\tau_a}$ is given by

$$\begin{aligned} \mathbb{E}_x\left[e^{-q\tau_a - r\overline{G}_{\tau_a}}\right] \quad &= \quad \frac{1}{\lambda(a, q+r)}\Bigg[\Big(-\frac{\frac{c}{2}\lambda(a,q) + q + \frac{c^2}{\sigma^2}}{\sqrt{c^2 + 2q\sigma^2}} + \frac{1}{2}\lambda(a,q) - \frac{c^2}{\sigma^2}\Big)e^{(\sqrt{c^2+2q\sigma^2}-c)\frac{a}{\sigma^2}} \\ &\quad + \Big(\frac{\frac{c}{2}\lambda(a,q) + q + \frac{c^2}{\sigma^2}}{\sqrt{c^2 + 2q\sigma^2}} - \frac{1}{2}\lambda(a,q) + \frac{c^2}{\sigma^2}\Big)e^{-(\sqrt{c^2+2q\sigma^2}-c)\frac{a}{\sigma^2}}\Bigg] \, , \end{aligned} \tag{5.5.51}$$

with $\lambda(a, q)$ given in (5.5.43).

(2) the Laplace transform of the speed of depletion $\tau_a - \overline{G}_{\tau_a}$ is given by

$$\mathbb{E}_x\left[e^{-q(\tau_a - \overline{G}_{\tau_a})}\right] \quad = \quad \frac{\sigma^2(1 - e^{-\frac{2c}{\sigma^2}a})}{2c\, e^{-\frac{2c}{\sigma^2}a}}\Bigg[\Big(-\frac{\frac{c}{2}\lambda(a,q) + q + \frac{c^2}{\sigma^2}}{\sqrt{c^2 + 2q\sigma^2}} + \frac{1}{2}\lambda(a,q) - \frac{c^2}{\sigma^2}\Big)e^{(\sqrt{c^2+2q\sigma^2}-c)\frac{a}{\sigma^2}}$$

$$+(\frac{\frac{c}{2}\lambda(a,q)+q+\frac{c^2}{\sigma^2}}{\sqrt{c^2+2q\sigma^2}}-\frac{1}{2}\lambda(a,q)+\frac{c^2}{\sigma^2})e^{-(\sqrt{c^2+2q\sigma^2}-c)\frac{a}{\sigma^2}}\Big]\,. \tag{5.5.52}$$

PROOF. (1) As the risk process X has no jumps, then Proposition 5.4.2 will be reduced to

$$\mathbb{E}_x\left[e^{-q\tau_a-r\overline{G}_{\tau_a}}\right]=\frac{\lambda(a,q)}{\lambda(a,q+r)}\Delta^{(q)}(a). \tag{5.5.53}$$

We can then compute equation (5.5.53) using the given representation for $\Delta^{(q)}(a)$ in Proposition 5.5.13. That is,

$$\begin{aligned}\mathbb{E}_x\left[e^{-q\tau_a-r\overline{G}_{\tau_a}}\right] &= \frac{1}{\lambda(a,q+r)}\Big[(-\frac{\frac{c}{2}\lambda(a,q)+q+\frac{c^2}{\sigma^2}}{\sqrt{c^2+2q\sigma^2}}+\frac{1}{2}\lambda(a,q)-\frac{c^2}{\sigma^2})e^{(\sqrt{c^2+2q\sigma^2}-c)\frac{a}{\sigma^2}}\\ &+ (\frac{\frac{c}{2}\lambda(a,q)+q+\frac{c^2}{\sigma^2}}{\sqrt{c^2+2q\sigma^2}}-\frac{1}{2}\lambda(a,q)+\frac{c^2}{\sigma^2})e^{-(\sqrt{c^2+2q\sigma^2}-c)\frac{a}{\sigma^2}}\Big]\,,\end{aligned} \tag{5.5.54}$$

where in the last equality we have substituted $W^{(q)}(a)$ by its expression given by (5.5.42). Thus (5.5.51) holds.

(2) In a similar way, specializing Theorem 5.4.3 to this case and using the expression in (5.5.49) yield the result.

□

5.5.5. Meromorphic Risk Process (Beta Risk Process)

The beta risk model was introduced in [60]. It is of the form

$$R_t = x + X_t, \tag{5.5.55}$$

where the net aggregate claim process X is a beta process. That is,

$$X_t = ct - Z_t\ , \tag{5.5.56}$$

where the aggregate claims process $(Z_t)_{t\geqslant 0}$ has the following Lévy measure

$$\nu(dx)=\xi\,\beta\frac{e^{-(1+\alpha)\beta x}}{(1-e^{-\beta x})^{\lambda}}\,\text{x},\quad x>0\ , \tag{5.5.57}$$

with $\alpha,\beta,\xi>0$ and $\lambda\in(1,2)\cup(2,3)$.
Parameters α and β are responsible for the rate of decay of the tail of the Lévy measure and for the shape of this measure, parameter ξ controls the overall "intensity" of jumps, while the parameter λ describes the singularity of the Lévy measure at zero and therefore controls the intensity of the small jumps. Indeed, if $\lambda\in(1,2)$, then the process has jump part of infinite activity and finite variation while if $\lambda\in(2,3)$, the jump part of the process will be of infinite variation. For a thorough discussion of the beta family of processes we refer to [59].

The loaded premium c is again of the form $c = (1+\theta)\mathbb{E}[Z_1]$ for some safety loading factor $\theta > 0$. In turn, the Laplace exponent in (5.3.1) becomes,

$$\psi(s) = c\,s + \frac{1}{2}\sigma^2 s^2 + \xi\,\mathrm{B}(1+\alpha+s/\beta, 1-\lambda) - \xi\,\mathrm{B}(1+\alpha, 1-\lambda)\ . \qquad s > 0\ , \tag{5.5.58}$$

where $\mathrm{B}(x,y) = \Gamma(x)\Gamma(y)/\Gamma(x+y)$ is the Beta function.
Notice that this family has a Gaussian component that can be switched on and off using the σ parameter in (5.5.58) making our risk model a member of the so-called perturbed family of models.
The beta process is in fact an example of a member of a larger family of processes having either a rational or meromorphic Laplace exponent which is at the heart of their tractability. As was shown in [61], many results of fluctuation theory can be given rather explicitly for this family of processes. These processes are defined by requiring that their Lévy density $\Pi(\d x) = \pi(x)\d x$ is essentially a "mixture" of exponential distributions, and in the spectrally negative case this translates into the following definition

$$\pi(x) = \mathbb{1}_{\{x>0\}} \sum_{m=1}^{\infty} b_m e^{-\rho_m x}, \tag{5.5.59}$$

where all the coefficients b_m and ρ_m are positive and $\rho_1 < \rho_2 < \ldots$. The Laplace exponent $\psi(s)$ for a process with the Lévy measure (5.5.59) is a meromorphic ($M = +\infty$) function (see [61]). In the case of the beta process, it has been shown (see [59]) that its Lévy density (5.5.57) is of the form in (5.5.59) with coefficients

$$b_m = \xi\,\beta \binom{m+\lambda-2}{m-1}, \quad \rho_m = \beta(\alpha+m)\ . \tag{5.5.60}$$

In the following, we give the form of the q-scale function for beta process and study depletion quantities for this process. The following lemma is taken from [60]

Lemma 5.5.1. *Consider the process $(X_t)_{t\geq 0}$ in (5.5.56). Then, for $q \geq 0$*

$$W^{(q)}(x) = \frac{e^{\Phi(q)x}}{\psi'(\Phi(q))} + \sum_{n\geq 1} \frac{e^{-\zeta_n x}}{\psi'(-\zeta_n)},\ \ x > 0, \tag{5.5.61}$$

where $\{\Phi(q), -\zeta_1, -\zeta_2, \ldots\}$ is the set of simple poles of the function $\frac{e^{zx}}{\psi(z)-q}$.
As we have $\psi(s) = \mathbb{E}(X_1)s + O(s^2)$ when $s \to 0$, it gives $\Phi(0) = 0$. By setting $K(x) = \sum_{n\geq 1} \frac{e^{-\zeta_n x}}{\psi'(-\zeta_n)}$, we have $W^{(0)}(x) = \frac{1}{\mathbb{E}(X_1)} + K(x)$. Taking derivative of $W^{(0)}(x)$ provides us $\lambda(a,0)$. That is,

$$\lambda(a,0) = \frac{W'^{(0)}(x)}{W^{(0)}(x)} = \frac{\mathbb{E}(X_1)K'(x)}{1+\mathbb{E}(X_1)K(x)}. \tag{5.5.62}$$

Using (5.5.62) and $W''^{(0)}(x) = K''(x) = \sum_{n\geq 1} \frac{\zeta_n^2 e^{-\zeta_n x}}{\psi'(-\zeta_n)}$, we can also find expressions for $F_{0,0,a}(x), R_a^{(0)}(dy)$ and $\Delta^{(0)}(a)$.

Proposition 5.5.16. *Consider a meromorphic beta spectrally negative process $(X_t)_{t\geqslant 0}$ with an initial surplus $x \geqslant 0$ and let $a > 0$ be a fixed critical drawdown size. Then*

(1)

$$\mathbb{P}_x(\underline{X}_{\tau_a} < 0) = 1 - \frac{1+\mathbb{E}(X_1)K(x\wedge a)}{1+\mathbb{E}(X_1)K(a)} + \frac{1+\mathbb{E}(X_1)K(x\wedge a)}{1+\mathbb{E}(X_1)K(a)}\left[G(b,\rho,0) - G(b,\rho,\lambda(a,0))\right], \tag{5.5.63}$$

(2) the distribution of $Y_{\tau_a -}$, the largest drawdown observed before the critical drawdown of size a, is:

$$\mathbb{P}_x(Y_{\tau_a -} \in dy) = \sum_{m=1}^{\infty} \frac{b_m}{\rho_m} e^{-\rho_m(a-y)} R_a^{(0)}(dy) + \Delta^{(0)}(a)\delta_a(dy),$$

(3) the overshoot of critical drawdown over level a is:

$$\mathbb{P}_x(Y_{\tau_a} - a \in dh) = \sum_{m=1}^{\infty} b_m e^{-\rho_m(a+h)} dh \int_0^a e^{\rho_m y} R_a^{(0)}(dy) + \Delta^{(0)}(a)\delta_0(dy),$$

where $G(c,d,k)$ is defined in (5.5.65).

Proof. To show the first part of this proposition we need to use Theorem 5.4.1. We just need to compute the integrals in the expression given for $\mathbb{P}_x[\underline{X}_{\tau_a} < 0]$ in Theorem 5.4.1. The Lévy measure for the process X_t is given by (5.5.59). So, by replacing this into $\int_{h>0}\left(1 - e^{-\lambda(a,0)h}\right)\nu(x\vee a - y + dh)$ we have

$$\begin{aligned}\int_{h>0}\left(1 - e^{-\lambda(a,0)h}\right)\nu(x\vee a - y + dh) &= \int_{h>0}\left(1 - e^{-\lambda(a,0)h}\right)\sum_{m=1}^{\infty} b_m e^{-\rho_m(x\vee a - y + h)} dh \\ &= \sum_{m=1}^{\infty}\left[\frac{b_m}{\rho_m} - \frac{b_m}{\lambda(a,0)+\rho_m}\right] e^{-\rho_m(x\vee a - y)}.\end{aligned} \tag{5.5.64}$$

For two sequences $c = (c_m)_{m\geq 1}$ and $d = (d_m)_{m\geq 1}$ and constant $k \in \mathbb{R}$ define

$$G(c,d,k) = \sum_{m=1}^{\infty} \frac{c_m}{d_m + k} \int_0^a e^{-\rho_m(x\vee a - y)} R_a^{(0)}(dy). \tag{5.5.65}$$

We can compute (5.5.65) numerically. Therefore, using (5.5.64) and (5.5.65) we have

$$\int_0^a \int_{h>0}\left(1 - e^{-\lambda(a,0)h}\right)\nu(x\vee a - y + dh) R_a^{(0)}(dy) = G(b,\rho,0) - G(b,\rho,\lambda(a,0)). \tag{5.5.66}$$

To finish the proof we need to substitute (5.5.66) in equation (5.4.1) in Theorem 5.4.1. That is,

$$\mathbb{P}_x[\underline{X}_{\tau_a} < 0] = 1 - \frac{1+\mathbb{E}(X_1)K(x\wedge a)}{1+\mathbb{E}(X_1)K(a)} + \frac{1+\mathbb{E}(X_1)K(x\wedge a)}{1+\mathbb{E}(X_1)K(a)}\left[G(b,\rho,0) - G(b,\rho,\lambda(a,0))\right].$$

To prove the second part of the proposition we use equation (5.4.4) in Theorem 5.4.2. We have

$$\int_{h>0} \nu(a-y+dh)R_a^{(0)}(dy) = \sum_{m=1}^{\infty} \frac{b_m}{\rho_m} e^{-\rho_m(a-y)} R_a^{(0)}(dy). \tag{5.5.67}$$

By substituting (5.5.67) into (5.4.4) we get

$$\mathbb{P}_x(Y_{\tau_a -} \in dy) = \sum_{m=1}^{\infty} \frac{b_m}{\rho_m} e^{-\rho_m(a-y)} R_a^{(0)}(dy) + \Delta^{(0)}(a)\delta_a(dy).$$

To show the third part of the proposition we follow the same procedure as applied for the first and second parts. In fact, we use equation (5.4.5) in Theorem 5.4.2. In the other words, we have

$$\int_{y\in[0,a]} \nu(a-y+dh)R_a^{(0)}(dy) = \sum_{m=1}^{\infty} b_m e^{-\rho_m(a+h)}dh \int_0^a e^{\rho_m y} R_a^{(0)}(dy). \tag{5.5.68}$$

By substituting (5.5.68) into (5.4.5) we get

$$\mathbb{P}_x(Y_{\tau_a} - a \in dh) = \sum_{m=1}^{\infty} b_m e^{-\rho_m(a+h)}dh \int_0^a e^{\rho_m y} R_a^{(0)}(dy) + \Delta^{(0)}(a)\delta_0(dy).$$

$\square$

Remark 5.5.2. *It can be seen from Proposition 5.5.16 that we come up with expressions included series for depletion quantities. These expressions might be not straightforward to compute but we can use numerical method to compute these expressions. We can also apply numerical methods to study other depletion quantities like* $\mathbb{E}_x\left[e^{-q(\tau_a - \overline{G}_{\tau_a})}\right]$.

5.6. Simulations

5.6.1. Classical Cramer-Lundberg mode with Exponential Claims

We consider a drifted compound Poisson process, introduced in Subsection 5.5.1, with $\lambda = 5$, exponential claims with parameter $\mu = 1$ starting from an initial capital $x > 0$ with a loading factor $\theta > 0$.

We first consider a sample of size $n = 10000$ of a such process on the time interval $[0, 1000]$ and compute the depletion quantities by Monte-Carlo method with $a = 6$, the initial capital $x = 5$ and the loading factor $\theta = 1$. Note that $c = (1+\theta)\lambda/\mu = 10$.

In Figure 5.3, we observe the convergence of the empirical probability of ruin before the first critical drawdown (the black line) to its theoretical value (the red line), obtained by Formula (5.5.6), when n is increasing.
Figure 5.4 illustrates the second point of Proposition 5.5.4 by comparing the empirical distribution of the overshoot $Y_{\tau_a} - a$ over the critical drawdown and the exponential density with parameter μ (the red line).

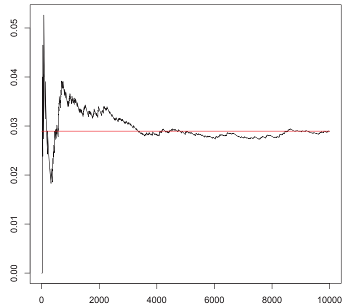


FIG. 5.3. Empirical probability of ruin before the first critical drawdown.

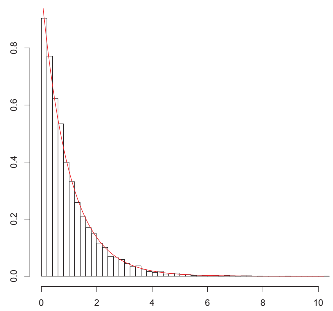
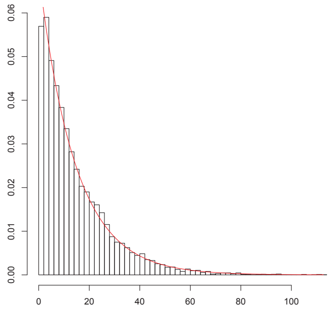
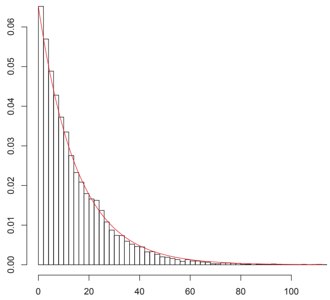


FIG. 5.4. Empirical distribution of the overshoot $Y_{\tau_a} - a$.

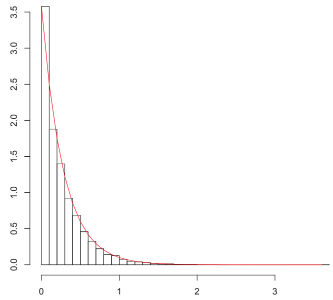
In Figure 5.5, we give the empirical distribution of τ_a, G_{τ_a} and the speed of depletion $\tau_a - G_{\tau_a}$ respectively and compare them to an exponential density (the red line) which seems to be a good estimation for the two first variables. The histogram of $\tau_a - G_{\tau_a}$ suggests that with a positive probability τ_a is equal to G_{τ_a}.



Histogram of τ_a.

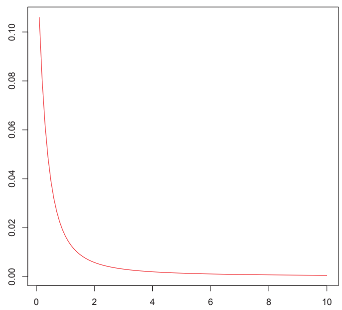


Histogram of G_{τ_a}

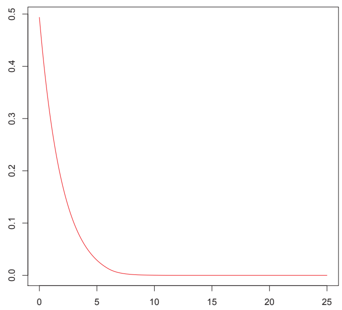


Histogram of $\tau_a - G_{\tau_a}$.

FIG. 5.5. Empirical distributions for a drifted compound Poisson process.



Evolution in $\theta \in (0, 10]$, with $x = 5$.



Evolution in $x \in [0, 25]$, with $\theta = 1$.

FIG. 5.6. Evolution of the probability of ruin before the first critical drawdown.

We now observe, in Figure 5.6, the evolution of the probability of ruin before the first critical drawdown in the loading factor θ and in the initial surplus x. The probability was computed using Formula 5.5.6.

5.6.2. Gamma Risk Process

We now consider a Gamma risk process, introduced in Subsection 5.5.2, with $\alpha = \beta = 1$ starting from the initial capital $x = 5$ with the loading factor $\theta = 0.2$. The loaded premium is then equal to $c = (1+\theta)\alpha/\beta = 1.2$.

We first consider a sample of size $n = 10000$ of a such process on the time interval $[0, T]$ with $T = 1000$ and compute the depletion quantities by Monte-Carlo method with $a = 3$.

We use a time step $h = T * 2^{-15}$ to generate the increments of a Gamma process on $[0, T]$.

In Figure 5.7, we observe the convergence of the empirical probability of ruin before the first critical drawdown (the black line) to a positive value which is equal to 0.97% (the red line), when n is increasing.

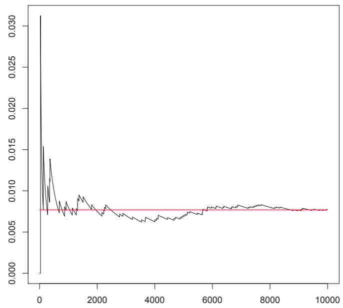
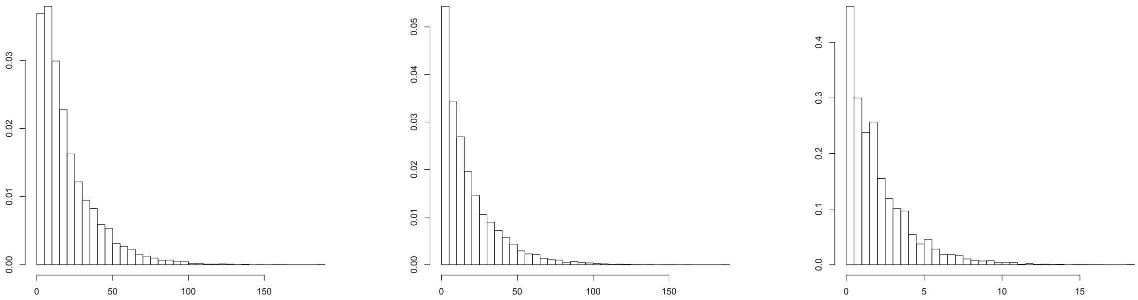


FIG. 5.7. Empirical probability of ruin before the first critical drawdown.

In Figure 5.8, we give the empirical distribution of τ_a, G_{τ_a} and the speed of depletion $\tau_a - G_{\tau_a}$ respectively. We observe that they are more distinct of an exponential density than in the Cramer-Lundberg model.



Histogram of τ_a. Histogram of G_{τ_a} Histogram of $\tau_a - G_{\tau_a}$.

FIG. 5.8. Empirical distributions for a Gamma process.

Chapter 6

CONCLUSIONS AND FUTURE RESEARCH DIRECTIONS

In this chapter, we will review the research contributions of this thesis, as well as discuss directions for future research.

6.1. Contributions

The following are the main research contributions of this thesis.

The first contribution the author of this thesis has made was designing a new class of coherent risk measures on space of stochastic processes. Both designing risk measures on space of stochastic processes and applying them to practical problems turn out to be a hard task. We circumvented these challenges by using a coherent risk measure defined on space of random variables as a benchmark risk measure which yields to come up with a class of desired risk measures on space of stochastic processes. This was done by multiplying the benchmark risk measure by a suitable weight function and integrating over a fixed period of time. The elements of this class satisfy in an axiomatic definition proposed in [24] and also capture the risk associated with the path-properties of a financial or an insurance model. Moreover, they have smooth explicit representations which enable us to easily compute without having to make use of the advanced mathematical tools. Finally, they can be enough tractable to apply them in practice. For instance, We applied them to the problem of capital allocation and derived interesting explicit results for some models in insurance context.

Next contribution the author of this thesis has made in this thesis was studying and designing a class of risk measures defined on data sets instead of dealing with random vectors. A risk measure on the space of random vectors assigns a set of vectors, instead of a real single value to a random vector. These vector-valued risk measures defined on space of random vectors are not enough tractable and straightforward to apply them into practice. We circumvented this problem by proposing multivariate data sets instead of using random vectors and defining a multivariate data-based risk measure. By inspiring from the definition of natural risk statistics proposed in [53] and its properties, we proposed to use a couple ordering which yields to define a class of risk measures, defined on multivariate data sets, satisfying in an axiomatic definition.

Introducing and proposing new path-related concepts in the collective risk theory which have not yet been studied in an insurance management context was another contribution the author has made in this thesis. These quantities are quantities in collective risk theory which are related to the so-called Lévy insurance risk processes and the first passage problem. These quantities are drawdowns and the speed of depletion, the speed at which an insurance reserve depletes as a consequence of the risk exposure of the company. We have derived expressions for many drawdown-related quantities in some cases of Lévy insurance risk processes for which they can be calculated, in particular for the classical Cramer-Lundberg model.

6.2. Future Research Directions

In the following we will point out several research directions which can be explored in different ways.

One of such directions would be to investigate possible ways to define another class of risk measures on space of stochastic processes using the obtained results in Chapter 5. One possibility could be to seek the density functions (if exist) for depletion-related quantities explored in Chapter 5 and apply as weight functions in risk measure introduced in (2.2.1). Deriving a new coherent risk measure defined on space of random variables rather than the one given in (2.2.2) and plug it in (2.2.1) would also be of interest. Next, it could be interesting to study the problem of capital allocation and the problem of portfolio optimization for such risk measures.

Another possible research direction could be to generalize the method used in Chapter 4 to define multivariate data-based risk measures. In the other words, one can study this class of data-based risk measures under another couple ordering than the one introduced in Chapter 4. In this case, applying obtained class of risk measures to practice and investigating different properties of them would also be of interest.

Bibliography

[1] Acerbi, C., Simonetti, P. *Portfolio optimization with spectral measures of risk*. arXiv:cond-mat/0203607v1, 2002.

[2] Ahmadi-Javid A. *Entropic Value-at-Risk: A New Coherent Risk Measure.* Journal of Optimization Theory and Application, Springer. 2011.

[3] Ahmed, D., Filipovic, S., Svindland, G. *A note on natural risk statistics.* Oper. Res. Lett., 36(6):662-664, 2008.

[4] Aliprantis, C.D., Border, K.C. *Infinite dimensional analysis*. Springer, Berlin, third edition, 2006. A hitchhiker's guide.

[5] Applebaum D. *Lévy processes and Stochastic Calculus* (second edition), Cambridge University Press, 2009.

[6] Artzner P., Delbaen F., Eber J.-M., and Heath D. *Coherent measures of risk*. Math. Finance, 1999.

[7] Asmussen S. *Ruin probabilities*, volume 2 of Advanced Series on Statistical Science & Applied Probability. World Scientific Publishing Co. Inc., River Edge, NJ, 2000.

[8] Assa, H., Morales, M. *Risk measures on the space of infinite sequences*. Math. Financ. Econ, 2(4):253-275, 2010.

[9] Assa H. *On Some Aspects of Coherent Risk Measures and their Applications* Ph.D.Thesis, 2011.

[10] Assa H. *Lebesgue property of risk measures for bounded cádlág processes and applications.* Methods and Applications of Analysis, 18(3):335-350, 2012.

[11] Assa H., Morales M., Omidi Firouzi H. *On the Capital Allocation Problem for a New Coherent Risk Measure in Collective Risk Theory.* submitted.

[12] Avram, F., Kyprianou, A.E., and Pistorius., M.R.*Exit problems for spectrally negative Lévy processes and applications to (Canadized) Russian options.* Ann. Appl. Probab., 14(1):215-238, 2004.

[13] Basak, S., Shapiro, A., *Value-at-risk based risk management: optimal policies and asset prices.* Rev. Financ. Stud. 14, 371–405, 2001.

[14] Belegundu A. D., Chandrupatla R. T., *Optimization Concepts and Applications in Engineering*. Cambridge University Press, 2nd edition, 2011.

[15] Ben Salah, Z., Guérin, H., Morales, M., Omidi Firouzi, H. *On the Depletion Problem for an Insurance Risk Process: New Non-ruin Quantities in Collective Risk Theory.* submitted.

[16] Bentahar, I. *Tail conditional expectation for vector-valued risks*. SFB 649 Discussion Papers, 2006.

[17] Bertoin, J. *Lévy Processes*, volume 121 of Cambridge Tracts in Mathematics. Cambridge University Press, 1996.

[18] Billera, L. J., Heath, D. C., *Allocation of shared costs: a set of axioms yielding a unique procedure*. Math. Oper. Res., 7(1):32–39, 1982.

[19] Biffis, E., Morales, M. *On a generalization of the Gerber-Shiu function to path-dependent penalties* Insurance: Mathematics and Economics. Volume 46, Issue 1, Pages 92–97, 2010.

[20] Biffis, E., Kyprianou, A. *A Note on Scale Functions and the Time Value of Ruin for Lévy Insurance Risk Processes* Insurance: Mathematics and Economics. 46, no. 1, 85–91, 2010.

[21] Cahuich, L.D., Hernandez, D.H. *Quantile Portfolio Optimization Under Risk Measure Constraints*. Appl Math Optim 68:157–179, 2013.

[22] Cascos, I., Molchanov, I. *Multivariate risks and depth-trimmed regions*. Finance and Stochastics, 11(3):373-397, 2007.

[23] Chambers, C. L., Mallows, J. M., Stuck, B. W. *A method for simulating stable random variables*. Journal of the American Statistical Association, 71(354):340-344, 1976.

[24] Cheridito, P., Delbaen F., Kupper M. *Coherent and convex monetary risk measures for bounded càdlàg processes.* Stochastic Process. Appl., 112(1):1–22, 2004.

[25] Cheridito, P., Delbaen F., Kupper M. *Coherent and convex monetary risk measures for unbounded càdlàg processes.* Finance Stoch., 9(3):369–387, 2005.

[26] Cheridito, P., Delbaen F., Kupper M. *Dynamic monetary risk measures for bounded discrete-time processes*. Electron. J. Probab., 11:no. 3, 57–106 (electronic), 2006.

[27] Cherny, A. *Pricing with coherent risk*. translation in Theory Probab. Appl. 52, no. 3, 2008.

[28] Cherubini, U., Luciano, L., Vecchiato, W. *Copula Methods in Finance.* ISBN: 978-0-470-86344-2. John Wiley & Sons Ltd, 2004.

[29] Chicheportiche, R., Bouchaud, J. P. *The joint distribution of stock returns is not elliptical.* International Journal of Theoretical and Applied Finance,

15(03), 2012.

[30] Cramér, H. *Historical review of Filip Lundberg's works on risk theory.* Scandinavian Actuarial Journal, 6–9, 1969.

[31] Cont, R., Deguest, R., Scandolo, G. *Robustness and sensitivity analysis of risk measurement procedures.* Quantitative Finance, 10(6), 2010.

[32] Cossette, H., Mailhot, M., Marceau, É., Mesfioui, M. *Bivariate lower and upper orthant value-at-risk.* European Actuarial Journal, 3(2), 2013.

[33] Cossette, H., Mailhot, M., Marceau, É., Mesfioui, M. *Vector-valued Tail Value-at-Risk and capital allocation*. Submitted, 2014.

[34] Delbaen, F. *Coherent risk measures.* Cattedra Galileiana. [Galileo Chair]. Scuola Normale Superiore, Classe di Scienze, Pisa, 2000.

[35] Delbaen, F. *Coherent risk measures on general probability spaces.* In Advances in finance and stochastics, pages 1-37. Springer, Berlin, 2002.

[36] Delbaen, F. *Coherent Monetary Utility Functions*, Preprint, Available at http://www.math.ethz.ch/$\sim$ delbaen under the name "Pisa Lecture Notes", 2005.

[37] Denault, M. *Coherent allocation of risk capital.* Journal of Risk, 4:1-34, 2001.

[38] Dufresne, D., Gerber, H., and Shiu., E. *Risk theory with the gamma process.* ASTIN Bulletin, 21(2), 1991.

[39] Embrechts, P., McNeil, A., Straumann, D. *Correlation and dependency in risk management: properties and pitfalls. In Risk Managemenmt: Value at Risk and Beyond, ed. by M. Dempster, and H. Moffatt.* Cambridge University Press, 2001.

[40] Embrechts, P., Puccetti, G. *Bounds for functions of multivariate risks.* J. Multivariate Anal. 97(2), 526-547, 2006.

[41] Feller, W. *An introduction to probability theory and its applications.* Vol. I. John Wiley and Sons, Inc., New York, 1957. 2nd ed.

[42] Fischer, T. *Risk Capital Allocation by Coherent Risk Measures Based on One-Sided Moments*, Insurance.: Math. Econ. 32(1), 135–146, 2003.

[43] Follmer, H., Schied, A. *Convex Measures of Risk and Trading Constraints.* Finance and Stochastics, 2002.

[44] Follmer, H., Schied, A. *Robust Preferences and Convex Measures of Risk.* Finance and Stochastics, 2002.

[45] Follmer H., Schied, A. *Stochastic finance, An introduction in discrete time.* volume 27 of de Gruyter Studies in Mathematics. Walter de Gruyter & Co., Berlin, extended edition, 2004.

[46] Follmer, H., Knispel, T. *Entropic risk measures: coherence vs. convexity, model ambiguity, and robust large deviations*. Stoch. Dyn. 11, no. 2-3, 333-351, 2011.

[47] Fritelli, M., Gianin, E. *Putting Order in Risk Measures*. Journal of Banking and Finance, 26(7), 2000.

[48] Furrer, H.J. *Risk Processes Perturbed by a α-stable Lévy Motion*. Scandinavian Actuarial Journal.**1.** pp. 59–74, 1998.

[49] Garrido, J., Morales, M. *On the expected discounted penalty function for Lévy risk processes*. North American Actuarial Journal, 10(4):196-217, 2006.

[50] Gerber, H.U., Shiu, E.S.W. *On the time value of ruin*. North American Actuarial Journal, 2(1):48-78, 1998.

[51] Grandell, J. A *Class of Approximations of Ruin Probabilities*. Scandinavian Actuarial Journal. **Supp**. pp. 37–52, 1977.

[52] Hamel, A. *Translative sets and functions and their applications to risk measure theory and nonlinear separation*. Working paper, 2006.

[53] Heyde, C. C., Kou, S. G., Peng, X. H. *What is a Good External Risk Measure: Bridging the Gaps between Robustness, Subadditivity, and Insurance Risk Measures*, (2007)

[54] Hu, W. *Calibration of multivariate generalized hyperbolic distributions using the EM algorithm, with applications in risk management, portfolio optimization and portfolio credit risk*. Thesis (Ph.D.)-The Florida State University. 115 pp. ISBN: 978-0542-66575-2, 2005.

[55] Huzak, M., Perman, M., Sikic, H., Vondracek, Z. *Ruin probabilities and decompositions for general perturbed risk processes*. Ann. Appl. Probab., 14(3):1378-1397, 2004.

[56] Jouini, E., Meddeb, M., Touzi, N. *Vector-valued coherent risk measures*. Finance Stoch., 8(4):531-552, 2004.

[57] Kalkbrener, M. *An axiomatic characterization of capital allocations of coherent risk measures*. Quantitative Finance, 9:8, 961–965, 2009.

[58] Kluppelberg, C., Kyprianou, A.E., Maller, R.A. *Ruin probabilities for general Lévy insurance risk processes*. Ann. Appl. Probab. 14, 1766-1801, 2004.

[59] Kuznetsov, A. *Wiener-Hopf factorization and distribution of extrema for a family of Lévy processes*. Ann. Appl. Prob., 20(5):1801-1830, 2010.

[60] Kuznetsov, A., Morales, M. *Computing the finite-time expected discounted penalty function for a family of Lévy risk processes*. Scandinavian Actuarial Journal, 2011.

[61] Kuznetsov, A., Kyprianou, A.E., Pardo, J.C. *Meromorphic Lévy processes and their fluctuation identities*. Ann. Appl. Probab., to appear, 2010.

[62] Kuznetsov, A., Kyprianou, A.E., Rivero, V. *The theory of scale functions for spectrally negative Lévy processes.* arXiv, 2011.

[63] Kyprianou, A.E. *Introductory Lectures on Fluctuations of Lévy processes with Applications.* Springer-Verlag, 2006.

[64] Kyprianou, A.E., Rivero, V. *Special, conjugate and complete scale functions for spectrally negative Lévy processes.* Electronic Journal of Probability, 13:1672-1701, 2008.

[65] Loisel, S., Trufin, J. *Properties of a risk measure derived from the expected area in red.* Insurance: Mathematics and Economics, Volume 55, 191-199, 2014.

[66] Morales, M. *Risk theory with the generalized inverse Gaussian Lévy process.* ASTIN Bulletin 34, 361–377. 2004,

[67] Morales, M. *On the expected discounted penalty function for a perturbed risk process driven by a subordinator.* 2007.

[68] Lundberg, F. *Approximerad framställning av sannolikhetsfunktionen. aterförsäkring av kollektivrisker.* Akad. Afhandling. Almqvist. och Wiksell, 1903.

[69] Mailhot, M., Morales M., Omidi Firouzi, H. *Multivariate Data-Based Natural Risk Statistics.* submitted.

[70] Markowitz, H.M. *Portfolio Selection.* The Journal of Finance 7 (1): 77-91, 1952.

[71] Mijatovic, A., Pistorius, M. *On the drawdown of completely asymmetric Lévy processes.* Stochastic Processes and Their Applications, 22:3812-3836, 2012.

[72] Morales, M., Schoutens, W. *A risk model driven by Lévy processes.* Appl. Stoch. Models Bus. Ind., 19:147-167, 2003.

[73] Nolan, J.P. *Stable Distributions: Models for Heavy Tailed Data.* Retrieved 2009-02-21

[74] Omidi Firouzi, H., Luong, A. *Optimal Portfolio Problem Using Entropic Value at Risk: When the Underlying Distribution is Non-Elliptical.* submitted.

[75] Pistorius, M.R. *On exit and ergodicity of the spectrally one-sided Lévy process reflected at its infimum.* J. Theoret. Probab., 17(1):183-220, 2004.

[76] Press, S. J. *A Compound Events Model for Security Prices.* Journal of Business, Vol. 40, No. 3, 1967.

[77] Press, S. J. *A Compound Poisson Process For Multiple Security Analysis.* Random Counts in Scientific Work: Random Counts in Physical Science, Geo Science and Business v. 3, 1971.

[78] Rachev, S.T., Stoyanov, S.V., Fabozzi, F.J. *Advanced Stochastic Models, Risk Assessment, and Portfolio Optimization: The Ideal Risk, Uncertainty, and*

Performance Measures. Wiley, ISBN: 978-0-470-05316-4, 2008.

[79] Rockafellar, R. *Convex Analysis*. Princeton University Press, Princeton, 1970.

[80] Ruijter M. J., Oosterlee C. W. *Two-dimensional Fourier cosine series expansion method for pricing financial options*. SIAM J. Sci. Comput. 34, no. 5, 2012, P. 642-671

[81] Sato, K.-I. *Lévy processes and Infinitely Divisible Distributions*. Cambridge University Press (1999).

[82] Tasche, D. *Risk contributions and performance measurement.* Technical document, Munich University of Technology, Munich, 1999.

[83] Tasche, D. *Euler allocation: theory and practice*. Technical document. Fitch Ratings, London, 2007.

[84] Trufin, J., Albrecher, H., Denuit, M. *Properties of a Risk Measure Derived from Ruin Theory*. The geneva risk and insurance review palgrave, 36, 174–188, 2011.

[85] Tsai, C.C.L., Willmot, G.E. *A generalized defective renewal equation for the surplus process perturbed by diffusion*. Insurance: Mathematics and Economics 30, 51–66, 2002.

[86] Millossovich, P., Tsanakas, A. *Sensitivity Analysis of Internal Risk Models.* submitted.

[87] Tsukahara, H. *One-parameter families of distortion risk measures*. Math. Finance 19, 691–705, 2009.

[88] Urban, M., Dittrich, J., Kluppelberg, C., Stolting, R. *Allocation of risk capital to insurance portfolios*. Blatter der DGVFM, Springer Berlin / Heidelberg. Pages 389–406, 2004.

[89] Vonesh, E., Chinchilli, V.M. *Linear and Nonlinear Models for the Analysis of Repeated Measurements*. CRC Press, 1996.

[90] Wang, S., Young, V., Panjer, H. *Axiomatic characterization of insurance prices*. Insurance Mathematics and Economics, 21:173-183, 1997.

[91] Weron, R. *On the chambers-mallows-stuck method for simulating skewed stable random variables*. Statistics & Probability Letters, 28:165-171, 1996.

[92] Weron, R. *Correction to: On the chambers-mallows-stuck method for simulating skewed stable random variables*.

[93] Zhang, H., Hadjiliadis, O. *Drawdowns and the speed of market crash*, Methodology and Computing in Applied Probability. 14(3):739-752, 2012.

Appendix A

MEASURE THEORY AND STOCHASTIC PROCESSES

A.1. Measure and probability space

The aim of this section is to recall key notions of measure theory and probability that will be used extensively throughout this thesis. The following discussion is mostly taken from Section 1.1.1 of [5].

Definition A.1.1. *([5]) Let Ω be a non-empty set and $\mathcal{F}$ a collection of subsets of Ω . We call $\mathcal{F}$ a σ-algebra if the following hold:*

(1) Non-empty. $\Omega \in \mathcal{F}$.

(2) Closed under complement. If $A \in \mathcal{F}$, then $A^c \in \mathcal{F}$.

(3) Closed under countable union. If $(A_n)_{n\in\mathbb{N}}$ is a sequence of subsets in $\mathcal{F}$, then $\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$.

The pair $(\Omega, \mathcal{F})$ is called a measurable space.

Definition A.1.2. *([5]) A measure $\mathbb{P}$ defined on $(\Omega, \mathcal{F})$ is a mapping $\mathbb{P} : \mathcal{F} \to [0, \infty)$ that satisfies*

(1) $\mathbb{P}(\emptyset) = 0$.

(2) For every sequence $(A_n)_{n\in\mathbb{N}}$ of mutually disjoint sets in $\mathcal{F}$

$$\mathbb{P}(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mathbb{P}(A_n).$$

The triple $(\Omega, \mathcal{F}, \mathbb{P})$ is called a measure space. If $\mathbb{P}(\Omega) = 1$ then the triple $(\Omega, \mathcal{F}, \mathbb{P})$ is called a probability space.

Definition A.1.3. *([5]) A Borel σ-algebra of $\mathbb{R}$ is the smallest σ-algebra of subsets of $\mathbb{R}$ that contains all the open sets (or equivalently, all closed sets) and is denoted by $\mathcal{B}(\mathbb{R})$. Any measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ is called a Borel measure.*

Definition A.1.4. *([5]) Let $(\Omega, \mathcal{F})$ be a measurable space. The function $X : \Omega \to \mathbb{R}$ is a random variable if $\{\omega : X(\omega) \leq r\} \in \mathcal{F}$ for all $r \in \mathbb{R}$.*

Definition A.1.5. *([5]) let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability measure and X be a random variable. Its law (or distribution) is the Borel probability measure $\mathbb{P}_X$ on $\mathbb{R}$ defined by $\mathbb{P}_X = \mathbb{P}oX^{-1}$ where "o" is the composition operator between functions $\mathbb{P}$ and X^{-1}.*

Now, let $(\Omega, \mathcal{F}, \mathbb{P})$ contain all null sets (i.e. is complete). The expectation $\mathbb{E}$ is a functional defined as $\mathbb{E}(X) = \int_\Omega |X| d\mathbb{P}$, where $|.|$ is the absolute value function.

In this thesis we use different stochastic models for studying the random behavior of surplus of financial and insurance companies. These models use stochastic processes for which we provide the following definition.

Definition A.1.6. *([5]) Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. A stochastic process $(X_t)_{t\in[0,T]}$ is a function from $\Omega \times [0,T]$ to $\mathbb{R}$ which is measurable with respect to the σ-algebra $\sigma(\mathcal{F} \times \mathcal{B})$ where $\mathcal{B}$ is the Borel σ- algebra on $[0,T]$.*

An important class of stochastic processes which we use widely in this thesis is the class of Lévy processes. This class of processes meet interesting path properties in both analytical and practical point of view (see [63]). One of the important tools which enables us to deeply study this class of processes is characteristic functions. This function finds a lot of applications in probability and is used to prove several important theorems in probability. For instance, the central limit theorem and Lévy continuity theorem. Characteristic functions are also used to study an important decomposition theorem for Lévy processes so-called the Lévy-Ito decomposition theorem. This theorem will be discussed in the next section. For the sake of completeness, we recall the definition of characteristic function for a random variable and give some of its properties.

The following definition and lemma are taken from Subsection 1.1.6 of [5].

Definition A.1.7. *([5]) Let X be a random variable defined on $(\Omega, \mathcal{F}, \mathbb{P})$ and taking values in $\mathbb{R}$ with probability law $\mathbb{P}_X$(i.e. $\mathbb{P}_X = \mathbb{P}oX^{-1}$). Its characteristic function $\phi_X : \mathbb{R} \to \mathbb{C}$ is defined by*

$$\phi_X(u) = \mathbb{E}e^{(iuX)} = \int_{\mathbb{R}} e^{iuy} \mathbb{P}_X(dy),$$

for each $u \in \mathbb{R}$.

Some properties of ϕ_X are collected in the following lemma.

Lemma A.1.1. *([5]) Let X be a random variable defined on $(\Omega, \mathcal{F}, \mathbb{P})$ with characteristic function ϕ_X. Then*

- $\phi_X(0) = 1$ *and* $|\phi_X(u)| \leq 1$,
- $\phi_X(-u) = \overline{\phi_X(u)}$,
- *X is symmetric if and only if ϕ_X is real-valued,*
- *The map $u \to \phi_X(u)$ is continuous at the origin.*

A.2. Banach and L^p spaces

The theory of risk measures has been defined on different spaces of random variables (or stochastic processes at each fixed time) and extensively studied (see [45]). Important classes for which this theory has been introduced and studied are Banach and $L^p(\Omega)$ spaces. In this thesis, the idea we put forward to study coherent risk measures defined on subspaces of the space of stochastic processes requiring to know Banach and L^p spaces. The following definitions are taken from [4].

Definition A.2.1. *([4]) Let B be a linear space over the field $\mathbb{R}$. A norm $||.||$ on B is a function from B to $[0, \infty)$ such that*

(1) $\forall X \in B$, $||X|| = 0$ if and only if $X = 0$;

(2) $\forall X \in B$, $t \geq 0$, $||tX|| = t||X||$;

(3) $\forall X, Y \in B$, $||X + Y|| \leq ||X|| + ||Y||$.

The linear space B is a normed space if its topology is induced by metric $d(X, Y) = ||X - Y||$.

Definition A.2.2. *([4]) A normed space $(B, ||.||)$ is called a Banach space if it is complete. i.e. every Cauchy sequence in B converges to an element in B.*

In the following we provide definition for L^p spaces as an example of Banach spaces.

Definition A.2.3. *([4]) Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a measure space. For $1 \leq p < \infty$ we define the space $L^p(\Omega)$ as the space of all measurable functions X on Ω such that $||X||_p = (\int_\Omega |X|^p d\mathbb{P})^{\frac{1}{p}}$ is finite. The space $L^\infty(\Omega)$ is defined as the space of all measurable functions X on Ω such that $||X||_\infty = esssup|X| = \inf\{C \geq 0 : |X(\omega)| \leq C$ for almost every $\omega\}$ is finite.*

A.3. Lévy-Ito Decomposition

One of the most important result in the theory of Lévy process is Lévy-Ito decomposition. This results characterizes this class of processes. Roughly speaking, Lévy-Ito decomposition says every Lévy process has a continuous part, a compound Poisson part and a pure jump part. This characterization helps to have a better sight about path properties of Lévy processes (see [63]). In the following, we provide some preliminaries related to this decomposition and then recall the main theorem which gives Lévy-Ito decomposition.

Definition A.3.1. *(**Poisson Random Measure**) ([63]) Let $(E, \mathcal{A}, \mu)$ be some measure space with σ-finite measure μ. The Poisson random measure with intensity measure μ is a family of random variables $\{M_A\}_{A \in \mathcal{A}}$ defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ such that*

(1) $\forall A \in \mathcal{A},\ M_A$ *is a Poisson random variable with rate* $\mu(A)$*.*

(2) If sets $A_1, A_2, \ldots, A_n \in \mathcal{A}$ *don't intersect then the corresponding random variables from* (1) *are mutually independent.*

(3) $\forall \omega \in \Omega,\ M_\bullet(\omega)$ *is a measure on* (E, A)*.*

Assume that $(X_t)_{t\geq 0}$ is a cádlág process. Let $\Delta X_t = X_t - X(t-)$ (it is well defined due to existence of $X(t-)$) and A be bounded below, i.e. $0 \in \bar{A}$. For $t \geq 0$ define,

$$N(t, A) = \#\{\Delta X(s) \in A : s \in [0, t]\}. \tag{A.3.1}$$

It is obvious that $N(t, A)$ is a stochastic process. In the following proposition we recall some properties if this process.

Proposition A.3.1. *([5]) Let* $N(t, A)$ *be the process given by* (A.3.1)*. Then,*

(1) For each A *bounded below,*$(N(t, .), t \geq 0)$ *is a Poisson process with intensity*

$$\mu(A) = \mathbb{E}[N(1, A)]. \tag{A.3.2}$$

(2) The measure μ *given by* (A.3.2) *is a Lévy measure.*

(3) The compensator $\tilde{N}(t, .), t \geq 0$ *is a martingale-valued measure where*

$$\tilde{N}(t, A) = N(t, A) - t\mu(A),$$

for A *bounded below. i.e. for a fixed* A *bounded below,* $\tilde{N}(t, A)$ *is a martingale.*

Let f be a Borel measurable function $f : \mathbb{R} \to \mathbb{R}$ and let A be bounded below, then for each $t > 0$, $\omega \in \Omega$ the Poisson integral of f can be defined in the following way.

$$\int_A f(x)N(t, dx)(\omega) = \sum_{x\in A} f(x)N(t, \{x\})(\omega). \tag{A.3.3}$$

Note that each $\int_A f(x)N(t, dx)$ is a $\mathbb{R}$-valued random variable and gives rise to a càdlàg stochastic process as we vary t. This leads to define the compensated Poisson integral associated to $\tilde{N}(t, dx)$, (see [5]), in the following way.

$$\int_A f(x)\tilde{N}(t, dx) = \int_A f(x)N(t, dx) - t\int_A f(x)\mu(dx), \tag{A.3.4}$$

where f is an integrable function with respect to (w.r.t.) the measure μ restricted to A. It can be shown that if f is a square integrable function w.r.t. μ restricted to A, then $\int_A f(x)\tilde{N}(t, dx)$ is a square integrable martingale (see [5]).

Now, we are in a position to recall Lévy-Ito decomposition.

Theorem A.3.1. *(***Lévy-Ito decomposition***) ([5]) If X is a Lévy process, then there exists $b \in \mathbb{R}$, $\sigma \geq 0$, Brownian motion W, and an independent Poisson random measure N on $\mathbb{R}^+ \times (\mathbb{R}\backslash\{0\})$ such that for each $t \geq 0$,*

$$X_t = bt + \sigma W(t) + \int_{|x|\geq 1} xN(t,dx) + \int_{|x|<1} x\tilde{N}(t,dx). \tag{A.3.5}$$

Remark A.3.1. *It can be seen from Theorem A.3.1 that each Lévy process X can be represented as $X = X^{(1)} + X^{(2)} + X^{(3)}$ where*

(1) $X^{(1)}$ is a Brownian motion with drift;
(2) $X^{(2)}$ is a compound Poisson process;
(3) $X^{(3)}$ is a square integrable pure jump martingale.