Expected Risk-Adjusted Return for Insurance Based Models

Tatiana Solcà

Diploma thesis in mathematics at
ETH Zürich under the supervision of
Prof. Dr. P. Embrechts and Dr. U. Schmock

Spring 2000

Ai miei genitori:
per avermi dato la possibilita'
di raggiungere questo traguardo.

Acknowledgements

It is my pleasure to thank Prof. Paul Embrechts for giving me the opportunity to work in this very interesting field.

I am very grateful to Dr. Uwe Schmock for his helpful suggestions and support. He always managed to find time to discuss problems related to my work.

Next I would like to thank Francesco for his encouragement, constant support and patience.

Let me finally thank all my friends and schoolmates for their advice and all-round help during this work. In particular, I thank Nicola for his pleasant and constructive "pineletters" and "talks" from Paris, Giacomo and Andrea for answering my tedious questions concerning LaTeX and Carlo for the pleasant time spent working in the computer room.

Hereby, I express my sincere gratitude to all these people.

Contents

Chapter 1

Introduction

An insurance company, nowadays, takes an interest in the applications of actuarial techniques for measuring risk and for assessing profitable areas of business. Its management is faced continually with the task of reconciling the conflicting interests of policyholders and shareholders. The former are interested in strong financial strength, while the latter are more concerned with a return on equity that is commensurate with the risk inherent in their investment. In order to satisfy these needs, the management selects profitable business and limits the company's risk.

Obviously, an insurance company's liabilities cannot be entirely foreseen. If companies are to maintain a high degree of financial security, they must efficiently manage asset and liability portfolios, as well as understand and keep control of the underlying risks.

One difficulty is that an insurance company faces many different types of risk, and they are not at all easy to model. For a thorough understanding of them it is essential to have quantitative models, but even though there are many different measures of risk, up to today, none of them can be considered the "best".

In this paper, we focus on the kinds of risks which can be represented by random variables. In particular, we analyze a model denoting the risk portfolio of an insurance company. We suppose that the management steers the company by choosing the numbers of independent risks so as to improve the overall expected risk-adjusted return, defined as the expected return divided by the assigned risk capital.

In other words, with respect to the defined model, we try to determine if an optimal portfolio exists. For these purposes, we organized this paper in four parts.

In Chapter 2, we first develop some approaches to measuring risk by considering the statement of axioms on a risk measure done by Artzner et al., which leads to the concept of coherent risk measure. Moreover, we briefly introduce some basic concepts of probability theory (like convergence of random variables and ergodic theorem), which will be seen to be useful for our purpose.

Then, in Section 2.5, we give a description of the model considered in this whole work and we indicate which methods will be used to measure risk. In more detail, we suppose that the whole profit R of an insurance company consisting of n units can be denoted by the sum of the stochastic gains R_i of every unit $i \in \{1, \dots, n\}$. R_i is defined as the revenues minus the costs and the losses in the form illustrated in (2.3). Then we will analyze this model using the expected shortfall risk measure, the coherent risk measure suggested by Artzner et al. (1998) and, later, by means of the standard deviation risk measure, which is very popular in practice.

In Chapter 3 we discuss the model using the expected shortfall for quantifying risk. We will estimate the performance of the company examining the expected risk-adjusted return r, i.e., $r = E[R]/\rho(R)$, where ρ denotes the risk measure. In fact, the company's aim will be to invest its resources optimally and maximize r. We therefore will try to determine an optimal portfolio by choosing the values of $N_1, \dots, N_n$, which denote the number of contracts of the respective business units $1, \dots, n$, such that a maximum for r is attained. In particular, we focus on a proposition which shows the existence of the limit of the upper bounds for the expected risk-adjusted return r.

In Chapter 4 we repeat the same approaches using the standard deviation risk measure. In this case we concentrate on the optimization problem defined in (4.2), i.e., we will try to determine the optimal number of contracts of every unit $i \in \{1, \dots, n\}$ in order to maximize $E[R]/C$ subject to the constraint $\rho(R) \leq C$, where C denotes the capital the company wants to invest. We will examine three different variants of the model defined in Section 2.5 and we will show that this optimization problem has a solution for every variant.

In Chapter 5 we consider two different capital allocation principles, namely the covariance principle and the expected shortfall principle. We will calculate the covariance principle for $\varrho(R) = -E[R] + \kappa\, \sigma(R)$ with $\kappa > 0$, and for two variants of the model R considered in Chapter 4. Moreover, for simplicity we consider a company consisting only of two units, but the same results can be computed in a similar way for the general case, too, and we calculate $E[R_i \,|\, R \leq c]$ for the multivariate normal case.

In the Appendix we briefly recall some useful technical rules to calculate the expected shortfall.

Chapter 2

Preliminaries

2.1 Risk and Risk Measures

It is not easy to define risk, and we will avoid attempting to give an exact definition. Nevertheless in a recent paper, Artzner et al. (1998) have come up with an appropriate description of what risk actually is. In this paper, we consider risk related to the variability of the future value of a position due to uncertain events. Therefore, we treat those kinds of risks which can be represented by random variables, and which indicate the possible future values of positions.

2.1.1 Notation and properties

Let Ω be the set of possible states of nature, and assume it is finite. By a random variable X we denote the final net worth of a position for each element of Ω. Let $\mathcal{G}$ be the set of all risk, i.e., the set of all real valued functions on Ω. Remark that $\mathcal{G}$ can be identified with $\mathbb{R}^n$, where $n = \text{card}(\Omega)$.

Definition 2.1. A *measure of risk* is a mapping ρ from $\mathcal{G}$ into $\mathbb{R}$.

Then the real number $\rho(X)$ can be interpreted, when positive, as the minimum extra cash to add to the risky position X, or when negative, as the cash amount that can be subtracted from the position.

We now consider some properties for a risk measure ρ defined on $\mathcal{G}$ listed in the form of axioms.

Axiom (Translation invariance).
For all $X \in \mathcal{G}$ and all real numbers α: $\rho(X + \alpha r) = \rho(X) - \alpha$,
where r is the rate of return on a reference riskless investment.

Axiom (Subadditivity).
For all $X_1, X_2 \in \mathcal{G}$: $\rho(X_1 + X_2) \leq \rho(X_1) + \rho(X_2)$.

Axiom (Positive homogeneity).
For all $\lambda \geq 0$ and all $X \in \mathcal{G}$: $\rho(\lambda X) = \lambda \rho(X)$.

Axiom (Monotonicity).
For all X and $Y \in \mathcal{G}$ with $X \leq Y$: $\rho(Y) \leq \rho(X)$.

Axiom (Relevance).
For all $X \in \mathcal{G}$ with $X \leq 0$ and $X \neq 0$: $\rho(X) \geq 0$.

Remarks.

- Translation invariance means that adding (resp. subtracting) the sure initial amount α to (from) the initial position, and investing it in the reference instrument (with rate of return r) the risk measure only decreases (resp. increases) by α.
- Subadditivity reflects the diversification of portfolios and ensures that the risk measure behaves reasonably when adding two positions; we can say: "a merger does not create extra risk".

These axioms on measures of risk are related to the axioms on acceptance sets, but we won't treat this topic (for more details see Artzner et al. (1998)). We are interested in the following definition then we argue that any risk measure which is to be used to effectively regulate or manage risks satisfies these axioms.

Definition 2.2. A risk measure satisfying the four axioms of translation invariance, subadditivity, positive homogeneity and monotonicity is called *coherent.*

In their paper Artzner et al. suggest a specific coherent measure called *tail conditional expectation* and in the following chapter we will study the model of a portfolio using this risk measure.

2.1.2 The "tail conditional expectation" measure of risk

In practice there are various methods of measuring risk, and these axioms are not restrictive enough to specify a unique risk measure. The choice of precisely which measure to use should be made on the basis of additional considerations.

In this work we consider tail conditional expectation (expected shortfall), which, under some assumptions, is the least expensive among those which are coherent and accepted by regulators [1] since they are more conservative than the value-at-risk measurement. Managers and regulators are primarily interested in setting "minimal requirements" or a maximal limit on the potential losses. With a shortfall approach, one can answer the question "how bad is bad?" by measuring the negative of the average future net worth X of a position, given that X is below the quantile $c \leq 0$, i.e.,

$$\rho(X) = E[-X \,|\, X \leq c\,], \tag{2.1}$$

provided that $P(X \leq c) > 0$.

In the following sections we will focus on definitions and theorems, which will be useful later for our purposes.

2.2 Convergence of random variables

Definition 2.3. Let $X_1, X_2, \ldots$ and X be real-valued random variables on some probability space $(\Omega, \mathcal{F}, P)$. We say:

i) $\{X_n\}_{n\in\mathbb{N}}$ converges to X *almost surely* (a.s.) if

$$P(\{\omega \,|\, \lim_{n\to\infty} X_n(\omega) = X(\omega)\}) = 1,$$

ii) $\{X_n\}_{n\in\mathbb{N}}$ converges to X *in r-th mean* ($\xrightarrow{r}$) for $r > 0$ if

$$E\,|\,X_n^r\,| < \infty \text{ for all } n \text{ and } E(\,|\,X_n - X\,|^r) \xrightarrow{n\to\infty} 0,$$

iii) $\{X_n\}_{n\in\mathbb{N}}$ converges to X *in probability* ($\xrightarrow{P}$) if for every $\varepsilon > 0$

$$P(\,|\,X_n - X\,| > \varepsilon) \xrightarrow{n\to\infty} 0\,.$$

iv) $\{X_n\}_{n\in\mathbb{N}}$ converges to X *in distribution* ($\xrightarrow{D}$) if

$$P(X_n \leq x) \xrightarrow{n\to\infty} P(X \leq x)$$

for all points x at which $F_X(x) = P(X \leq x)$ is continuous.

[1] A regulator is a supervisor who takes into account the unfavorable states when allowing a risky position.

Remark. The following implications hold in general:

i) $X_n \xrightarrow{\text{a.s.}} X \quad \Longrightarrow \quad X_n \xrightarrow{P} X \quad \Longrightarrow \quad X_n \xrightarrow{D} X$,

ii) $X_n \xrightarrow{r} X \quad \Longrightarrow \quad X_n \xrightarrow{P} X \quad \Longrightarrow \quad X_n \xrightarrow{D} X$.

We now recall the theorem of bounded convergence.

Theorem 2.4 (Lebesgue bounded convergence theorem). Consider a sequence $\{X_n\}_{n\in\mathbb{N}}$ of variables with $X_n \xrightarrow{a.s.} X$. If there is a random variable Y such that $E|Y| < \infty$ and $|X_n| \leq Y$ for all n, then

$$E[X_n] \xrightarrow{n\to\infty} E[X]\,.$$

Proof. See Grimmet and Stirzaker (1992), Chapter 5.6. □

Remark. It is appropriate to specify that by the convergence in r-th mean the values $r = 1$ and $r = 2$ are of most use. Therefore, in these cases, we write, respectively

i) $X_n \longrightarrow X$ in mean, instead of $X_n \xrightarrow{1} X$,

ii) $X_n \longrightarrow X$ in mean square, instead of $X_n \xrightarrow{2} X$.

2.3 Stationary process and ergodic theorem

Definition 2.5. A real-valued process $X_1, X_2, \dots$ is called *stationary* if for every $x_1, \dots, x_n \in \mathbb{R}$ and integer $k > 0$

$$P[X_1 \leq x_1, \dots, X_n \leq x_n] = P[X_{1+k} \leq x_1, \dots, X_{n+k} \leq x_n]\,.$$

Consider a probability space $(\Omega, \mathcal{F}, P)$ and a transformation $T : \Omega \to \Omega$.

Definition 2.6. A transformation $T : \Omega \to \Omega$ will be called *measurable* if for all $A \in \mathcal{F}$: $\quad T^{-1}(A) = \{\omega \in \Omega \mid T(\omega) \in A\} \in \mathcal{F}$.

Definition 2.7. A measurable transformation $T : \Omega \to \Omega$ will be called *measure-preserving* if for all $A \in \mathcal{F}$: $\quad P(T^{-1}(A)) = P(A)$.

Let T be a measure-preserving transformation on $(\Omega, \mathcal{F}, P)$.

Definition 2.8. A set $A \in \mathcal{F}$ is called *invariant* if $T^{-1}(A) = A$.

We denote by $\mathcal{J}$ the set of all invariant $A \in \mathcal{F}$. Note that $\mathcal{J}$ is a σ-field.

Definition 2.9. T is called *ergodic* if $P(A) \in \{0,1\}$ for every $A \in \mathcal{J}$.

Theorem 2.10 (Ergodic theorem). Let T be a measure-preserving transformation on $(\Omega, \mathcal{F}, P)$. Then for any random variable X such that $E|X| < \infty$:

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} X(T^k(\omega)) = E[X \,|\, \mathcal{J}] \quad \text{a.s. and in mean.}$$

Proof. See Breiman (1968), pp. 113–115. □

Corollary 2.11. Let T be a measure-preserving and ergodic transformation on $(\Omega, \mathcal{F}, P)$. Then for any random variable X such that $E|X| < \infty$:

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} X(T^k(\omega)) = E[X] \quad \text{a.s. and in mean.}$$

Proof. See Breiman (1968), p. 115. □

These general definitions and results concerning invariance and ergodicity can be applied to the original stationary process $X_1, X_2, \ldots$ by considering a shift transformation T, i.e., if $x = (x_0, x_1, \ldots)$ is a real sequence of values of the stationary process then $Tx = (x_1, x_2, \ldots)$. For more details see Grimmet and Stirzaker (1992). The corresponding form of the ergodic theorem for stationary processes is:

Theorem 2.12 (Ergodic theorem for stationary processes). Let X_1, X_2, … be a stationary process such that $E|X_1| < \infty$, then

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} X_k = E[X|\mathcal{J}] \quad \text{a.s. and in mean.}$$

Proof. See Grimmet and Stirzaker (1992), Chapter 9.5. □

Furthermore, we define an ergodic stationary process and we write down the corresponding version of the ergodic theorem.

Definition 2.13. Let T be the shift-operator and A a set of real sequences. A stationary process is said to be *ergodic* if $P\{(X_0, X_1, \ldots) \in A\} = 0$ or 1, whenever A is shift-invariant, i.e., A is invariant with respect to the shift-operator T.

Remark. Let $\{X_n\}_{n\in\mathbb{N}_0}$ be a real-valued stationary process. Then the following conditions are equivalent:

(a) $\{X_n\}_{n\in\mathbb{N}_0}$ is ergodic,

(b) $P\{(X_0, X_1, \dots) \in A\} = 0$ or 1, for every invariant set A,

(c) $\lim_{n\to\infty} \frac{1}{n}\sum_{j=1}^{n} \phi(X_j, X_{j+1}, \dots) = E[\phi(X_0, X_1, \dots)]$, for every measurable function ϕ of real sequences, provided the expectation exists,

(d) $\lim_{n\to\infty} \frac{1}{n}\sum_{j=1}^{n} \phi(X_j, \dots, X_{j+k}) = E[\phi(X_0, \dots, X_k)]$, for every $k \in \mathbb{N}_0$ and every measurable function ϕ of $k+1$ variables, provided the expectation exists.

For more details see Karlin and Taylor (1975), Chapter 9.5.

So, since a stationary process $X_1, X_2, \dots$ is ergodic if every shift-invariant event has probability zero or one, if $\mathcal{J}$ has this zero-one property, of course the average converges to $E[X_1]$. Hence, we restrict ourselves to:

Theorem 2.14 (Ergodic theorem for ergodic stationary processes). If $X_1, X_2, \dots$ is a stationary ergodic process such that $E|X_1| < \infty$, then

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} X_k = E[X_1] \quad \text{a.s. and in mean.}$$

Proof. See Grimmet and Stirzaker (1992), Chapter 9.5. □

2.4 Properties of weak convergence

Let S be a metric space and $\mathcal{S}$ the Borel σ-field in S, i.e., $\mathcal{S}$ is the smallest σ-field containing all the open sets. Let $C_b(S)$ be the set of all bounded continuous real functions f on S, and P be a probability measure on $\mathcal{S}$, i.e., a nonnegative, countably additive set function with $P(S) = 1$.

Definition 2.15. Let P_n for $n \in \mathbb{N}$ and P be probability measures on $(S, \mathcal{S})$. We say that $\{P_n\}_{n\in\mathbb{N}}$ *converges weakly* to P ($P_n \xrightarrow{w} P$) if for all $f \in C_b(S)$:

$$\int_S f\, dP_n \longrightarrow \int_S f\, dP.$$

Definition 2.16. A set A in $\mathcal{S}$, whose boundary ∂A satisfies $P(\partial A) = 0$, is called a *P-continuity* set.

Theorem 2.17 (Portmanteau Theorem). Let P_n for $n \in \mathbb{N}$ and P be probability measures on $(S, \mathcal{S})$. These five conditions are equal:

i) $P_n \xrightarrow{w} P$,

ii) $\limsup_{n\to\infty} \int_S f \, dP_n = \int_S f \, dP$, for all bounded uniformly continuous functions $f : S \to \mathbb{R}$,

iii) $\limsup_{n\to\infty} P_n(F) \leq P(F)$, for all closed $F \subset S$,

iv) $\liminf_{n\to\infty} P_n(G) \geq P(G)$, for all open $G \subset S$,

v) $\lim_{n\to\infty} P_n(A) = P(A)$, for all P-continuity sets A.

Proof. See Billingsley (1968), pp. 11–14. □

2.5 The model

2.5.1 The idea and the notation

The aim of this work is to study a model for the portfolio of an insurance company by means of two different risk measures which take the whole profit of the company into consideration. We investigate some possible risk portfolios, which are directly dependent on the number of the contracts to be stipulated and on the expected losses, so as to determine, if possible, the optimal number of contracts in order to maximize the profit. Hence, we are going to assume that an insurance company consists of n organizational units or business segments. The whole profit of the company (positive value means gain, negative value means loss) will be denoted by

$$R = \sum_{i=1}^{n} R_i, \tag{2.2}$$

with R_i describing the stochastic gain of the unit i during a fixed time period (usually one year). Moreover, we suppose that for every unit $i \in \{1, \dots, n\}$

$$R_i = \nu_i N_i - \sum_{j=1}^{N_i} X_{i,j} - Y_i N_i \tag{2.3}$$

where:

- ν_i is the premium income for one contract of unit i,

- N_i is the number of contracts of unit i, $N_1 + \cdots + N_n$ is the whole number of contracts of the insurance company,
- $\{X_{i,j}\}_{j\in\mathbb{N}}$ is a sequence of random variables for $i \in \{1, \ldots, n\}$: $X_{i,j}$ represents the loss associated with the risk in the j-th contract of unit i and $\sum_{j=1}^{N_i} X_{i,j}$ is the total (annual) claim amount of unit i,
- Y_i is a random variable which represents, for one contract of unit i, a safety loading needed to obviate both the possible approximation error in the calculation of the optimal premium and unforeseeable, catastrophic events.

Therefore, we will examine how to determine an optimal portfolio, which guarantees a maximal profit. We will discuss this problem for two risk measures: the expected shortfall, the coherent risk measure suggested by Artzner et al. (1998), and the standard deviation, which is very popular in practice. We will analyze the portfolio by means of a risk adjusted performance measurement, this means that we compute the return in the way commonly called RORAC (return on risk adjusted capital). In particular, considering any risk measure ρ and provided that $\rho(R) \neq 0$, we define the expected risk-adjusted return for a risk R as

$$r(R, \rho) = \frac{E[R]}{\rho(R)}\,. \tag{2.4}$$

In practice, the company tries hard to improve its results and this means that its aim is to maximize r. Therefore, we will try to determine the optimal value for the numbers of contracts N_i of every unit $i \in \{1, \ldots, n\}$, such that this maximum is attained.

We now introduce our running examples of risk measure. We start considering the expected shortfall defined by

$$\rho(R) = E[-R \mid R \le c] \qquad \text{with} \quad c \le 0\,. \tag{2.5}$$

The expected shortfall is an alternative risk measure to the quantile which overcomes some of the theorical deficiencies of the latter. In particular, this risk measure gives some information about the size of potential losses, given that a loss bigger than c has occurred.

Then, the same consideration will be repeated using the standard deviation risk measure defined by

$$\rho(R) = -E[R] + \kappa\, \sigma(R), \tag{2.6}$$

where $\kappa > 0$ is some positive constant and $\sigma(R)$ denotes the standard deviation operator, i.e., $\sigma(R) = \sqrt{\mathrm{Var}(R)}$.

2.5.2 The different variants of the model

In Chapter 4 we will examine the expected risk-adjusted return for a risk R considering the standard deviation risk measure. In particular, we will study three different variants of the model representing the whole profit of a company consisting of n units, previously defined as

$$R = \sum_{i=1}^{n} \left(\nu_i N_i - \sum_{j=1}^{N_i} X_{i,j} - Y_i N_i \right). \tag{2.7}$$

Moreover, we make the following assumptions, which are valid through all of Chapter 4.

- $\{X_{i,j}\}_{j \in \mathbb{N}}$, for all $i \in \{1, \dots, n\}$, are sequences of independent identically distributed (i.i.d.) random variables, with $X_{i,j}$ having a finite mean denoted by μ_i, i.e,

$$\mu_i = E[X_{i,j}] \qquad \text{for all } i \in \{1, \dots, n\} \text{ and } j \in \mathbb{N},$$

- $Y_1, \dots, Y_n$ are random variables having finite mean denoted by $\tilde{\mu}_i$, i.e,

$$\tilde{\mu}_i = E[Y_i] \qquad \text{for all } i \in \{1, \dots, n\},$$

- all the sequences $\{X_{i,j}\}_{j \in \mathbb{N}}$, with $i \in \{1, \dots, n\}$, and the random variables $Y_1, \dots, Y_n$ are independent.

Remark. In Chapter 3, we will examine the above-defined model with the aid of the expected shortfall as risk measure, but for the moment we do not need such strong assumptions. In fact it is not necessary to require that $\{X_{i,j}\}_{j \in \mathbb{N}}$ are sequences of i.i.d. random variables, but rather we will prove some statements for which it is enough to assume the existence of real constants $\mu_1, \dots, \mu_n$ such that, for all $i \in \{1, \dots, n\}$,

$$\frac{1}{N} \sum_{j=1}^{N} X_{i,j} \xrightarrow{N \to \infty} \mu_i \qquad \text{in mean.}$$

Therefore, for a model of the form of (2.7) with the above-mentioned assumptions, there are different variants depending on the choice of the distributions of the random variables and on the nature of the parameters N_i for $i \in \{1, \dots, n\}$.

Then, in Chapter 4 we will examine three different cases which result from (2.7) if we choose N_i for $i \in \{1, \dots, n\}$ in different ways. In particular, we consider the N_i's first as positive integers, then as Poisson-distributed random variables and, finally, as the sum of both.

More precisely, these possibilities can be represented as follows:

1. $R = \sum_{i=1}^{n} \left(\nu_i N_i - \sum_{j=1}^{N_i} X_{i,j} - Y_i N_i\right)$,
   with N_i a positive integer for all $i \in \{1, \ldots, n\}$,

2. $R = \sum_{i=1}^{n} \left(\nu_i N_i - \sum_{j=1}^{N_i} X_{i,j} - Y_i N_i\right)$,
   with $N_i \sim \text{POIS}(\lambda_i),\ \lambda_i > 0$ for all $i \in \{1, \ldots, n\}$,

3. $R = \sum_{i=1}^{n} \left(\nu_i N_i - \sum_{j=1}^{N_i} X_{i,j} - Y_i N_i\right)$,
   with $N_i = N_i^{\text{fix}} + N_i^{\text{pois}}$, where N_i^{fix} are positive integers and N_i^{pois} are Poisson-distributed random variables, i.e., $N_i^{\text{pois}} \sim \text{POIS}(\lambda_i),\ \lambda_i > 0$, for all $i \in \{1, \ldots, n\}$.

Chapter 3

Results for the expected shortfall risk measure

In this chapter, we will examine the portfolio of a company, analyze the profit represented by the model previously described, and consider the expected shortfall as an aid to quantifying risk.

Without loss of generality and, in order to simplify the following calculations, we start by examining a one-dimensional model. This means we assume that an insurance company consists of only one business unit. Hence, we denote the whole profit by

$$R(N) = \nu N - \sum_{j=1}^{N} X_j - YN, \tag{3.1}$$

where $X_1, X_2, \dots$ is a real-valued process which represents the claim sizes; Y is a random variable having finite mean $\tilde{\mu}$ and N is some positive integer. Recall that it is assumed that Y and the sequence $\{X_j\}_{j\in\mathbb{N}}$ are independent, and that we still do not require any particular properties for the process $\{X_j\}_{j\in\mathbb{N}}$. Moreover, we assume $\nu - \mu - \tilde{\mu} > 0$, i.e., the company chooses a premium income rate greater than the expected losses.

In order to estimate the performance of the company, we consider the expected risk-adjusted return for the risk $R(N)$ which, in this case, can be represented by

$$r_N = \frac{E\big[R(N)\big]}{E\big[-R(N)\,|\,R(N) \le c\big]}, \quad c \le 0, \tag{3.2}$$

provided that $P\big(R(N) < c\big) > 0$.

More generally, in Section 3.2 we will take the same approach considering an n-dimensional model

$$R(N_1, \dots, N_n) = \sum_{i=1}^{n} \Big(\nu_i N_i - \sum_{j=1}^{N_i} X_{i,j} - Y_i N_i \Big) \tag{3.3}$$

with all the assumptions listed in Section 2.5.1.

Now, we begin with the following lemma which is valid in general for any risk represented by a random variable $R \in L^1(\Omega, \mathcal{F}, P)$.

Lemma 3.1. Let R be an integrable random variable on a probability space $(\Omega, \mathcal{F}, P)$, this means $R \in L^1(\Omega, \mathcal{F}, P)$, and let c_0 denote the infimum of the support of the distribution of R, i.e., $c_0 = \inf\{c \in \mathbb{R} \,|\, P(R \le c) > 0\}$.
Then the map

$$(c_0, \infty) \ni c \longmapsto E[-R \,|\, R \le c]$$

is non-increasing.

Proof. First, we consider $E[-R \,|\, R \le c]$. This conditional expectation is defined by

$$E[-R \,|\, R \le c] = \frac{E\big[-R\, 1_{\{R \le c\}}\big]}{P[R \le c]}\,.$$

Then we consider any constants $c_1, c_2 \in (c_0, \infty)$ such that $c_1 < c_2$.

Given that, since $\{R \le c_1\}$ and $\{c_1 < R \le c_2\}$ are disjoint, it holds that

$$1_{\{R \le c_2\}} = 1_{\{R \le c_1\}} + 1_{\{c_1 < R \le c_2\}},$$

we can write:

$$E[-R \,|\, R \le c_2] = \frac{E[-R \,|\, R \le c_1]\; P[R \le c_1] + E[-R \,|\, c_1 < R \le c_2]\; P[c_1 < R \le c_2]}{P[R \le c_2]}\,.$$

Then, in order to prove the monotony, it suffices to show that

$$E[-R \,|\, R \le c_2] \le E[-R \,|\, R \le c_1]\,.$$

We split the proof into two cases.

i) If $P(c_1 < R \le c_2) = 0$, then it holds that $1_{\{R \le c_2\}} = 1_{\{R \le c_1\}}$ P-a.s. and $P(R \le c_1) = P(R \le c_2)$. Therefore, it follows that

$$E[-R \,|\, R \le c_2] = \frac{E[-R\, 1_{\{R \le c_2\}}]}{P[R \le c_2]} = \frac{E[-R\, 1_{\{R \le c_1\}}]}{P[R \le c_1]} = E[-R \,|\, R \le c_1]\,.$$

ii) If $P(c_1 < R \leq c_2) > 0$, then we can write

$$\begin{aligned}
&E[-R\,|\,R \leq c_2] \\
&= \frac{E[-R\,|\,R \leq c_1]\ P[R \leq c_1] + E[-R\,|\,c_1 < R \leq c_2]\ P[c_1 < R \leq c_2]}{P[R \leq c_2]} \\
&= \frac{E[-R\,|\,R \leq c_1]\ \big(P[R \leq c_2] - P[c_1 < R \leq c_2]\big)}{P[R \leq c_2]} \\
&\qquad\qquad + \frac{E[-R\,|\,c_1 < R \leq c_2]\ P[c_1 < R \leq c_2]}{P[R \leq c_2]} \\
&= E[-R\,|\,R \leq c_1] \\
&\qquad + \underbrace{\big(\underbrace{E[-R\,|\,c_1 < R \leq c_2]}_{-c_2 \leq \cdots < -c_1} - \underbrace{E[-R\,|\,R \leq c_1]}_{\geq -c_1}\big)}_{<0} \underbrace{\frac{P[c_1 < R \leq c_2]}{P[R \leq c_2]}}_{0 < \cdots \leq 1} \\
&< E[-R\,|\,R \leq c_1]\,.
\end{aligned}$$

Thus, $E[-R\,|\,R \leq c_2] \leq E[-R\,|\,R \leq c_1]$, as desired. □

3.1 The asymptotical limit of upper bounds

In this section we focus on the portfolio of a company represented by the model defined in (3.1). In particular, we will study the expected risk-adjusted return r_N for the risk $R(N)$. The main question is whether it is possible to find a maximal value for r_N thereby determining an optimal value for N such that this maximum is attained. We will show that the limit of the upper bounds for the expected risk-adjusted return exists under quite general conditions. We also derive an explicit formula for the limit. Recall that, considering the expected shortfall risk measure, r_N is defined by (3.2).

Now, for any $c \leq 0$ we can write

$$r_N = \frac{E\big[R(N)\big]}{E\big[-R(N)\,|\,R(N) \leq c\big]} = \frac{\frac{E[R(N)]}{N}}{E\big[-\frac{R(N)}{N}\,|\,\frac{R(N)}{N} \leq \frac{c}{N}\big]}\,.$$

Therefore, if we define

$$R_N = \nu - \frac{1}{N}\sum_{j=1}^{N} X_j - Y, \tag{3.4}$$

since $E[R(N)] = N(\nu - \mu - \tilde{\mu})$, it holds that

$$r_N = \frac{\nu - \mu - \tilde{\mu}}{E\big[-R_N\,|\ R_N \leq \frac{c}{N}\big]}\,. \tag{3.5}$$

In particular, the aim of this chapter is to prove the following proposition.

Proposition 3.2. Let $X_1, X_2, \dots$ be a real-valued process and Y be a integrable random variable. Let $R_N = \nu - \frac{1}{N}\sum_{j=1}^{N} X_j - Y$ and assume that a real constant μ exists such that

$$\frac{1}{N}\sum_{j=1}^{N} X_j \xrightarrow{N\to\infty} \mu \qquad \text{in mean}\,.$$

Moreover, suppose that $P(\nu-\mu-Y=0)=0$. Then, it holds that

$$E[R_N \,|\, R_N \le 0] \xrightarrow{N\to\infty} E[\nu-\mu-Y \,|\, \nu-\mu-Y \le 0]\,.$$

For the proof of this proposition we need the next lemma.

Lemma 3.3. Let R_N be defined as above and assume that a real constant μ exists such that

$$\frac{1}{N}\sum_{j=1}^{N} X_j \xrightarrow{N\to\infty} \mu \qquad \text{in mean.}$$

Let Z be an integrable random variable and f be a bounded uniformly Lipschitz continuous function. Then, it holds that

$$E\big[Z\, f(R_N)\big] \xrightarrow{N\to\infty} E\big[Z\, f(\nu-\mu-Y)\big]\,.$$

Proof. Let $\varepsilon > 0$. Since $Z \in L^1$, because of the theorem of Lebesgue, a real constant $K > 0$ exists such that

$$E\big[\,|\,Z\,|\,1_{\{|\,Z\,|>K\}}\big] \le \varepsilon\,.$$

Moreover, f is Lipschitz continuous, i.e., a real constant α exists such that

$$|\,f(R_N) - f(\nu-\mu-Y)\,| \le \alpha\,|\,R_N - \big(\nu-\mu-Y\big)\,|\,.$$

Then,

$$
\begin{aligned}
&E\big[\,|\,Z\,f(R_N) - Z\,f(\nu-\mu-Y)\,|\,\big] \\
&\quad\le\; E\big[\,|\,Z\,|\;\;|\,f(R_N) - f(\nu-\mu-Y)\,|\,\big] \\
&\quad\le\; E\big[\,|\,Z\,|\,1_{\{|\,Z\,|>K\}}\;\underbrace{|\,f(R_N) - f(\nu-\mu-Y)\,|}_{\le 2\,\sup|f|}\,\big] \\
&\qquad + E\big[\,\underbrace{|\,Z\,|\,1_{\{|\,Z\,|\le K\}}}_{\le K}\;\;|\,f(R_N) - f(\nu-\mu-Y)\,|\,\big] \\
&\quad\le\; \underbrace{E\big[\,|\,Z\,|\,1_{\{|\,Z\,|>K\}}\big]}_{\le\varepsilon}\,2\,\sup|f| + \,K\,E\big[\,\underbrace{|\,f(R_N) - f(\nu-\mu-Y)\,|}_{\le\alpha\,|R_N-(\nu-\mu-Y)|}\,\big] \\
&\quad\le 2\,\varepsilon\sup|f| + K\,\alpha\;E\big[\,\underbrace{|\,R_N - \big(\nu-\mu-Y\big)\,|}_{=|\frac{1}{N}\sum_{j=1}^{N}X_j-\mu|}\,\big] \xrightarrow{N\to\infty} 2\,\varepsilon\sup|f|
\end{aligned}
$$

Therefore, since ε is arbitrary, we can conclude that, as desired,

$$
E\big[\,|\,Z\,f(R_N) - Z\,f(\nu-\mu-Y)\,|\,\big] \xrightarrow{N\to\infty} 0\,.
$$

□

Proof (Proposition 3.2). We remember that $E[R_N\,|\,R_N\le 0]$ is defined by

$$
E[R_N\,|\,R_N\le 0] = \frac{E\big[R_N 1_{\{R_N\le 0\}}\big]}{P[R_N\le 0]}\,. \tag{3.6}
$$

We first consider $E[R_N 1_{\{R_N\le 0\}}]$. Clearly, it holds that

$$
\begin{aligned}
E\big[R_N 1_{\{R_N\le 0\}}\big] &= E\Bigg[\Big(\nu - \frac{1}{N}\sum_{j=1}^{N}X_j - Y\Big)1_{\{R_N\le 0\}}\Bigg] \qquad (3.7)\\
&= \nu P[R_N\le 0] - E\Bigg[\Big(\frac{1}{N}\sum_{j=1}^{N}X_j\Big)1_{\{R_N\le 0\}}\Bigg] - E\big[Y 1_{\{R_N\le 0\}}\big]\,.
\end{aligned}
$$

Since

$$
R_N \xrightarrow{N\to\infty} \nu-\mu-Y \quad \text{in mean}\,,
$$

and because the convergence in mean implicates the convergence in distribution, it holds that

$$
P(R_N\le 0) \xrightarrow{N\to\infty} P(\nu-\mu-Y\le 0)\,.
$$

Therefore, since it is assumed that $P(\nu-\mu-Y=0)=0$, we can prove, on the one hand, that

$$E\Big[\Big(\frac{1}{N}\sum_{j=1}^{N}X_j\Big)1_{\{R_N\le 0\}}\Big] \xrightarrow{N\to\infty} \mu\ P(\nu-\mu-Y\le 0) \tag{3.8}$$

and, on the other, that

$$E\big[Y1_{\{R_N\le 0\}}\big] \xrightarrow{N\to\infty} E[Y\,|\,\nu-\mu-Y\le 0]\ P(\nu-\mu-Y\le 0)\,. \tag{3.9}$$

From

$$E\Big[\Big|\frac{1}{N}\sum_{j=1}^{N}X_j-\mu\Big|\Big] \xrightarrow{N\to\infty} 0$$

it follows that

$$\begin{aligned} E\Big[\Big|\Big(\frac{1}{N}\sum_{j=1}^{N}X_j\Big)1_{\{R_N\le 0\}}-\mu\,1_{\{R_N\le 0\}}\Big|\Big] &= E\Big[\Big|\Big(\frac{1}{N}\sum_{j=1}^{N}X_j\Big)-\mu\Big|\ \underbrace{1_{\{R_N\le 0\}}}_{\le 1}\Big] \\ &\le E\Big[\Big|\frac{1}{N}\sum_{j=1}^{N}X_i-\mu\Big|\Big] \xrightarrow{N\to\infty} 0\,. \end{aligned}$$

Thus, it still remains to be calculated

$$\lim_{N\to\infty} E[\mu\,1_{\{R_N\le 0\}}] \qquad \text{and} \qquad \lim_{N\to\infty} E[Y\,1_{\{R_N\le 0\}}]\,.$$

Let Z be any integrable random variable. We will determine

$$\lim_{N\to\infty} E[Z\,1_{\{R_N\le 0\}}]$$

and then we will employ the result for $Z=\mu$ and $Z=Y$. As shown in Figure 3.1, let f_n and g_n bounded continuous functions defined by

$$f_n(x)=\begin{cases}1\,, & \text{if } x\in(-\infty,-\frac{1}{n}]\,,\\ -nx\,, & \text{if } x\in(-\frac{1}{n},0)\,,\\ 0\,, & \text{if } x\in[0,\infty)\,,\end{cases}$$

and

$$g_n(x)=\begin{cases}1\,, & \text{if } x\in(-\infty,0]\,,\\ 1-nx\,, & \text{if } x\in(0,\frac{1}{n})\,,\\ 0\,, & \text{if } x\in[\frac{1}{n},\infty)\,.\end{cases}$$

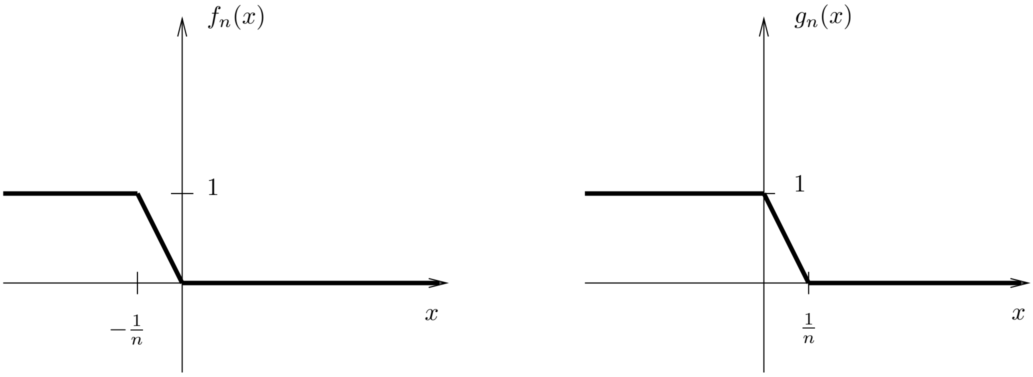

Figure 3.1: Continuous approximation of the indicator function $1_{(-\infty,0]}$ from below by f_n and from above by g_n.

Instead of $1_{\{R_N \le 0\}}$ it is useful to write $1_{(-\infty,0]} \circ R_N$, then it follows that

$$1_{(-\infty,-\frac{1}{n}]} \circ R_N \le f_n(R_N) \le 1_{(-\infty,0]} \circ R_N \le g_n(R_N) \le 1_{(-\infty,\frac{1}{n}]} \circ R_N$$

and clearly for $Z \ge 0$ it holds that

$$E\big[Z\, f_n(R_N)\big] \le E\big[Z\, 1_{(-\infty,0]} \circ R_N\big] \le E\big[Z\, g_n(R_N)\big]\,. \tag{3.10}$$

For a general integrable random variable Z, consider the decomposition $Z = Z^+ - Z^-$ with $Z^+ = \max\{Z, 0\}$ and $Z^- = \max\{-Z, 0\}$.

Then, considering the left and right side respectively of the inequality (3.10) separately, since f_n and g_n are Lipschitz continuous functions, i.e.,

$$|\, f_n(R_N) - f_n(\nu - \mu - Y)\,| \le n\,|\,R_N - \big(\nu - \mu - Y\big)\,|$$

and analogously

$$|\, g_n(R_N) - g_n(\nu - \mu - Y)\,| \le n\,|\,R_N - \big(\nu - \mu - Y\big)\,|,$$

it follows from Lemma 3.3 that

$$E\big[Z\, f_n(R_N)\big] \xrightarrow{N\to\infty} E\big[Z\, f_n(\nu - \mu - Y)\big]$$

and

$$E\big[Z\, g_n(R_N)\big] \xrightarrow{N\to\infty} E\big[Z\, g_n(\nu - \mu - Y)\big]\,.$$

Moreover, since $f_n(\nu - \mu - Y) \xrightarrow{n\to\infty} 1_{(-\infty,0)} \circ (\nu - \mu - Y)$ pointwise, from the Lebesgue theorem 2.4 it follows that

$$E\big[Z\, f_n(\nu - \mu - Y)\big] \xrightarrow{n\to\infty} E\big[Z\, 1_{(-\infty,0)} \circ (\nu - \mu - Y)\big]$$

and analogously

$$E\big[Z\, g_n(\nu-\mu-Y)\big] \xrightarrow{n\to\infty} E\big[Z\, 1_{(-\infty,0]}\circ(\nu-\mu-Y)\big]\,.$$

Now, by the assumption $P[\nu-\mu-Y=0]=0$,

$$E\big[Z\, 1_{(-\infty,0)}\circ(\nu-\mu-Y)\big] = E\big[Z\, 1_{(-\infty,0]}\circ(\nu-\mu-Y)\big]$$

and consequently we obtain the statements (3.8) and (3.9) replacing $Z=\mu$ and $Z=Y$, respectively, i.e.,

$$E\big[\mu 1_{\{R_N\le 0\}}\big] \xrightarrow{N\to\infty} E\big[\mu 1_{\{\nu-\mu-Y\le 0\}}\big] = \mu\; P\big(\nu-\mu-Y\le 0\big),$$

and

$$\begin{aligned} E\big[Y 1_{\{R_N\le 0\}}\big] &\xrightarrow{N\to\infty} E\big[Y 1_{\{\nu-\mu-Y\le 0\}}\big] \\ &\qquad = E[Y\,|\,\nu-\mu-Y\le 0]\; P(\nu-\mu-Y\le 0)\,. \end{aligned}$$

Thus, to conclude,

$$E[R_N 1_{\{R_N\le 0\}}] \xrightarrow{N\to\infty} \big(\nu-\mu-E[Y\,|\,\nu-\mu-Y\le 0]\big)\; P(\nu-\mu-Y\le 0)$$

and given that this last quantity equals

$$E[\nu-\mu-Y\,|\,\nu-\mu-Y\le 0]\; P(\nu-\mu-Y\le 0)\,,$$

the statement in the proposition follows directly from the definition of the conditional expectation. □

Consequently, from Lemma 3.1 and from Proposition 3.2 we obtain the limit of the upper bounds for the expected risk-adjusted return r_N. In fact, we also have

$$r_N \xrightarrow{N\to\infty} r_\infty,$$

as $R_N \xrightarrow{N\to\infty} \nu-\mu-Y$ in mean and in law and hence

$$P\big(R_N\le \frac{c}{N}\big) \xrightarrow{N\to\infty} P\big(\nu-\mu-Y\le 0\big)$$

and

$$E\big[R_N\, 1_{R_N\le c/N}\big] \xrightarrow{N\to\infty} E\big[(\nu-\mu-Y)\, 1_{\nu-\mu-Y\le 0}\big]\,.$$

Remark. We obtain the same result if we assume a model of the form of $R(N) = \nu N - \sum_{j=1}^{N} X_j - YN$ such that $X_1, X_2, \ldots$ is a real stationary ergodic process having finite mean $\mu = E[X_1]$, and $P(\nu - \mu - Y = 0) = 0$. In fact, as a result of the Ergodic Theorem 2.14, it holds that

$$\frac{1}{N}\sum_{j=1}^{N} X_j \xrightarrow{N\to\infty} \mu \qquad \text{a.s. and in mean.}$$

Therefore, the requirements of the Proposition 3.2 are satisfied and it follows that, if $N \to \infty$, the expected risk-adjusted return for $R(N)$ converges to r_∞.

We would like to illustrate Proposition 3.2 by concentrating on normally distributed random variables and by giving a numerical example.

Example 1. Assume that an insurance company consists of one business unit, i.e., $n = 1$. Then we consider

$$R(N) = \nu N - \sum_{j=1}^{N} X_j - YN.$$

Suppose that $\{X_j\}_{j\in\mathbb{N}}$ is an i.i.d. sequence of random variables having normal distribution and that Y is a normally distributed random variable independent of the sequence $\{X_j\}_{j\in\mathbb{N}}$. In particular, we write

$$X_j \sim \mathcal{N}\big(\mu, \sigma^2\big) \quad \text{for all } j \in \mathbb{N}\,,$$
$$Y \sim \mathcal{N}\big(\tilde{\mu}, \tilde{\sigma}^2\big).$$

Because of the particular properties of the normal distribution it holds that

$$\frac{1}{N}\sum_{j=1}^{N} X_j \sim \mathcal{N}\Big(\mu, \frac{\sigma^2}{N}\Big)$$

and therefore

$$R_N = \nu - \frac{1}{N}\sum_{j=1}^{N} X_j - Y \sim \mathcal{N}\Big(\nu - \mu - \tilde{\mu}, \frac{\sigma^2}{N} + \tilde{\sigma}^2\Big).$$

For more details see Johnson et al. (1994), Chapter 13, Section 3.

Based on R_N, we can calculate the expected risk-adjusted return r_N. Given $c \le 0$, define

$$c_N = \frac{c/N - E[R_N]}{\sqrt{\operatorname{Var}(R_N)}} = \frac{c/N - \nu + \mu + \tilde{\mu}}{\sqrt{\sigma^2/N + \tilde{\sigma}^2}}\,.$$

Using (A.3) from the Appendix, we obtain

$$\begin{aligned} r_N &= \frac{E\big[R(N)\big]}{E\big[-R(N)\,|\,R(N) \le c\big]} \\ &= \frac{E[R_N]}{E[-R_N\,|\,R_N \le c/N]} \\ &= \frac{\nu - \mu - \tilde{\mu}}{-\nu + \mu + \tilde{\mu} + \sqrt{\sigma^2/N + \tilde{\sigma}^2}\big(\log\Phi\big)'(c_N)} \,. \end{aligned}$$

Note that the last expression is mathematically meaningful even if N is not an integer. If $c = 0$, the above expression simplifies to

$$r_N = \frac{-c_N}{c_N + \big(\log\Phi\big)'(c_N)} \,.$$

For every $c \le 0$ we get

$$c_\infty := \lim_{N\to\infty} c_N = \frac{-\nu + \mu + \tilde{\mu}}{\tilde{\sigma}} \,,$$

hence

$$\begin{aligned} r_\infty = \lim_{N\to\infty} r_N &= \frac{\nu - \mu - \tilde{\mu}}{-\nu + \mu + \tilde{\mu} + \tilde{\sigma}\big(\log\Phi\big)'(c_\infty)} \\ &= \frac{-c_\infty}{c_\infty + \big(\log\Phi\big)'(c_\infty)} \,. \end{aligned}$$

Let us now choose the parameter as follows

$$\nu = 5, \quad \mu = 1, \quad \tilde{\mu} = 2, \quad \sigma = 2, \quad \tilde{\sigma} = 1$$

and we consider different values for c. In this particular case we choose

$$c = 0, \quad c = -1, \quad c = -2, \quad c = -5, \quad \text{respectively.}$$

In Figure 3.2, a graphical representation of this situation is shown.

In Figure 3.3, we can observe that the returns $\{r_N\}_{N\in\mathbb{N}}$ converge to r_∞ for any $c \le 0$, as asserted in Proposition 3.2 . In fact,

$$r_\infty = \frac{2}{-2 + \big(\log\Phi\big)'(-2)} \cong 5.359 \,.$$

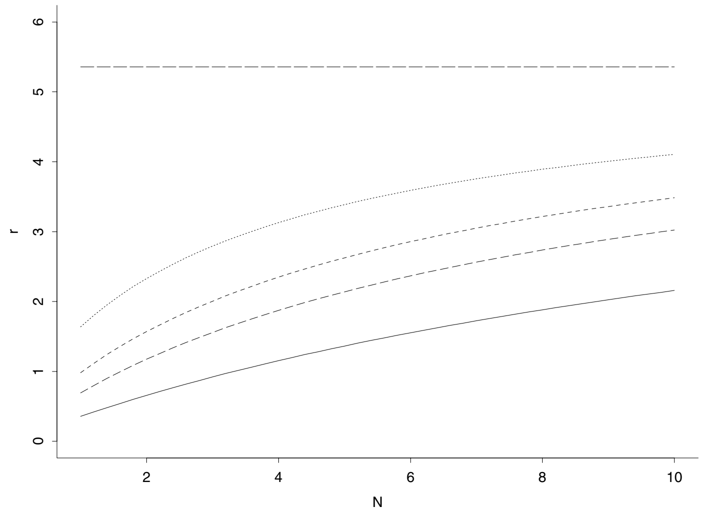

Figure 3.2: This figure shows the risk-adjusted returns r_N of the form mentioned above for the parameter values under the assumptions that parameter $c = 0$ (dotted line), $c = -1$ (short-dashed line), $c = -2$ (dashed line), $c = -5$ (solid line), respectively, and the limit r_∞ (long-dashed line).

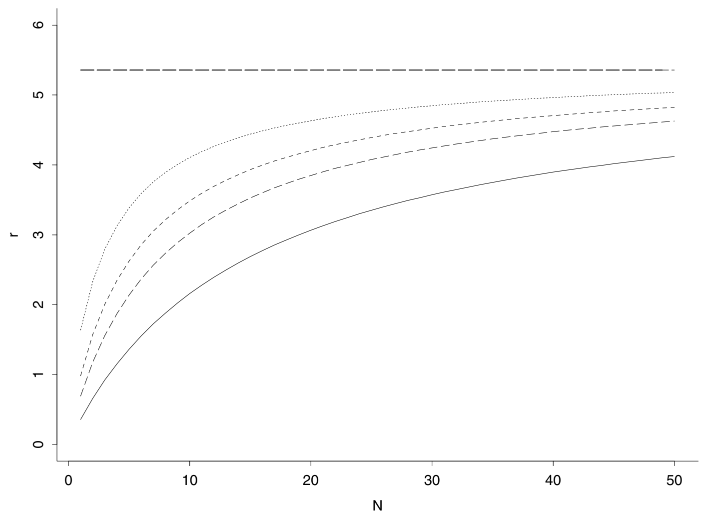

Figure 3.3: Plot of the same returns r_N shown in Figure 3.2 for greater values of N.

3.2 The n-dimensional model

The previous considerations are also valid for an n-dimensional model. In fact, in general we can denote the whole profit of a company consisting of n units by

$$R(N_1,\dots,N_n) = N_1\Big(\nu_1 - \frac{1}{N_1}\sum_{j=1}^{N_1} X_{1,j} - Y_1\Big) + \cdots \\ + N_n\Big(\nu_n - \frac{1}{N_n}\sum_{j=1}^{N_n} X_{n,j} - Y_n\Big),$$

where N_i are some positive integers for all $i \in \{1,\dots,n\}$ and $Y_1,\dots,Y_n$ are integrable random variables. Moreover, assume that real constants $\mu_1,\dots,\mu_n$ exist such that, for all $i \in \{1,\dots,n\}$,

$$\frac{1}{N_i}\sum_{j=1}^{N_i} X_{i,j} \xrightarrow{N_i\to\infty} \mu_i \qquad \text{in mean.}$$

Then we define

$$\tilde{R}(N_1,\dots,N_n) = \frac{R(N_1,\dots,N_n)}{N_1+\cdots+N_n}$$

and we obtain that the expected risk-adjusted return for all $c \le 0$ can be written as

$$\begin{aligned} r(N_1,\dots,N_n) &= \frac{E\big[R(N_1,\dots,N_n)\big]}{E\big[-R(N_1,\dots,N_n)\,|\,R(N_1,\dots,N_n)\le c\big]} \\ &= \frac{\frac{E\left[R(N_1,\dots,N_n)\right]}{N_1+\cdots+N_n}}{E\left[\frac{-R(N_1,\dots,N_n)}{N_1+\cdots+N_n}\,\middle|\,\frac{R(N_1,\dots,N_n)}{N_1+\cdots+N_n}\le\frac{c}{N_1+\cdots+N_n}\right]} \\ &= \frac{E\big[\tilde{R}(N_1,\dots,N_n)\big]}{E\left[-\tilde{R}(N_1,\dots,N_n)\,\middle|\,\tilde{R}(N_1,\dots,N_n)\le\frac{c}{N_1+\cdots+N_n}\right]}. \end{aligned}$$

Moreover, we assume that $N_1,\dots,N_n\to\infty$, such that

$$\frac{N_i}{N_1+\cdots+N_n} \longrightarrow t_i \quad \text{exists, for all } i=1,\dots,n\,.$$

Then, clearly, it follows that

$$(t_1, \dots, t_n) \in [0,1]^n \quad \text{and} \quad \sum_{i=1}^{n} t_i = 1 \tag{3.11}$$

and that

$$\tilde{R}(N_1, \dots, N_n) \xrightarrow{N_1,\dots,N_n \to \infty} \tilde{R}_\infty(t) := \sum_{i=1}^{n} t_i(\nu_i - \mu_i - Y_i) \quad \text{in mean,}$$

for all $t = (t_1, \dots, t_n)$ satisfying (3.11). Therefore assuming that

$$P\big(\tilde{R}_\infty(t) = 0\big) = 0,$$

it holds that

$$E[\tilde{R}(N_1, \dots, N_n) \,|\, \tilde{R}(N_1, \dots, N_n) \leq 0] \xrightarrow{N_1,\dots,N_n \to \infty} E\big[\tilde{R}_\infty(t) \,|\, \tilde{R}_\infty(t) \leq 0\big] \,.$$

Note that this is the same statement we arrived at (3.2) for the simplest model. For this reason, we omit the proof in this case, because it suffices to adapt the arguments considered there.

Analogously, as before, we obtain that the limit of the upper bounds exists and will be denoted by

$$r\big(N_1, \dots, N_n\big) \xrightarrow{N_1,\dots,N_n \to \infty} r_\infty\big(t_1, \dots, t_n\big)\,, \tag{3.12}$$

where

$$r_\infty\big(t_1, \dots, t_n\big) = \frac{\sum_{i=1}^{n} t_i(\nu_i - \mu_i - \tilde{\mu}_i)}{E\big[-\tilde{R}_\infty(t_1, \dots t_n) \,|\, \tilde{R}_\infty(t_1, \dots t_n) \leq 0\big]}\,.$$

We can observe that the limit of the upper bounds obtained for the n-dimensional model is directly dependent on $(t_1, \dots t_n)$. If $P\big(\tilde{R}_\infty(t) = 0\big) = 0$ for every $t = \big(t_1, \dots, t_n\big)$ satisfying (3.11), then (3.12) holds for all these t and it makes sense to determine the maximum of $r_\infty\big(t_1, \dots, t_n\big)$ for $(t_1, \dots, t_n)$ satisfying (3.11). In fact, if r_∞ is upper semi-continuous, then this maximum is attained by $t_1^*, \dots, t_n^*$, which can be calculated through

$$(t_1^*, \dots, t_n^*) = \underset{\substack{(t_1,\dots,t_n) \\ t_i \in [0,1]}}{\operatorname{argmax}} \frac{\sum_{i=1}^{n} t_i(\nu_i - \mu_i - \tilde{\mu}_i)}{E\big[-\tilde{R}_\infty(t_1, \dots t_n) \,|\, \tilde{R}_\infty(t_1, \dots t_n) \leq 0\big]}\,.$$

Then, for all $(t_1, \dots, t_n)$ satisfying (3.11), it holds that

$$r_\infty\big(t_1, \dots, t_n\big) \leq r_\infty\big(t_1^*, \dots, t_n^*\big) < \infty\,.$$

Note that, if $Y_1, \dots, Y_n$ are independent with continuous distribution functions, then $P\big(\tilde{R}_\infty(t) = 0\big) = 0$ certainly holds for all $(t_1, \dots, t_n)$ satisfying (3.11).

Example 2. Assume that an insurance company consists of two business units, i.e., $n = 2$. Then

$$R(N_1, N_2) = \nu_1 N_1 - \sum_{j=1}^{N_1} X_{1,j} - Y_1 N_1 + \nu_2 N_2 - \sum_{j=1}^{N_2} X_{2,j} - Y_2 N_2 \,.$$

Assume that $X_{i,j} \sim \mathcal{N}\big(\mu_i, \sigma_i^2\big)$ and $Y_i \sim \mathcal{N}\big(\tilde{\mu}_i, \tilde{\sigma}_i^2\big)$ for all $i \in \{1, 2\}$ and $j \in \mathbb{N}$. Furthermore, assume that all the random variables are independent. For simplicity, we denote $\hat{\mu}_i = \nu_i - \mu_i - \tilde{\mu}_i$ for $i \in \{1, 2\}$,

$$\hat{\mu} = N_1\,\hat{\mu}_1 + N_2\,\hat{\mu}_2$$

and

$$\hat{\sigma}^2 = N_1\,\sigma_1^2 + N_1^2\,\tilde{\sigma}_1^2 + N_2\,\sigma_2^2 + N_2^2\,\tilde{\sigma}_2^2 \,.$$

Then we have

$$R(N_1, N_2) \sim \mathcal{N}\big(\hat{\mu}, \hat{\sigma}^2\big) \,.$$

Using (A.3) as in the previous example, we obtain for $c \leq 0$

$$\begin{aligned} r(N_1, N_2) &= \frac{E[R(N_1, N_2)]}{E[-R(N_1, N_2)\,|\,R(N_1, N_2) \leq c]} \\ &= \frac{\hat{\mu}}{-\hat{\mu} + \hat{\sigma}\big(\log\Phi\big)'\Big((c - \hat{\mu})/\hat{\sigma}\Big)} \,. \end{aligned}$$

Assume that $N_1, N_2 \to \infty$ such that

$$\frac{N_1}{N_1 + N_2} \xrightarrow{N_1, N_2 \to \infty} t \quad \text{exists.}$$

Then

$$\frac{\hat{\mu}}{N_1 + N_2} \xrightarrow{N_1, N_2 \to \infty} t\,\hat{\mu}_1 + (1 - t)\,\hat{\mu}_2$$

and

$$\frac{\hat{\sigma}}{N_1 + N_2} \xrightarrow{N_1, N_2 \to \infty} \sqrt{t^2\,\tilde{\sigma}_1^2 + (1 - t)^2\,\tilde{\sigma}_2^2} \,.$$

Hence

$$r(N_1, N_2) \xrightarrow{N_1, N_2 \to \infty} r_\infty(t) \,,$$

with

$$\begin{aligned} r_\infty(t) &= \frac{t\,\hat{\mu}_1 + (1 - t)\,\hat{\mu}_2}{-t\,\hat{\mu}_1 - (1 - t)\,\hat{\mu}_2 + \sqrt{t^2\,\tilde{\sigma}_1^2 + (1 - t)^2\,\tilde{\sigma}_2^2}\big(\log\Phi\big)'\big(c(t)\big)} \\ &= \frac{-c(t)}{c(t) + \big(\log\Phi\big)'\big(c(t)\big)} \,, \end{aligned}$$

where

$$c(t) = -\frac{t\,\hat{\mu}_1 + (1-t)\,\hat{\mu}_2}{\sqrt{t^2\,\tilde{\sigma}_1^2 + (1-t)^2\,\tilde{\sigma}_2^2}}\,.$$

We now want to find the $t^* \in [0,1]$ which maximizes the limiting expected risk-adjusted return $[0,1] \ni t \mapsto r_\infty(t)$. We start with the following lemma.

Lemma 3.4. The function

$$(-\infty, 0] \ni c \mapsto -\frac{c}{c + \big(\log\Phi\big)'(c)}$$

is monotonely decreasing.

Proof. Using the substitution $z = x^2/2$, we get for every $c \in (-\infty, 0)$

$$\Phi(c) = \int_{-\infty}^{c} \varphi(x)\,dx < \int_{-\infty}^{c} \frac{x}{c}\,\varphi(x)\,dx = \frac{e^{-z}}{c\,\sqrt{2\pi}}\,\Big|_{c^2/2}^{\infty} = -\frac{\varphi(c)}{c}\,,$$

hence

$$\frac{1}{c}\big(\log\Phi\big)'(c) < -1\,.$$

Therefore, it suffices to show that $g : (-\infty, 0) \to \mathbb{R}$ with

$$g(c) = \frac{1}{c}\big(\log\Phi\big)'(c), \qquad c \in (-\infty, 0)\,,$$

is monotonely decreasing.

Since

$$g'(c) = \Big(\frac{\varphi(c)}{c\,\Phi(c)}\Big)' = \frac{-c^2\,\varphi(c)\,\Phi(c) - \varphi(c)\,\Phi(c) - c\,\varphi(c)^2}{c^2\,\Phi(c)^2}\,,$$

we obtain that for every $c \in (-\infty, 0)$

$$\begin{aligned} g'(c) \le 0 \iff & (c^2+1)\Phi(c) + c\,\varphi(c) \ge 0 \\ \iff & \Big(1 + \frac{1}{c^2}\Big)\Phi(c) \ge -\frac{\varphi(c)}{c}\,. \end{aligned}$$

Since for $c < 0$

$$\begin{aligned} \Big(1 + \frac{1}{c^2}\Big)\Phi(c) &= \int_{-\infty}^{c} \Big(1 + \frac{1}{c^2}\Big)\varphi(x)\,dx \\ &\ge \int_{-\infty}^{c} \Big(1 + \frac{1}{x^2}\Big)\varphi(x)\,dx = -\frac{\varphi(x)}{x}\,\Big|_{-\infty}^{c} = -\frac{\varphi(c)}{c}\,, \end{aligned}$$

the lemma is proved. □

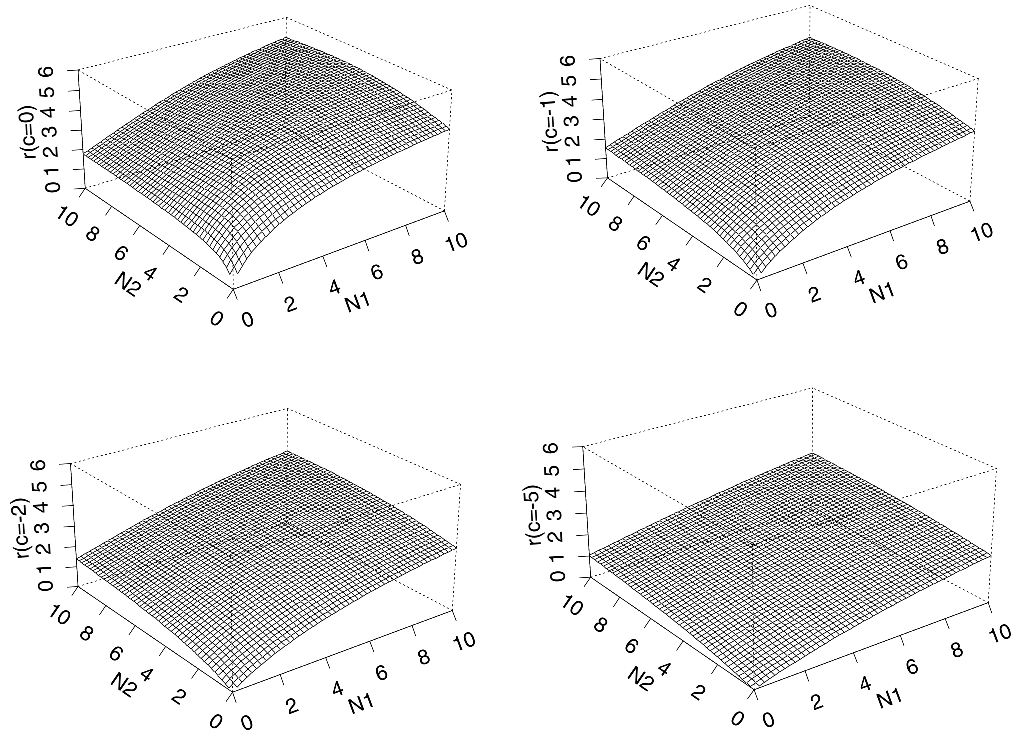

Figure 3.4: Plot of the risk-adjusted return $r(N_1, N_2)$: at the top left for the parameter $c = 0$, at the top right for $c = -1$, at the bottom left for $c = -2$, at the bottom right for $c = -5$.

Due to the lemma, it remains for us to find the $t^* \in [0,1]$ which minimizes $[0,1] \ni t \mapsto c(t)$. Solving the equation $c'(t^*) = 0$ leads to

$$t^* = \frac{\hat{\mu}_1 \, \tilde{\sigma}_2^2}{\hat{\mu}_1 \, \tilde{\sigma}_2^2 + \hat{\mu}_2 \, \tilde{\sigma}_1^2} \, ,$$

and

$$c(t^*) = -\sqrt{\frac{\hat{\mu}_1^2}{\tilde{\sigma}_1^2} + \frac{\hat{\mu}_2^2}{\tilde{\sigma}_2^2}} \, .$$

This is the minimum in $[0,1]$, because

$$c(1) = -\frac{\hat{\mu}_1}{\tilde{\sigma}_1} > c(t^*) \quad \text{and} \quad c(0) = -\frac{\hat{\mu}_2}{\tilde{\sigma}_2} > c(t^*) \, .$$

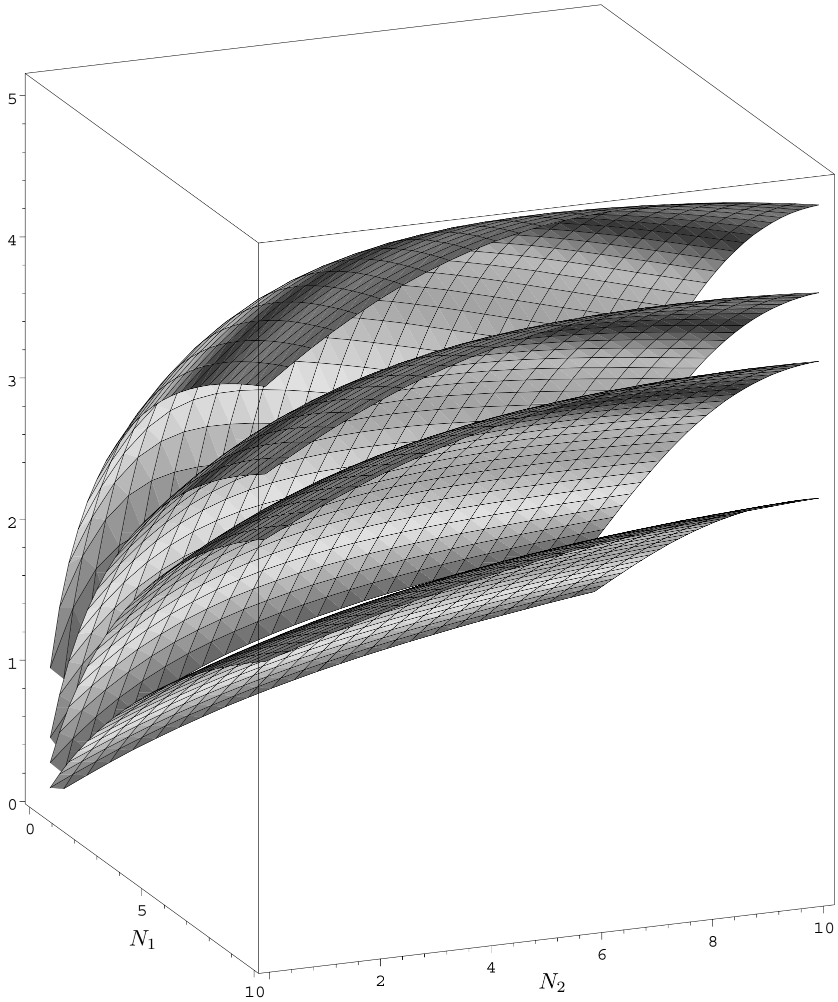

Figure 3.5: Plot of the risk-adjusted return $r(N_1, N_2)$ for the parameters $c = -5$, $c = -2$, $c = -1$, $c = 0$, respectively.

In Figures 3.4–3.7 we give a graphical representation assuming that

$$\nu_1 = 5, \quad \mu_1 = 1, \quad \tilde{\mu}_1 = 2, \quad \sigma_1 = 2, \quad \tilde{\sigma}_1 = 1,$$
$$\nu_2 = 3, \quad \mu_2 = 1, \quad \tilde{\mu}_2 = 1, \quad \sigma_2 = 1, \quad \tilde{\sigma}_2 = 1,$$

and choosing

$$c = 0, \quad c = -1, \quad c = -2, \quad c = -5, \quad \text{respectively.}$$

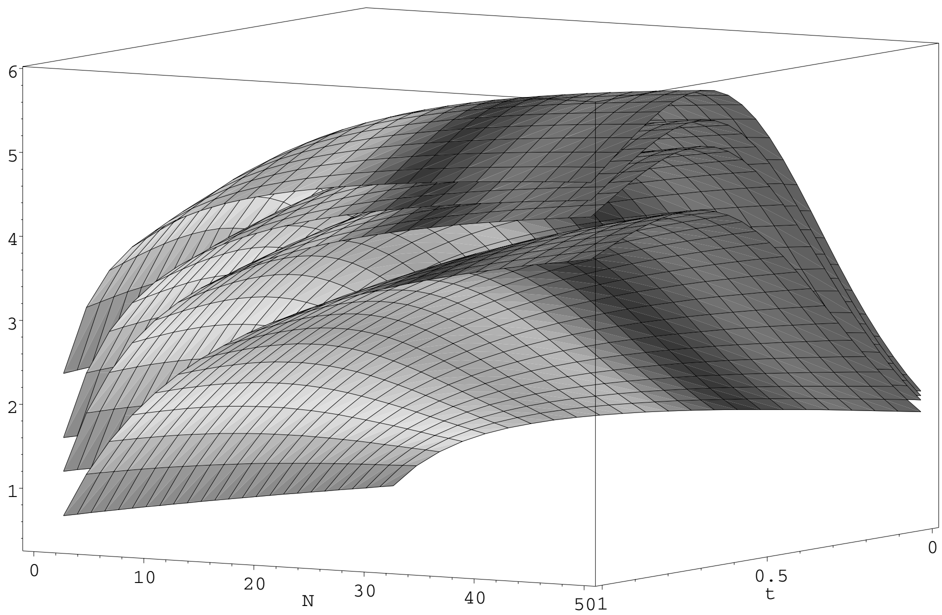

Figure 3.6: Plot of the risk-adjusted return $r(N_1, N_2)$ for $N_1 = t\,N$ and $N_2 = (1-t)\,N$ and for the parameters $c = -5$, $c = -2$, $c = -1$, $c = 0$, respectively.

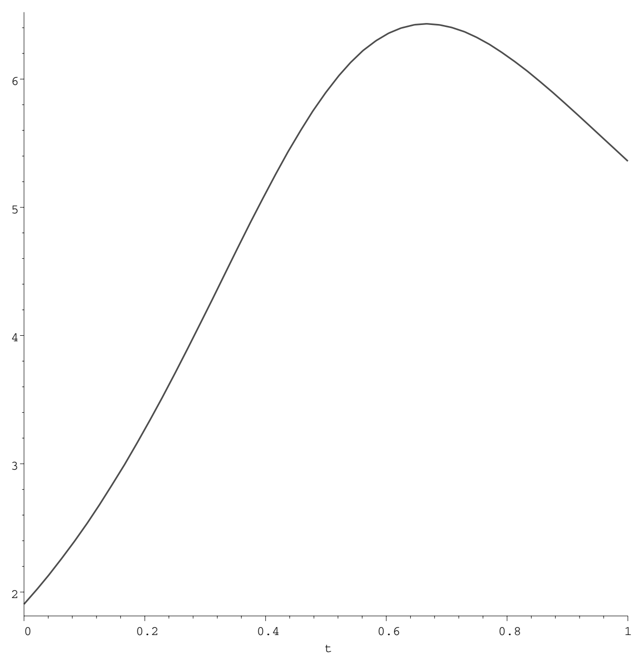

Figure 3.7: Plot of the limit $r_\infty(t)$.

In this case, it holds that

$$r_\infty(t) = \frac{-c(t)}{c(t) + \big(\log \Phi\big)'\big(c(t)\big)},$$

where

$$c(t) = \frac{t+1}{\sqrt{t^2 - 2\,t + 1}}.$$

Moreover, the $t^* \in [0,1]$ which maximises $r_\infty(t)$ is

$$t^* = \frac{2}{3}$$

and therefore

$$r_\infty(t^*) \cong 6.429\,.$$

We can observe that, for the same values of N_1 and N_2, increasing c, $r(N_1, N_2)$ becomes greater. This is a direct consequence of Lemma 3.1 which shows that the map $c \mapsto E[-R\,|\,R \leq c]$ decreases monotonously for any $c \in (c_0, \infty)$, where $c_0 = \inf\{c \in \mathbb{R}\,|\,P(R \leq c) > 0\}$. Moreover, assuming that $N_1, N_2 \to \infty$ such that $N_1/(N_1+N_2) \to t$, we can observe in the following figures that in any case – for every $c \leq 0$ – the return $r(N_1, N_2)$ converges to $r_\infty(t)$. In particular, in Figure 6, we can observe that $r_\infty(t) \leq r_\infty(t^*) \cong 6.429$ for all $t \in [0,1]$.

3.3 Conclusion

If we examine the portfolio of an insurance company represented by the model (3.1) and we try to optimize it by determining the number of contracts, we can conclude that, for the expected risk-adjusted return r_N defined by (3.2), a limit exists.

Considering the n-dimensional model we obtain a limit of the upper bounds which depends directly on the partition of the contracts between the different n units of the company. An optimal partition can be calculated in the case of normal distributions. In general, it has to be done numerically. Nevertheless, we cannot find the optimal number of contracts which should assure a maximal risk-adjusted return to the company.

Usually, the company has a fixed capital C at its disposal to invest. So, considering that the insurance company wants to invest this capital C, it is possible to calculate an optimal solution to the following problem:

$$\begin{aligned} &\text{maximize} && \frac{E\big[R(N_1,\dots N_n)\big]}{C} \\ &\text{subject to} && \rho\big(R(N_1,\dots,N_n)\big) \leq C \quad \text{and} \quad N_i \geq 0\,. \end{aligned}$$

In this case, using the expected shortfall as risk measure, we do not obtain a general solution, but it exists and can be found numerically if the distributions of Y_i for $i = 1, \dots, n$ are known. If the random variables Y_i for $i = 1, \dots, n$ are independent and normally distributed, this numerical calculation is not very complicated, because of the particular properties of the normal distribution.

However, we can observe that the obtained solution is not always a good one, because frequently the optimal values for N_i for $i = 1, \dots, n$ are not integers. For this reason we have to consider once more which optimal integer values we shall choose.

In the following chapters we will show that we obtain better solutions to such a problem if we take the standard deviation risk measure into consideration.

Chapter 4

Results for the standard deviation risk measure

In this chapter, we want to study the model of the portfolio of an insurance company defined in Section 2.5 once again. We will now examine the risk-adjusted performance of the company by studying the return

$$r = \frac{E[R]}{\rho(R)}$$

and using the standard deviation risk measure of R defined by

$$\rho(R) = -E[R] + \kappa\sigma(R),$$

where $\kappa > 0$ is some positive constant and $\sigma(R)$ denotes the standard deviation of the risk R.

In consequence, we assume the existence of second moments for the following, because the definition of $\sigma(R)$ clearly requires this.

Considering the portfolio of the form $R = \sum_{i=1}^{n} R_i$, let R_i be real-valued random variables on $(\Omega, \mathcal{F}, P)$ having finite expectation and finite second moment for all $i \in \{1, \ldots, n\}$

$$R_i = \nu_i N_i - \sum_{j=1}^{N_i} X_{i,j} - Y_i N_i\,, \tag{4.1}$$

where N_i are some positive integers for all $i \in \{1, \ldots, n\}$. For the sequences $\{X_{i,j}\}_{j\in\mathbb{N}}$ and for the random variables Y_i with $i \in \{1, \ldots, n\}$ the assumptions mentioned in Section 2.5.2 are still valid. Moreover, we assume that $(R_1, \ldots, R_n)$ are non-trivial and this means that $\rho(R)$ takes values other than zero, where R is the portfolio of $(R_1, \ldots, R_n)$.

Let $C > 0$ be the capital that the company wants to invest. Now we look at the following problem:

$$\begin{aligned} \text{maximize} \quad & \frac{E[R]}{C} \\ \text{subject to} \quad & \rho(R) \leq C \\ & N_i \geq 0 \quad \text{integers for } i \in \{1, \ldots, n\} \\ & \qquad\quad \text{but not all of them equal to zero.} \end{aligned} \tag{4.2}$$

We try to determine the optimal number of contracts in order to obtain a maximal return r, which means a better performance for the company.

First, we will examine the simplest cases with $n = 1$ and $n = 2$, and then we will take the same approach for a general n-dimensional model. In Section 4.3 we will repeat the same procedure, again analyzing the same model, but this time with the assumption that N_i are Poisson-distributed random variables with parameters $\lambda_i > 0$ for all $i = 1, \ldots, n$. This last assumption is convenient for us, because in this case the optimal solution of the optimization problem consists of real positive values which are no longer required to be integers.

4.1 The simplest cases: $n = 1$ and $n = 2$

We start considering $n = 1$. Let the portfolio be represented by

$$R(N) = \nu N - \sum_{j=1}^{N} X_j - YN \tag{4.3}$$

and take the assumptions mentioned in Section 2.5. We then have that N is a positive integer, $\{X_j\}_{j\in\mathbb{N}}$ are uncorrelated random variables with finite mean μ, Y has finite mean $\tilde{\mu}$ and the sequence $\{X_j\}_{j\in\mathbb{N}}$ and Y are uncorrelated. Moreover, we assume that both X_j for all $j \in \mathbb{N}$ and Y have finite variances denoted by $\sigma^2 = \text{Var}(X_j)$, independent of $j \in \mathbb{N}$, and $\tilde{\sigma}^2 = \text{Var}(Y)$ respectively.

Then we can compute

$$\begin{aligned} E\big[R(N)\big] &= N(\nu - \mu - \tilde{\mu}) \\ \rho\big(R(N)\big) &= -N(\nu - \mu - \tilde{\mu}) + \kappa\sqrt{N\sigma^2 + N^2\tilde{\sigma}^2}. \end{aligned}$$

We get that the optimization problem (4.2) can be written as

$$\begin{aligned} &\text{maximize} \quad && \frac{N(\nu-\mu-\tilde{\mu})}{C} \\ &\text{subject to} \quad && -N(\nu-\mu-\tilde{\mu}) + \kappa\sqrt{N\sigma^2+N^2\tilde{\sigma}^2} \le C\,, \\ & && N>0 \quad \text{integer.} \end{aligned} \tag{4.4}$$

For the sake of practicality we define

$$a = \frac{\nu-\mu-\tilde{\mu}}{C}.$$

Note that a is a positive constant because $\nu-\mu-\tilde{\mu}$ is assumed to be positive and the capital C to invest is positive too. Moreover, from now on, we assume

$$\frac{\kappa}{C}\tilde{\sigma} > a. \tag{4.5}$$

This means that even for the case $\sigma^2 = 0$, we get $\rho\big(R(N)\big) > 0$ for every $N \in \mathbb{N}$, hence, accepting contracts involves real risk. Furthermore, for any $\sigma^2 \ge 0$,

$$\lim_{N\to\infty} \frac{\rho\big(R(N)\big)}{N} > 0\,,$$

i.e., the risk per contract stays positive even in the limit $N \to \infty$.

Returning to the optimization problem; if we square the constraint (4.4) we reduce the previous formulation to

$$\begin{aligned} &\text{maximize} \quad && aN \\ &\text{subject to} \quad && N^2\Big(\frac{\kappa^2}{C^2}\tilde{\sigma}^2 - a^2\Big) + N\Big(\frac{\kappa^2}{C^2}\sigma^2 - 2a\Big) - 1 \le 0\,, \\ & && N>0 \quad \text{integer.} \end{aligned} \tag{4.6}$$

Given that $\kappa\tilde{\sigma}/C$ is assumed to be greater than a, we obtain that the constraint (4.6) is fulfilled if $N \in [z_1, z_2] \cap \mathbb{N}$, where with z_1 and z_2 we denote the zeros of the constraint itself. So, since a is positive and the function to maximize becomes greater with increasing N, the value of N which assures all requirements is the greatest positive integer which satisfies the condition (4.6). We have that the greatest zero of (4.6) has the following form:

$$\begin{aligned} z_2 &= \frac{-\Big(\frac{\kappa^2}{C^2}\sigma^2 - 2a\Big) + \sqrt{\Big(\frac{\kappa^2}{C^2}\sigma^2 - 2a\Big)^2 + 4\Big(\frac{\kappa^2}{C^2}\tilde{\sigma}^2 - a^2\Big)}}{2\Big(\frac{\kappa^2}{C^2}\tilde{\sigma}^2 - a^2\Big)} \\ &= \frac{2aC^2 - \kappa^2\sigma^2 + \kappa\sqrt{\kappa^2\sigma^4 - 4a\sigma^2C^2 + 4\tilde{\sigma}^2C^2}}{2\big(\kappa^2\tilde{\sigma}^2 - a^2C^2\big)}\,. \end{aligned}$$

We observe that z_2 is clearly positive but not necessarily an integer. So, in consequence of these arguments we can conclude that the solution of the optimization problem (4.4) is

$$N^* = \left\lfloor \frac{2aC^2 - \kappa^2\sigma^2 + \kappa\sqrt{\kappa^2\sigma^4 - 4a\sigma^2C^2 + 4\tilde{\sigma}^2C^2}}{2\big(\kappa^2\tilde{\sigma}^2 - a^2C^2\big)} \right\rfloor,$$

where with $\lfloor x \rfloor$ we denote the greatest integer smaller than x, i.e.,

$$\lfloor x \rfloor = \max\{k \in \mathbb{Z} \,|\, k \le x\}\,.$$

Now let $n = 2$. The portfolio has the form

$$R(N_1, N_2) = \nu_1 N_1 - \sum_{j=1}^{N_1} X_{1,j} - Y_1 N_1 + \nu_2 N_2 - \sum_{j=1}^{N_2} X_{2,j} - Y_2 N_2\,, \tag{4.7}$$

where N_1, N_2 are positive integers, $\{X_{1,j}\}_{j\in\mathbb{N}}$ and $\{X_{2,j}\}_{j\in\mathbb{N}}$ are sequences of uncorrelated random variables which are uncorrelated to Y_1 and Y_2. We suppose that the sequences $\{X_{1,j}\}_{j\in\mathbb{N}}$ and $\{X_{2,j}\}_{j\in\mathbb{N}}$ are uncorrelated and Y_1 and Y_2 are uncorrelated, too. Furthermore, we assume that the first two moments of $\{X_{1,j}\}_{j\in\mathbb{N}}$ and $\{X_{2,j}\}_{j\in\mathbb{N}}$ do not depend on $j \in \mathbb{N}$. Moreover, we define

$$\mu_i = E[X_{i,j}] \;\text{ and }\; \sigma_i^2 = \mathrm{Var}(X_{i,j}), \quad \text{for all } j \in \mathbb{N} \text{ and for all } i = 1, 2,$$
$$\tilde{\mu}_i = E[Y_i] \;\text{ and }\; \tilde{\sigma}_i^2 = \mathrm{Var}(Y_i), \quad \text{for } i = 1, 2,$$

where both means and variances are finite.

Then, we calculate

$$\begin{aligned} E\big[R(N_1, N_2)\big] &= N_1(\nu_1 - \mu_1 - \tilde{\mu}_1) + N_2(\nu_2 - \mu_2 - \tilde{\mu}_2)\,, \\ \rho\big(R(N_1, N_2)\big) &= -N_1(\nu_1 - \mu_1 - \tilde{\mu}_1) - N_2(\nu_2 - \mu_2 - \tilde{\mu}_2) \\ &\quad + \kappa\sqrt{N_1\sigma_1^2 + N_1^2\tilde{\sigma}_1^2 + N_2\sigma_2^2 + N_2^2\tilde{\sigma}_2^2}\,. \end{aligned}$$

In order to simplify the following calculations we define two constants as above

$$a_1 = \frac{\nu_1 - \mu_1 - \tilde{\mu}_1}{C} \quad \text{and} \quad a_2 = \frac{\nu_2 - \mu_2 - \tilde{\mu}_2}{C}.$$

Once again, we assume that these constants are positive, because we consider contracts with positive expectation.

The initial optimization problem (4.2) therefore turns out to have the following form

$$\begin{aligned}
&\text{maximize} \quad a_1 N_1 + a_2 N_2 \\
&\text{subject to} \quad -a_1 N_1 - a_2 N_2 + \frac{\kappa}{C}\sqrt{N_1\sigma_1^2 + N_1^2\tilde{\sigma}_1^2 + N_2\sigma_2^2 + N_2^2\tilde{\sigma}_2^2} \le 1\,, \qquad (4.8)\\
&\qquad\qquad N_1, N_2 \ge 0 \quad \text{integers, not both of them equal to zero.}
\end{aligned}$$

It is useful, analogously, as before, to square the constraint (4.8), so we obtain

$$\begin{aligned}
&\text{maximize} \quad a_1 N_1 + a_2 N_2 \\
&\text{subject to} \quad N_1\Big(\frac{\kappa^2}{C^2}\sigma_1^2 - 2a_1\Big) + N_2\Big(\frac{\kappa^2}{C^2}\sigma_2^2 - 2a_2\Big) + N_1^2\Big(\frac{\kappa^2}{C^2}\tilde{\sigma}_1^2 - a_1^2\Big) \\
&\qquad + N_2^2\Big(\frac{\kappa^2}{C^2}\tilde{\sigma}_2^2 - a_2^2\Big) - 2a_1a_2N_1N_2 - 1 \le 0\,, \qquad (4.9)\\
&\qquad N_1, N_2 \ge 0 \quad \text{integers, not both of them equal to zero.}
\end{aligned}$$

Note that, both here and in the n-dimensional case, we make the same assumption as in (4.5) and, in general, this implies

$$\frac{\kappa}{C}\tilde{\sigma}_i > a_i \qquad \text{for all } i \in \{1, \dots, n\}.$$

In particular in this case, where $n = 2$, it is assumed that

$$\frac{\kappa}{C}\tilde{\sigma}_1 > a_1 \quad \text{and} \quad \frac{\kappa}{C}\tilde{\sigma}_2 > a_2. \qquad (4.10)$$

Now we consider only the constraint (4.9). It can be written as

$$N_1\Big(\frac{\kappa^2}{C^2}\sigma_1^2 - 2a_1\Big) + N_2\Big(\frac{\kappa^2}{C^2}\sigma_2^2 - 2a_2\Big) + q \le 1\,, \qquad (4.11)$$

where with q we denote a real quadratic form

$$q = \mathbf{X}^T\mathbf{A}\mathbf{X}, \qquad (4.12)$$

with

$$\mathbf{X} = \begin{pmatrix} N_1 \\ N_2 \end{pmatrix}$$

and $\mathbf{A}$ representing the symmetric matrix of q relative to N_1 and N_2

$$\mathbf{A} = \begin{pmatrix} \frac{\kappa^2}{C^2}\tilde{\sigma}_1^2 - a_1^2 & -a_1a_2 \\ -a_1a_2 & \frac{\kappa^2}{C^2}\tilde{\sigma}_2^2 - a_2^2 \end{pmatrix}.$$

This notation is useful, since any real quadratic form q, with $q = \mathbf{X}^T\mathbf{A}\mathbf{X}$ as in (4.12) can be reduced to a diagonalized representation, and by an orthogonal change of variables the expression (4.11) can be rewritten in a form related to an ellipse. It therefore follows that the optimization problem can be solved directly.

So we denote the eigenvalues of $\mathbf{A}$ by α_1 and α_2 and the associated orthonormal eigenvectors by $\mathbf{P_1}$ and $\mathbf{P_2}$, then we can represent q in the form

$$q = \begin{pmatrix} x \\ y \end{pmatrix}^T \mathbf{B} \begin{pmatrix} x \\ y \end{pmatrix} = \alpha_1 x^2 + \alpha_2 y^2, \tag{4.13}$$

where

$$\mathbf{B} = \mathbf{P}^T\mathbf{A}\mathbf{P} = \begin{pmatrix} \alpha_1 & 0 \\ 0 & \alpha_2 \end{pmatrix}$$

and

$$\mathbf{P} = \Big(\mathbf{P_1}, \mathbf{P_2}\Big) = \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix},$$

i.e., $\mathbf{P}$ is the orthogonal matrix with the orthonormal eigenvectors of $\mathbf{A}$ in its columns. For more details see Gilbert and Gilbert (1995) Chapter 8.5.

From now on, we denote $\mathbf{Y} = (x, y)^T$. Then, by a change of variables from N_1, N_2 to x, y according to the rule that $\mathbf{X} = \mathbf{P}\mathbf{Y}$, the constraint considered until now can be re-written as follows:

$$\beta_1 x + \beta_2 y + \alpha_1 x^2 + \alpha_2 y^2 \leq 1. \tag{4.14}$$

Recall that if $\mathbf{A}$ is a real and symmetric matrix, its eigenvalues are real, so α_1, α_2, β_1 and β_2 are real constants dependent on a_1, a_2, $\tilde{\sigma}_1$, $\tilde{\sigma}_2$, κ and C. In particular, α_1 and α_2 have the following forms:

$$\alpha_i = \frac{1}{2}\left(\frac{\kappa^2}{C^2}\big(\tilde{\sigma}_1^2 + \tilde{\sigma}_2^2\big) - a_1^2 - a_2^2 \pm \sqrt{\Big(\frac{\kappa^2}{C^2}\big(\tilde{\sigma}_1^2 - \tilde{\sigma}_2^2\big) - a_1^2 + a_2^2\Big)^2 + 4a_1^2 a_2^2}\right),$$

where we choose α_1 to correspond to the plus sign and α_2 to the minus sign. Both constants are positive due to (4.10).

Then the values of the constants β_1 and β_2 can be calculated too, but for reasons of space, we will not list these formulae at full length. However it holds that

$$\beta_i = p_{1i}\Big(\frac{\kappa^2}{C^2}\sigma_1^2 - 2a_1\Big) + p_{2i}\Big(\frac{\kappa^2}{C^2}\sigma_2^2 - 2a_2\Big)$$

for $i = 1, 2$.

The normalized eigenvector $\mathbf{P_i}$ associated to α_i can be represented as follows

$$\mathbf{P_i} = \frac{\mathbf{v_i}}{\|\mathbf{v_i}\|}, \qquad \text{with} \quad \mathbf{v_i} = \begin{pmatrix} \frac{\kappa^2}{C^2}\tilde{\sigma}_1^2 - a_1^2 - \alpha_i + a_1 a_2 \\ -a_1 a_2 - \frac{\kappa^2}{C^2}\tilde{\sigma}_2^2 - a_2^2 - \alpha_i \end{pmatrix}.$$

Consequently, the form of the constraint mentioned in (4.14) can be transformed once again, so as to get the following form related to an ellipse

$$\begin{aligned} & \beta_1 x + \beta_2 y + \alpha_1 x^2 + \alpha_2 y^2 \leq 1 \\ \Longleftrightarrow \quad & \alpha_1\Big[(x-\gamma_1)^2 - \gamma_1^2\Big] + \alpha_2\Big[(x-\gamma_2)^2 - \gamma_2^2\Big] \leq 1, \quad \text{with } \gamma_i = \frac{\beta_i}{2\alpha_i}, \\ \Longleftrightarrow \quad & \alpha_1\big(x-\gamma_1\big)^2 + \alpha_2\big(x-\gamma_2\big)^2 - \alpha_1\gamma_1^2 - \alpha_2\gamma_2^2 \leq 1 \\ \Longleftrightarrow \quad & \frac{\alpha_1\big(x-\gamma_1\big)^2}{1+\alpha_1\gamma_1^2+\alpha_2\gamma_2^2} + \frac{\alpha_2\big(y-\gamma_2\big)^2}{1+\alpha_1\gamma_1^2+\alpha_2\gamma_2^2} \leq 1. \end{aligned} \tag{4.15}$$

To make it easier, we define

$$\varepsilon_i = \sqrt{\frac{1+\alpha_1\gamma_1^2+\alpha_2\gamma_2^2}{\alpha_i}} \qquad \text{for} \quad i = 1, 2\,.$$

Clearly, ε_1 and ε_2 represent the two half axes of the ellipse given in (4.15).

By the same change of variable, the function to maximize

$$a_1 N_1 + a_2 N_2$$

becomes

$$\tilde{a}_1 x + \tilde{a}_2 y \tag{4.16}$$

with

$$\tilde{a}_i = a_1 p_{1i} + a_2 p_{2i} \qquad \text{for} \quad i = 1, 2\,.$$

To recapitulate, we now want to solve the following problem:

$$\begin{array}{ll} \text{maximize} & f(x,y) \\ \text{subject to the condition} & g(x,y) \leq 0\,, \end{array} \tag{4.17}$$

where

$$\begin{aligned} f(x,y) &= \tilde{a}_1 x + \tilde{a}_2 y\,, \\ g(x,y) &= \frac{\big(x-\gamma_1\big)^2}{\varepsilon_1^2} + \frac{\big(y-\gamma_2\big)^2}{\varepsilon_2^2} - 1\,. \end{aligned}$$

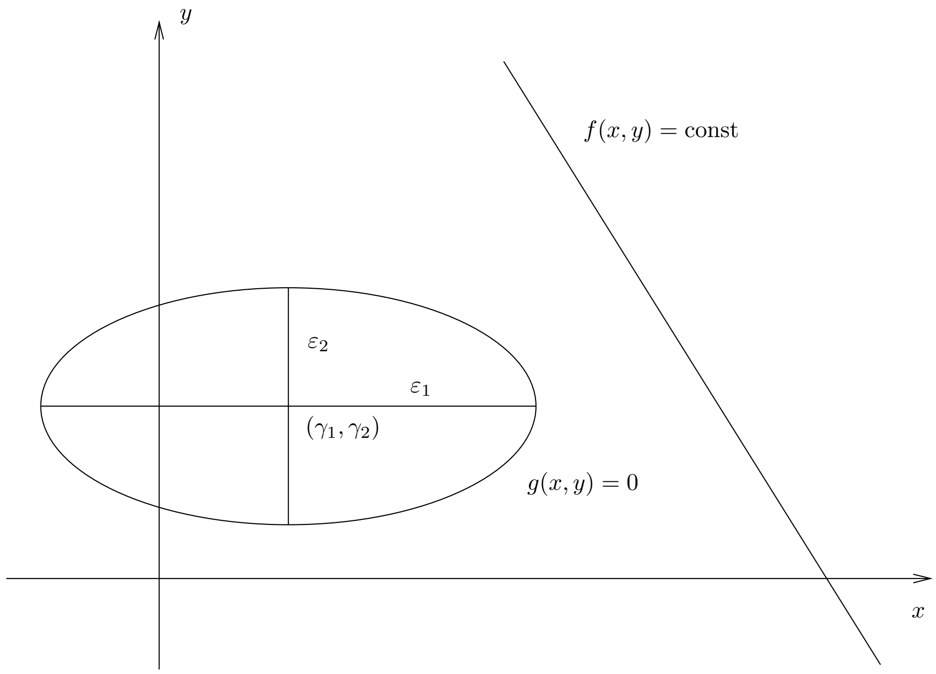

Figure 4.1: Illustration of the optimization problem 4.17

Let us ignore the restrictions coming from $N_1, N_2 \in \mathbb{N}_0$ with $\big(N_1, N_2\big) \neq \big(0, 0\big)$ for the moment. An optimization problem of this form can be illustrated by Figure 4.1, where the line $f(x, y) = \text{const}$ is orthogonal to $(\tilde{a}_1, \tilde{a}_2)$.

We see that the required maximal value for $f(x, y)$ is reached on the boundary of the ellipse described by $g(x, y) = 0$, at the point on the ellipse whose tangent is parallel to the straight line. The optimal solution is (x^*, y^*) and it can be calculated with the aid of Lagrange multipliers. In fact, by introducing suitable multipliers, the constrained extremum problem can be treated as one of an ordinary extremum. More precisely, as follows from the Lagrange multipliers rule, it holds that:
if f has a relative extremum in (x^*, y^*) subject to the constraint $g(x^*, y^*) = 0$ and if $g_y(x^*, y^*) \neq 0$, then a real number δ called the Lagrange multiplier exists such that (x^*, y^*, δ) is the critical point of the function H defined by $H(x, y, \delta) = f(x, y) + \delta g(x, y)$. For more details see Appendix B.

Therefore, the pertinent problem to resolve is

$$\begin{cases} H_x(x, y, \delta) = f_x(x, y) + \delta g_x(x, y) = 0 \\ H_y(x, y, \delta) = f_y(x, y) + \delta g_y(x, y) = 0 \\ H_\delta(x, y) = g(x, y) = 0. \end{cases} \tag{4.18}$$

We get a system of three equations and in our case we find a solution with the aid of the computer. Then, through the inverse change of variables, we can finally determine the values for N_1 and N_2, respectively, that solve the initial problem (4.8). We obtain that the number of contracts which guarantee a maximal risk-adjusted return are

$$N_1^* = \frac{a_1\tilde{\sigma}_2^2\kappa\sqrt{s} + C^2\left(a_1a_2\sigma_2^2 - a_2^2\sigma_1^2 - 2a_1\tilde{\sigma}_2^2\right) + \kappa^2\sigma_1^2\tilde{\sigma}_2^2}{2\Big(C^2\left(a_1^2\tilde{\sigma}_2^2 + a_2^2\tilde{\sigma}_1^2\right) - \kappa^2\tilde{\sigma}_1^2\tilde{\sigma}_2^2\Big)}$$

and

$$N_2^* = \frac{a_2\tilde{\sigma}_1^2\kappa\sqrt{s} + C^2\left(a_1a_2\sigma_1^2 - a_1^2\sigma_2^2 - 2a_2\tilde{\sigma}_1^2\right) + \kappa^2\sigma_2^2\tilde{\sigma}_1^2}{2\Big(C^2\left(a_1^2\tilde{\sigma}_2^2 + a_2^2\tilde{\sigma}_1^2\right) - \kappa^2\tilde{\sigma}_1^2\tilde{\sigma}_2^2\Big)},$$

where

$$s = \frac{C^2\Big(4\left(\tilde{\sigma}_1^2\tilde{\sigma}_2^2 - a_1\sigma_1^2\tilde{\sigma}_2^2 - a_2\sigma_2^2\tilde{\sigma}_1^2\right) - \left(a_1\sigma_2^2 - a_2\sigma_1^2\right)^2\Big) + \kappa^2\left(\sigma_1^4\tilde{\sigma}_2^2 + \sigma_2^4\tilde{\sigma}_1^2\right)}{a_1^2\tilde{\sigma}_2^2 + a_2^2\tilde{\sigma}_1^2}.$$

Note that N_1^* and N_2^* defined above are not necessarily integers, so they have to be rounded in practice.

Remark. We can proceed analogously even if we consider a model of the same form as (4.7) and we allow correlation between the two random variables Y_1, Y_2. In fact, it holds that

$$E\big[R(N_1, N_2)\big] = N_1\left(\nu_1 - \mu_1 - \tilde{\mu}_1\right) + N_2\left(\nu_2 - \mu_2 - \tilde{\mu}_2\right)$$

and

$$\begin{aligned}\operatorname{Var}\big(R(N_1, N_2)\big) &= \operatorname{Var}(R_1 + R_2)\\ &= \operatorname{Var}(R_1) + \operatorname{Var}(R_2) + 2\operatorname{Cov}(R_1, R_2)\\ &= N_1\tilde{\sigma}_1^2 + N_1^2\tilde{\sigma}_1^2 + N_2\tilde{\sigma}_2^2 + N_2^2\tilde{\sigma}_2^2 + 2\,N_1N_2\,\rho_Y\tilde{\sigma}_1\tilde{\sigma}_2\,,\end{aligned}$$

because

$$\begin{aligned}\operatorname{Cov}(R_1, R_2) &= N_1\,N_2\operatorname{Cov}(Y_1, Y_2)\\ &= N_1\,N_2\,\rho_Y\tilde{\sigma}_1\tilde{\sigma}_2\,,\end{aligned}$$

where $\rho_Y = \operatorname{corr}(Y_1, Y_2)$. Therefore, if we want to solve the optimization problem of the form of (4.2) for such a model R, we must find an optimal solution in the same way as before. As a matter of fact, in this case the optimization problem turns out to be very similar to (4.9). In particular, the constraint has the following form

$$N_1\Big(\frac{\kappa^2}{C^2}\sigma_1^2 - 2a_1\Big) + N_2\Big(\frac{\kappa^2}{C^2}\sigma_2^2 - 2a_2\Big) + N_1^2\Big(\frac{\kappa^2}{C^2}\tilde{\sigma}_1^2 - a_1^2\Big)$$
$$+ N_2^2\Big(\frac{\kappa^2}{C^2}\tilde{\sigma}_2^2 - a_2^2\Big) - 2\,N_1N_2\,\big(a_1a_2 - \rho_Y\tilde{\sigma}_1\tilde{\sigma}_2\big) - 1 \leq 0$$

and can be written as

$$N_1\Big(\frac{\kappa^2}{C^2}\sigma_1^2 - 2a_1\Big) + N_2\Big(\frac{\kappa^2}{C^2}\sigma_2^2 - 2a_2\Big) + q \leq 0,$$

where with q we denote a real quadratic form

$$q = \begin{pmatrix} N_1 \\ N_2 \end{pmatrix}^T \mathbf{A} \begin{pmatrix} N_1 \\ N_2 \end{pmatrix}$$

with

$$\mathbf{A} = \begin{pmatrix} \frac{\kappa^2}{C^2}\tilde{\sigma}_1^2 - a_1^2 & -a_1a_2 + \rho_Y\tilde{\sigma}_1\tilde{\sigma}_2 \\ -a_1a_2 + \rho_Y\tilde{\sigma}_1\tilde{\sigma}_2 & \frac{\kappa^2}{C^2}\tilde{\sigma}_2^2 - a_2^2 \end{pmatrix}.$$

Thus, adapting the previously considered argumentations (such as change of variables, form related to an ellipse, Lagrange multipliers) we can compute the optimal solution.

4.2 The general case

If we consider a general model related to a firm consisting of n units, i.e.,

$$R = \sum_{i=1}^{n} R_i, \qquad \text{with } R_i \text{ of the form of (4.1)},$$

we can proceed analogously, as before. In fact, similar arguments can be adapted to solve an n-dimensional optimization problem, too.

Starting from the initial problem of the form of (4.2), we square the constraint $\rho(R) \leq C$ and then we represent it with help of a quadratic form defined from an $n \times n-$matrix $\mathbf{A}$. Diagonalizing $\mathbf{A}$ and by an orthogonal change of variables, we can rewrite the constraint in a form related to a conic section. Therefore, we get a problem with new variables, i.e., $x_1, \ldots, x_n$ instead of $N_1, \ldots, N_n$ and it has the following form

$$\begin{array}{ll} \text{maximize} & f(x_1, \ldots, x_n) \\ \text{subject to the condition} & g(x_1, \ldots, x_n) \leq 0, \end{array}$$

where, as before, the maximal value for $f(x_1, \ldots, x_n)$ is reached on the boundary described by $g(x_1, \ldots, x_n) = 0$.

This means that the optimal solution $(x_1^*, \ldots, x_n^*)$ can be found employing the Lagrange multipliers rule (see Appendix B for more details). As a consequence, we have to solve a system of $n+1$ equations with respect to the unknowns $x_1, \ldots, x_n, \delta$

$$\begin{cases} H_{x_i}(x_1, \ldots, x_n, \delta) = f_{x_i}(x_1, \ldots, x_n) + \delta g_{x_i}(x_1, \ldots, x_n) = 0\,, \quad i = 1, \ldots, n \\ H_\delta(x_1, \ldots, x_n) = g(x_1, \ldots, x_n) = 0\,. \end{cases} \tag{4.19}$$

Then, by the inverse orthogonal change of variables, we can determine the solution of the initial problem with respect to the original variables $N_1, \ldots, N_n$. Due to the length of the expression related to the two-dimensional case, we avoid listing the values of $N_1^*, \ldots, N_n^*$, i.e., the optimal solution of the initial problem.

4.3 The second variant of the model: the N_i's are random variables

In this section, we still consider the same model representing the whole portfolio of a company – but with one difference. In fact, we assume that the number of contracts for every unit i, denoted by N_i with $i \in \{1, \ldots, n\}$, are Poisson-distributed random variables with parameter $\lambda_i > 0$.

To resume, we have

$$R = \sum_{i=1}^{n} \Big(\nu_i N_i - \sum_{j=1}^{N_i} X_{i,j} - Y_i N_i \Big)\,, \tag{4.20}$$

with

- $N_i \sim \mathrm{POIS}(\lambda_i)$ for all $i \in \{1, \ldots, n\}$,
- $\{X_{i,j}\}_{j \in \mathbb{N}}$ are independent[1] sequences of i.i.d. random variables with finite means and finite variances for all $i \in \{1, \ldots, n\}$

$$\mu_i = E[X_{i,j}] \;\text{ and }\; \sigma_i^2 = \mathrm{Var}(X_{i,j}),$$

[1] Note that these independence assumptions are introduced for simplicity. In fact, even if we consider a model which does not fulfil these supplementary requirements, we can proceed in a similar way.

- $Y_1, \ldots, Y_n$ are independent[1] random variables with both means and variances finite

$$\tilde{\mu}_i = E[Y_i] \quad \text{and} \quad \tilde{\sigma}_i^2 = \mathrm{Var}(Y_i),$$

- the sequences $\{X_{i,j}\}_{j\in\mathbb{N}}$, $i \in \{1, \ldots, n\}$, are independent of the random variables $Y_1, \ldots, Y_n$.

Moreover, the number of contracts $N_1, \ldots, N_n$ are assumed to be independent of the sequences $\{X_{i,j}\}_{j\in\mathbb{N}}$ for all $i \in \{1, \ldots, n\}$ and of the random variables $Y_1, \ldots, Y_n$.

Recall that the aforementioned assumption

$$\frac{\kappa}{C}\tilde{\sigma}_i > \frac{\nu_i - \mu_i - \tilde{\mu}_i}{C}$$

is still valid for all $i \in \{1, \ldots, n\}$.

We are now interested in the solution of the optimization problem

$$\begin{aligned} \text{maximize} \quad & \frac{E[R]}{C} \\ \text{subject to} \quad & \rho(R) \le C \\ & \lambda_i > 0 \quad \text{for all } i \in \{1, \ldots, n\}. \end{aligned} \tag{4.21}$$

We observe that this problem is analogous to (4.2), but that this time we have weaker conditions, in fact the values of the solution are no longer required to be integers.

Before considering this problem, we must do some calculations needed in the following. First, in the next Proposition we recall two particular equalities concerning expectation and variance.

Proposition 4.1. Let X be an $\mathbb{R}$-valued random variable defined on a probability space $(\Omega, \mathcal{F}, P)$ with a finite second moment and let $\mathcal{G}$ be a sub-σ-field of $\mathcal{F}$. Then

$$\begin{aligned} E[X] &= E\big[E[X \mid \mathcal{G}]\big] \\ \mathrm{Var}(X) &= E\big[\mathrm{Var}(X \mid \mathcal{G})\big] + \mathrm{Var}\big(E[X \mid \mathcal{G}\,]\big). \end{aligned}$$

Proof. See Fristedt and Gray (1997), Chapter 23. □

Then, since N_i is independent of the sequence $\{X_{i,j}\}_{j\in\mathbb{N}}$ and of Y_i and because of the Wald Identity we obtain

$$\begin{aligned}
E[R] &= \sum_{i=1}^{n} E\Big[N_i\nu_i - \sum_{j=1}^{N_i} X_{i,j} - N_iY_i\Big] \\
&= \sum_{i=1}^{n}\Big(E[N_i\nu_i] - E\Big[\sum_{j=1}^{N_i} X_{i,j}\Big] - E[N_iY_i]\Big) \\
&= \sum_{i=1}^{n}\Big(\nu_iE[N_i] - E[N_i]E[X_{i,j}] - E[N_i]E[Y_i]\Big) \\
&= \sum_{i=1}^{n}\lambda_i\big(\nu_i - \mu_i - \tilde{\mu}_i\big)
\end{aligned}$$

and since all the sequences $\{X_{i,j}\}_{j\in\mathbb{N}}$ with $i \in \{1,\dots,n\}$ are independent of $Y_1,\dots,Y_n$ it holds that

$$\begin{aligned}
\operatorname{Var}(R) &= \sum_{i=1}^{n}\operatorname{Var}\Big(N_i\nu_i - \sum_{j=1}^{N_i} X_{i,j} - N_iY_i\Big) \\
&= \sum_{i=1}^{n}\Bigg(\operatorname{Var}(N_i\nu_i) + \operatorname{Var}\Big(\sum_{j=1}^{N_i} X_{i,j}\Big) + \operatorname{Var}(N_iY_i) - 2\operatorname{Cov}\Big(N_i\nu_i, N_iY_i\Big) \\
&\qquad - 2\operatorname{Cov}\Big(N_i\nu_i, \sum_{j=1}^{N_i} X_{i,j}\Big) + 2\operatorname{Cov}\Big(\sum_{j=1}^{N_i} X_{i,j}, N_iY_i\Big)\Bigg).
\end{aligned}$$

First, we calculate the single variances and covariances separately:

$$\begin{aligned}
\operatorname{Var}(N_i\nu_i) &= \nu_i^2\operatorname{Var}(N_i) = \lambda_i\nu_i^2, \\
\operatorname{Var}\Big(\sum_{j=1}^{N_i} X_{i,j}\Big) &= E[N_i\sigma_i^2] + \operatorname{Var}(N_i\mu_i) = \lambda_i\big(\sigma_i^2 + \mu_i^2\big), \\
\operatorname{Var}(N_iY_i) &= E[N_i^2\tilde{\sigma}_i^2] + \operatorname{Var}(N_i\tilde{\mu}_i) = (\lambda_i^2 + \lambda_i)\tilde{\sigma}_i^2 + \lambda_i\tilde{\mu}_i^2, \\
\operatorname{Cov}\Big(N_i\nu_i, \sum_{j=1}^{N_i} X_{i,j}\Big) &= \nu_i\operatorname{Cov}\Big(N_i, \sum_{j=1}^{N_i} X_{i,j}\Big) \\
&= \nu_i\Big(E\Big[N_i\sum_{j=1}^{N_i} X_{i,j}\Big] - E[N_i]E\Big[\sum_{j=1}^{N_i} X_{i,j}\Big]\Big) \\
&= \nu_i\Big(E\big[N_i^2\mu_i\big] - E[N_i]E[N_i]\mu_i\Big) \\
&= \lambda_i\nu_i\mu_i,
\end{aligned}$$

$$\begin{aligned}
\operatorname{Cov}\Big(\sum_{j=1}^{N_i} X_{i,j}, N_iY_i\Big) &= E\Big[N_iY_i\sum_{j=1}^{N_i} X_{i,j}\Big] - E\Big[\sum_{j=1}^{N_i} X_{i,j}\Big]E\big[N_iY_i\big] \\
&= E\big[N_i^2\big]\mu_i\tilde{\mu}_i - E\big[N_i\big]^2\mu_i\tilde{\mu}_i \\
&= \lambda_i\mu_i\tilde{\mu}_i, \\
\operatorname{Cov}\Big(N_i\nu_i, N_iY_i\Big) &= \nu_i \operatorname{Cov}\Big(N_i, N_iY_i\Big) \\
&= \nu_i\Big(E[N_i^2Y_i] - E[N_i]^2E[Y_i]\Big) \\
&= \lambda_i\nu_i\tilde{\mu}_i.
\end{aligned}$$

Then, if we resume, we obtain

$$\begin{aligned}
\operatorname{Var}(R) &= \sum_{i=1}^{n}\Big(\lambda_i\nu_i^2 + \lambda_i\big(\sigma_i^2 + \mu_i^2\big) + (\lambda_i^2 + \lambda_i)\tilde{\sigma}_i^2 + \lambda_i\tilde{\mu}_i^2 + 2\lambda_i\mu_i\tilde{\mu}_i \\
&\qquad - 2\lambda_i\nu_i\mu_i - 2\lambda_i\nu_i\tilde{\mu}_i\Big) \\
&= \sum_{i=1}^{n}\Big(\lambda_i^2\tilde{\sigma}_i^2 + \lambda_i\big(\nu_i^2 + \mu_i^2 + \tilde{\mu}_i^2 + 2\mu_i\tilde{\mu}_i - 2\nu_i\mu_i - 2\nu_i\tilde{\mu}_i \\
&\qquad + \sigma_i^2 + \tilde{\sigma}_i^2\big)\Big) \\
&= \sum_{i=1}^{n}\Big(\lambda_i^2\tilde{\sigma}_i^2 + \lambda_i\big((\nu_i - \mu_i - \tilde{\mu}_i)^2 + \sigma_i^2 + \tilde{\sigma}_i^2\big)\Big).
\end{aligned}$$

Remark. Note that, if we allow dependence between the sequences $\{X_{i,j}\}_{j\in\mathbb{N}}$, for $i \in \{1, \dots, n\}$, and also between the random variables $Y_1, \dots, Y_n$, it holds that

$$\begin{aligned}
\operatorname{Var}(R) &= \sum_{i=1}^{n}\operatorname{Var}\Big(N_i\nu_i - \sum_{j=1}^{N_i} X_{i,j} - N_iY_i\Big) \\
&\quad + 2\sum_{i=1}^{n-1}\sum_{k>i}^{n}\operatorname{Cov}\Big(N_i\nu_i - \sum_{j=1}^{N_i} X_{i,j} - N_iY_i,\ N_k\nu_k - \sum_{j=1}^{N_k} X_{k,j} - N_kY_k\Big),
\end{aligned}$$

but we will not deal with this case in detail.

In order to determine the solution of (4.21) we take the same approach as in the previous section. This means that we consider the simplest cases with $n = 1$ and $n = 2$. Because of the length of the expressions we will not reiterate the general case, but all the following argumentations can be adapted in order to solve an n-dimensional optimization problem with respect to the model (4.20).

Let $n = 1$ and

$$R(N) = \nu N - \sum_{j=1}^{N} X_j - YN$$

with the above-mentioned assumptions. Recall that in this section N is a Poisson-distributed random variable with parameter $\lambda > 0$.

From the previous calculations we obtain

$$\begin{aligned} E\big[R(N)\big] &= \lambda\nu - \mu - \tilde{\mu} \\ \mathrm{Var}\big(R(N)\big) &= \lambda^2\tilde{\sigma}^2 + \lambda\big((\nu - \mu - \tilde{\mu})^2 + \sigma^2 + \tilde{\sigma}^2\big). \end{aligned}$$

We now have the standard deviation risk measure for $R(N)$ which is

$$\rho\big(R(N)\big) = -\lambda(\nu - \mu - \tilde{\mu}) + \kappa\sqrt{\lambda^2\tilde{\sigma}^2 + \lambda\big((\nu - \mu - \tilde{\mu})^2 + \sigma^2 + \tilde{\sigma}^2\big)}.$$

Consequently, in this case the initial problem (4.21) has the following form

$$\begin{aligned} &\text{maximize} && \frac{\lambda(\nu - \mu - \tilde{\mu})}{C} \\ &\text{subject to} && -\lambda(\nu - \mu - \tilde{\mu}) + \kappa\sqrt{\lambda^2\tilde{\sigma}^2 + \lambda\big((\nu - \mu - \tilde{\mu})^2 + \sigma^2 + \tilde{\sigma}^2\big)} \le C\,, \\ & && \lambda > 0. \end{aligned}$$

Analogously, as before, it is useful to introduce some positive constants in order to get shorter expressions. We define

$$\begin{aligned} a &= \frac{\nu - \mu - \tilde{\mu}}{C}, \\ b &= \big(\nu - \mu - \tilde{\mu}\big)^2 + \sigma^2 + \tilde{\sigma}^2. \end{aligned}$$

We remember that κ/C is assumed to be greater than a, because of the argumentations already mentioned.

Moreover, if we square the constraint we obtain a new formulation of the problem to solve, which is very similar to (4.6).

$$\begin{aligned} &\text{maximize} && a\lambda \\ &\text{subject to} && \lambda^2\Big(\frac{\kappa^2}{C^2}\tilde{\sigma}^2 - a^2\Big) + \lambda\Big(\frac{\kappa^2}{C^2}b - 2a\Big) - 1 \le 0\,, \\ & && \lambda > 0\,. \end{aligned} \tag{4.22}$$

We observe that, with respect to the problem (4.6), the only difference is the value of the constant related to the variable λ. In this case we will not rewrite all the passages to calculate the solution, because those in the previous section are easily adaptable. Therefore, we can determine that the optimal solution of (4.22) is

$$\lambda^* = \frac{2aC^2 - \kappa^2 b + \kappa\sqrt{\kappa^2 b^2 - 4abC^2 + 4\tilde{\sigma}^2 C^2}}{2\big(\kappa^2\tilde{\sigma}^2 - a^2 C^2\big)}.$$

Now, let $n = 2$ and

$$R(N_1, N_2) = \nu_1 N_1 - \sum_{j=1}^{N_1} X_{1,j} - Y_1 N_1 + \nu_2 N_2 - \sum_{j=1}^{N_2} X_{2,j} - Y_2 N_2\,,$$

with the assumptions cited at the beginning of this section. From the previous calculation we obtain

$$\begin{aligned} E\big[R(N_1, N_2)\big] &= \lambda_1\big(\nu_1 - \mu_1 - \tilde{\mu}_1\big) + \lambda_2\big(\nu_2 - \mu_2 - \tilde{\mu}_2\big) \\ \mathrm{Var}\big(R(N_1, N_2)\big) &= \lambda_1^2\,\tilde{\sigma}_1^2 + \lambda_1\big((\nu_1 - \mu_1 - \tilde{\mu}_1)^2 + \sigma_1^2 + \tilde{\sigma}_1^2\big) \\ &\quad + \lambda_2^2\,\tilde{\sigma}_2^2 + \lambda_2\big((\nu_2 - \mu_2 - \tilde{\mu}_2)^2 + \sigma_2^2 + \tilde{\sigma}_2^2\big), \end{aligned}$$

and introducing the constants

$$\begin{aligned} a_1 &= \frac{\nu_1 - \mu_1 - \tilde{\mu}_1}{C}\,, \\ a_2 &= \frac{\nu_2 - \mu_2 - \tilde{\mu}_2}{C}\,, \\ b_1 &= (\nu_1 - \mu_1 - \tilde{\mu}_1)^2 + \sigma_1^2 + \tilde{\sigma}_1^2\,, \\ b_2 &= (\nu_2 - \mu_2 - \tilde{\mu}_2)^2 + \sigma_2^2 + \tilde{\sigma}_2^2\,, \end{aligned}$$

it follows that the problem to be solved in this case is

$$\begin{aligned} &\text{maximize} && a_1\lambda_1 + a_2\lambda_2 \\ &\text{subject to} && -a_1\lambda_1 - a_2\lambda_2 + \frac{\kappa}{C}\sqrt{\lambda_1 b_1 + \lambda_1^2\tilde{\sigma}_1^2 + \lambda_2 b_2 + \lambda_2^2\tilde{\sigma}_2^2} \le 1, \\ &&& \lambda_1, \lambda_2 > 0\,. \end{aligned} \tag{4.23}$$

Remember that we assume

$$\frac{\kappa}{C}\tilde{\sigma}_1 > a_1 \qquad \text{and} \qquad \frac{\kappa}{C}\tilde{\sigma}_2 > a_2\,.$$

Once again, squaring the main constraint, we get a new formulation very similar to (4.9) with the exception of the constants related to λ_1 and to λ_2 respectively, i.e.,

$$\begin{aligned} \text{maximize} \quad & a_1\lambda_1 + a_2\lambda_2 \\ \text{subject to} \quad & \lambda_1\Big(\frac{\kappa^2}{C^2}b_1 - 2a_1\Big) + \lambda_2\Big(\frac{\kappa^2}{C^2}b_2 - 2a_2\Big) + \lambda_1^2\Big(\frac{\kappa^2}{C^2}\tilde{\sigma}_1^2 - a_1^2\Big) \\ & \quad + \lambda_2^2\Big(\frac{\kappa^2}{C^2}\tilde{\sigma}_2^2 - a_2^2\Big) - 2a_1a_2\lambda_1\lambda_2 - 1 \le 0 \qquad (4.24) \\ & \lambda_1, \lambda_2 > 0\,. \end{aligned}$$

We can calculate the solution for this problem in the same way as in the previous section. Without rewriting all the passages concerning the respectively real quadratic form q, its matrix and the orthogonal change of variables – which bring us to a form of the constraint corresponding to an ellipse – we obtain that the optimal value of the function to maximize is reached by

$$\lambda_1^* = \frac{a_1\tilde{\sigma}_2\kappa\sqrt{s} + \kappa^2\tilde{\sigma}_2 b_1 + C^2\big(a_1a_2b_2 - a_2^2b_1 - 2a_1\tilde{\sigma}_2\big)}{2\Big(C^2\big(a_1^2\tilde{\sigma}_2 + a_2^2\tilde{\sigma}_1\big) - \kappa^2\tilde{\sigma}_1\tilde{\sigma}_2\Big)}$$

and

$$\lambda_2^* = \frac{a_2\tilde{\sigma}_1\kappa\sqrt{s} + \kappa^2\tilde{\sigma}_1 b_2 + C^2\big(a_1a_2b_1 - a_1^2b_2 - 2a_2\tilde{\sigma}_1\big)}{2\Big(C^2\big(a_1^2\tilde{\sigma}_2 + a_2^2\tilde{\sigma}_1\big) - \kappa^2\tilde{\sigma}_1\tilde{\sigma}_2\Big)}\,,$$

where we use s as short notation for

$$s = \frac{\kappa^2\big(\tilde{\sigma}_2b_1^2 + \tilde{\sigma}_1b_2^2\big) + C^2\Big(4\big(\tilde{\sigma}_1\tilde{\sigma}_2 - a_2\tilde{\sigma}_1b_2 - a_1\tilde{\sigma}_2b_1\big) - \big(a_1b_2 - a_2b_1\big)^2\Big)}{a_1^2\tilde{\sigma}_2 + a_2^2\tilde{\sigma}_1}\,.$$

All these argumentations are also valid for an n-dimensional problem. Consequently, the initial optimization problem (4.21) has optimal solutions for every model of the form of (4.20), but because of the length of the expression we will not list general solutions.

4.4 The third variant of the model: N_i is the sum of a positive integer and a random variable for $i \in \{1, \ldots, n\}$

In this section, we consider a different variation of the model examined in this work and defined in Section 2.5. This new variation is based on the definition of N_i, i.e. the variables which represent the number of contracts of every unit i of the company, for $i \in \{1, \ldots, n\}$. We define

$$N_i = N_i^{\text{fix}} + N_i^{\text{pois}} \qquad \text{for } i \in \{1, \ldots, n\}.$$

This means that the number of contracts of any unit i is defined as the sum of a fixed integer N_i^{fix} and a Poisson-distributed random variable N_i^{pois} with parameter $\lambda_i \geq 0$.

We observe that this variant of the model includes both versions we considered previously. In fact, if we set $N_i^{\text{fix}} = 0$ for all $i \in \{1, \ldots, n\}$, we obtain the same variant as those examined in Section 4.3 or, if we set $\lambda_i = 0$ for all $i \in \{1, \ldots, n\}$, we have the initial model considered in Sections 4.1 and 4.2.

We represent the whole portfolio through

$$R = \sum_{i=1}^{n} \left(\nu_i \big(N_i^{\text{fix}} + N_i^{\text{pois}}\big) - \sum_{j=1}^{N_i^{\text{fix}} + N_i^{\text{pois}}} X_{i,j} - Y_i \big(N_i^{\text{fix}} + N_i^{\text{pois}}\big) \right),$$

with the assumptions mentioned. In particular, we remember that $\{X_{i,j}\}_{j \in \mathbb{N}}$ with $i \in \{1, \ldots, n\}$ are independent sequences of i.i.d. random variables, independent of the random variables $Y_1, \ldots, Y_n$ and that $Y_1, \ldots, Y_n$ are independent, too. Moreover, $N_1, \ldots, N_n$ are assumed to be independent of all the sequences $\{X_{i,j}\}_{j \in \mathbb{N}}$ with $i \in \{1, \ldots, n\}$ and of $Y_1, \ldots, Y_n$. Analogously, we define the finite means and finite variances of $X_{i,j}$ and Y_i, respectively, by

$$\mu_i = E[X_{i,j}] \text{ and } \sigma_i^2 = \operatorname{Var}(X_{i,j}), \quad \text{for all } j \in \mathbb{N} \text{ and all } i \in \{1, \ldots, n\},$$
$$\tilde{\mu}_i = E[Y_i] \text{ and } \tilde{\sigma}_i^2 = \operatorname{Var}(Y_i), \quad \text{for all } i \in \{1, \ldots, n\}.$$

Our interest turns to the following optimization problem in order to determine the portfolio which guarantees a maximal return.

$$\begin{aligned} \text{maximize} \quad & \frac{E[R]}{C}, \\ \text{subject to} \quad & \rho(R) \leq C, \\ & \lambda_i \geq 0 \quad \text{for all } i \in \{1, \ldots, n\}, \\ & N_i^{\text{fix}} \geq 0, \text{ integers for all } i \in \{1, \ldots, n\}. \end{aligned} \tag{4.25}$$

We will consider only the simplest case with $n = 1$, because the form of N_i already entails solving a two-dimensional optimization problem. A solution for a more general problem can be computed with the same argumentations, which can be easily adapted. For the sake of practicality, we rewrite the risk R as follows

$$R = \nu \tilde{N} - \sum_{j=1}^{\tilde{N}} X_j - Y \tilde{N},$$

where

$$\tilde{N} = N + N^{\mathrm{p}}\,,$$

with N denoting some positive integer and N^{p} a Poisson-distributed random variable with parameter $\lambda \geq 0$. Using the formulae cited in the previous section, we compute

$$\begin{aligned}
E[R] &= (N+\lambda)(\nu - \mu - \tilde{\mu})\,,\\
\mathrm{Var}(R) &= \mathrm{Var}\big(\nu\ (N+N^{\mathrm{p}})\big) + \mathrm{Var}\Big(\sum_{j=1}^{N+N^{\mathrm{p}}} X_j\Big) + \mathrm{Var}\big(Y(N+N^{\mathrm{p}})\big)\\
&\quad - 2\ \mathrm{Cov}\Big(\nu(N+N^{\mathrm{p}}), \sum_{j=1}^{N+N^{\mathrm{p}}} X_j\Big) + 2\ \mathrm{Cov}\Big(\sum_{j=1}^{N+N^{\mathrm{p}}} X_j, Y(N+N^{\mathrm{p}})\Big)\\
&\quad - 2\ \mathrm{Cov}\big(\nu(N+N^{\mathrm{p}}), Y(N+N^{\mathrm{p}})\big)\\
&= \lambda\big((\nu-\mu-\tilde{\mu})^2 + \sigma^2 + \tilde{\sigma}^2\big) + N\sigma^2 + (N^2+\lambda^2)\tilde{\sigma}^2\,.
\end{aligned}$$

To simplify the calculations we introduce the constants

$$\begin{aligned}
a &= (\nu-\mu-\tilde{\mu})/C\,,\\
b &= (\nu-\mu-\tilde{\mu})^2 + \sigma^2 + \tilde{\sigma}^2\,.
\end{aligned}$$

Moreover, as in previous sections we assume $\kappa\tilde{\sigma}/C > a$. Therefore, the optimization problem can be written as

$$\begin{aligned}
&\text{maximize} && a(N+\lambda) && (4.26)\\
&\text{subject to} && -a(N+\lambda) + \frac{\kappa}{C}\sqrt{\lambda b + N\sigma^2 + (N^2+\lambda^2)\,\tilde{\sigma}^2} \leq 1\,,\\
& && \lambda \geq 0,\ N \geq 0\,.
\end{aligned}$$

If we square the main constraint we obtain

$$\begin{aligned}
&\text{maximize} && a(N+\lambda) && (4.27)\\
&\text{subject to} && \lambda^2\Big(2\frac{\kappa^2}{C^2}\tilde{\sigma}^2 - a^2\Big) + \lambda\Big(\frac{\kappa^2}{C^2}b - 2a\Big) + N^2\Big(\frac{\kappa^2}{C^2}\tilde{\sigma}^2 - a^2\Big)\\
& && + N\Big(\frac{\kappa^2}{C^2}\sigma^2 - 2a\Big) - 2N\lambda\Big(\frac{\kappa^2}{C^2}\tilde{\sigma}^2 - 2a\Big) - 1 \leq 0,\\
& && \lambda \geq 0,\ N \geq 0.
\end{aligned}$$

Then, if we consider only the constraint, we can rewrite it in a new form related to an ellipse. In fact, introducing a quadratic form

$$q = \lambda^2\Big(2\frac{\kappa^2}{C^2}\tilde{\sigma}^2 - a^2\Big) + N^2\Big(\frac{\kappa^2}{C^2}\tilde{\sigma}^2 - a^2\Big) - 2N\lambda\Big(\frac{\kappa^2}{C^2}\tilde{\sigma}^2 - 2a\Big)$$

and computing its corresponding matrix $\mathbf{A}$, we can determine an orthogonal matrix $\mathbf{P}$ with the orthonormal eigenvectors of $\mathbf{A}$ in its columns. Then, by a change of variables from N, λ to x, y according to

$$\begin{pmatrix} N \\ \lambda \end{pmatrix} = \mathbf{P} \begin{pmatrix} x \\ y \end{pmatrix},$$

we obtain that the constraint can be written as

$$\beta_1 x + \beta_2 y + \alpha_1 x^2 + \alpha_2 y^2 \leq 1.$$

We observe that we get the same result as in the previous sections. The only differences are the values of the constants, but, in any case, the way to proceed is similar. Without writing all the calculations needed to compute these values, we can now solve the optimization problem as we did in Section 4.2, and we obtain that the optimal solution is

$$\begin{aligned} N^* = & \frac{\kappa^4\tilde{\sigma}^2\big(2\sigma^2+b\big) - aC^2\kappa^2\big(6\tilde{\sigma}^2 - a(b+\sigma^2)\big) + 4a^3C^4}{2\kappa^2\tilde{\sigma}^2\big(a^2C^2-\kappa^2\tilde{\sigma}^2\big)} \\ & + \frac{\big(3\kappa^2\tilde{\sigma}^2 - 2a^2C^2\big)\sqrt{s}}{2\kappa^2\tilde{\sigma}^2\big(a^2C^2-\kappa^2\tilde{\sigma}^2\big)} \end{aligned}$$

and

$$\lambda^* = \frac{4aC^2 - \kappa^2\big(\sigma^2+b\big) - 2\sqrt{s}}{2\kappa^2\tilde{\sigma}^2},$$

where we use s as short notation for

$$\begin{aligned} s = & \frac{16a^4C^6 - 8\big(a^3(\sigma^2+b) + 2a^2\tilde{\sigma}^2\big)\kappa^2C^4 - \big((\sigma^2+b)^2+\sigma^4\big)\tilde{\sigma}^2\kappa^6}{4a^2C^2 - 5\kappa^2\tilde{\sigma}^2} \\ & + \frac{\big(a^2(\sigma^2+b)^2 + 4a(2b+3\sigma^2)\tilde{\sigma}^2 - 4\tilde{\sigma}^4\big)\kappa^4C^2}{4a^2C^2-5\kappa^2\tilde{\sigma}^2}. \end{aligned}$$

Chapter 5

Capital allocation

An insurance company has a total target return and spreads down the total return to different portfolios when establishing a business plan. The company is also interested in comparing the individual returns $r_i = E[R_i]/C_i$ of the various business units $i \in \{1, \ldots, n\}$.

The vital question concerning this is: how can one choose C_i? The idea is to introduce a capital allocation principle and to split up the risk capital amongst the various business units, but there is no general answer to the question of how the risk capital should be allocated. There are different classes of capital allocation methodologies, and a formal description can be found in Albrecht (1997).

In the following sections, we consider two different capital allocation principles, namely the covariance principle and the expected shortfall principle.

5.1 Covariance principle

We now want to calculate the well-known covariance principle for the special case of $\mathrm{RAC}(R) = -E[R] + \kappa\sigma(R)$ with $\kappa > 0$ and with R having the following form

$$R = \sum_{i=1}^{n} R_i = \sum_{i=1}^{n} \left(\nu_i N_i - \sum_{j=1}^{N_i} X_{i,j} - Y_i N_i \right)$$

where

- $N_1, \ldots, N_n$ are positive integers,
- $\{X_{i,1}, X_{i,2}, \ldots, X_{i,N_i}\}$ are finite sequences of i.i.d. random variables with finite means and finite variances for all $i \in \{1, \ldots, n\}$ with

$$\mu_i = E[X_{i,j}] \quad \text{and} \quad \sigma_i^2 = \mathrm{Var}(X_{i,j}),$$

- the linear correlation coefficients satisfy
$$\rho_{ik} = \operatorname{corr}(X_{i,j}, X_{k,l})$$
for all $i, k \in \{1, \dots, n\}$, $i \neq k$ and $j \in \{1, \dots, N_i\}$, $l \in \{1, \dots, N_k\}$,
- $Y_1, \dots, Y_n$ are random variables with finite means and finite variances
$$\tilde{\mu}_i = E[Y_i] \quad \text{and} \quad \tilde{\sigma}_i^2 = \operatorname{Var}(Y_i),$$
- the sequences $\{X_{i,1}, \dots, X_{i,N_i}\}$, for every $i \in \{1, \dots, n\}$, are independent of the random variables $Y_1, \dots, Y_n$.

Remark. There are restrictions on the correlation coefficients ρ_{ik} with i, k in $\{1, \dots, n\}$, $i \neq k$. The covariance matrix of the $(N_1 + \dots + N_n)$-dimensional random vector
$$(X_{1,1}, \dots, X_{1,N_1}, X_{2,1}, \dots, X_{2,N_2}, \dots, X_{n,1}, \dots, X_{n,N_n})$$
has to be positive semidefinite. If this is the case, then there exist at least multivariate normally distributed random variables with the prescribed dependence structure.

In general it holds that
$$C_i = C\, \frac{\operatorname{Cov}(R_i, R)}{\operatorname{Var}(R)},$$
where C denotes the capital the company wants to invest and C_i is the capital the company will allocate to business unit i.

Let us compute the enumerator and the denominator separately. On one side, we have that, for the enumerator, it holds that
$$\begin{aligned}
\operatorname{Cov}(R_i, R) &= \operatorname{Cov}\Big(R_i, \sum_{k=1}^{n} R_k\Big) \\
&= \operatorname{Var}(R_i) + \operatorname{Cov}\Big(R_i, \sum_{\substack{k=1 \\ k\neq i}}^{n} R_k\Big) \\
&= \operatorname{Var}\Big(\nu_i N_i - \sum_{j=1}^{N_i} X_{i,j} - Y_i N_i\Big) + \sum_{\substack{k=1 \\ k\neq i}}^{n} \operatorname{Cov}\big(R_i, R_k\big) \\
&= N_i \sigma_i^2 + N_i^2 \tilde{\sigma}_i^2 + \sum_{\substack{k=1 \\ k\neq i}}^{n} N_i N_k \; \big(\rho_{ik}\, \sigma_i\, \sigma_k + \tilde{\rho}_{ik}\, \tilde{\sigma}_i\, \tilde{\sigma}_k\big),
\end{aligned}$$

since, for any $k \neq i$,

$$\begin{aligned}\operatorname{Cov}\big(R_i, R_k\big) &= \operatorname{Cov}\Big(\nu_i N_i - \sum_{j=1}^{N_i} X_{i,j} - Y_i N_i,\ \nu_k N_k - \sum_{l=1}^{N_k} X_{k,l} - Y_k N_k\Big) \\ &= \operatorname{Cov}\Big(\sum_{j=1}^{N_i} X_{i,j} + Y_i N_i,\ \sum_{l=1}^{N_k} X_{k,l} + Y_k N_k\Big) \\ &= \sum_{j=1}^{N_i}\sum_{l=1}^{N_k} \underbrace{\operatorname{Cov}\big(X_{i,j},\ X_{k,l}\big)}_{=\rho_{ik}\,\sigma_i\,\sigma_k} + N_i N_k \underbrace{\operatorname{Cov}\big(Y_i,\ Y_k\big)}_{=\tilde\rho_{ik}\,\tilde\sigma_i\,\tilde\sigma_k},\end{aligned}$$

where $\tilde\rho_{ik}$ denotes the linear correlation coefficient, in particular for i, k in $\{1, \ldots, n\}$ with $i \neq k$,

$$\tilde\rho_{ik} = \operatorname{corr}\big(Y_i,\ Y_k\big).$$

On the other side, for the denominator, it holds that

$$\begin{aligned}\operatorname{Var}(R) &= \operatorname{Var}\Big(\sum_{j=1}^{n} R_j\Big) \\ &= \sum_{j=1}^{n} \operatorname{Var}\big(R_j\big) + 2\sum_{j=1}^{n-1}\sum_{k=j+1}^{n} \operatorname{Cov}\big(R_j, R_k\big) \\ &= \sum_{j=1}^{n} \big(N_j\sigma_j^2 + N_j^2\tilde\sigma_j^2\big) + 2\sum_{j=1}^{n-1}\sum_{k=j+1}^{n} N_j N_k\ \big(\rho_{jk}\,\sigma_j\,\sigma_k + \tilde\rho_{jk}\,\tilde\sigma_j\,\tilde\sigma_k\big).\end{aligned}$$

Therefore, it follows that

$$C_i = C\ \frac{N_i\sigma_i^2 + N_i^2\tilde\sigma_i^2 + \sum_{k\neq i}^{n} N_i N_k\ \big(\rho_{ik}\,\sigma_i\,\sigma_k + \tilde\rho_{ik}\,\tilde\sigma_i\,\tilde\sigma_k\big)}{\sum_{j=1}^{n}\big(N_j\sigma_j^2 + N_j^2\tilde\sigma_j^2\big) + 2\ \sum_{j=1}^{n-1}\sum_{k=j+1}^{n} N_j N_k\ \big(\rho_{jk}\,\sigma_j\,\sigma_k + \tilde\rho_{jk}\,\tilde\sigma_j\,\tilde\sigma_k\big)}.$$

Remark. In the case where $\{X_{i,j}\}_{j\in\mathbb{N}}$ are independent sequences consisting of i.i.d. random variables and $Y_1, \ldots, Y_n$ are independent too, it holds that all correlation coefficients are zero. Therefore, it follows that

$$C_i = C\ \frac{N_i\sigma_i^2 + N_i^2\tilde\sigma_i^2}{\sum_{j=1}^{n}\big(N_j\sigma_j^2 + N_j^2\tilde\sigma_j^2\big)}.$$

We now consider the second variant of the initial model. In fact, we assume that all N_i, with $i \in \{1, \ldots, n\}$, which denote the number of contracts for every unit i, are Poisson-distributed random variables with parameters

$\lambda_i > 0$. To recapitulate, we will calculate the covariance principle for R having the following form

$$R = \sum_{i=1}^{n} R_i = \sum_{i=1}^{n} \left(\nu_i N_i - \sum_{j=1}^{N_i} X_{i,j} - Y_i N_i \right),$$

where

- $N_1, \ldots, N_n$ are Poisson-distributed random variables, i.e.,

$$N_i \sim \mathrm{POIS}(\lambda_i) \quad \text{with } \lambda_i > 0,$$

- $\{X_{i,j}\}_{j \in \mathbb{N}}$ are uncorrelated sequences of i.i.d. random variables with finite means and finite variances for all $i \in \{1, \ldots, n\}$

$$\mu_i = E[X_{i,j}] \text{ and } \sigma_i^2 = \mathrm{Var}(X_{i,j}),$$

- $Y_1, \ldots, Y_n$ are random variables with finite means and finite variances

$$\tilde{\mu}_i = E[Y_i] \text{ and } \tilde{\sigma}_i^2 = \mathrm{Var}(Y_i),$$

- all the sequences $\{X_{i,j}\}_{j \in \mathbb{N}}$ with $i \in \{1, \ldots, n\}$ are independent of the random variables $Y_1, \ldots, Y_n$,

- $N_1, \ldots, N_n$ are assumed to be independent of all the sequences $\{X_{i,j}\}_{j \in \mathbb{N}}$ with $i \in \{1, \ldots, n\}$ and of $Y_1, \ldots, Y_n$.

Remark. Consider sequences $\{U_i\}_{i \in \mathbb{N}}$ and $\{V_i\}_{i \in \mathbb{N}}$ of i.i.d. random variables. For simplicity, assume that $E[U_i] = E[V_i] = 0$ and $\mathrm{Var}(U_i) = \mathrm{Var}(V_i) = 1$ for all $i \in \mathbb{N}$. If the correlation $\rho = \mathrm{Cov}(U_i, V_j)$ does not depend on $i, j \in \mathbb{N}$, then ρ has to be zero. To see this, fix $n \in \mathbb{N}$. Then

$$\begin{aligned} 0 &\le \mathrm{Var}\left(\sum_{i=1}^{n} (U_i + V_i) \right) \\ &= \sum_{i=1}^{n} \big(\mathrm{Var}(U_i) + \mathrm{Var}(V_i) \big) + 2 \sum_{i,j=1}^{n} \mathrm{Cov}(U_i, V_j) \\ &= 2n + 2n^2 \rho \end{aligned}$$

implies $\rho \ge -1/n$. Similarly,

$$0 \le \mathrm{Var}\left(\sum_{i=1}^{n} (U_i - V_i) \right) = 2n - 2n^2 \rho,$$

hence $\rho \le 1/n$. Since $n \in \mathbb{N}$ was arbitrary, $\rho = 0$. For this reason, the above sequences $\{X_{i,j}\}_{j \in \mathbb{N}}$ for $i \in \{1, \ldots, n\}$ are assumed to be uncorrelated.

In the above model it holds that

$$C_i = C\,\frac{\operatorname{Cov}(R_i, R)}{\operatorname{Var}(R)},$$

with

$$\begin{aligned}\operatorname{Cov}(R_i, R) &= \operatorname{Var}(R_i) + \operatorname{Cov}\Big(R_i, \sum_{\substack{k=1\\k\neq i}}^{n} R_k\Big)\\ &= \lambda_i^2\tilde{\sigma}_i^2 + \lambda_i\big((\nu_i - \mu_i - \tilde{\mu}_i)^2 + \sigma_i^2 + \tilde{\sigma}_i^2 + \mu_i^2\big) + \sum_{\substack{k=1\\k\neq i}}^{n} \operatorname{Cov}(R_i, R_k)\end{aligned}$$

and

$$\operatorname{Var}(R) = \sum_{i=1}^{n} \operatorname{Var}(R_i) + 2\sum_{i=1}^{n-1}\sum_{k=i+1}^{n} \operatorname{Cov}(R_i, R_k)\,.$$

In detail, we first calculate the covariance between R_i and R_k for $i \neq k$ separately:

$$\begin{aligned}\operatorname{Cov}(R_i, R_k) &= \operatorname{Cov}\Big(\nu_i N_i - \sum_{j=1}^{N_i} X_{i,j} - Y_i N_i, \nu_k N_k - \sum_{l=1}^{N_k} X_{k,l} - Y_k N_k\Big)\\ &= \operatorname{Cov}(\nu_i N_i,\ \nu_k N_k) - \operatorname{Cov}\Big(\nu_i N_i, \sum_{l=1}^{N_k} X_{k,l}\Big) - \operatorname{Cov}(\nu_i N_i, Y_k N_k)\\ &\quad - \operatorname{Cov}\Big(\sum_{j=1}^{N_i} X_{i,j}, \nu_k N_k\Big) + \operatorname{Cov}\Big(\sum_{j=1}^{N_i} X_{i,j}, \sum_{l=1}^{N_k} X_{k,l}\Big)\\ &\quad + \operatorname{Cov}\Big(\sum_{j=1}^{N_i} X_{i,j}, Y_k N_k\Big) - \operatorname{Cov}(Y_i N_i, \nu_k N_k)\\ &\quad + \operatorname{Cov}\Big(Y_i N_i, \sum_{l=1}^{N_k} X_{k,l}\Big) + \operatorname{Cov}(Y_i N_i, Y_k N_k)\,,\end{aligned}$$

with

$$\operatorname{Cov}(\nu_i N_i, \nu_k N_k) = \nu_i\,\nu_k\,\hat{\rho}_{ik}\,\sqrt{\lambda_i\,\lambda_k}\,,$$

$$
\begin{aligned}
\operatorname{Cov}\Big(\nu_i N_i, \sum_{l=1}^{N_k} X_{k,l}\Big) &= \nu_i \operatorname{Cov}\Big(N_i, \sum_{l=1}^{N_k} X_{k,l}\Big) \\
&= \nu_i \left(E\Big[N_i \sum_{l=1}^{N_k} X_{k,l}\Big] - E\big[N_i\big] E\Big[\sum_{l=1}^{N_k} X_{k,l}\Big] \right) \\
&= \nu_i \left(E\big[N_i\, N_k\, \mu_k\big] - \lambda_i\, \lambda_k\, \mu_k \right) \\
&= \nu_i\, \mu_k \operatorname{Cov}(N_i, N_k) = \nu_i\, \mu_k\, \hat{\rho}_{ik} \sqrt{\lambda_i\, \lambda_k}\,,
\end{aligned}
$$

$$
\begin{aligned}
\operatorname{Cov}(\nu_i N_i, Y_k N_k) &= \nu_i \Big(E\big[N_i\, Y_k\, N_k\big] - E\big[N_i\big] E\big[Y_k\, N_k\big] \Big) \\
&= \nu_i\, \tilde{\mu}_k \operatorname{Cov}(N_i, N_k) = \nu_i\, \tilde{\mu}_k\, \hat{\rho}_{ik} \sqrt{\lambda_i\, \lambda_k}\,,
\end{aligned}
$$

$$
\begin{aligned}
\operatorname{Cov}\Big(\sum_{j=1}^{N_i} X_{i,j}, \nu_k N_k\Big) &= \nu_k \left(E\Big[N_k \sum_{j=1}^{N_i} X_{i,j}\Big] - E\big[N_k\big] E\Big[\sum_{j=1}^{N_i} X_{i,j}\Big] \right) \\
&= \nu_k\, \mu_i\, \hat{\rho}_{ik} \sqrt{\lambda_i\, \lambda_k}\,,
\end{aligned}
$$

$$
\begin{aligned}
\operatorname{Cov}\Big(\sum_{j=1}^{N_i} X_{i,j}, \sum_{l=1}^{N_k} X_{k,l}\Big) &= E\Big[\sum_{j=1}^{N_i} X_{i,j} \sum_{l=1}^{N_k} X_{k,l}\Big] - E\Big[\sum_{j=1}^{N_i} X_{i,j}\Big] E\Big[\sum_{l=1}^{N_k} X_{k,l}\Big] \\
&= E\big[N_i\, N_k\big] \mu_i \mu_k - \mu_i\, \mu_k\, \lambda_i\, \lambda_k \\
&= \mu_i\, \mu_k \operatorname{Cov}\big(N_i, N_k\big) \\
&= \hat{\rho}_{ik} \sqrt{\lambda_i\, \lambda_k}\, \mu_i\, \mu_k,
\end{aligned}
$$

$$
\begin{aligned}
\operatorname{Cov}\Big(\sum_{j=1}^{N_i} X_{i,j}, Y_k N_k\Big) &= E\big[N_i\, N_k\, \mu_i\, \tilde{\mu}_k\big] - \lambda_i\, \mu_i\, \lambda_k\, \tilde{\mu}_k \\
&= \mu_i\, \tilde{\mu}_k \operatorname{Cov}(N_i, N_k) = \mu_i\, \tilde{\mu}_k\, \hat{\rho}_{ik} \sqrt{\lambda_i\, \lambda_k}\,,
\end{aligned}
$$

$$
\begin{aligned}
\operatorname{Cov}(Y_i N_i, \nu_k N_k) &= \nu_k \Big(E\big[Y_i\, N_i\, N_k\big] - E\big[Y_i\, N_i\big] E\big[N_k\big] \Big) \\
&= \nu_k\, \tilde{\mu}_i \operatorname{Cov}(N_i, N_k) = \nu_k\, \tilde{\mu}_i\, \hat{\rho}_{ik} \sqrt{\lambda_i\, \lambda_k}\,,
\end{aligned}
$$

$$
\begin{aligned}
\operatorname{Cov}\Big(Y_i N_i, \sum_{l=1}^{N_k} X_{k,l}\Big) &= E\big[N_i\, N_k\, \tilde{\mu}_i\, \mu_k\big] - \lambda_i\, \tilde{\mu}_i\, \lambda_k\, \mu_k \\
&= \tilde{\mu}_i\, \mu_k \operatorname{Cov}(N_i, N_k) = \tilde{\mu}_i\, \mu_k\, \hat{\rho}_{ik} \sqrt{\lambda_i\, \lambda_k}\,,
\end{aligned}
$$

$$\begin{aligned}\mathrm{Cov}(Y_i N_i, Y_k N_k) &= E\big[Y_i\, N_i\, Y_k\, N_k\big] - \tilde{\mu}_i\, \tilde{\mu}_k\, \lambda_i\, \lambda_k \\ &= E\big[Y_i\, Y_k\big]\, E\big[N_i\, N_k\big] - \tilde{\mu}_i\, \tilde{\mu}_k\, \lambda_i\, \lambda_k \\ &= \mathrm{Cov}\big(Y_i, Y_k\big)\, \mathrm{Cov}\big(N_i, N_k\big) + \tilde{\mu}_i\, \tilde{\mu}_k\, \mathrm{Cov}\big(N_i, N_k\big) \\ &\qquad + \lambda_i\, \lambda_k\, \mathrm{Cov}\big(Y_i, Y_k\big) \\ &= \tilde{\rho}_{ik}\, \tilde{\sigma}_i\, \tilde{\sigma}_k\, \hat{\rho}_{ik}\, \sqrt{\lambda_i\, \lambda_k} + \hat{\rho}_{ik}\, \sqrt{\lambda_i\, \lambda_k}\, \tilde{\mu}_i\, \tilde{\mu}_k \\ &\qquad + \tilde{\rho}_{ik}\, \tilde{\sigma}_i\, \tilde{\sigma}_k\, \lambda_i\, \lambda_k\,,\end{aligned}$$

where with $\hat{\rho}_{ik}$ and $\tilde{\rho}_{ik}$ we denote the linear correlation coefficients between the dependent random variables, i.e.,

$$\hat{\rho}_{ik} = \mathrm{corr}\big(N_i, N_k\big) \quad \text{and} \quad \tilde{\rho}_{ik} = \mathrm{corr}\big(Y_i, Y_k\big),$$

respectively. It follows that

$$\begin{aligned}\mathrm{Cov}(R_i, R_k) = \hat{\rho}_{ik}\sqrt{\lambda_i \lambda_k}\Big(\big(\nu_i - \mu_i - \tilde{\mu}_i\big)\big(\nu_k - \mu_k - \tilde{\mu}_k\big) + \tilde{\rho}_{ik}\, \tilde{\sigma}_i\, \tilde{\sigma}_k\Big) \\ + \lambda_i\, \lambda_k\, \tilde{\rho}_{ik}\, \tilde{\sigma}_i\, \tilde{\sigma}_k.\end{aligned}$$

A similar calculation with $i = k$ for the nine terms given above gives similar results with $\tilde{\rho}_{ik}\sqrt{\lambda_i \lambda_k}$ replaced by λ_i and $\tilde{\rho}_{ik}\tilde{\sigma}_i\tilde{\sigma}_k$ replaced by $\tilde{\sigma}_i^2$. However, for the fifth term we get

$$\begin{aligned}\mathrm{Var}\Bigg(\sum_{j=1}^{N_i} X_{i,j}\Bigg) &= E\Bigg[\sum_{\substack{j,l=1 \\ j \neq l}}^{N_i} X_{i,j} X_{i,l}\Bigg] + E\Bigg[\sum_{j=1}^{N_i} X_{i,j}^2\Bigg] - \Bigg(E\Bigg[\sum_{j=1}^{N_1} X_{i,j}\Bigg]\Bigg)^2 \\ &= \mu_i^2\, E[N_i(N_i - 1)] + E[N_i]\, E[X_{i,1}^2] - (E[N_i]\, \mu_i)^2 \\ &= \lambda_i^2 \mu_i^2 + \lambda_i\, E[X_{i,1}^2] - \lambda_i^2 \mu_i^2 \\ &= \lambda_i \sigma_i^2 + \lambda_i \mu_i^2,\end{aligned}$$

hence

$$\mathrm{Var}(R_i) = \lambda_i^2 \tilde{\sigma}_i^2 + \lambda_i\big((\nu_i - \mu_i - \tilde{\mu}_i)^2 + \sigma_i^2 + \tilde{\sigma}_i^2 + \mu_i^2\big).$$

Therefore, we can conclude that

$$\begin{aligned}\mathrm{Cov}(R_i, R) = \lambda_i^2 \tilde{\sigma}_i^2 + \lambda_i\big((\nu_i - \mu_i - \tilde{\mu}_i)^2 + \sigma_i^2 + \tilde{\sigma}_i^2 + \mu_i^2\big) \\ + \sum_{\substack{k=1 \\ k \neq i}}^{n} \Bigg(\hat{\rho}_{ik}\sqrt{\lambda_i \lambda_k}\Big(\big(\nu_i - \mu_i - \tilde{\mu}_i\big)\big(\nu_k - \mu_k - \tilde{\mu}_k\big) + \tilde{\rho}_{ik}\, \tilde{\sigma}_i\, \tilde{\sigma}_k\Big) \\ + \lambda_i\, \lambda_k\, \tilde{\rho}_{ik}\, \tilde{\sigma}_i\, \tilde{\sigma}_k\Bigg)\,,\end{aligned}$$

and

$$\begin{aligned}
\operatorname{Var}(R) &= \sum_{i=1}^{n} \operatorname{Var}(R_i) + 2 \sum_{i=1}^{n-1} \sum_{k=i+1}^{n} \operatorname{Cov}(R_i, R_k) \\
&= \sum_{i=1}^{n} \left(\lambda_i^2 \tilde{\sigma}_i^2 + \lambda_i \big((\nu_i - \mu_i - \tilde{\mu}_i)^2 + \sigma_i^2 + \tilde{\sigma}_i^2 + \mu_i^2 \big) \right) + \\
&\quad + 2 \sum_{i=1}^{n-1} \sum_{k=i+1}^{n} \Bigg(\lambda_i \, \lambda_k \, \tilde{\rho}_{ik} \, \tilde{\sigma}_i \, \tilde{\sigma}_k \\
&\quad + \hat{\rho}_{ik} \sqrt{\lambda_i \lambda_k} \Big(\big(\nu_i - \mu_i - \tilde{\mu}_i\big) \big(\nu_k - \mu_k - \tilde{\mu}_k\big) + \tilde{\rho}_{ik} \, \tilde{\sigma}_i \, \tilde{\sigma}_k \Big) \Bigg) .
\end{aligned}$$

Remark. Note that where $N_1, \ldots, N_n$ are independent random variables, all the $\{X_{i,j}\}_{j\in\mathbb{N}}$ with $i \in \{1, \ldots, n\}$ are uncorrelated sequences and the random variables $Y_1, \ldots Y_n$ are also independent, it holds that all correlation coefficients are zero. Thus, it follows that

$$C_i = C \, \frac{\lambda_i^2 \tilde{\sigma}_i^2 + \lambda_i \big((\nu_i - \mu_i - \tilde{\mu}_i)^2 + \sigma_i^2 + \tilde{\sigma}_i^2 + \mu_i^2 \big)}{\sum_{j=1}^{n} \left(\lambda_j^2 \tilde{\sigma}_j^2 + \lambda_j \big((\nu_j - \mu_j - \tilde{\mu}_j)^2 + \sigma_j^2 + \tilde{\sigma}_j^2 + \mu_j^2 \big) \right)} .$$

5.2 Expected-shortfall principle

An alternative to the covariance principle for the purpose of allocation of risk capital are the conditional expectations.

Let $R_1, \ldots, R_n$ be the stochastic gains of the business units and $R = \sum_{i=1}^{n} R_i$ the whole profit and loss of an insurance company. Let $c \leq 0$ be the capital loss threshold, for example, the α-quantile r_α of R. If a total loss $R \leq c$ occurs, we consider the expected shares $E[-R_i \,|\, R \leq c]$ of the single losses with respect to the total loss. Obviously, it holds that

$$E[-R \,|\, R \leq c] = \sum_{i=1}^{n} E[-R_i \,|\, R \leq c] ,$$

where $E[-R \,|\, R \leq c]$ denotes the risk capital of the entire company, while $E[-R_i \,|\, R \leq c]$ denotes the risk capital assigned to business unit i.

This risk capital allocation principle is additive, as is the covariance principle, but in addition it can be applied for integrable random variable, so the existence of second moments is no longer required. In order to compute these conditional expectations we can use the procedures proposed in Appendix A.

As an example, we consider R having the usual form for a company consisting of two units, i.e.

$$R = R_1 + R_2 = \Big(\nu_1 N_1 - \sum_{j=1}^{N_1} X_{1,j} - Y_1 N_1\Big) + \Big(\nu_2 N_2 - \sum_{j=1}^{N_2} X_{2,j} - Y_2 N_2\Big),$$

with the following assumptions:

- N_1 and N_2 are positive integers,
- the sequences $\{X_{1,j}\}_{j\in\mathbb{N}}$ and $\{X_{2,j}\}_{j\in\mathbb{N}}$ are independent and consist of independent, identically normally distributed random variables, i.e.,

$$X_{i,j} \sim \mathcal{N}\big(\mu_i, \sigma_i^2\big), \qquad \text{for} \quad i = 1, 2 \quad \text{and} \quad j \in \mathbb{N},$$

- Y_1 and Y_2 are independent normally distributed random variables, i.e.,

$$Y_i \sim \mathcal{N}\big(\tilde{\mu}_i, \tilde{\sigma}_i^2\big), \qquad \text{for} \quad i = 1, 2,$$

- the sequences $\{X_{1,j}\}_{j\in\mathbb{N}}$ and $\{X_{2,j}\}_{j\in\mathbb{N}}$ are independent of Y_1 and Y_2.

It is known that, given the independent random variables $X_{i,j}$ that are normally distributed for $i = 1, 2$ and all $j \in \mathbb{N}$, the random variables R_1, R_2 and R are also normally distributed, i.e.,

$$R_1 \sim \mathcal{N}\big(\hat{\mu}_1, \hat{\sigma}_1^2\big), \quad R_2 \sim \mathcal{N}\big(\hat{\mu}_2, \hat{\sigma}_2^2\big) \quad \text{and} \quad R \sim \mathcal{N}\big(\mu_R, \sigma_R^2\big),$$

where we define

$$\hat{\mu}_i = N_i\left(\nu_i - \mu_i - \tilde{\mu}_i\right), \quad \hat{\sigma}_i^2 = N_i\,\sigma_i^2 + N_i^2\,\tilde{\sigma}_i^2, \quad \text{for } i = 1, 2,$$

and

$$\mu_R = \hat{\mu}_1 + \hat{\mu}_2, \quad \sigma_R^2 = \hat{\sigma}_1^2 + \hat{\sigma}_2^2.$$

We denote the density of the standard normal distribution by

$$\varphi(u) = \frac{1}{\sqrt{2\pi}}\, e^{-\frac{1}{2}u^2}, \quad u \in \mathbb{R},$$

and the standard normal distribution function by

$$\Phi(t) = \int_{-\infty}^{t} \varphi(u)\, du, \quad t \in \mathbb{R}.$$

Note that

$$P\big(R \le t\big) = \Phi\Big(\frac{t-\mu_R}{\sigma_R}\Big).$$

We are now interested in the measures of risk

$$E\big[R_i \,|\, R \le c\big]$$

for a given value $c \in \mathbb{R}$ and $i = 1, 2$. Using formula (A6) from Appendix A, we obtain

$$E[R_i \,|\, R \le c] = \hat{\mu}_i - \frac{\hat{\sigma}_i^2}{\sigma_R}(\log \Phi)'\bigg(\frac{c-\mu_R}{\sigma_R}\bigg).$$

Appendix A

Calculating expected shortfall

In the following pages we will introduce some rules which will be seen to be useful to calculate the expected shortfall.

Let X be an integrable random variable with distribution function F_X. Then, for all $c \in \mathbb{R}$ such that $F_X(c) < 1$, it holds that

$$E\big[X \,|\, X > c\big] = \frac{1}{1 - F_X(c)} \int_c^\infty x \; F_X(dx). \tag{A.1}$$

Let $X_1, \dots, X_n$ be exchangeable[1] integrable random variables and let $X = \sum_{i=1}^n X_i$. Then, for all $c \in \mathbb{R}$ such that $P(X > c) > 0$ and for all $i, j \in \{1, \dots, n\}$, it holds that

$$E\big[X_i \,|\, X > c\big] = E\big[X_j \,|\, X > c\big].$$

Since the conditional expectation is linear, for all $i \in \{1, \dots, n\}$, it follows that

$$E\big[X_i \,|\, X > c\big] = \frac{1}{n} E\big[X \,|\, X > c\big]$$

and this last conditional expectation can be solved by means of (A.1).

Let X and Y be independent integrable random variables with distribution functions F_X and F_Y, respectively, and let $c \in \mathbb{R}$ such that $P(X + Y > c) > 0$. It then holds that

$$E\big[X \,|\, X + Y > c\big] = \frac{1}{1 - \big(F_X * F_Y\big)(c)} \int_{\mathbb{R}} x \; \big(1 - F_Y(c - x)\big) F_X(dx) \; , \tag{A.2}$$

[1] Let $\mathcal{I}$ be a countable set. A sequence $\big(X_i, i \in \mathcal{I}\big)$ of random variables on a probability space $\big(\Omega, \mathcal{F}, P\big)$ is exchangeable if, for every permutation ρ of $\mathcal{I}$, the distributions of $\big(X_{\rho(i)}, i \in \mathcal{I}\big)$ and $\big(X_i, i \in \mathcal{I}\big)$ are identical. Note that a finite or infinite i.i.d. sequence is exchangeable.

where $F_X * F_Y$ denotes the convolution of the distribution functions F_X and F_Y, that is, the distribution function of the sum $X+Y$, i.e.,

$$P(X+Y>c) = 1 - \big(F_X * F_Y\big)(c)\,.$$

This result is generalizable to independent random variables $X_1, \dots, X_n$ with distribution functions $F_1, \dots, F_n$, respectively. Let $X = X_1$ and $Y = X_2 + \cdots + X_n$. X and Y are likewise independent with distribution functions $F_X = F_1$ and $F_Y = F_2 * \cdots * F_n$, respectively. The conditional expectation $E\big[X_1 \,|\, X_1 + \cdots + X_n > c\big]$ can be then solved by means of (A.2).

Let

$$\varphi(t) = \frac{1}{\sqrt{2\pi}} e^{-t^2/2}\,, \qquad t \in \mathbb{R}\,,$$

and

$$\Phi(x) = \int_{-\infty}^{x} \varphi(t)\,dt\,, \qquad x \in \mathbb{R}\,,$$

denote the density and the distribution function of the standard normal distribution, respectively.

Given $X \sim \mathcal{N}(0,1)$ and $c \in \mathbb{R}$, we obtain with the substitution $z = x^2/2$

$$\begin{aligned} E[X \,|\, X \le c] &= \frac{1}{\Phi(c)} \int_{-\infty}^{c} x\varphi(x)\,dx \\ &= -\frac{1}{\sqrt{2\pi}\,\Phi(c)} \int_{c^2/2}^{\infty} e^{-z} dz \\ &= -\frac{e^{-c^2/2}}{\sqrt{2\pi}\,\Phi(c)} \\ &= -\frac{\varphi(c)}{\Phi(c)} = -\big(\log\Phi\big)'(c). \end{aligned}$$

If $Y \sim \mathcal{N}(\mu, \sigma^2)$, then

$$X := \frac{Y-\mu}{\sigma} \sim \mathcal{N}(0,1)$$

and with the above result we obtain

$$\begin{aligned} E[Y \,|\, Y \le c] &= E[\mu + \sigma X \,|\, \mu + \sigma X \le c] \\ &= \mu + \sigma E[X \,|\, X \le (c-\mu)/\sigma] \\ &= \mu - \sigma\big(\log\Phi\big)'\Big(\frac{c-\mu}{\sigma}\Big). \end{aligned} \tag{A.3}$$

In particular, for $c = 0$,

$$\frac{E[Y]}{E[-Y\,|\,Y \le 0]} = \frac{\mu}{-\mu + \sigma\big(\log\Phi\big)'\big(\frac{-\mu}{\sigma}\big)} = \frac{1}{-1 + \frac{\sigma}{\mu}\big(\log\Phi\big)'\big(\frac{-\mu}{\sigma}\big)}\,.$$

Now consider independent random variables $X, Y \sim \mathcal{N}(0,1)$ and constants $c \in \mathbb{R}$, $\sigma > 0$. To compute $E[X\,|\,\sigma X + Y \le c]$, define

$$\gamma = \frac{c}{\sqrt{1+\sigma^2}}$$

and note that

$$P(\sigma X + Y \le c) = \Phi(\gamma)$$

because $(\sigma X + Y)/\sqrt{1+\sigma^2} \sim \mathcal{N}(0,1)$. By conditioning on the σ-algebra generated by X,

$$E[X\,|\,\sigma X + Y \le c] = \frac{1}{\Phi(\gamma)} E\big[E[X 1_{\{\sigma X + Y \le c\}}\,|\,X]\big] = \frac{1}{\Phi(\gamma)} E[X P(Y \le c - \sigma X\,|\,X)]\,.$$

Since X and Y are independent,

$$P(Y \le c - \sigma X\,|\,X) = \Phi(c - \sigma X) \quad P\text{-almost surely.}$$

Since $x\varphi(x) = -\varphi'(x)$, partial integration gives

$$E[X\Phi(c - \sigma X)] = \int_{\mathbb{R}} x\varphi(x)\Phi(c - \sigma x)\,dx = -\sigma \int_{\mathbb{R}} \varphi(x)\varphi(c - \sigma x)\,dx\,.$$

Since

$$x^2 + (c - \sigma x)^2 = (1+\sigma^2)x^2 - 2c\sigma x + c^2 = \big(\sqrt{1+\sigma^2}\,x - \sigma\gamma\big)^2 + \gamma^2\,,$$

the substitution $u = \sqrt{1+\sigma^2}x - \sigma\gamma$ yields

$$\int_{\mathbb{R}} \varphi(x)\varphi(c - \sigma x)\,dx = \varphi(\gamma) \int_{\mathbb{R}} \varphi\big(\sqrt{1+\sigma^2}\,x - \sigma\gamma\big)\,dx = \frac{\varphi(\gamma)}{\sqrt{1+\sigma^2}} \int_{\mathbb{R}} \varphi(u)\,du = \frac{\varphi(\gamma)}{\sqrt{1+\sigma^2}}\,.$$

Therefore,

$$\begin{aligned} E[X \mid \sigma X + Y \leq c] &= -\frac{\sigma}{\sqrt{1+\sigma^2}} \frac{\varphi(\gamma)}{\Phi(\gamma)} \\ &= -\frac{\sigma}{\sqrt{1+\sigma^2}} (\log \Phi)' \left(\frac{c}{\sqrt{1+\sigma^2}} \right). \end{aligned} \tag{A5}$$

More generally, for independent $X \sim \mathcal{N}(\mu_X, \sigma_X^2)$ and $Y \sim \mathcal{N}(\mu_Y, \sigma_Y^2)$ with $\sigma_X > 0$ and $\sigma_Y > 0$, we obtain

$$\begin{aligned} &E[X \mid X + Y \leq c] \\ &\qquad = \mu_X + \sigma_X E\left[\frac{X-\mu_X}{\sigma_X} \,\middle|\, \frac{\sigma_X}{\sigma_Y} \frac{X-\mu_X}{\sigma_X} + \frac{Y-\mu_Y}{\sigma_Y} \leq \frac{c-\mu_X-\mu_Y}{\sigma_Y} \right]. \end{aligned}$$

Since

$$\frac{X-\mu_X}{\sigma_X},\ \frac{Y-\mu_Y}{\sigma_Y} \sim \mathcal{N}(0,1)\,,$$

equation (A5) gives

$$\begin{aligned} &E[X \mid X + Y \leq c] \\ &\quad = \mu_X - \sigma_X \frac{\sigma_X/\sigma_Y}{\sqrt{1+(\sigma_X/\sigma_Y)^2}} (\log \Phi)' \left(\frac{c-\mu_X-\mu_Y}{\sigma_Y \sqrt{1+(\sigma_X/\sigma_Y)^2}} \right) \\ &\quad = \mu_X - \frac{\sigma_X^2}{\sqrt{\sigma_X^2+\sigma_Y^2}} (\log \Phi)' \left(\frac{c-\mu_X-\mu_Y}{\sqrt{\sigma_X^2+\sigma_Y^2}} \right). \end{aligned} \tag{A6}$$

Appendix B

The Lagrange multipliers rule

We consider the general case. Let m and n be natural numbers and U be an open subset of $\mathbb{R}^{n+m}$. Suppose $x \in \mathbb{R}^n$, $y \in \mathbb{R}^m$, $f : U \to \mathbb{R}$ and $g : U \to \mathbb{R}^m$. Let us examine the problem of finding the extrema of $f(x, y) = f(x_1, \ldots, x_n; y_1, \ldots, y_m)$ subject to the m constraints:

$$\begin{aligned} g_1(x_1, \ldots, x_n; y_1, \ldots, y_m) &= 0\,, \\ \vdots \qquad\qquad & \vdots \\ g_m(x_1, \ldots, x_n; y_1, \ldots, y_m) &= 0\,. \end{aligned}$$

Theorem. Let m, n and U be defined as above.
Suppose that $f \in C^1(U, \mathbb{R})$ and $g \in C^1(U, \mathbb{R}^m)$. Moreover, assume that $(\xi, \eta) \in M = \{(x, y) \in U \,|\, g(x, y) = 0\}$ is such that the restriction $f|_M$ of f to M has a local extremum at (ξ, η); this means that there is some neighborhood V of (ξ, η) such that either $f(x, y) \le f(\xi, \eta)$ for all $(x, y) \in V \cap M$ or $f(x, y) \ge f(\xi, \eta)$ for all $(x, y) \in V \cap M$. Suppose also that the n by n matrix $\partial g(\xi, \eta)/\partial y$ has nonzero determinant. Then, a vector $\lambda = (\lambda_1, \ldots, \lambda_m) \in \mathbb{R}^m$ exists such that the function

$$\begin{aligned} H(x, y, z) &= f(x, y) + \langle z, g(x, y)\rangle \\ &= f(x, y) + \sum_{j=1}^{m} z_j g_j(x, y)\,, \quad (x, y, z) \in U \times \mathbb{R}^m\,, \end{aligned}$$

has a critical point in (ξ, η, λ).

Proof. See Walter (1990), pp. 130–133. □

As consequence, in the point (ξ, η, λ) it holds that

$$\begin{aligned} H_{x_i} &= f_{x_i} + \langle \lambda, g_{x_i} \rangle = 0\,, & i &= 1, \ldots, n\,, \\ H_{y_k} &= f_{y_k} + \langle \lambda, g_{y_k} \rangle = 0\,, & k &= 1, \ldots, m\,, \\ H_{\lambda_k} &= g_k = 0\,, & k &= 1, \ldots, m\,. \end{aligned} \tag{B.1}$$

The variables $\lambda_1, \ldots, \lambda_n$ are called Lagrange multipliers. The virtue of this result is that if one seeks points (ξ, η) at which $f|_M$ has local extrema, then one need only search among those $(\xi, \eta) \in M$ for which the system (B.1) has a solution λ.

Bibliography

Albrecht, P., (1997). *Risk Based Capital Allocation and Risk Adjusted Performance Management in Property/Liability Insurance: A Risk Theoretical Framework.* Joint Day Proocedings of the XXVIIIth International AFIR Colloquium & the 7th International AFIR Colloquium. Cairns, Australia, 13 August 1997.

Arztner, P., Delbaen, F., Eber, J. M., Heath, D., (1998). *Coherent measure of risk.* Mathematical Finance, Volume 9/3, July 1999.

Billingsley, P. (1968). *Convergence of Probability Measures.* John Wiley & Sohns, New York.

Billingsley, P. (1979). *Probability and Measures.* John Wiley & Sohns, Inc. New York.

Breiman, L. (1968). *Probability.* Addison-Wesley Publishing Company.

Crouhy, M., Turnbull, S. M., Wakeman, L. M. (1999). Measuring risk-adjusted performance. *Journal of Risk.* Volume 2/Number 1, Fall 1999.

Denault, M. (1999). *Coherent allocation of risk capital.* RiskLab Research Paper. Available under http://www.risklab.ch/Papers.html .

Embrechts, P., Küppelberg, C. and Mikosch, T.(1999). *Modelling Extremal Events for Insurance and Finance.* Springer

Feller, W., (1966). *An Introduction to Probability Theory and its Applications.* Volume II. John Wiley & Sohns, Inc.

Fristedt, B., Gray, L. (1997). *A Modern Approach to Probability Theory.* Birkhäuser, Boston.

Gilbert, J. and Gilbert, L. (1995). *Linear Algebra and Matrix Theory.* Academic Press, Inc.

Grimmet, G. R. and Stirzaker, D. R. (1992). *Probability and random processes.* Oxford Science Publications, Oxford.

Johnson, N. L., Kotz, S., Balakrishnan, N. (1994). *Continuous univariate Distribution.* Volume 1, 2^{nd} Edition. John Wiley & Sohn, Inc..

Karlin, S., Taylor, H. M. (1975). *A first course in stochastic processes.* 2^{nd} Edition, Academic Press, Inc. New York.

Schmock, U. and Straumann, D. *Allocation of risk capital and performance measurement.* Working paper, private communication.

Stromberg, K. R. (1981). *An introduction to classical real analysis.* Wadsworth International Group, Belmont, California.

Tasche, D. (1999). Risk contributions and performance measurement. Research paper, Zentrum Mathematik (SCA), TU München.

Walter, W. (1990). *Analysis II.* Springer-Verlag.