UvA-DARE (Digital Academic Repository)

Comonotonicity: from risk measurement to risk management

Vanduffel, S.

Publication date

2005

Document Version

Final published version

[Link to publication](#)

Citation for published version (APA):

Vanduffel, S. (2005). *Comonotonicity: from risk measurement to risk management*. [Thesis, externally prepared, Universiteit van Amsterdam].

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

Comonotonicity:

From Risk Measurement to Risk Management

Steven Vanduffel

Comonotonicity:

From Risk Measurement to Risk Management

Comonotonicity: From Risk Measurement to Risk Management

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor

aan de Universiteit van Amsterdam

op gezag van de Rector Magnificus

prof. mr. P.F. van der Heijden

ten overstaan van een door het college voor promoties ingestelde
commissie, in het openbaar te verdedigen in de Aula der Universiteit

op donderdag 27 januari 2005, te 12.00 uur

door

Steven Vanduffel

geboren te Genk (België)

Promotores:

Prof. dr. J.L.M. Dhaene

Prof. dr. M.J. Goovaerts

Leden van de commissie:

Prof. dr. M. Denuit

Ir. L. Henrard

Prof. dr. F.C.J.M. de Jong

Prof. dr. R. Kaas

Prof. dr. C. Van Hulle

Faculteit der Economische Wetenschappen en Econometrie

Voor Sonja, Line en Julie

Contents

Contents	i
Acknowledgement	v
Introduction	vii
 Comonotonicity and Risk Measurement	 1
1 Ordering of risks	3
1.1 Introduction	3
1.2 Actuarial ordering concepts	4
1.3 Inverse distribution functions	8
 2 Comonotonicity	 11
2.1 Introduction	11
2.2 Comonotonic sets and random vectors	12
2.3 The comonotonic sum	15
2.4 Comonotonicity and linear correlation	21
2.5 Convex bounds for $\sum_{i=1}^n X_i$	23
2.6 Example	25
 3 Risk measures	 29
3.1 Introduction	29
3.2 Some well-known risk measures	30
3.3 Risk measures and ordering of risks	34
3.3.1 General results	34
3.3.2 Risk measures and comonotonicity	36
3.4 Distortion risk measures	40
3.4.1 Definition, examples and properties	40
3.4.2 Concave distortion risk measures	44
3.4.3 Risk measures for comonotonic sums	48
3.5 Theories of choice under risk	52
3.6 Final remarks	55

4	Comparing approximations for sums of lognormals	57
4.1	Introduction	57
4.2	Comonotonic approximations	59
4.2.1	General results	59
4.2.2	The ‘maximal variance’ lower bound approach	60
4.3	Well-known moment matching approximations	61
4.3.1	The Reciprocal Gamma approximation	61
4.3.2	The lognormal approximation	63
4.4	Comparing the approximations	63
4.5	Final remarks	67
5	Constant continuous annuities	69
5.1	Introduction	69
5.2	Closed-form comonotonic approximations	72
5.2.1	Upper bound approach	72
5.2.2	Lower bound approaches	73
5.3	Application on perpetuities	76
5.4	Final remarks	77
	Applications of Comonotonicity in Risk Management	79
6	Optimal portfolio selection problems	81
6.1	Introduction	81
6.2	Stochastic return processes	84
6.2.1	The Black & Scholes setting	84
6.2.2	Constant mix strategies	86
6.2.3	Markowitz mean-variance analysis	88
6.3	Saving and terminal wealth	91
6.3.1	General problem description	91
6.3.2	The case of a single investment	93
6.3.3	Comonotonic approximations for the general problem	97
6.3.4	Determining the investment strategy that maximizes the target capital for a given probability level	99
6.3.5	Determining the investment strategy that maximizes the probability level for a given target capital	104
6.4	Reserves for future obligations	106
6.4.1	General problem description	106
6.4.2	The case of a single obligation	108
6.4.3	Comonotonic approximations for the general case	109
6.4.4	Determining the investment strategy that minimizes the p -quantile initial reserve for a given probability level	111
6.4.5	Determining the investment strategy that minimizes the CTE_p -quantile initial reserve for a given probability level	114
6.4.6	Determining the investment strategy that maximizes the probability level for a given initial reserve	117

6.5	Final remarks	118
7	The hurdle-race problem	121
7.1	Introduction	121
7.2	Determining the provision	122
7.3	Approximations for $\overline{R}_0$	125
7.4	Numerical illustrations	127
7.4.1	Example 1	127
7.4.2	Example 2	128
8	The savings-retirement problem	133
8.1	Introduction	133
8.2	Problem description	134
8.3	Approximations for the quantiles of W_{n+m}	137
8.4	Numerical illustration	140
	Bibliography	143
	Samenvatting	149

Acknowledgement

The last 5 years of my professional career have been a hell of a life but... also a lot of fun. I want to start by thanking my 2 advisors, Jan Dhaene and Marc Goovaerts. They gave me the opportunity to combine research activities with other professional ambitions. Jan has an unquenchable curiosity and love for science. Moreover, he has the rare talent to transform embryonic and sometimes very vague ideas into relatively easy-to-read papers. He was always there for me, supported me in difficult days and challenged me when it was necessary. Without him, this thesis would not have been possible. Marc played a crucial role in bringing Actuarial Science in the University of Amsterdam and K.U. Leuven to a worldwide top level. I can only confirm that these research centres are an ideal environment for every young actuarial researcher. I am also grateful to Rob Kaas for giving me the opportunity to complete this thesis at the University of Amsterdam. Many teachers and professors influenced my thinking and the choices I made, in particular J. Moors, F. Dumortier and M. Fannes. I also really want to thank my parents. They were, and still are, always there for me in case I needed them. When I was a student, they gave me the space I needed, even though it was not always clear for them what I was doing. A lot of sympathy goes to all my friends, especially Geert for carefully reading this manuscript, my colleagues and several Fortis people who supported me while doing this thesis. Finally, all my love goes to my wife Sonja and the other joy in my life: my daughters Line and Julie. Sonja allowed me to be a bad husband and a lousy father. In difficult times, when I was down, she encouraged me to go on. Her enthusiasm and the happiness of Line and Julie were of great help to me.

Steven Vanduffel, 30 June 2004

“Dum vivimus, vivamus” - Anonymous.

Introduction

All information about a stand-alone random variable X is contained in its cumulative distribution function F_X . It is therefore natural that any decision-taking regarding X will center on the evaluation of F_X and it is essentially the task of any actuary to compute or to estimate this function. In case of an insurance company, the random variable X might represent the risk related to a particular policy, a specific line-of-business or the entire insurance portfolio over a specified time horizon. The risk X can also appear in a personal finance context, when a decision-maker targets a given series of future consumptions or when he faces, for instance, the risk that his fixed contribution pension plan does not meet his retirement financial goal.

The actuary will use mathematical and statistical techniques to derive the cumulative distribution function F_X , but in many financial and actuarial situations he will encounter random variables of the type $S = \sum_{i=1}^n X_i$, for which he knows the marginal cumulative distribution functions of the random variables X_i , but the dependency structure of the random vector $\underline{X} = (X_1, X_2, \dots, X_n)$ is not specified, or too difficult to be fruitfully exploitable.

From a mathematical point of view it is often convenient to assume that the random variables X_i are mutually independent, because powerful and accurate computation methods such as Panjer's recursion and the technique of convolution can then be applied. In this case, one can also take advantage of the Central Limit Theorem which states that S will be approximately normally distributed if the number of risks is sufficiently high. In fact, the mere existence of Insurance is based on the assumption of mutual independency between the insured risks, and sometimes this complies, approximately, with reality. In the majority of cases however, the different risks will be interrelated to a certain extent. Hence, the cumulative distribution function of S cannot be easily specified and this is highly jeopardizing the whole decision-taking process. Such a sum S of dependent risks occurs for instance, when considering the aggregate claims amount of a non-life insurance portfolio because the insured risks are subject to some common factors such as geography, climate or economical environment.

In this thesis, we will essentially deal with sums S of dependent lognormals:

$$S = \sum_{i=1}^n \alpha_i e^{Z_i}. \quad (1)$$

Here, the α_i are real numbers and $(Z_1, Z_2, \dots, Z_n)$ is a multivariate normal distributed random vector. The sum (1) plays a central role in the actuarial and financial theory. This is because the accumulated value at time n of a series of future deterministic saving amounts α_i can be written in the form (1), where Z_i denotes the random accumulation factor over the period $[i, n]$. Also the present value of a series of future deterministic payments α_i can be written in the form (1), where now Z_i denotes the random discount factor over the period $[0, i]$.

Hence, the valuation of Asian or basket options in a Black & Scholes model and the setting of provisions and required capitals in an insurance context, boils down to the evaluation of *risk measures* related to the cumulative distribution function of a random variable S as defined in (1), and this underlines the importance of the random variable S .

We define a risk measure as a mapping from the set of random variables, representing the risks at hand, to the real line. Risk measures can be a helpful tool for decision-making since they resume the information available about the random variable S into one single number $\rho[S]$. Common risk measures in Actuarial Science are premium principles, but in this work we will mainly concentrate on risk measures that can be used for reserving and solvency purposes.

The random variable S defined as in (1) will be, at least in case the Z_i represent stochastic discounting or accumulation factors, a sum of strongly dependent lognormal random variables. This is because there is a natural overlapping process when discounting (or compounding) over the different time periods. Unfortunately, the cumulative distribution function and most risk measures of S cannot be determined analytically. In such cases, to be able to make decisions, it may be helpful to find the dependency structure for the random vector $(Z_1, Z_2, \dots, Z_n)$ that entails a less favorable or ‘larger’ sum S for the given marginals. This is because this will lead to more prudent and conservative decision-taking. One has then to decide upon what ‘larger’ means, i.e. one has to define how we can rank risks.

We will essentially rely on convex order. This ordering concept represents, in both the expected utility theory and the competing Yaari’s dual theory of choice under risk, the common preferences of risk averse decision-takers who have to choose between risks with the same expectation. The ‘largest’ sum will occur when the random variables $\alpha_i e^{Z_i}$ are comonotonic, meaning that these are monotonic functions of a common random variable, which also explains the word comonotonic (common monotonic). The comonotonic upper bound, denoted by S^c , completely avoids simulation as analytical expressions for several important actuarial quantities such as quantiles and stop-loss premiums are readily available for this sum. We will also show that it often makes sense to replace S by the more favorable lower bound S^l , that is obtained by computing the conditional expectation $E[S \mid \Lambda] = \sum_{i=0}^n \alpha_i E[e^{Z_i} \mid \Lambda]$ with respect to a specific random variable Λ . In case the $\alpha_i E[e^{Z_i} \mid \Lambda]$ are monotonic functions of Λ , we obtain, just like in case of the comonotonic upper bound, a comonotonic sum and many risk measures can be easily obtained. It looks counter-intuitive for the risk averse actuary to use S^l instead of S^c , but numerical comparisons

revealed that the risk measures of S^l can, statistically speaking, almost not be distinguished from the risk measures of the random variable S , obtained by simulation. Depending on the judgement of the actuary, this might outweigh the fact that the lower bound S^l in principle leads to more risk-taking.

In literature, the assumption of normal distributed investment returns is often made and is supported by empirical evidence, at least if one considers weekly (and longer period) returns. Therefore, as the time unit that we consider in our practical applications is long (typically 1 year), assuming a Gaussian model for the $(Z_1, Z_2, \dots, Z_n)$ seems to be appropriate.

We are ready now to provide the outline of this thesis.

In the first part of this thesis we relate the concept of comonotonicity to the measurement of risk. First, in Chapter 1, we summarize the most important actuarial ordering concepts and the relations that exist between them. Next, in Chapter 2, we recall the concept of comonotonicity and we also show how the comonotonic upper bounds for sums of random variables can be derived. We will also demonstrate through the technique of conditioning, how one can also obtain a convex lower bound. Chapter 3 deals with (distribution based) risk measures and shows that, for comonotonic sums, many of them can be obtained rather easily by summing corresponding risk measures for each of the terms in the sum. Chapter 4 investigates, for a sum of dependent lognormals like in (1), the accuracy of the comonotonic lower and upper bound approximations and two popular moment matching approximations for the risk measures of (1), by comparing these approximations with the estimates obtained from extensive Monte Carlo simulations. The chapter concludes that especially the ‘maximal variance’ lower bound approximation provides an excellent fit. In Chapter 5, which ends the first part of this thesis, we derive closed-form expressions for the upper bound and lower bound approximations for several risk measures of constant continuous annuities, which can be seen to be a continuous equivalent of the discrete sum (1).

In the second part of this thesis, we investigate some important problems in Finance and Actuarial Science that involve random variables, defined as in (1).

In Chapter 6, which is the heart of this thesis, we use comonotonic approximations to solve multi-period portfolio selection problems in a Black & Scholes lognormal setting. Strategic portfolio selection is the process used to identify the best allocation of wealth among a basket of securities for an investor with a given consumption/saving behavior over a given investment horizon. The basket of available securities will typically be a selection of risky assets such as stocks, bonds and real estate, and risk-free components such as cash and money market instruments. The individual investor or the asset manager chooses an initial asset mix and a particular tactical trading strategy within a given set of strategies, according to which he will buy and sell risky and risk-free assets, during the whole time period under consideration. We will assume that the

investor has to choose the optimal investment strategy for a given consumption or savings pattern, within the class of constant mix strategies. We will consider two general types of problems, which will be referred to as the terminal wealth problem and the reserving problem respectively.

In the *terminal wealth problem*, the decision-maker will invest a given series of positive saving amounts $\alpha_0, \alpha_1, \dots, \alpha_{n-1}$ at predetermined times $0, 1, \dots, n-1$ such that his terminal wealth at time n will reach or exceed some target capital K with a sufficiently large probability. As terminal wealth is a sum of dependent lognormal random variables, its cumulative distribution function cannot be determined exactly and is too cumbersome to work with. Therefore, we will use the comonotonic upper and lower bound approximations. The approximations that we propose have the tremendous advantage that for any given investment strategy, they provide an accurate and easy to compute approximation for any risk measure that is additive for comonotonic risks, such as distortion risk measures (VaR and TailVaR for instance). The comonotonic approximations reduce the multivariate randomness of the multiperiod problem to a univariate randomness. The optimal investment mix could be defined as the one that requires the smallest constant amount α that has to be invested from period to period in order to reach a final wealth of at least K with a probability of at least $1 - \epsilon$. Or, one could define the optimal mix as the one that maximizes the probability of reaching terminal wealth of at least K for a given investment of α per period.

The proposed methodology can be used to solve several personal finance problems. An example of such a problem is what one could call the ‘saving for retirement problem’. In this case, one wants to retire in n years with a ‘nest egg’ of K . How much does one have to save monthly in order to assure a $(1 - \epsilon)$ chance to reach the retirement financial goal? Clearly the answer will depend on the investment mix. The theory on comonotonicity gives a quick, elegant and accurate answer to this question.

In the *reserving problem*, which is in some sense dual to the final wealth problem, the decision-maker targets a series of future consumptions $\alpha_1, \alpha_2, \dots, \alpha_n$ at times $1, 2, \dots, n$. For this type of investment problems, the optimal investment mix could be defined as the one that leads to the largest survival probability, given the initial reserve. Or, one could fix the required survival probability and determine the optimal investment strategy as the one that minimizes the required initial reserve.

In Chapter 7 we tackle the *hurdle-race problem*. We will present another way of determining the initial provision, which does not only take into account the goal of ‘reaching the finish’, like in the *reserving problem*, but also the conditions that year-to-year the available provision R_j is larger than a predefined value V_j . These additional requirements V_j are the ‘hurdles’ that have to be taken with probabilities $1 - \varepsilon_1, 1 - \varepsilon_2, \dots, 1 - \varepsilon_n$. In case of an insurer establishing his reserve, they might be imposed by a supervisory authority or by internal policy. Determining the initial provision in this way allows one to make the probability of taking the hurdles time-dependent. In situations where year-to-year adjustments of the level of the reserve are possible, the required probabilities for taking

the hurdles in the first years could be chosen larger than the probabilities for the later years. This can lead to situations in which the binding constraint is not ‘reaching the finish’ but ‘taking an intermediate hurdle’ and our method will reflect this in the required level for the initial reserve.

The previous chapters and the various numerical illustrations herein, will reveal that, especially if one uses the ‘maximal variance’ lower bound, one obtains in general a very accurate and easy to compute approximate solution for the problem at hand. We remark that these problems typically involve a cash-flow pattern in which all α_i have the same sign. This property is crucial for making the computations straightforward. Indeed, in case of changing signs of the α_i , it is not possible to find a relevant conditioning random variable Λ such that $E[S \mid \Lambda] = \sum_{i=0}^n \alpha_i E[e^{Z_i} \mid \Lambda]$ is a sum of non-decreasing functions of Λ . This implies that, for any good choice of Λ , the distortion risk measures related to $E[S \mid \Lambda]$ cannot be obtained by simply summing the corresponding risk measures of the individual terms in the sum, as is the case when all α_i are positive.

In Chapter 8, that concludes this thesis, we show how one can still use the comonotonic lower bound in order to determine accurate and easy computable approximations for a sum S as defined in (1), in case one first has positive cash-flows (savings) followed by negative ones (withdrawals). An important situation in which we encounter this particular cash-flow pattern and, hence, where this result can be applied, is the ‘savings-retirement problem’. Take as an example the so-called 20/65/95 pension plan: in a defined contribution pension scheme, a person of age 20 intends to save money for 45 consecutive years (until retirement). After his retirement, he wants to withdraw money from this pension account on a regular basis and this for a period of 30 years. Assume that his yearly savings are constant and equal to α , while his yearly withdrawal rate is constant and equal to 1. A relevant question that we can answer in an elegant way is: “For each Euro of pension income, what is the minimal yearly savings effort α such that this person, with a probability of at least $(1 - \varepsilon)$, will be able to meet his consumption pattern during the 30 year withdrawal period?”

In Finance and Actuarial Science, one often encounters random variables of the type $S = \sum_{i=1}^n X_i$, for which most common risk measures cannot be computed because the multivariate structure is not known or too difficult to work with. Comonotonicity essentially reduces this multivariate randomness to a univariate randomness. In this thesis we show how this concept enables to derive approximate results for common risk measures for S . Next, we demonstrate how the ‘comonotonic approach’, in a Black & Scholes setting, allows us to derive accurate and easy to compute solutions for several important risk management problems.

Comonotonicity and Risk Measurement

Ordering of risks

1.1 Introduction

¹All information about a stand-alone random variable (r.v.) X is contained in its cumulative distribution function (c.d.f.) F_X . It is therefore natural that any decision-taking regarding X will center on the evaluation of F_X and it is essentially the task of any actuary to compute or to estimate this function. The actuary will use mathematical and statistical techniques to do so, but in many financial and actuarial situations he will encounter random variables of the type $S = \sum_{i=1}^n X_i$, for which he knows the marginal distribution functions of the random variables X_i , but the dependency structure of the random vector $\underline{X} = (X_1, X_2, \dots, X_n)$ is not specified or too cumbersome to work with. Consequently the cumulative distribution function of S cannot be specified exactly and this is of course jeopardizing the whole decision-taking process. Such sums S occur for instance when considering the aggregate claims amount of a non-life insurance portfolio or when computing the stochastic present value of future payments for a life policy or portfolio.

In such cases, to be able to make decisions, it may be helpful to find the dependency structure for the random vector $(X_1, \dots, X_n)$ that entails a less favorable or ‘larger’ sum S for the given marginals. This is because this will lead to more prudent and conservative decision-taking. One has then to decide upon what ‘larger’ means, i.e. one has to define how we can rank risks.

We will essentially use the expected utility theory framework to describe how decision makers rank random variables. In the utility paradigm, every decision-taker asserts a utility $u(w)$ to each possible wealth-level w , see von Neumann & Morgenstern (1947). This real-valued function $u(\cdot)$ is called the utility function. When such a decision-taker, with initial wealth w , has to choose between random losses X and Y , he compares the expected utilities $E[u(w - X)]$ with $E[u(w - Y)]$, provided these exist, for some specific utility

¹“All of life is the management of risk, not its elimination” - Walter Wriston (former chairman of Citicorp).

function u and chooses the loss which gives rise to the highest expected utility. As a rational decision maker is likely to prefer more to less, it is generally accepted that a utility function is non-decreasing. Also, it makes sense to accept the hypothesis that every reasonable decision-taker prefers a sure loss above a random loss with the same mean. By virtue of Jensen's inequality this means that the utility function is concave. Decision-takers with an increasing concave utility function will be called risk averse. In this important class of decision-takers the random loss X is preferred above the random loss Y if and only if $E[u(-X)] \geq E[u(-Y)]$ for all concave increasing utility functions u . It can be shown that the common preferences of all risk averse decision-takers can also be reflected by the well-known actuarial concept of stop-loss order and hence this hints a parallel between the 'economic' utility framework for comparing risks and the 'actuarial' ordering concepts. The next sections are essentially based on the paper of Dhaene, Denuit, Goovaerts, Kaas & Vyncke (2002a).

1.2 Actuarial ordering concepts

In the sequel we will consider random variables X with finite mean and will denote the cumulative distribution function (c.d.f.) of X by $F_X(x) = \Pr[X \leq x]$. Using the technique of integration by parts on both terms of the right-hand side in $E[X] = \int_{-\infty}^0 x dF_X(x) - \int_0^{\infty} x d(1 - F_X(x))$, one immediately finds the following expression for the mean $E[X]$:

$$E[X] = - \int_{-\infty}^0 F_X(x) dx + \int_0^{\infty} (1 - F_X(x)) dx. \quad (1.1)$$

The stop-loss premium with retention d of the random variable X is defined by $E[(X - d)_+]$, with the notation $(x - d)_+ = \max(x - d, 0)$. Again by using partial integration we obtain:

$$E[(X - d)_+] = \int_d^{\infty} (1 - F_X(x)) dx, \quad -\infty < d < +\infty, \quad (1.2)$$

from which we see that the stop-loss premium with retention d can be considered as the weight of an upper tail of (the cumulative distribution function of) X : it is the surface between the c.d.f. F_X of X and the constant function 1, from d on. Also useful is the observation that $E[(X - d)_+]$ is a decreasing continuous function of d , with derivative $F_X(d) - 1$ at d , which vanishes when d reaches infinity.

We are able now to define stochastic and stop-loss order between random variables.

Definition 1.2.1 (Stochastic order) *Consider two random variables X and Y . Then X is said to precede Y in the stochastic order, notation $X \leq_{st} Y$, if and only if the c.d.f. of X always exceeds that of Y :*

$$F_X(x) \geq F_Y(x), \quad -\infty < x < +\infty. \quad (1.3)$$

Definition 1.2.2 (Stop-loss order) *Consider two random variables X and Y . Then X is said to precede Y in the stop-loss order sense, notation $X \leq_{sl} Y$, if and only if X has lower stop-loss premiums than Y :*

$$E[(X - d)_+] \leq E[(Y - d)_+], \quad -\infty < d < +\infty. \quad (1.4)$$

It follows from (1.1) that a stochastic order between two random variables X and Y implies a corresponding ordering of the means and from (1.2) we also have that stochastic order induces stop-loss order. Both ordering concepts also have a natural interpretation in terms of utility theory. Indeed, it can be shown easily that $X \leq_{st} Y$ if and only if $E[u(-X)] \geq E[u(-Y)]$ holds for all non-decreasing real functions u for which the expectations exist. This exactly means that stochastic order reflects the preferences of all decision-makers with an increasing utility function.

On the other hand, stop-loss order reflects the common preferences of all risk averse decision-makers. To see this, note that every non-decreasing convex function can be obtained as a uniform limit of a sequence of piecewise linear functions each of them written as a linear combination of functions $(x - t)_+$. Hence, it is not difficult to show that $X \leq_{sl} Y$ if and only if $E[v(X)] \leq E[v(Y)]$ holds for all non-decreasing convex functions v for which the expectations exist. Consequently we find that $X \leq_{sl} Y$ if and only if $E[u(-X)] \geq E[u(-Y)]$ holds for all non-decreasing concave real functions u for which the expectations exist. For more details and properties of stop-loss order see Shaked & Shanthikumar (1994), Kaas, van Heerwaarden & Goovaerts (1994) or Kaas, Goovaerts, Dhaene & Denuit (2001).

Like stochastic order, stop-loss order between random variables X and Y implies a corresponding ordering of their means. To prove this, assume that $d < 0$. From the expression (1.2) of stop-loss premiums as upper tails, we immediately find the following equality:

$$d + E[(X - d)_+] = - \int_d^0 F_X(x) dx + \int_0^\infty (1 - F_X(x)) dx, \quad (1.5)$$

and also, letting $d \rightarrow -\infty$,

$$\lim_{d \rightarrow -\infty} (d + E[(X - d)_+]) = E[X].$$

Hence, adding d to both members of the inequality in Definition 1.2.2 and taking the limit for $d \rightarrow -\infty$, we obtain that $E[X] \leq E[Y]$.

A sufficient condition for $X \leq_{sl} Y$ to hold is that $E[X] \leq E[Y]$, together with the condition that their c.d.f.'s only cross once, i.e. there exists a real number c such that $F_X(x) \leq F_Y(x)$ for $x < c$, but $F_X(x) \geq F_Y(x)$ for $x \geq c$. Indeed, considering the function $f(d) = E[(Y - d)_+] - E[(X - d)_+]$, we have that $\lim_{d \rightarrow -\infty} f(d) = E[Y] - E[X] \geq 0$, and $\lim_{d \rightarrow +\infty} f(d) = 0$. Further, $f(d)$ first increases, and then decreases (from c on) but remains non-negative.

Recall from the introduction that our original problem was to replace a random payment X by a less favorable random payment Y , for which the c.d.f.

is easier to obtain. For the class of risk averse decision-takers we know that if $X \leq_{sl} Y$, then also $E[X] \leq E[Y]$ and it is intuitively clear that the best approximations arise in the borderline case where $E[X] = E[Y]$. This leads to the so-called convex order.

Definition 1.2.3 (Convex order) Consider two random variables X and Y . Then X is said to precede Y in the convex order sense, notation $X \leq_{cx} Y$, if and only if

$$\begin{aligned} E[X] &= E[Y], \\ E[(X - d)_+] &\leq E[(Y - d)_+], \quad -\infty < d < +\infty. \end{aligned} \quad (1.6)$$

From $E[(X - d)_+] - E[(d - X)_+] = E[X] - d$, we find that convex order can also be characterized by

$$X \leq_{cx} Y \Leftrightarrow \begin{cases} E[X] = E[Y], \\ E[(d - X)_+] \leq E[(d - Y)_+], \quad -\infty < d < +\infty. \end{cases} \quad (1.7)$$

Using integration by parts we obtain

$$E[(d - X)_+] = \int_{-\infty}^d F_X(x) dx, \quad (1.8)$$

which means that $E[(d - X)_+]$ can be interpreted as the weight of a lower tail of X : it is the surface between the constant function 0 and the c.d.f. of X , from $-\infty$ to d . We have seen that stop-loss order entails uniformly heavier upper tails. The additional condition of equal means implies that convex order also leads to uniformly heavier lower tails.

Let $d > 0$. From (1.8) we find

$$d - E[(d - X)_+] = - \int_{-\infty}^0 F_X(x) dx + \int_0^d (1 - F_X(x)) dx,$$

and also

$$\lim_{d \rightarrow +\infty} (d - E[(d - X)_+]) = E[X]. \quad (1.9)$$

Now, (1.5) and (1.9) together with (1.6) and (1.7) imply another characterization for convex order:

$$X \leq_{cx} Y \Leftrightarrow \begin{cases} E[(X - d)_+] \leq E[(Y - d)_+], & -\infty < d < +\infty, \\ E[(d - X)_+] \leq E[(d - Y)_+], & -\infty < d < +\infty. \end{cases} \quad (1.10)$$

Consequently we have that $X \leq_{cx} Y \Leftrightarrow X \leq_{sl} Y$ and $-X \leq_{sl} -Y$. We also find that $X \leq_{cx} Y$ is equivalent to $-X \leq_{cx} -Y$. Hence, the interpretation of the random variables as payments or as incomes is irrelevant for the convex order. Note that with stop-loss order, we are concerned with large values of a random loss and call the random variable Y less attractive than X if the expected values

of all top parts $(Y - d)_+$ are larger than those of X . Negative values for these random variables are actually gains. With stability in mind, excessive gains might also be unattractive for the decision-maker, for instance for tax reasons. In this situation, X could be considered to be more attractive than Y if both the top parts $(X - d)_+$ and the bottom parts $(d - X)_+$ have a lower expected value than for Y . Both conditions just define the convex order introduced above.

A sufficient condition for $X \leq_{cx} Y$ or $(Y \leq_{cx} X)$ to hold is that $E[X] = E[Y]$, together with the condition that their c.d.f.'s only cross once. It can also be proven that $X \leq_{cx} Y$ if and only if $E[v(X)] \leq E[v(Y)]$ for all convex functions v , provided the expectations exist. This explains the name 'convex order'. Note that when characterizing stop-loss order, the convex functions v are additionally required to be non-decreasing. Hence, stop-loss order (also called increasing convex order) is weaker: more pairs of random variables are ordered. We also find that $X \leq_{cx} Y$ if and only if $E[X] = E[Y]$ and $E[u(-X)] \geq E[u(-Y)]$ for all non-decreasing concave functions u , provided the expectations exist. Hence, in a utility context, convex order represents the common preferences of all risk averse decision-makers regarding random variables with equal mean.

As the function v defined by $v(x) = x^2$ is a convex function, it follows immediately that $X \leq_{cx} Y$ implies $Var[X] \leq Var[Y]$. The reverse implication does not hold in general.

Comparing variances is meaningful when comparing stop-loss premiums of convex ordered random variables, see, e.g. Kaas, van Heerwaarden & Goovaerts (1994, p. 68). The following relation links variances and stop-loss premiums:

$$\frac{1}{2} Var[X] = \int_{-\infty}^{\infty} (E[(X - t)_+] - (E[X] - t)_+) dt. \quad (1.11)$$

To prove this relation, write

$$\int_{-\infty}^{\infty} (E[(X - t)_+] - (E[X] - t)_+) dt = \int_{-\infty}^{E[X]} E[(t - X)_+] dt + \int_{E[X]}^{\infty} E[(X - t)_+] dt.$$

Interchanging the order of the integrations and using integration by parts, one finds

$$\int_{-\infty}^{E[X]} E[(t - X)_+] dt = \int_{-\infty}^{E[X]} \int_{-\infty}^t F_X(x) dx dt = \frac{1}{2} \int_{-\infty}^{E[X]} (x - E[X])^2 dF_X(x).$$

Similarly,

$$\int_{E[X]}^{\infty} E[(X - t)_+] dt = \frac{1}{2} \int_{E[X]}^{\infty} (x - E[X])^2 dF_X(x).$$

This proves (1.11). We deduce from this that if $X \leq_{cx} Y$,

$$\int_{-\infty}^{\infty} |E[(Y - t)_+] - E[(X - t)_+]| dt = \frac{1}{2} \{Var[Y] - Var[X]\} \quad (1.12)$$

Thus, if $X \leq_{cx} Y$, their stop-loss distance, i.e. the integrated absolute difference of their respective stop-loss premiums, equals half the variance difference between these two random variables. As the integrand in (1.12) is non-negative, we find that if $X \leq_{cx} Y$ whilst $Var[X] = Var[Y]$, then this means that X and Y must have equal stop-loss premiums and hence the same c.d.f. . We also find that if $X \leq_{cx} Y$, and X and Y are not equal in distribution, then $Var[X] < Var[Y]$ must hold.

This makes clear that if we want to replace X by a less favorable random payment Y in the sense of convex order, the best approximation is obtained when the variance of Y is ‘as close as possible’ to the variance of X . Hence, a measure for the global goodness-of-fit of such a Y can be defined as $z = \frac{Var(X)}{Var(Y)}$. Note however that a value of z that is close to one does not necessarily mean that for every $t > 0$, the stop-loss premium $E[(Y - t)_+]$ will be close to $E[(X - t)_+]$.

Further in this work we will show that it often makes sense to replace X by a more favorable random payment Z and not by a less favorable Y . This looks counter-intuitive for the risk averse actuary, but there are many situations where the goodness-of-fit of the more favorable Z is closer to one than in case one uses the less favorable Y and depending on the judgement of the actuary this might outweigh the fact that replacing X by the less convex Z leads in principle to more risk-taking.

Remark that (1.11) and (1.12) have been derived under the additional conditions that both $\lim_{x \rightarrow \infty} x^2 (1 - F_X(x))$ and $\lim_{x \rightarrow -\infty} x^2 F_X(x)$ are equal to 0 (and similar for Y). A sufficient condition for these requirements to hold is that X and Y have finite second moments.

1.3 Inverse distribution functions

The c.d.f. $F_X(x) = P[X \leq x]$ of a random variable X is a right-continuous (abbreviated as r.c.) non-decreasing function with

$$F_X(-\infty) = \lim_{x \rightarrow -\infty} F_X(x) = 0, \quad F_X(+\infty) = \lim_{x \rightarrow +\infty} F_X(x) = 1.$$

The classical definition of the inverse of a distribution function is the non-decreasing and left-continuous (l.c.) function $F_X^{-1}(p)$ defined by

$$F_X^{-1}(p) = \inf \{x \in \mathbb{R} \mid F_X(x) \geq p\}, \quad p \in [0, 1],$$

with $\inf \emptyset = +\infty$ by convention. For all $x \in \mathbb{R}$ and $p \in [0, 1]$, we have

$$F_X^{-1}(p) \leq x \Leftrightarrow p \leq F_X(x). \quad (1.13)$$

In this work, we will also use a slightly more sophisticated definition for inverses of distribution functions, see also Dhaene, Denuit, Goovaerts, Kaas & Vyncke (2002a). For any real $p \in [0, 1]$, a possible choice for the inverse of F_X in p is *any* point in the closed interval

$$[\inf \{x \in \mathbb{R} \mid F_X(x) \geq p\}, \sup \{x \in \mathbb{R} \mid F_X(x) \leq p\}],$$

where, as before, $\inf \emptyset = +\infty$ and also $\sup \emptyset = -\infty$. Taking the left-hand border of this interval to be the value of the inverse c.d.f. at p , we get $F_X^{-1}(p)$. Similarly, we define $F_X^{-1+}(p)$ as the right-hand border of the interval:

$$F_X^{-1+}(p) = \sup \{x \in \mathbb{R} \mid F_X(x) \leq p\}, \quad p \in [0, 1],$$

which is a non-decreasing and r.c. function. Note that $F_X^{-1}(0) = -\infty$ and $F_X^{-1+}(1) = +\infty$. All the outcomes of X are contained in the closed interval $[F_X^{-1+}(0), F_X^{-1}(1)]$. Also note that $F_X^{-1}(p)$ and $F_X^{-1+}(p)$ are finite for all $p \in (0, 1)$. In the sequel, we will always use p as a variable ranging over the open interval $(0, 1)$, unless stated otherwise.

For any $\alpha \in [0, 1]$, we will define now the α -mixed inverse function of F_X as follows:

$$F_X^{-1(\alpha)}(p) = \alpha F_X^{-1}(p) + (1 - \alpha) F_X^{-1+}(p), \quad p \in (0, 1),$$

which is a non-decreasing function. In particular, we find $F_X^{-1(0)}(p) = F_X^{-1+}(p)$ and $F_X^{-1(1)}(p) = F_X^{-1}(p)$. We also have that for all $\alpha \in [0, 1]$,

$$F_X^{-1}(p) \leq F_X^{-1(\alpha)}(p) \leq F_X^{-1+}(p), \quad p \in (0, 1).$$

Note that only values of p corresponding to a horizontal segment of F_X lead to different values of $F_X^{-1}(p)$, $F_X^{-1+}(p)$ and $F_X^{-1(\alpha)}(p)$. Now let d be such that $0 < F_X(d) < 1$. Then $F_X^{-1}(F_X(d))$ and $F_X^{-1+}(F_X(d))$ are finite, and $F_X^{-1}(F_X(d)) \leq d \leq F_X^{-1+}(F_X(d))$. So for some value $\alpha_d \in [0, 1]$, d can be expressed as $d = \alpha_d F_X^{-1}(F_X(d)) + (1 - \alpha_d) F_X^{-1+}(F_X(d)) = F_X^{-1(\alpha_d)}(F_X(d))$. This implies that for any random variable X and any d with $0 < F_X(d) < 1$, there exists an $\alpha_d \in [0, 1]$ such that $F_X^{-1(\alpha_d)}(F_X(d)) = d$.

In the following lemma, we state the relation between the inverse distribution functions of the random variables X and $g(X)$ for a monotone function g .

Lemma 1.3.1 (Inverse distribution function of $g(X)$) *Let X and $g(X)$ be real-valued random variables and let $0 < p < 1$.*

(a) *If g is non-decreasing and l.c., then*

$$F_{g(X)}^{-1}(p) = g(F_X^{-1}(p)).$$

(b) *If g is non-decreasing and r.c., then*

$$F_{g(X)}^{-1+}(p) = g(F_X^{-1+}(p)).$$

(c) *If g is non-increasing and l.c., then*

$$F_{g(X)}^{-1+}(p) = g(F_X^{-1}(1 - p)).$$

(d) If g is non-increasing and r.c., then

$$F_{g(X)}^{-1}(p) = g(F_X^{-1+}(1-p)).$$

Proof. We will prove (a). The other results can be proven similarly. Let $0 < p < 1$ and consider a non-decreasing and left-continuous function g . For any real x we find from (1.13) that

$$F_{g(X)}^{-1}(p) \leq x \Leftrightarrow p \leq F_{g(X)}(x).$$

As g is l.c., we have that

$$g(z) \leq x \Leftrightarrow z \leq \sup \{y \mid g(y) \leq x\},$$

holds for all real z and x . Hence,

$$p \leq F_{g(X)}(x) \Leftrightarrow p \leq F_X[\sup \{y \mid g(y) \leq x\}].$$

If $\sup \{y \mid g(y) \leq x\}$ is finite then we find from (1.13) and the equivalence above

$$p \leq F_X[\sup \{y \mid g(y) \leq x\}] \Leftrightarrow F_X^{-1}(p) \leq \sup \{y \mid g(y) \leq x\}.$$

In case $\sup \{y \mid g(y) \leq x\}$ is $+\infty$ or $-\infty$, we cannot use (1.13), but one can verify that the equivalence above also holds in this case. Indeed, if the supremum equals $-\infty$, then the equivalence becomes $p \leq 0 \Leftrightarrow F_X^{-1}(p) \leq -\infty$. If the supremum equals $+\infty$, then the equivalence becomes $p \leq 1 \Leftrightarrow F_X^{-1}(p) \leq +\infty$. Because g is non-decreasing and l.c., we get that

$$F_X^{-1}(p) \leq \sup \{y \mid g(y) \leq x\} \Leftrightarrow g(F_X^{-1}(p)) \leq x.$$

Combining the equivalences, we finally find that

$$F_{g(X)}^{-1}(p) \leq x \Leftrightarrow g(F_X^{-1}(p)) \leq x,$$

holds for all values of x , which means that (a) must hold. ■

In this work we will reserve the notation U for a uniform(0, 1) random variable, i.e. $F_U(p) = p$ and $F_U^{-1}(p) = p$ for all $0 < p < 1$. One can prove that for all $\alpha \in [0, 1]$,

$$X \stackrel{d}{=} F_X^{-1}(U) \stackrel{d}{=} F_X^{-1+}(U) \stackrel{d}{=} F_X^{-1(\alpha)}(U). \quad (1.14)$$

The first distributional equality is known as the quantile transform theorem and follows immediately from (1.13). It implies that a sample of random numbers from a general cumulative distribution function F_X can be generated from a sample of uniform random numbers. Note that F_X has at most a countable number of horizontal segments, implying that the last three random variables in (1.14) only differ in a null-set of values of U . This means that these random variables are equal with probability one.

Comonotonicity

2.1 Introduction

²From the introduction of Chapter 1 we recall that in financial and actuarial applications one is often confronted with random variables of the form $S = \sum_{i=1}^n X_i$ with the property that one only knows the marginal cumulative distribution functions of the individual random variables X_i , but not the multivariate cumulative distribution function of the joint random vector $\underline{X} = (X_1, X_2, \dots, X_n)$, because the dependencies between the components of $\underline{X}$ are not known or too difficult to work with. Therefore, in order to enhance prudent and conservative decision-taking we will look for the dependency structure of $\underline{X}$ that yields the largest sum, in the convex order sense. More specifically, we will prove that the convex-largest sum of the components of a random vector with given marginals will be obtained in the case that the random vector $(X_1, \dots, X_n)$ has the *comonotonic* distribution, which means that each two possible outcomes $(x_1, \dots, x_n)$ and $(y_1, \dots, y_n)$ of $(X_1, \dots, X_n)$ are ordered componentwise. Using the technique of conditioning we will also derive a convex lower bound. Considering a more attractive random variable than S will help to give an idea of the degree of overestimation of the risk involved, when replacing S by the less attractive convex upper bound. An example will illustrate that, in the for practical applications very important case of lognormal variables X_i , the (c.d.f. of the) convex lower bound is often very close to (the c.d.f. of) the original r.v. S , and this makes the convex lower bound also a valuable candidate for decision-taking even if it entails a slightly less conservative decision-taking process. This chapter is based on Dhaene, Denuit, Goovaerts, Kaas & Vyncke (2002a).

² “Nothing is so practical as a good theory” - Anonymous.

2.2 Comonotonic sets and random vectors

We start by defining comonotonicity of a set A of n -vectors in $\mathbb{R}^n$. A n -vector $(x_1, x_2, \dots, x_n)$ will be denoted by $\underline{x}$. For two n -vectors $\underline{x}$ and $\underline{y}$, the notation $\underline{x} \leq \underline{y}$ will be used for the componentwise order which is defined by $x_i \leq y_i$ for all $i = 1, 2, \dots, n$. We will denote the (i, j) -projection of a set A in $\mathbb{R}^n$ by $A_{i,j}$. It is formally defined by $A_{i,j} = \{(x_i, x_j) \mid \underline{x} \in A\}$.

Definition 2.2.1 (Comonotonic set) *The set $A \subseteq \mathbb{R}^n$ is comonotonic if for any $\underline{x}$ and $\underline{y}$ in A , either $\underline{x} \leq \underline{y}$ or $\underline{y} \leq \underline{x}$ holds.*

A set $A \subseteq \mathbb{R}^n$ is comonotonic if for any $\underline{x}$ and $\underline{y}$ in A , if $x_i < y_i$ for some i , then $\underline{x} \leq \underline{y}$ must hold. Hence, a comonotonic set is simultaneously non-decreasing in each component. Notice that a comonotonic set is a ‘thin’ set: it cannot contain any subset of dimension larger than 1. Any subset of a comonotonic set is also comonotonic. The proof of the following lemma is straightforward.

Lemma 2.2.1 (Pairwise comonotonicity) *$A \subseteq \mathbb{R}^n$ is comonotonic if and only if $A_{i,j}$ is comonotonic for all $i \neq j$ in $\{1, 2, \dots, n\}$.*

For a general set A , comonotonicity of the $(i, i+1)$ -projections $A_{i,i+1}$, ($i = 1, 2, \dots, n-1$), will not necessarily imply that A is comonotonic. As a counter example, consider the set $A = \{(x_1, 1, x_3) \mid 0 < x_1, x_3 < 1\}$. This set is not comonotonic, although $A_{1,2}$ and $A_{2,3}$ are comonotonic. Now we will characterize any n -dimensional random vector $\underline{X} = (X_1, \dots, X_n)$ through the notion of support.

Definition 2.2.2 (Support of a random vector) *Any subset $A \subseteq \mathbb{R}^n$ will be called a support of the n -dimensional random vector $\underline{X} = (X_1, \dots, X_n)$ if $\Pr[\underline{X} \in A] = 1$.*

We are ready now to introduce the concept of comonotonicity for random vectors.

Definition 2.2.3 (Comonotonic random vector) *The random vector $\underline{X}$ is comonotonic if it has a comonotonic support.*

From Definition 2.2.3 we can conclude that comonotonicity is a very strong positive dependency structure. Indeed, if $\underline{x}$ and $\underline{y}$ are elements of the comonotonic support of $\underline{X}$, i.e. $\underline{x}$ and $\underline{y}$ are possible outcomes of $\underline{X}$, then they must be ordered component by component. This explains the term comonotonic (common monotonic). Comonotonicity of a random vector $\underline{X}$ implies that the higher the value of one component X_j , the higher the value of any other component X_k . This means that comonotonicity entails that no X_j is in any way a ‘hedge’, for another component X_k .

In the following theorem, some equivalent characterizations are given for comonotonicity of the random vector $\underline{X}$.

Theorem 2.2.1 (Characterizations for comonotonicity) *A random vector $\underline{X} = (X_1, X_2, \dots, X_n)$ is comonotonic if and only if one of the following equivalent conditions holds:*

- (1) $\underline{X}$ has a comonotonic support ;
- (2) For all $\underline{x} = (x_1, x_2, \dots, x_n)$, we have

$$F_{\underline{X}}(\underline{x}) = \min \{F_{X_1}(x_1), F_{X_2}(x_2), \dots, F_{X_n}(x_n)\}; \quad (2.1)$$

- (3) For $U \sim \text{Uniform}(0,1)$, we have

$$\underline{X} \stackrel{d}{=} (F_{X_1}^{-1}(U), F_{X_2}^{-1}(U), \dots, F_{X_n}^{-1}(U)); \quad (2.2)$$

- (4) There exists a random variable Z and non-decreasing functions f_i , ($i = 1, 2, \dots, n$), such that

$$\underline{X} \stackrel{d}{=} (f_1(Z), f_2(Z), \dots, f_n(Z)). \quad (2.3)$$

Proof. (1) $\Rightarrow$ (2): Assume that $\underline{X}$ has comonotonic support B . Let $\underline{x} \in \mathbb{R}^n$ and let A_j be defined by

$$A_j = \{\underline{y} \in B \mid y_j \leq x_j\}, \quad j = 1, 2, \dots, n.$$

Because of the comonotonicity of B , there exists an i such that $A_i = \cap_{j=1}^n A_j$. Hence, we find

$$\begin{aligned} F_{\underline{X}}(\underline{x}) &= \Pr(\underline{X} \in \cap_{j=1}^n A_j) = \Pr(\underline{X} \in A_i) = F_{X_i}(x_i) \\ &= \min \{F_{X_1}(x_1), F_{X_2}(x_2), \dots, F_{X_n}(x_n)\}. \end{aligned}$$

The last equality follows from $A_i \subset A_j$ so that $F_{X_i}(x_i) \leq F_{X_j}(x_j)$ holds for all values of j .

(2) $\Rightarrow$ (3): Now assume that $F_{\underline{X}}(\underline{x}) = \min \{F_{X_1}(x_1), F_{X_2}(x_2), \dots, F_{X_n}(x_n)\}$ for all $\underline{x} = (x_1, x_2, \dots, x_n)$. Then we find by (1.13)

$$\begin{aligned} \Pr[F_{X_1}^{-1}(U) \leq x_1, \dots, F_{X_n}^{-1}(U) \leq x_n] \\ &= \Pr[U \leq F_{X_1}(x_1), \dots, U \leq F_{X_n}(x_n)] \\ &= \Pr[U \leq \min_{j=1, \dots, n} \{F_{X_j}(x_j)\}] \\ &= \min_{j=1, \dots, n} \{F_{X_j}(x_j)\}. \end{aligned}$$

(3) $\Rightarrow$ (4): straightforward.

(4) $\Rightarrow$ (1): Assume that there exists a random variable Z with support B , and non-decreasing functions f_i , ($i = 1, 2, \dots, n$), such that

$$\underline{X} \stackrel{d}{=} (f_1(Z), f_2(Z), \dots, f_n(Z)).$$

The set of possible outcomes of $\underline{X}$ is $\{(f_1(z), f_2(z), \dots, f_n(z)) \mid z \in B\}$ which is obviously comonotonic, which implies that $\underline{X}$ is indeed comonotonic. ■

It is a well-known fact that for any random vector $(X_1, \dots, X_n)$, not necessarily comonotonic, the following inequality holds:

$$\Pr [X_1 \leq x_1, X_2 \leq x_2, \dots, X_n \leq x_n] \leq \min \{F_{X_1}(x_1), F_{X_2}(x_2), \dots, F_{X_n}(x_n)\}, \quad (2.4)$$

and since Hoeffding (1940) and Fréchet (1951) it is known that the right-hand side of this inequality, i.e. $\min \{F_{X_1}(x_1), F_{X_2}(x_2), \dots, F_{X_n}(x_n)\}$ is indeed the multivariate c.d.f. of the random vector $(F_{X_1}^{-1}(U), F_{X_2}^{-1}(U), \dots, F_{X_n}^{-1}(U))$, which has the same marginals as $(X_1, \dots, X_n)$. The inequality (2.4) states that in the class of all random vectors $(X_1, \dots, X_n)$ with the same marginals, the probability that all X_i simultaneously realize ‘small’ values is maximized if the vector is comonotonic, suggesting that comonotonicity is indeed a very strong positive dependency structure.

As the random vectors $(F_{X_1}^{-1}(U), F_{X_2}^{-1}(U), \dots, F_{X_n}^{-1}(U))$ and $(F_{X_1}^{-1(\alpha_1)}(U), F_{X_2}^{-1(\alpha_2)}(U), \dots, F_{X_n}^{-1(\alpha_n)}(U))$ are equal with probability one, we find that comonotonicity of $\underline{X}$ can also be characterized by

$$\underline{X} \stackrel{d}{=} (F_{X_1}^{-1(\alpha_1)}(U), F_{X_2}^{-1(\alpha_2)}(U), \dots, F_{X_n}^{-1(\alpha_n)}(U)), \quad (2.5)$$

for $U \sim \text{Uniform}(0,1)$ and given real numbers $\alpha_i \in [0, 1]$. If $U \sim \text{Uniform}(0,1)$, then also $1 - U \sim \text{Uniform}(0,1)$. This implies that another characterization for the comonotonicity of $\underline{X}$ is given by:

$$\underline{X} \stackrel{d}{=} (F_{X_1}^{-1}(1 - U), F_{X_2}^{-1}(1 - U), \dots, F_{X_n}^{-1}(1 - U)). \quad (2.6)$$

In the sequel, for any random vector $(X_1, \dots, X_n)$, the notation $(X_1^c, \dots, X_n^c)$ will be used to indicate a comonotonic random vector with the same marginals as $(X_1, \dots, X_n)$. From (2.2) we find that for any random vector $\underline{X}$ the outcome of its comonotonic counterpart $\underline{X}^c = (X_1^c, \dots, X_n^c)$ lies with probability 1 in the following set:

$$\{(F_{X_1}^{-1}(p), F_{X_2}^{-1}(p), \dots, F_{X_n}^{-1}(p)) \mid 0 < p < 1\}. \quad (2.7)$$

This support of $\underline{X}^c$ is not necessarily a connected curve. Indeed, all horizontal segments of the c.d.f. of X_i lead to ‘missing pieces’ in this curve. This support can be seen as a series of ordered connected curves. Now, by connecting the endpoints of consecutive curves by straight lines, we obtain a comonotonic connected curve in $\mathbb{R}^n$. Hence, it may be traversed in a direction which is upwards for all components simultaneously. We will call this set the connected support of $\underline{X}^c$. It might be parameterized as follows:

$$\{(F_{X_1}^{-1(\alpha)}(p), F_{X_2}^{-1(\alpha)}(p), \dots, F_{X_n}^{-1(\alpha)}(p)) \mid 0 < p < 1, 0 \leq \alpha \leq 1\}. \quad (2.8)$$

Note that this parameterization is not necessarily unique: there may be elements in the connected support which can be characterized by different values of α .

The following theorem states essentially that the comonotonicity of a random vector is equivalent with pairwise comonotonicity.

Theorem 2.2.2 (Pairwise comonotonicity for random vectors) *The random vector $\underline{X}$ is comonotonic if and only if (X_i, X_j) is comonotonic for all $i \neq j$ in $\{1, 2, \dots, n\}$.*

Proof. The proof of the “ $\Rightarrow$ ” implication is straightforward.

For the proof of the “ $\Leftarrow$ ” implication, consider the set A in $\mathbb{R}^n$ defined by

$$A = \left\{ (F_{X_1}^{-1}(p), F_{X_2}^{-1}(p), \dots, F_{X_n}^{-1}(p)) \mid 0 < p < 1 \right\}.$$

Its (i, j) -projections are given by

$$A_{i,j} = \left\{ (F_{X_i}^{-1}(p), F_{X_j}^{-1}(p)) \mid 0 < p < 1 \right\}.$$

The event ‘ $\underline{X} \in A$ ’ is equivalent with the event ‘ $(X_i, X_j) \in A_{i,j}$ for all (i, j) ’. Because of the comonotonicity of the pairs (X_i, X_j) , we have that $\Pr[(X_i, X_j) \in A_{i,j}] = 1$ and hence we find that $\Pr[\underline{X} \in A] = 1$, so that the comonotonic set A is a support of $\underline{X}$. This implies that $\underline{X}$ is a comonotonic random vector. ■

For a general random vector $\underline{X}$, comonotonicity of the pairs (X_i, X_{i+1}) , $(i = 1, 2, \dots, n-1)$, will not necessarily imply comonotonicity of $\underline{X}$. A counterexample is the random vector $(U, 1, V)$ with U and V mutually independent random variables that are both uniformly distributed on the unit-interval $(0, 1)$. It is clear that $(U, 1)$ and $(1, V)$ are both comonotonic pairs, but $(U, 1, V)$ isn’t comonotonic.

2.3 The comonotonic sum

For any random vector $\underline{X} = (X_1, X_2, \dots, X_n)$, not necessarily comonotonic, we will call its comonotonic counterpart any random vector with the same marginal distributions and with the comonotonic dependency structure. The comonotonic counterpart of $\underline{X} = (X_1, X_2, \dots, X_n)$ will be denoted by $\underline{X}^c = (X_1^c, X_2^c, \dots, X_n^c)$. Recall from Theorem 2.2.1 that

$$(X_1^c, X_2^c, \dots, X_n^c) \stackrel{d}{=} (F_{X_1}^{-1}(U), F_{X_2}^{-1}(U), \dots, F_{X_n}^{-1}(U)).$$

In the sequel, the notation S^c will be used for the sum of the components of the comonotonic counterpart $\underline{X}^c = (X_1^c, X_2^c, \dots, X_n^c)$

$$S^c = X_1^c + X_2^c + \dots + X_n^c.$$

We will prove now that the stop-loss premiums and the cumulative distribution function of S^c can be obtained from the corresponding quantities for each of the terms in the sum. We will also prove that S^c is the largest convex sum that can be obtained in the class of all random vectors $\underline{X}$ with given marginals. This two facts make it an excellent candidate for prudent decision-taking in case the joint distribution of the random vector $\underline{X}$ is either unspecified or too cumbersome to work with.

Theorem 2.3.1 (Inverse c.d.f. of a comonotonic sum) *The α -inverse distribution function $F_{S^c}^{-1(\alpha)}$ for the sum S^c of the comonotonic random variables $\underline{X}^c = (X_1^c, X_2^c, \dots, X_n^c)$ is given by*

$$F_{S^c}^{-1(\alpha)}(p) = \sum_{i=1}^n F_{X_i^c}^{-1(\alpha)}(p), \quad 0 < p < 1, \quad 0 \leq \alpha \leq 1. \quad (2.9)$$

Proof. Consider the random vector $(X_1, X_2, \dots, X_n)$ and its comonotonic counterpart $(X_1^c, X_2^c, \dots, X_n^c)$. Then $S^c = X_1^c + X_2^c + \dots + X_n^c \stackrel{d}{=} g(U)$, with U uniformly distributed on $(0, 1)$ and with the function g defined by

$$g(u) = \sum_{i=1}^n F_{X_i}^{-1}(u), \quad 0 < u < 1.$$

It is clear that g is non-decreasing and left-continuous. Application of Lemma 1.3.1 (a) leads to

$$F_{S^c}^{-1}(p) = F_{g(U)}^{-1}(p) = g(F_U^{-1}(p)) = g(p), \quad 0 < p < 1$$

and the inverse distribution function of S^c can be computed from

$$F_{S^c}^{-1}(p) = \sum_{i=1}^n F_{X_i}^{-1}(p), \quad 0 < p < 1.$$

Similarly, from Lemma 1.3.1(b), we find that

$$F_{S^c}^{-1+}(p) = \sum_{i=1}^n F_{X_i}^{-1+}(p), \quad 0 < p < 1.$$

Multiplying the last two equalities by α and $1 - \alpha$ respectively, and adding up, we find the desired result. ■

Note that

$$S^c \stackrel{d}{=} \sum_{i=1}^n F_{X_i}^{-1(\alpha)}(U). \quad (2.10)$$

By the theorem above, we find that the connected support of S^c is given by

$$\begin{aligned} & \left\{ F_{S^c}^{-1(\alpha)}(p) \mid 0 < p < 1, 0 \leq \alpha \leq 1 \right\} \\ &= \left\{ \sum_{i=1}^n F_{X_i}^{-1(\alpha)}(p) \mid 0 < p < 1, 0 \leq \alpha \leq 1 \right\}. \end{aligned}$$

This implies

$$F_{S^c}^{-1+}(0) = \sum_{i=1}^n F_{X_i}^{-1+}(0), \quad (2.11)$$

$$F_{S^c}^{-1}(1) = \sum_{i=1}^n F_{X_i}^{-1}(1). \quad (2.12)$$

Hence, the minimal value of the comonotonic sum equals the sum of the minimal values of each term. Similarly, the maximal value of the comonotonic sum equals the sum of the maximal values of each term. The number $\sum_{i=1}^n F_{X_i}^{-1+}(0)$, which is either finite or $-\infty$ (if any of the terms in the sum is $-\infty$), is the minimum possible value of S^c , and likewise $\sum_{i=1}^n F_{X_i}^{-1}(1)$ is the maximum. Also note that

$$\begin{aligned} F_{S^c}^{-1+}(1) &= \sum_{i=1}^n F_{X_i}^{-1+}(1) = +\infty, \\ F_{S^c}^{-1}(0) &= \sum_{i=1}^n F_{X_i}^{-1}(0) = -\infty. \end{aligned}$$

For any $(X_1, X_2, \dots, X_n)$, we have that $S = X_1 + X_2 + \dots + X_n \geq \sum_{i=1}^n F_{X_i}^{-1+}(0)$ must hold with probability 1. This implies

$$\sum_{i=1}^n F_{X_i}^{-1+}(0) \leq F_S^{-1+}(0). \quad (2.13)$$

Similarly, we find

$$F_S^{-1}(1) \leq \sum_{i=1}^n F_{X_i}^{-1}(1). \quad (2.14)$$

So, the sum S of the components of the random vector $(X_1, X_2, \dots, X_n)$ has a support that is contained in the interval $[\sum_{i=1}^n F_{X_i}^{-1+}(0), \sum_{i=1}^n F_{X_i}^{-1}(1)]$. The minimal value of S is larger than or equal to the one of S^c , since by comonotonicity all terms of the latter are simultaneously small.

Given the inverse functions $F_{X_i}^{-1}$, the c.d.f. of $S^c = X_1^c + X_2^c + \dots + X_n^c$ can be determined as follows:

$$\begin{aligned} F_{S^c}(x) &= \sup \{p \in (0, 1) \mid F_{S^c}(x) \geq p\} \\ &= \sup \{p \in (0, 1) \mid F_{S^c}^{-1}(p) \leq x\} \\ &= \sup \left\{ p \in (0, 1) \mid \sum_{i=1}^n F_{X_i}^{-1}(p) \leq x \right\}. \end{aligned} \quad (2.15)$$

In the sequel, for any random variable X , the expression “ F_X is strictly increasing” has to be interpreted as: “ F_X is strictly increasing on $(F_X^{-1+}(0), F_X^{-1}(1))$ ”. Observe that for any random variable X , the following equivalences hold:

$$F_X \text{ is strictly increasing} \iff F_X^{-1} \text{ is continuous on } (0, 1), \quad (2.16)$$

and also

$$F_X \text{ is continuous} \iff F_X^{-1} \text{ is strictly increasing on } (0, 1). \quad (2.17)$$

Now assume that the marginal cumulative distribution functions F_{X_i} , $i = 1, \dots, n$ of the comonotonic random vector $(X_1^c, X_2^c, \dots, X_n^c)$ are strictly increasing and

continuous. Then each inverse distribution function $F_{X_i}^{-1}$ is continuous on $(0, 1)$, which implies that $F_{S^c}^{-1}$ is continuous on $(0, 1)$ because $F_{S^c}^{-1}(p) = \sum_{i=1}^n F_{X_i}^{-1}(p)$ holds for $0 < p < 1$. This in turn implies that F_{S^c} is strictly increasing on $(F_{S^c}^{-1+}(0), F_{S^c}^{-1}(1))$. By a similar reasoning one finds that F_{S^c} is continuous. Hence, in case of strictly increasing and continuous marginals, for any $F_{S^c}^{-1+}(0) < x < F_{S^c}^{-1}(1)$, the probability $F_{S^c}(x)$ is uniquely determined by $F_{S^c}^{-1}(F_{S^c}(x)) = x$, or equivalently,

$$\sum_{i=1}^n F_{X_i}^{-1}(F_{S^c}(x)) = x, \quad F_{S^c}^{-1+}(0) < x < F_{S^c}^{-1}(1). \quad (2.18)$$

It suffices thus to solve the latter equation to get $F_{S^c}(x)$.

In the following theorem, we prove that also the stop-loss premiums of a sum of comonotonic random variables can be obtained from the stop-loss premiums of the terms.

Theorem 2.3.2 (Stop-loss premium of comonotonic sum) *The stop-loss premiums of the sum S^c of the components of the comonotonic random vector $(X_1^c, X_2^c, \dots, X_n^c)$ are given by*

$$E[(S^c - d)_+] = \sum_{i=1}^n E[(X_i - d_i)_+], \quad (F_{S^c}^{-1+}(0) < d < F_{S^c}^{-1}(1)), \quad (2.19)$$

with the d_i given by

$$d_i = F_{X_i}^{-1(\alpha_d)}(F_{S^c}(d)), \quad (i = 1, 2, \dots, n) \quad (2.20)$$

and $\alpha_d \in [0, 1]$ determined by

$$F_{S^c}^{-1(\alpha_d)}(F_{S^c}(d)) = d. \quad (2.21)$$

Proof. Let $d \in (F_{S^c}^{-1+}(0), F_{S^c}^{-1}(1))$, hence $0 < F_{S^c}(d) < 1$.

As the connected support of $\underline{X}^c$ as defined in (2.8) is comonotonic, it can have at most one point of intersection with the hyperplane $\{\underline{x} \mid x_1 + \dots + x_n = d\}$. This is because the hyperplane cannot contain vectors $\underline{x}$ and $\underline{y}$ that are ordered componentwise, i.e. such that $\underline{x} \leq \underline{y}$ or $\underline{x} \geq \underline{y}$ holds.

Now we will prove that the vector $\underline{d} = (d_1, d_2, \dots, d_n)$ as defined above is the unique point of this intersection. As $0 < F_{S^c}(d) < 1$ must hold, we know from Section(1.3) that there exists an $\alpha_d \in [0, 1]$ that fulfils condition (2.21). Note that by Theorem 2.3.1, we have that $\sum_{i=1}^n d_i = d$. Hence, the vector $\underline{d}$ with the d_i defined in (2.20) and (2.21) is an element of both the connected support of $\underline{X}^c$ and the hyperplane $\{\underline{x} \mid x_1 + \dots + x_n = d\}$ so that we can conclude that $\underline{d}$ is the unique element of the intersection of the connected support and the hyperplane.

Let $\underline{x}$ be an element of the connected support of $\underline{X}^c$. Then the following equality holds:

$$(x_1 + x_2 + \dots + x_n - d)_+ \equiv (x_1 - d_1)_+ + (x_2 - d_2)_+ + \dots + (x_n - d_n)_+.$$

This is because $\underline{x}$ and $\underline{d}$ are both elements of the connected support of $\underline{X}^c$, and hence, if there exists any j such that $x_j > d_j$ holds, then we also have $x_k \geq d_k$ for all k , and the left-hand side equals the right-hand side because $\sum_{i=1}^n d_i = d$. On the other hand, when all $x_j \leq d_j$, obviously the left-hand side is 0 as well. Now, replacing constants by the corresponding random variables in the equality above and taking expectations, we find (2.19). ■

In case the marginal c.d.f.'s F_{X_i} are strictly increasing, (2.20) reduces to

$$E[(S^c - d)_+] = \sum_{i=1}^n E[(X_i - F_{X_i}^{-1}(F_{S^c}(d)))_+], \quad d \in (F_{S^c}^{-1+}(0), F_{S^c}^{-1}(1)). \quad (2.22)$$

From Theorem 2.3.2, we can conclude that any stop-loss premium of a sum of comonotonic random variables can be written as the sum of stop-loss premiums for the individual random variables involved. The theorem provides an algorithm for directly computing stop-loss premiums of sums of comonotonic random variables, without having to compute the entire cumulative distribution function of the sum itself. Indeed, in order to compute the stop-loss premium with retention d , we only need to know $F_{S^c}(d)$, which can be computed directly from (2.15).

Application of the relation $E[(X - d)_+] = E[(d - X)_+] + E[X] - d$ for S^c and the X_i in expression (2.19) leads to the following expression for the lower tails of a sum of comonotonic random variables:

$$E[(d - S^c)_+] = \sum_{i=1}^n E[(d_i - X_i)_+], \quad F_{S^c}^{-1+}(0) < d < F_{S^c}^{-1}(1), \quad (2.23)$$

with the d_i as defined in (2.20) and (2.21).

Theorem 2.3.3 (Upper bound for a sum of random variables) *For any random vector $(X_1, X_2, \dots, X_n)$ we have*

$$X_1 + X_2 + \dots + X_n \leq_{cx} X_1^c + X_2^c + \dots + X_n^c. \quad (2.24)$$

Proof. It suffices to prove stop-loss order, since it is obvious that the means of these two sums are equal. Hence, we have to prove that

$$E[(X_1 + X_2 + \dots + X_n - d)_+] \leq E[(X_1^c + X_2^c + \dots + X_n^c - d)_+]$$

holds for all retentions d with $d \in (F_{S^c}^{-1+}(0), F_{S^c}^{-1}(1))$, since the stop-loss premiums can be seen to be equal for other values of d .

The following holds for all $(x_1, x_2, \dots, x_n)$ when $d_1 + d_2 + \dots + d_n = d$:

$$\begin{aligned} & (x_1 + x_2 + \dots + x_n - d)_+ \\ &= ((x_1 - d_1) + (x_2 - d_2) + \dots + (x_n - d_n))_+ \\ &\leq ((x_1 - d_1)_+ + (x_2 - d_2)_+ + \dots + (x_n - d_n)_+)_+ \\ &= (x_1 - d_1)_+ + (x_2 - d_2)_+ + \dots + (x_n - d_n)_+. \end{aligned}$$

Now replacing constants by the corresponding random variables in the inequality above and taking expectations, we get that

$$E[(X_1 + X_2 + \dots + X_n - d)_+] \leq E[(X_1 - d_1)_+] + E[(X_2 - d_2)_+] + \dots + E[(X_n - d_n)_+] \quad (2.25)$$

holds for all d and d_i such that $\sum_{i=1}^n d_i = d$.

By choosing $d \in (F_{S^c}^{-1+}(0), F_{S^c}^{-1}(1))$ and the d_i as in Theorem 2.3.2, the above inequality becomes the one that had to be proven. ■

The theorem above states that the least attractive random vector $(X_1, \dots, X_n)$ with given marginals F_i , in the sense that the sum of their components is largest in the convex order, has the *comonotonic* joint distribution, which means that it has the joint distribution of $(F_1^{-1}(U), F_2^{-1}(U), \dots, F_n^{-1}(U))$. The components of this random vector are maximally dependent, all components being non-decreasing functions of the same random variable. See also Denneberg (1994), Dhaene & Goovaerts (1997), Müller (1997), Dhaene, Wang, Young & Goovaerts (2000) or Dhaene, Denuit, Goovaerts, Kaas & Vyncke (2002a) for various proofs of this result. Note also that the inequality (2.23) holds in particular if $(X_1, \dots, X_n)$ is comonotonic. From the Theorems 2.3.2 and 2.3.3, we find that for any random vector $\underline{X}$ the inequalities

$$\begin{aligned} E[(X_1 + X_2 + \dots + X_n - d)_+] &\leq \sum_{i=1}^n E[(X_i - F_{X_i}^{-1(\alpha_d)}(F_{S^c}(d)))_+] \\ &\leq \sum_{i=1}^n E[(X_i - d_i)_+], \end{aligned} \quad (2.26)$$

hold for all $d \in (F_{S^c}^{-1+}(0), F_{S^c}^{-1}(1))$ such that $\sum_{i=1}^n d_i = d$. Hence, the smallest upper bound of the form $\sum_{i=1}^n E[(X_i - d_i)_+]$ with $\sum_{i=1}^n d_i = d$ for the stop-loss premium $E[(X_1 + X_2 + \dots + X_n - d)_+]$ is the comonotonic upper bound.

Corollary 2.3.1 (Comonotonicity and stop-loss ordered components)

Consider the random vectors $(X_1, X_2, \dots, X_n)$ and $(Y_1, Y_2, \dots, Y_n)$. If $X_i \leq_{sl} Y_i$ holds for all $i = 1, \dots, n$, then

$$X_1 + X_2 + \dots + X_n \leq_{sl} Y_1^c + Y_2^c + \dots + Y_n^c. \quad (2.27)$$

Proof. Since $Y_1^c + \dots + Y_n^c$ is comonotonic, we know since Theorem 2.3.2 that for any real d , one can find $d_1, \dots, d_n$ with $d = d_1 + \dots + d_n$ and

$$E[(Y_1^c + \dots + Y_n^c - d)_+] = E[(Y_1 - d_1)_+] + \dots + E[(Y_n - d_n)_+].$$

Hence

$$\begin{aligned} E[(X_1 + X_2 + \dots + X_n - d)_+] &\leq E[(X_1 - d_1)_+] + \dots + E[(X_n - d_n)_+] \\ &\leq E[(Y_1 - d_1)_+] + \dots + E[(Y_n - d_n)_+] \\ &= E[(Y_1^c + \dots + Y_n^c - d)_+] \end{aligned}$$

and this ends the proof ■

2.4 Comonotonicity and linear correlation

From Theorem 2.2.2 one knows that the comonotonicity of the random vector $(X_1, X_2, \dots, X_n)$ is equivalent with the comonotonicity of all pairs (X_i, X_j) with $i \neq j$. It is therefore interesting to compare comonotonicity of the pairs (X_i, X_j) with the well-known Pearson's measure of correlation between X_i and X_j .

Definition 2.4.1 (Pearson's correlation coefficient) *For a random couple (X, Y) , the Pearson's correlation coefficient is defined by*

$$r(X, Y) = \frac{\text{Cov}[X, Y]}{\sqrt{\text{Var}[X] \text{Var}[Y]}}, \quad (2.28)$$

with

$$\text{Cov}[X, Y] = E[(X - E[X])(Y - E[Y])],$$

the covariance between X and Y .

As a consequence of the Cauchy-Schwarz inequality we obtain the fact that $0 \leq r(X, Y) \leq 1$. The following lemma shows that the Pearson's correlation coefficient measures the linear dependency between 2 random variables and links a perfect positive linear correlation (i.e. when $r(X, Y) = +1$) to comonotonicity. A proof for this can be found in any statistical handbook.

Lemma 2.4.1 (Characterization of correlation) *$r(X, Y) = +1$ if and only if real numbers $a > 0$ and b exist such that $Y = aX + b$ holds almost surely.*

So, if $r(X, Y) = 1$ then the connected support is a straight line and hence the couple (X, Y) will be comonotonic. On the other hand, if the support is not a straight line, we always have that $r(X, Y) < 1$, hence comonotonicity of (X, Y) does not imply that $r(X, Y) = 1$. When generalizing this to n -dimensions, we find that if $r(X_i, X_j) = 1$ for all $i \neq j$ then this implies that $(X_1, X_2, \dots, X_n)$ is comonotonic, but the opposite statement is only true if there exists a random variable Y , positive real constants a_i and real constants b_i such that the relation $X_i \stackrel{d}{=} a_i Y + b_i$ holds for $i = 1, 2, \dots, n$. We can prove however that the comonotonicity of $(X_1, X_2, \dots, X_n)$ is equivalent with the maximization of the $r(X_i, X_j)$. In order to prove this result, we need an expression for the stop-loss premiums of a sum of two random variables in terms of the bivariate cumulative distribution function.

Lemma 2.4.2 (Stop-loss premiums for a bivariate sum) *For any bivariate random variable (X, Y) and any real number d , the stop-loss premium of $X + Y$ at retention d is given by*

$$E[(X + Y - d)_+] = E[X] + E[Y] - d + \int_{-\infty}^{+\infty} F_{X,Y}(x, d - x) dx \quad (2.29)$$

Proof. By reversing the order of the integration, we find

$$\begin{aligned}
E[(d - X - Y)_+] &= \int_{x=-\infty}^{+\infty} \int_{y=-\infty}^{d-x} \int_{t=x}^{d-y} dt \, dF_{X,Y}(x, y) \\
&= \int_{t=-\infty}^{+\infty} \int_{x=-\infty}^t \int_{y=-\infty}^{d-t} dF_{X,Y}(x, y) \, dx \\
&= \int_{t=-\infty}^{+\infty} F_{X,Y}(t, d-t) \, dt,
\end{aligned}$$

from which we find the desired result. ■

Theorem 2.4.1 (Comonotonicity and maximum correlation) *A random vector $\underline{X}$ is comonotonic if and only if $r(X_i, X_j)$ is maximal for all $i, j \in \{1, 2, \dots, n\}$.*

Proof. “ $\Rightarrow$ ” From Theorem 2.3.1 and Theorem 2.3.3 we know that for all $i, j \in \{1, 2, \dots, n\}$, $X_i + X_j$ will be the largest convex sum. From (1.12) we find that this implies that the variance of $X_i + X_j$ will be maximal, which ends the proof. For the “ $\Leftarrow$ ” statement, note that a maximal variance for $X_i + X_j$ together with $X_1 + X_2 \leq_{cx} X_1^c + X_2^c$ yields that $X_1 + X_2 \stackrel{d}{=} X_1^c + X_2^c$. This implies that for all real d we have that

$$E[(X_1 + X_2 - d)_+] = E[(X_1^c + X_2^c - d)_+]$$

Using Lemma 2.4.2 this means that

$$\int_{-\infty}^{\infty} [F_{X_1^c, X_2^c}(x, d-x) - F_{X_1, X_2}(x, d-x)] \, dx = 0.$$

From (2.3) we know that $F_{X_1, X_2}(x, d-x) \leq F_{X_1^c, X_2^c}(x, d-x)$ and this implies that $F_{X_1^c, X_2^c}(x, d-x) = F_{X_1, X_2}(x, d-x)$ for all real x and as this also has to hold for any real d this ends the proof. ■

From the proof of Theorem 2.4.1 we also deduce that a random vector $\underline{X}$ is comonotonic if and only if $X_i + X_j$ is comonotonic for all $i, j \in \{1, 2, \dots, n\}$. Denuit & Dhaene (2003) show that for other measures of dependency such as Kendall’s τ , Spearman’s ρ and Gini’s γ , a similar result holds, i.e. for any of these dependency measures we have that comonotonicity of the vector $\underline{X}$ is characterized by extremal values for these dependency measures for the couples (X_i, X_j) , for all $i, j \in \{1, 2, \dots, n\}$. As these measures are preserved under increasing transformations of the random variables X_i , one can show that, in contrast to the Pearson’s correlation coefficient, these extremal values coincide with 1. As a final remark, we notice that the properties of Pearson’s correlation coefficient are not always fully known by practitioners, see Dhaene, Vanduffel, Goovaerts, Koch, Olieslagers & Romijn (2004) for an illustration in case of a credit risk portfolio.

2.5 Convex bounds for $\sum_{i=1}^n X_i$

In the previous section we derived a convex upper bound for $\sum_{i=1}^n X_i$. We will derive now a convex lower bound for this sum, see also Kaas, Dhaene & Goovaerts (2000). This will be based on the well-known fact that for any variable X and Λ , the variable $E[X | \Lambda]$ is smaller in convex order than X . In practical situations it often turns out that the (c.d.f. of the) convex lower bound is often very close to (the c.d.f. of) the original r.v. $\sum_{i=1}^n X_i$ and this makes the convex lower bound a valuable candidate for decision-taking even though it might lead to a slightly less conservative decision taking-process

Lemma 2.5.1 (Lower bound for a random variable) *For any couple (X, Λ) we have that $E[X | \Lambda] \leq_{cx} X$.*

Proof. Because of the iterativity of the expectation operator, we have that the moments of $E[X | \Lambda]$ and X are equal. By applying Jensen's inequality we find for any convex function v that

$$\begin{aligned} E[v(E[X | \Lambda])] &\leq E[E[v(X) | \Lambda]] \\ &= E[v(X)] \end{aligned} \quad (2.30)$$

and this ends the proof. ■

As an immediate application of Lemma 2.5.1 and Corollary 2.3.1 we obtain the following important result.

Theorem 2.5.1 (Lower bound for a sum of random variables) *For any random vector $(X_1, X_2, \dots, X_n)$ and any random variable Λ , we have that*

$$\sum_{i=1}^n E[X_i | \Lambda] \leq_{cx} \sum_{i=1}^n X_i. \quad (2.31)$$

The following corollary follows now from Theorem 2.5.1 and Theorem 2.3.3 and will be often used in the remainder of this work.

Corollary 2.5.1 (Bounds for a sum of random variables) *For any random vector $(X_1, X_2, \dots, X_n)$ and any random variable Λ , we have that*

$$S^l \leq_{cx} S \leq_{cx} S^c. \quad (2.32)$$

with S^l, S, S^c defined as

$$\begin{aligned} S^l &= \sum_{i=1}^n E[X_i | \Lambda], \\ S^l &= \sum_{i=1}^n X_i, \\ S^c &= \sum_{i=1}^n F_{X_i}^{-1}(U). \end{aligned} \quad (2.33)$$

Approximating (the c.d.f. of) $S = X_1 + X_2 + \dots + X_n$ by (the c.d.f. of) the convex lower bound S^l will only be meaningful if we can easily determine the c.d.f. and the stop-loss premiums of S^l . A necessary condition for this, is that the c.d.f. of the conditioning variable Λ is known and the same holds for the expectations of the X_i , conditional on $\Lambda = \lambda$. Hence more information is needed than in case one uses the upper bound S^c . We also point out that the random vector $(E[X_1 | \Lambda], E[X_2 | \Lambda], \dots, E[X_n | \Lambda])$ will in general not have the same marginal distributions as $(X_1, X_2, \dots, X_n)$.

Nevertheless, if one can find a conditioning random variable Λ with the property that the variance of $\sum_{i=1}^n E[X_i | \Lambda]$ is close to the variance of $\sum_{i=1}^n X_i$, we might obtain excellent approximations for the c.d.f. of $\sum_{i=1}^n X_i$, see also Chapter 1. Moreover, if Λ is such that all $E[X_i | \Lambda]$ are non-increasing functions of Λ (or all are non-decreasing functions of Λ), the lower bound $S^l = \sum_{i=1}^n E[X_i | \Lambda]$ will be a comonotonic sum so that in this case we can apply the results derived in Section 2.3. to find the c.d.f. and the stop-loss premiums of S^l .

Also note that in case the lower bound yields a comonotonic sum, this is not a result of imposing a comonotonic dependency structure, but follows naturally if the random variables $E[X_i | \Lambda]$ are non-increasing functions of Λ (or all are non-decreasing functions of Λ) which is often the case in practical situations.

Let us now consider the general case where not all $E[X_i | \Lambda]$ are non-increasing (or not all are non-decreasing). In this case the lower bound is not a sum of comonotonic random variables, making the computation of the cumulative distribution function of the lower bound more complicated. The c.d.f. and the stop-loss premiums of S^l can be determined as follows in this case:

$$\begin{aligned} F_{S^l}(x) &= \int_{-\infty}^{+\infty} \Pr \left[\sum_{i=1}^n E[X_i | \Lambda] \leq x \mid \Lambda = \lambda \right] dF_{\Lambda}(\lambda) \\ &= \int_{-\infty}^{+\infty} I \left(\sum_{i=1}^n E[X_i | \Lambda = \lambda] \leq x \right) dF_{\Lambda}(\lambda), \end{aligned} \quad (2.34)$$

$$E \left[(S^l - d)_+ \right] = \int_{-\infty}^{+\infty} \left(\sum_{i=1}^n E[X_i | \Lambda = \lambda] - d \right)_+ dF_{\Lambda}(\lambda). \quad (2.35)$$

A somewhat different procedure can be used when F_{Λ} is continuous and strictly increasing. In this case, define the random variable $U \equiv F_{\Lambda}(\Lambda)$ which is uniformly distributed on the unit interval. We have that $U = u \Leftrightarrow \Lambda = F_{\Lambda}^{-1}(u)$ holds for all $0 < u < 1$. Hence, the c.d.f. and the stop-loss premiums of S^l then follow from

$$\begin{aligned} F_{S^l}(x) &= \int_0^1 \Pr \left[\sum_{i=1}^n E[X_i | \Lambda] \leq x \mid U = u \right] du \\ &= \int_0^1 I \left(\sum_{i=1}^n E[X_i | \Lambda = F_{\Lambda}^{-1}(u)] \leq x \right) du, \end{aligned} \quad (2.36)$$

$$E \left[(S^l - d)_+ \right] = \int_0^1 \left(\sum_{i=1}^n E [X_i | \Lambda = F_\Lambda^{-1}(u)] - d \right)_+ du. \quad (2.37)$$

The technique for deriving lower bounds as explained in this section is also considered (for some special cases) in Vyncke, Goovaerts & Dhaene (2001). The idea of this technique stems from mathematical physics, and was applied by Rogers & Shi (1995) to derive approximate values for the price of Asian options.

In the following example we will show how the upper and lower bounds can be applied in case one deals with sums of lognormals. In fact, the sum we consider here is a special case of general sums of lognormals that we will study in detail in Chapter 4. We present it here to fix some ideas. Sums of lognormals turn out to be important for many practical applications that we will discuss in the second part of this thesis. We refer to Dhaene, Denuit, Goovaerts, Kaas & Vyncke (2002a) for other applications of the theory on convex bounds.

2.6 Example

Take Y_1 and Y_2 independent $N(\mu, \sigma^2)$ random variables. We will examine the sum of $X_1 = e^{Y_1+Y_2} \sim \text{lognormal}(2\mu, 2\sigma^2)$ with $X_2 = e^{Y_2} \sim \text{lognormal}(\mu, \sigma^2)$ the sum $S = X_1 + X_2$ can be interpreted as the value at time 2 of a periodical savings plan that consists in investing a unit amount at time 0 and another one at time 1, where the (log)investment returns in year 1 and 2 are given by Y_1 and Y_2 respectively. Although the dependency structure is known, it is so cumbersome that we cannot fruitfully use it to determine the c.d.f. of S . Note that for determining the lower bound, we will consider conditioning variables $\Lambda = aY_1 + Y_2$ for the appropriate choice of a .

Using the expressions (2.33) we find for the upper bound approximation S^c

$$S^c = +e^{2\mu+\sqrt{2}\sigma Z} + e^{\mu+\sigma Z}, \quad (2.38)$$

whereas for the lower bound S^l we obtain

$$S^l = +e^{2\mu+(1-r_1^2)\sigma^2+r_1\sigma\sqrt{2}\frac{\Lambda-E[\Lambda]}{\sigma_\Lambda}} + e^{\mu+\frac{1}{2}(1-r_2^2)\sigma^2+r_2\sigma\frac{\Lambda-E[\Lambda]}{\sigma_\Lambda}}, \quad (2.39)$$

with

$$\begin{aligned} r_1 &= \frac{1+a}{\sqrt{2}\sqrt{1+a^2}}, \\ r_2 &= \frac{1}{\sqrt{1+a^2}}. \end{aligned} \quad (2.40)$$

Hence, we have that S^c and S^l respectively are distributed as

$$S^c \stackrel{d}{=} +e^{2\mu+\sqrt{2}\sigma\Phi^{-1}(U)} + e^{\mu+\sigma\Phi^{-1}(U)} \quad (2.41)$$

and

$$S^l \stackrel{d}{=} e^{2\mu + \frac{1+a^2-2a}{2(1+a^2)}\sigma^2 + \frac{1+a}{\sqrt{1+a^2}}\sigma\Phi^{-1}(U)} + e^{\mu + \frac{1}{2(1+a^2)}\sigma^2 + \frac{a}{\sqrt{1+a^2}}\sigma\Phi^{-1}(U)}, \quad (2.42)$$

with $U \sim \text{uniform}(0,1)$ and Φ the c.d.f. of the standard normal random variable. For the variances we obtain

$$\begin{aligned} \text{Var}(S) &= e^{2\mu+\sigma^2}(e^{\sigma^2}-1) + e^{4\mu+2\sigma^2}(e^{2\sigma^2}-1) \\ &\quad + 2e^{3\mu+\frac{3}{2}\sigma^2}(e^{\sigma^2}-1), \\ \text{Var}(S^c) &= e^{2\mu+\sigma^2}(e^{\sigma^2}-1) + e^{4\mu+2\sigma^2}(e^{2\sigma^2}-1) \\ &\quad + 2e^{3\mu+\frac{3}{2}\sigma^2}(e^{\sqrt{2}\sigma^2}-1), \\ \text{Var}(S^l) &= e^{2\mu+\sigma^2}(e^{\frac{1}{1+a^2}\sigma^2}-1) + e^{4\mu+2\sigma^2}(e^{\frac{(1+a)^2}{(1+a^2)}\sigma^2}-1) \\ &\quad + 2e^{3\mu+\frac{3}{2}\sigma^2}(e^{\frac{(1+a)}{(1+a^2)}\sigma^2}-1). \end{aligned} \quad (2.43)$$

Let us first take $a = 1$. We find that

$$S^l = e^\Lambda + e^{\frac{1}{4}\sigma^2 + \frac{\Lambda}{2}}, \quad (2.44)$$

We remark that S^l , conditional on the realization $\Lambda = \lambda$, depends on λ and σ^2 but *not* on μ . For the distribution of S^c one finds

$$S^l \stackrel{d}{=} e^{2\mu+\sqrt{2}\sigma\Phi^{-1}(U)} + e^{\frac{1}{4}\sigma^2 + \frac{\sqrt{2}}{2}\sigma\Phi^{-1}(U)}. \quad (2.45)$$

whereas for the distribution of S^c one obtains

$$S^c \stackrel{d}{=} e^{2\mu+\sqrt{2}\sigma\Phi^{-1}(U)} + e^{\mu+\frac{\sqrt{2}}{2}\sigma\Phi^{-1}(U)}. \quad (2.46)$$

For this choice for a , we see by comparing (2.45) with (2.46) that the c.d.f. of S^c will cross the c.d.f. of S^l exactly once. It is an open question if this result holds if the sum of lognormals involves more than 2 terms and/or if we consider other suitable choices for Λ . Denote now the risk-free rate by r and let us further compare S with S^l . We know from Chapter 1 that $E[u(S^l)] \geq E[u(S)]$ for all risk averse decision takers. In case X_1 and X_2 are prices of a non-dividend paying stock at time $t = 1$ and $t = 2$ respectively, and under the assumption that for all $t > 0$ there is an instantaneous risk-free rate δ_t , we can compare the arbitrage-free prices of S and S_l . Note that both random variables are indeed functions of the underlying stochastic process $\{X(t), t \geq 0\}$. We will assume that δ_t is constant through time ($\delta_t = \delta$). One can show now that in this Black & Scholes setting, the absence of arbitrage opportunities is equivalent with the existence of unique prices for S and S_l respectively, and that these prices have to be determined by taking the discounted expectation of S and S_l not with respect to the (initial) physical probability measure, but with respect to the equivalent probability measure 'Q', i.e. the probability measure on $\{X(t), t \geq 0\}$ such that the discounted price process $\{e^{-\delta t} X(t), t \geq 0\}$ is a martingale, which means that for all $t \geq 0$:

$$E^Q[e^{-\delta t} X(t)] = X(0), \quad t \geq 0. \quad (2.47)$$

We refer to e.g. Harrison & Kreps (1979) or Harrison & Pliska (1981) for extensive theory on arbitrage-free pricing. A reference in the actuarial literature is Gerber & Shiu (1996). In our case we have that for all $t > 0$, $X(t)$ is Lognormal $(\mu t, \sigma^2 t)$ distributed and from (2.47) it follows directly that, under the equivalent martingale measure, $X(t)$ will follow a Lognormal $(\delta t, \sigma^2 t)$ distribution. Loosely speaking, one can say now that the prices of S^l and S are given by $e^{-2\delta} E_r[S]$ and $e^{-2\delta} E_r[S^l]$ where the expectations are obtained by substituting μ by r in the parameterization of the stochastic process $\{X(t), t \geq 0\}$. One can then verify that these expectations, and hence the prices coincide, but this also follows directly from the comments made on (2.44), because this implies that $E[S] = E[S^l]$ for every μ -probability measure. We conclude that the prices of S and S^l are the same but that every risk averse decision-taker will always prefer S^l above S . Remark that S^l depends only on the final value X_1 of the underlying process, whereas S is path-dependent, and hence all this suggests that, at least in a Black & Scholes setting, path-dependent financial structures can be outperformed by path-independent structures. The optimality of path-dependent structures has been considered by Cox & Leeland (1982) and Dybvig (1988). For the remainder of the example we will assume that $\mu = 0$ and $\sigma^2 = 1$. We find that

$$\begin{aligned} \text{Var}[S^l] &= 64.374, \\ \text{Var}[S] &= 67.281, \\ \text{Var}[S^e] &= 79.785. \end{aligned} \tag{2.48}$$

from which we see that the lower bound is likely generate the best approximation for the c.d.f. of S .

Recall now from Chapter 1 that the lower bound S^l will be optimal if its variance is as close as possible to the variance of S . One can intuitively expect that this will happen if the conditioning random variable Λ resembles S as much as possible. Approximating e^{Y_2} and $e^{Y_1+Y_2}$ by $1 + Y_2$ and $1 + Y_1 + Y_2$ respectively, we see that $S \approx 2 + Y_1 + 2Y_2$. Hence, we could expect that taking $\frac{Y_1}{2} + Y_2$ instead of $Y_1 + Y_2$ as our conditioning random variable might lead to a better lower bound. A bit surprisingly this is not true since the variance of the lower bound is only 61.440 in this case. This suggests that there might be better criteria for obtaining Λ than the one that consists in linearizing S . Of course one can maximize the variance of S^l but this will lead to numerical procedures which outweigh the advantage of the bounds, namely that these lead to easy to compute actuarial quantities. However, rather than linearizing S , one could linearize $\text{Var}(S)$, and then maximize it, which can be done explicitly, see also Chapter 4. Using such an approach, we obtain that $a = \frac{1}{1+e^{\frac{1}{2}}} = 0.622$ and the variance for S^l is then equal to 64.65. It proves that the optimal lower bound obtained by conditioning on random variables of type $Y_1 + aY_2$ is reached for $a = 0.788$ and the variance of S^l is then 66.082. We conclude that the choice for Λ that entails that the first order approximation for the variance of S_l is maximized, yields a choice that is much closer to the optimal choice as compared to the choice that follows from a linearization of S .

Risk measures

3.1 Introduction

³In order to make decisions, one has to evaluate and to compare the risk (random variable) one faces. We define a risk measure as a mapping from the set of random variables representing the risks at hand to the real line. A risk measure can be a helpful tool for decision-making since they resume the information available about the r.v. X into one single number $\rho[X]$. We will always consider random variables as losses, or payments that have to be made. A negative outcome for the loss variable means that a gain has occurred. Common risk measures in Actuarial Science are premium principles, see for instance Goovaerts, De Vijlder & Haezendonck (1984), or also Chapter 5 in Kaas, Goovaerts, Dhaene & Denuit (2001). Other risk measures are used for determining provisions and capital requirements of an insurer, in order to avoid insolvency. Then they measure the upper tails of distribution functions. Such measures of risk are considered in Artzner, Delbaen, Eber & Heath (1999), Wirth & Hardy (2000), Panjer (2002), Dhaene, Goovaerts & Kaas (2003), Tsanakas & Desli (2003), among others. In this work, we will mainly concentrate on risk measures that can be used for reserving and solvency purposes.

Let X now be the r.v. representing an insurance company's risks related to a particular policy, a particular line-of-business or to the entire insurance portfolio over a specified time horizon. We do not specify what kind of risk X is. It could be one specific risk type, such as credit risk for all assets. Or it could be a sum of dependent risks $X_1 + \dots + X_n$, where the X_i represent the different risk types such as market risk, event risk and so on, or where the X_i represent the claims related to the different policies of the portfolio. The risk X can also appear in a personal finance context, when a decision-maker targets a given series of future consumptions or when he faces, for instance, the risk that his fixed contribution pension plan does not meet his retirement financial goal.

³“Wiskunde is de noodzakelijke grondlaag die je niet ziet maar zonder die laag komt het papier los” - Kardinaal Daneels (15/2/03, De Standaard).

Ensuring that insurers have the financial means to meet their obligations to pay the present and future claims related to policyholders is the purpose of solvency. In order to avoid insolvency over the specified time horizon at some given level of risk tolerance, the insurer should hold assets of value $\rho[X]$ or more. Essentially, $\rho[X]$ should be such that $\Pr[X > \rho[X]]$ is ‘small enough’. Note that $\rho[X]$ is a risk measure expressed in monetary terms. It could be defined for instance as the 99-th percentile of the cumulative distribution function of X .

As mentioned above, the risk X will often be a sum of non-independent risks. Hence, we will consider the general problem of determining approximations for risk measures of sums of random variables of which the dependency structure is unknown or too cumbersome to work with.

In Section 3.2 we introduce several well-known risk measures and the relations that hold between them. Characterizations for ordering concepts in terms of risk measures are explored in Section 3.3. The class of distortion risk measures is examined in Section 3.4. In this section we also discuss approximations for distortion risk measures in case one deals with sums of non-independent random variables. Section 3.5 investigates the relationship between theories for choice under risk and distortion risk measures. Section 3.6 concludes the chapter. This chapter is based on Dhaene, Vanduffel, Tang, Goovaerts, Kaas and Vyncke (2004).

3.2 Some well-known risk measures

As a first example of a risk measure, consider the p -quantile risk measure, often called the ‘VaR’ (Value-at-Risk) at level p in the financial and actuarial literature. For any p in $(0, 1)$, the p -quantile risk measure for a random variable X , which will be denoted by $Q_p[X]$, is defined by

$$Q_p[X] = \inf \{x \in \mathbb{R} \mid F_X(x) \geq p\}, \quad p \in (0, 1), \quad (3.1)$$

where $F_X(x) = \Pr[X \leq x]$. We also introduce the risk measure $Q_p^+[X]$ which is defined by

$$Q_p^+[X] = \sup \{x \in \mathbb{R} \mid F_X(x) \leq p\}, \quad p \in (0, 1). \quad (3.2)$$

Remark that the definitions for $Q_p[X]$ and $Q_p^+[X]$ exactly coincide with the definitions for $F_X^{-1}(p)$ and $F_X^{-1+}(p)$ respectively, that we introduced in Chapter 1. The p -quantile risk measure $Q_p[X]$ is indeed often denoted by $F_X^{-1}(p)$ and likewise $Q_p^+[X]$ can also be denoted by $F_X^{-1+}(p)$. Note that expression (3.1) can also be used to define $Q_0[X]$ and $Q_1[X]$. For the latter quantile, we take the convention $\inf \emptyset = +\infty$. We then find that $Q_0(X) = -\infty$. For a bounded random variable X , we have that $Q_1[X] = \max(X)$. The quantile function $Q_p[X]$ is a non-decreasing and left-continuous function of p . Finally, from (1.13) we find for all $x \in \mathbb{R}$ and $p \in [0, 1]$ that

$$Q_p[X] \leq x \Leftrightarrow p \leq F_X(x). \quad (3.3)$$

Note that the equivalence relation (3.3) holds with equalities if F_X is continuous at this particular x .

Let X denote the aggregate claims of an insurance portfolio. The liabilities (provisions) for this portfolio are given by P . Assume the insurer establishes a solvency capital $K = Q_p[X] - P$ with p sufficiently large, e.g. $p = 0.99$. In this case, the capital can be interpreted as the ‘smallest’ capital such that the insurer becomes technically insolvent, i.e. claims exceed provisions and capital, with a (small) probability of at most $1 - p$:

$$K = \inf \{L \mid \Pr[X > P + L] \leq 1 - p\}. \quad (3.4)$$

Using the p -quantile risk measure for determining a solvency capital is meaningful in situations where the default event should be avoided, but the size of the shortfall is less important. For example, for shareholders or management, the quantile risk measure gives useful information since avoiding default is the primary concern, whereas the size of the shortfall is only secondary. We refer to Dhaene, Vanduffel, Goovaerts, Olieslagers & Koch (2003) for an example in which the p -quantile risk measure has been used to determine the solvency capital. However, a single quantile risk measure of a predetermined level p does not give any information about the thickness of the upper tail of the cumulative distribution function from $Q_p[X]$ on. A regulator for instance, is not only concerned with the frequency of default, but also about the severity of default. Also shareholders and management should be concerned with the question “how bad is bad?” when they want to evaluate the risks at hand in a consistent way.

Therefore, one often uses another risk measure which is called the Tail Value-at-Risk (TVaR) at level p . It is denoted by $\text{TVaR}_p[X]$, and defined by

$$\text{TVaR}_p[X] = \frac{1}{1-p} \int_p^1 Q_q[X] \, dq, \quad p \in (0, 1). \quad (3.5)$$

It is the arithmetic average of the quantiles of X , from p on. Note that the TVaR is always larger than the corresponding quantile. From (3.5) it follows immediately that the Tail Value-at-Risk is a non-decreasing function of p .

Let X again denote the aggregate claims of an insurance portfolio over a given reference period and P the provision for this portfolio. Setting the capital equal to $\text{TVaR}_p[X] - P$, we could define ‘bad times’ as those where X takes a value in the interval $[Q_p[X], \text{TVaR}_p[X]]$. Hence, ‘bad times’ are those where the aggregate claims exceed the threshold $Q_p[X]$, but not using up all available capital. The width of the interval is a ‘cushion’ that is used in case of ‘bad times’. For more details, see Overbeck (2000).

The Conditional Tail Expectation (CTE) at level p will be denoted by $\text{CTE}_p[X]$. It is defined as

$$\text{CTE}_p[X] = E[X \mid X > Q_p[X]], \quad p \in (0, 1). \quad (3.6)$$

Loosely speaking, the conditional tail expectation at level p is equal to the mean of the top $(1-p)\%$ losses. It can also be interpreted as the VaR at level p

augmented by the average exceeding of the claims X over that quantile, given that such exceeding occurs.

The Expected Shortfall (ESF) at level p will be denoted by $\text{ESF}_p[X]$, and is defined as

$$\text{ESF}_p[X] = \mathbb{E} \left[(X - Q_p[X])_+ \right], \quad p \in (0, 1). \quad (3.7)$$

This risk measure can be interpreted as the expected value of the shortfall in case the capital is set equal to $Q_p[X] - P$.

We will also need a risk measure that measures the left tail of the cumulative distribution function. Therefore, we introduce the Conditional Left Tail Expectation, denoted by $\text{CLTE}_p[X]$ and defined as

$$\text{CLTE}_p[X] = \mathbb{E} [X \mid X < Q_p^+[X]], \quad p \in (0, 1). \quad (3.8)$$

The following relations hold between the four risk measures defined above.

Theorem 3.2.1 (Relation between Quantiles, TVaR, CTE and ESF)

For $p \in (0, 1)$, we have that

$$\text{TVaR}_p[X] = Q_p[X] + \frac{1}{1-p} \text{ESF}_p[X], \quad (3.9)$$

$$\text{CTE}_p[X] = Q_p[X] + \frac{1}{1 - F_X(Q_p[X])} \text{ESF}_p[X], \quad (3.10)$$

$$\text{CTE}_p[X] = \text{TVaR}_{F_X(Q_p[X])}[X], \quad (3.11)$$

$$\text{CLTE}_p[X] = -\text{CTE}_{1-p}[-X] \quad (3.12)$$

Proof. Expression (3.9) follows from

$$\begin{aligned} \text{ESF}_p[X] &= \int_0^1 (Q_q[X] - Q_p[X])_+ dq \\ &= \int_p^1 Q_q[X] dq - Q_p[X] [1 - p]. \end{aligned}$$

Expression (3.10) follows from

$$\text{ESF}_p[X] = \mathbb{E} [X - Q_p[X] \mid X > Q_p[X]] (1 - F_X(Q_p[X])). \quad (3.13)$$

Expression (3.11) follows immediately from (3.9) and (3.10). Finally, expression (3.12) follows as a straight forward application from lemma (1.3.1). ■

About the Tail Value-at-Risk we find from Definition (3.5) the following elementary result, which will be applied later: if X has a finite expectation $\mathbb{E}[X]$, then

$$\lim_{p \searrow 0} \text{TVaR}_p[X] = \mathbb{E}[X]. \quad (3.14)$$

Note that if F_X is continuous then

$$\text{CTE}_p[X] = \text{TVaR}_p[X], \quad p \in (0, 1). \quad (3.15)$$

Example 3.2.1 (Normal losses)

Consider a random variable $X \sim N(\mu, \sigma^2)$ which is normally distributed with mean μ and variance σ^2 . From Lemma 1.3.1, it follows immediately that the quantiles of X are given by

$$Q_p[X] = \mu + \sigma \Phi^{-1}(p), \quad p \in (0, 1), \quad (3.16)$$

where Φ denotes the standard normal cumulative distribution function. The stop-loss premiums of X are given by

$$\mathbb{E}[(X - d)_+] = \sigma \phi\left(\frac{d - \mu}{\sigma}\right) - (d - \mu) \left[1 - \Phi\left(\frac{d - \mu}{\sigma}\right)\right], \quad -\infty < d < +\infty, \quad (3.17)$$

where $\phi(x) = \Phi'(x)$ denotes the density function of the standard normal distribution. For a proof, see e.g. Example 3.9.1 in Kaas, Goovaerts, Dhaene & Denuit (2001). From (3.17) we find the following expression for the Expected Shortfall:

$$\text{ESF}_p[X] = \sigma \phi(\Phi^{-1}(p)) - \sigma \Phi^{-1}(p) (1 - p), \quad p \in (0, 1). \quad (3.18)$$

Using (3.10), we find that the Conditional Tail Expectation is given by

$$\text{CTE}_p[X] = \mu + \sigma \frac{\phi(\Phi^{-1}(p))}{1 - p}, \quad p \in (0, 1). \quad (3.19)$$

The Conditional Left Tail Expectation follows from (3.12):

$$\text{CLTE}_p[X] = \mu - \sigma \frac{\phi(\Phi^{-1}(p))}{p}, \quad p \in (0, 1). \quad (3.20)$$

▽

Example 3.2.2 (Lognormal losses)

Consider a random variable X that is lognormally distributed. Hence, $\ln X \sim N(\mu, \sigma^2)$. The quantiles of X follow now from Lemma 1.3.1:

$$Q_p[X] = e^{\mu + \sigma \Phi^{-1}(p)}, \quad p \in (0, 1). \quad (3.21)$$

It is well-known that the stop-loss premiums of X are given by

$$\mathbb{E}[(X - d)_+] = e^{\mu + \sigma^2/2} \Phi(d_1) - d \Phi(d_2), \quad d > 0, \quad (3.22)$$

where $d_1 = \sigma + (\mu - \ln d)/\sigma$ and $d_2 = d_1 - \sigma$. The Black & Scholes (1973) call-option pricing formula is based on this expression for the stop-loss premium of a lognormal random variable. The Expected Shortfall is then given by

$$\text{ESF}_p[X] = e^{\mu+\sigma^2/2} \Phi(\sigma - \Phi^{-1}(p)) - e^{\mu+\sigma} \Phi^{-1}(p) (1-p), \quad p \in (0, 1). \quad (3.23)$$

The Conditional Tail Expectation is given by

$$\text{CTE}_p[X] = e^{\mu+\sigma^2/2} \frac{\Phi(\sigma - \Phi^{-1}(p))}{1-p}, \quad p \in (0, 1). \quad (3.24)$$

Finally, we also find

$$\text{CLTE}_p[X] = e^{\mu+\sigma^2/2} \frac{1 - \Phi(\sigma - \Phi^{-1}(p))}{p}, \quad p \in (0, 1). \quad (3.25)$$

▽

3.3 Risk measures and ordering of risks

3.3.1 General results

We recall from Chapter 1 that when comparing the random variables X and Y , X is said to precede Y in the stochastic dominance sense, notation $X \leq_{st} Y$, if and only if the cumulative distribution function of X always exceeds that of Y : The two variables are said to be stop-loss ordered, notation $X \leq_{sl} Y$, when X has lower stop-loss premiums than Y . Finally if $E(X)=E(Y)$ whilst $X \leq_{sl} Y$ then we say that X is convex smaller than Y denoted by $X \leq_{cx} Y$. In the definitions of stop-loss order and convex order, we always tacitly assume that the expectations exist.

In the following theorem it is stated that stochastic dominance can be characterized in terms of ordered quantiles. The proof is straightforward.

Theorem 3.3.1 (Stochastic dominance and ordered quantiles) *For any random pair (X, Y) we have that X is smaller than Y in stochastic dominance sense if and only if their respective quantiles are ordered:*

$$X \leq_{st} Y \Leftrightarrow Q_p[X] \leq Q_p[Y] \text{ for all } p \in (0, 1). \quad (3.26)$$

In the following theorem, we prove that stop-loss order can be characterized in terms of ordered TVaR's.

Theorem 3.3.2 (Stop-loss order and ordered TVaR's) *For any random pair (X, Y) we have that $X \leq_{sl} Y$ if and only if their respective TVaR's are ordered:*

$$X \leq_{sl} Y \Leftrightarrow \text{TVaR}_p[X] \leq \text{TVaR}_p[Y] \text{ for all } p \in (0, 1). \quad (3.27)$$

Proof. First we assume $X \leq_{sl} Y$ and let $p \in (0, 1)$. Consider the function $f(d)$ defined by

$$f(d) = (1 - p)d + E[(X - d)_+] = (1 - p)d + \int_d^\infty \bar{F}_X(x)dx,$$

where $\bar{F}_X(x) = 1 - F_X(x)$ is the decumulative distribution function of X . Observe that $\bar{F}_X(Q_p[X]) \leq 1 - p \leq \bar{F}_X(Q_p[X] - 0)$. So by the monotonicity of the function $\bar{F}_X(x)$, one easily sees that the function $f(d)$, and hence also the function $f(d)/(1 - p)$, is minimized for d equal to $Q_p[X]$. Hence, by choosing $d = Q_p[Y]$, we find

$$\begin{aligned} \text{TVaR}_p[X] &= Q_p[X] + \frac{1}{1 - p} E[(X - Q_p[X])_+] \\ &= \frac{f(Q_p[X])}{1 - p} \\ &\leq \frac{f(Q_p[Y])}{1 - p} \\ &= Q_p[Y] + \frac{1}{1 - p} E[(X - Q_p[Y])_+] \\ &\leq \text{TVaR}_p[Y]. \end{aligned}$$

To prove the other implication, we assume that the TVaR's are ordered for all $p \in (0, 1)$. Note that for any random variable X , we have that

$$\begin{aligned} E[(X - d)_+] &= E[(F_X^{-1}(U) - d)_+] \\ &= \int_{F_X(d)}^1 Q_q[X]dq - d(1 - F_X(d)). \end{aligned}$$

Hence, for d such that $0 < F_X(d) < 1$, we find

$$\begin{aligned} E[(X - d)_+] &= (\text{TVaR}_{F_X(d)}[X] - d)(1 - F_X(d)) \\ &\leq (\text{TVaR}_{F_X(d)}[Y] - d)(1 - F_X(d)) \\ &= \int_{F_X(d)}^1 Q_q[Y]dq - d(1 - F_X(d)) \\ &= \int_{F_Y(d)}^1 Q_q[Y]dq - d(1 - F_Y(d)) \\ &\quad + \int_{F_X(d)}^{F_Y(d)} Q_q[Y]dq - d(F_Y(d) - F_X(d)) \\ &= E[(Y - d)_+] + \int_{F_X(d)}^{F_Y(d)} (Q_q[Y] - d) dq. \end{aligned}$$

Using the equivalence $q \leq F_Y(d) \Leftrightarrow d \geq Q_q[Y]$, it is straightforward to prove that

$$\int_{F_X(d)}^{F_Y(d)} (Q_q[Y] - d) dq \leq 0.$$

This proves that the stop-loss premiums of X are smaller than that of Y for any retention d such that $0 < F_X(d) < 1$. If $F_X(d) = 1$, we find $E[(X - d)_+] = 0 \leq E[(Y - d)_+]$.

Recalling (3.14), the assumption that $\text{TVaR}_p[X] \leq \text{TVaR}_p[Y]$ for all $p \in (0, 1)$ immediately implies that $E[X] \leq E[Y]$. Thus $E[(X - d)_+] \leq E[(Y - d)_+]$ also holds for d such that $F_X(d) = 0$. Hence, we have proven that $X \leq_{sl} Y$. ■

Remark 3.3.1 (CTE does not preserve convex order)

Recall from the third item of Theorem 3.2.1. that the identity $\text{TVaR}_{F_X(d)}[X] = \text{CTE}_{F_X(d)}[X]$ holds for any d such that $0 < F_X(d) < 1$. Hence along the same line as the proof of (b) above, we can obtain the implication that

$$X \leq_{sl} Y \Leftrightarrow \text{CTE}_p[X] \leq \text{CTE}_p[Y] \text{ for all } p \in (0, 1).$$

However, the other implication is not true, in general. Actually, we make a somewhat stronger statement below:

$$X \leq_{cx} Y \not\Leftrightarrow \text{CTE}_p[X] \leq \text{CTE}_p[Y] \text{ for all } p \in (0, 1). \quad (3.28)$$

(3.28) can easily be illustrated as follows: Let X and Y be two random variables where F_Y is uniform over $[0, 1]$, and F_X is given by

$$F_X(x) = \begin{cases} x & \text{if } 0 \leq x < 0.85, \\ 0.85 & \text{if } 0.85 \leq x < 0.9, \\ 0.95 & \text{if } 0.9 \leq x < 0.95, \\ x & \text{if } 0.95 \leq x \leq 1. \end{cases} \quad (3.29)$$

Clearly, $F_X(x) \leq F_Y(x)$ for $x < 0.9$, and $F_X(x) \geq F_Y(x)$ for $x \geq 0.9$. We have that $E[X] = E[Y] = 0.5$ and $X \leq_{sl} Y$, hence that $X \leq_{cx} Y$. However, we easily check that $\text{CTE}_{0.9}[X] > \text{CTE}_{0.9}[Y]$ since $\text{CTE}_{0.9}[X] = 0.975$ and $\text{CTE}_{0.9}[Y] = 0.95$. ▽

3.3.2 Risk measures and comonotonicity

In Theorem 2.7 we proved that the quantile risk measure is additive for a sum of comonotonic random variables. In the next theorem we prove that this is also the case for Tail Value-at-Risk and the Expected Shortfall.

Theorem 3.3.3 (Additive risk measures for sums of comonotonic risks)

Consider a comonotonic random vector $(X_1^c, X_2^c, \dots, X_n^c)$, and let $S^c = X_1^c + X_2^c + \dots + X_n^c$. Then we have for all $p \in (0, 1)$ that

$$Q_p[S^c] = \sum_{i=1}^n Q_p[X_i], \quad (3.30)$$

$$\text{TVaR}_p[S^c] = \sum_{i=1}^n \text{TVaR}_p[X_i], \quad (3.31)$$

$$\text{ESF}_p[S^c] = \sum_{i=1}^n \text{ESF}_p[X_i]. \quad (3.32)$$

Proof. (a) We have that

$$S^c \stackrel{d}{=} F_{X_1}^{-1}(U) + F_{X_2}^{-1}(U) + \cdots + F_{X_n}^{-1}(U) = g(U),$$

with g a non-decreasing and left-continuous function. Hence, (3.30) follows from Lemma 1.3.1.

(b) Relation (3.31) follows immediately from (3.5) and (3.30).

(c) Relation (3.32) follows from (3.9), (3.30) and (3.31). ■

From the theorem above, we can conclude that the quantile, TVaR and ESF risk measure for a comonotonic sum can easily be obtained by summing the corresponding risk measures of the marginal distributions involved. Specifically, if all the random variables X_i above have the same distribution as that of X , then we find $Q_p[nX] = n Q_p[X]$, $\text{TVaR}_p[nX] = n \text{TVaR}_p[X]$ and $E[(nX - Q_p[nX])_+] = n E[(X - Q_p[X])_+]$. As we will see, the CTE risk measure is in general not additive for sums of comonotonic risks. Nevertheless, we immediately find that $\text{CTE}_p[nX] = n \text{CTE}_p[X]$. Another case where the additivity property does hold for CTE is given in the following remark.

Remark 3.3.2 (CTE is additive for sum of comonotonic cont. risks)

Consider a comonotonic random vector $(X_1^c, X_2^c, \dots, X_n^c)$ with continuous marginal distributions. For any random variable X , we have that $F_X(x)$ is continuous in $x \in (-\infty, \infty)$ if and only if $Q_p[X]$ is strictly increasing in $p \in (0, 1)$. This implies that the sum S^c is continuously distributed. Furthermore, the continuity of the cumulative distribution function of S^c implies that $\text{CTE}_p[S^c] = \text{TVaR}_p[S^c]$ for each $p \in (0, 1)$. Therefore it follows from (3.31) that

$$\text{CTE}_p[S^c] = \text{TVaR}_p[S^c] = \sum_{i=1}^n \text{TVaR}_p[X_i] = \sum_{i=1}^n \text{CTE}_p[X_i].$$

▽

For the case where the marginal distributions are not continuous and not the same, however, the CTE is, in general, not additive for comonotonic risks. Here we propose an illustration for this case.

Remark 3.3.3 (CTE is not additive for sum of comonotonic risks)

Consider the comonotonic random vector (X^c, Y^c) , where X has a distribution F_X given by

$$F_X(x) = \begin{cases} x & 0 \leq x < 0.85 \\ 0.85 & 0.85 \leq x < 0.9 \\ 0.95 & 0.9 \leq x < 0.95 \\ x & 0.95 \leq x \leq 1 \end{cases},$$

and Y is uniformly distributed in $(0, 1)$. We write

$$S^c = X^c + Y^c \stackrel{d}{=} F_X^{-1}(U) + U = g(U).$$

Since $F_X^{-1}(y)$ and $F_Y^{-1}(y)$ are non-decreasing in $y \in (0, 1)$ and the monotonicity of $F_Y^{-1}(y)$ is strict, the function $g(y)$ is strictly increasing. Hence, the sum $S^c = X^c + Y^c$ has a continuous distribution. Because of the additivity of the risk measures $Q_{0.9}(\cdot)$ and $\text{ESF}_{0.9}(\cdot)$, we find

$$\begin{aligned}
\text{CTE}_{0.9}[S^c] &= Q_{0.9}[S^c] + \frac{1}{1-0.9}\text{ESF}_{0.9}[S^c] \\
&= Q_{0.9}[X^c] + Q_{0.9}[Y^c] + \frac{1}{1-0.9}(\text{ESF}_{0.9}[X^c] + \text{ESF}_{0.9}[Y^c]) \\
&= \left(Q_{0.9}[X^c] + \frac{1}{1-0.95}\text{ESF}_{0.9}[X^c] \right) \\
&\quad + \left(Q_{0.9}[Y^c] + \frac{1}{1-0.9}\text{ESF}_{0.9}[Y^c] \right) \\
&\quad - \left(\frac{1}{1-0.95} - \frac{1}{1-0.9} \right) \text{ESF}_{0.9}[X^c] \\
&= \text{CTE}_{0.9}[X^c] + \text{CTE}_{0.9}[Y^c] \\
&\quad - \left(\frac{1}{1-0.95} - \frac{1}{1-0.9} \right) \text{ESF}_{0.9}[X^c] \\
&< \text{CTE}_{0.9}[X^c] + \text{CTE}_{0.9}[Y^c].
\end{aligned}$$

▽

A risk measure ρ is said to be subadditive if for any random variables X and Y , one has $\rho(X + Y) \leq \rho(X) + \rho(Y)$. Subadditivity of a risk measure ρ immediately implies

$$\rho\left(\sum_{i=1}^n X_i\right) \leq \sum_{i=1}^n \rho(X_i).$$

A risk measure is said to preserve stop-loss order if for any X and Y , one has that $X \leq_{sl} Y$ implies $\rho[X] \leq \rho[Y]$.

Theorem 3.3.4 (Subadditivity of risk measures) *Any risk measure that preserves stop-loss order and that is additive for comonotonic risks is subadditive.*

Proof. From Corollary 2.5.1, we have that a sum of random variables with given marginal distributions is largest in the convex order sense if these random variables are comonotonic:

$$X + Y \leq_{sl} X^c + Y^c.$$

If the risk measure ρ preserves stop-loss order and is additive for comonotonic risks, then

$$\rho[X + Y] \leq \rho[X^c + Y^c] = \rho[X] + \rho[Y],$$

which proves the stated result. ■

Recall Theorems 3.3.2 and 3.3.3. As a special case of Theorem 3.3.4, we find that TVaR is subadditive:

$$\text{TVaR}_p[X + Y] \leq \text{TVaR}_p[X] + \text{TVaR}_p[Y], \quad p \in (0, 1). \quad (3.33)$$

In the following remark we show that CTE is not subadditive.

Remark 3.3.4 (CTE is not subadditive)

Let X be a random variable uniformly distributed in $(0, 1)$, and let Y be another random variable defined by

$$Y = (0.95 - X)I_{(0 < X \leq 0.95)} + (1.95 - X)I_{(0.95 < X < 1)},$$

where I_A denotes the indicator function which equals 1 if condition A holds, and 0 otherwise. It is easy to see that Y is also uniformly distributed on $(0, 1)$ and

$$X + Y = 0.95 I_{(0 < X \leq 0.95)} + 1.95 I_{(0.95 < X < 1)}. \quad (3.34)$$

Equation (3.34) indicates that $X + Y$ follows a discrete law with only two jumps:

$$\Pr(X + Y = 0.95) = 1 - \Pr(X + Y = 1.95) = 0.95.$$

For $p = 0.90$, by formula (3.10) one easily checks that

$$\text{CTE}_p[X + Y] = 1.95, \quad \text{CTE}_p[X] = \text{CTE}_p[Y] = 0.95.$$

Hence

$$\text{CTE}_p[X + Y] > \text{CTE}_p[X] + \text{CTE}_p[Y].$$

▽

In the following remarks we show that both the quantile risk measure and ESF are not subadditive.

Remark 3.3.5 (VaR is not subadditive)

Let X and Y be i.i.d. random variables which are Bernoulli (0.02) distributed. We immediately find that $Q_{0.975}[X] = Q_{0.975}[Y] = 0$. On the other hand, $\Pr(X + Y = 0) = 0.9604$, which implies that $Q_{0.975}[X + Y] > 0$. As another illustration of the fact that the quantile risk measure is not subadditive, consider a bivariate normal random vector (X, Y) . One can easily prove that the cumulative distribution functions of $X + Y$ and $X^c + Y^c$ only cross once, in $(\mu_X + \mu_Y, 0.5)$. This implies that $Q_p[X + Y] > Q_p[X] + Q_p[Y]$ if $p < 0.5$, whereas $Q_p[X + Y] < Q_p[X] + Q_p[Y]$ if $p > 0.5$. ▽

Remark 3.3.6 (ESF is not subadditive)

Let X and Y be i.i.d. random variables which are Bernoulli (0.02) distributed. It is straightforward to prove that $\text{ESF}_{0.99}[X] = 0$ but $\text{ESF}_{0.99}[X + Y] > 0$. ▽

Remark 3.3.7 (Translation-scale invariant distributions)

The cumulative distribution functions of the risks $X_1, X_2, \dots, X_n$ are said to belong to the same translation-scale invariant family of distributions if there exists a random variable Y , positive real constants a_i and real constants b_i such that X_i has the same distribution as $a_i Y + b_i$ for each $i = 1, 2, \dots, n$. Examples of translation-scale invariant families of distributions are normal distributions, or more generally, elliptical distributions with the same characteristic generator, see e.g. Valdez & Dhaene (2003). Now assume that the risk measure ρ preserves stop-loss order and that $\rho[aX + b] = a \rho[X] + b$ for any positive real number a and any real number b . It is easy to prove that if the set of risks is restricted to a translation-scale invariant family, then the risk measure ρ is subadditive in this family. ∇

3.4 Distortion risk measures**3.4.1 Definition, examples and properties**

In this section we will consider the class of distortion risk measures, introduced by Wang (1996). The quantile risk measure and TVaR belong to this class. A number of the properties of these risk measures can be generalized to the class of distortion risk measures.

From (1.1) we have that the expectation of X , if it exists, can be written as

$$\mathbb{E}[X] = - \int_{-\infty}^0 [1 - \bar{F}_X(x)] dx + \int_0^{\infty} \bar{F}_X(x) dx. \quad (3.35)$$

Wang (1996) defines a family of risk measures by using the concept of distortion function as introduced in Yaari's dual theory of choice under risk, see also Wang & Young (1998). A distortion function is defined as a non-decreasing function $g : [0, 1] \rightarrow [0, 1]$ such that $g(0) = 0$ and $g(1) = 1$. The distortion risk measure associated with distortion function g is denoted by $\rho_g[\cdot]$ and is defined by

$$\rho_g[X] = - \int_{-\infty}^0 [1 - g(\bar{F}_X(x))] dx + \int_0^{\infty} g(\bar{F}_X(x)) dx, \quad (3.36)$$

for any random variable X . Note that the distortion function g is assumed to be independent of the cumulative distribution function of the random variable X . The distortion function $g(q) = q$ corresponds to $\mathbb{E}[X]$. Note that if $g(q) \geq q$ for all $q \in [0, 1]$, then $\rho_g[X] \geq \mathbb{E}[X]$. In particular this result holds in case g is a concave distortion function. Also note that $g_1(q) \leq g_2(q)$ for all $q \in [0, 1]$ implies that $\rho_{g_1}[X] \leq \rho_{g_2}[X]$.

One immediately finds that $g(\bar{F}_X(x))$ is a non-increasing function of x with values in the interval $[0, 1]$. However $\rho_g[X]$ cannot always be considered as the expectation of X under a new probability measure, because $g(\bar{F}_X(x))$ will not necessarily be right-continuous. For a general distortion function g , the risk

measure $\rho_g[X]$ can be interpreted as a ‘distorted expectation’ of X , evaluated with a ‘distorted probability measure’ in the sense of a Choquet-integral, see Denneberg (1994). Substituting $g(\bar{F}_X(x))$ by $\int_0^{\bar{F}_X(x)} dg(q)$ in (3.36) and reverting the order of the integrations, one finds that any distortion risk measure $\rho_g[X]$ can be written as

$$\rho_g[X] = \int_0^1 Q_{1-q}[X] dg(q). \quad (3.37)$$

Notice that when the distortion function g is differentiable, (3.37) can be rewritten as

$$\rho_g[X] = E[Q_U[X] g'(U)]. \quad (3.38)$$

From (3.37), one can easily verify that the quantile $Q_p[X]$, $p \in (0, 1)$, corresponds to the distortion function

$$g(x) = I_{(x > 1-p)}, \quad 0 \leq x \leq 1. \quad (3.39)$$

On the other hand, $\text{TVaR}_p[X]$, $p \in (0, 1)$, corresponds to the distortion function

$$g(x) = \min\left(\frac{x}{1-p}, 1\right), \quad 0 \leq x \leq 1. \quad (3.40)$$

Remark 3.4.1 (ESF is not a distortion risk measure)

Let us assume that the risk measure $\text{ESF}_p[X]$ can be expressed as (3.36) for some distortion function g . We first substitute to (3.36) a risk variable X with a uniform distribution on $(0, 1)$. Hence, for the given $p \in (0, 1)$, we have

$$\frac{1}{2}(1-p)^2 = \int_0^1 g(s) ds. \quad (3.41)$$

We then substitute to (3.36) a Bernoulli variable Y with

$$\Pr(Y = 0) = 1 - \Pr(Y = 1) = 1 - r,$$

for some arbitrarily but fixed $0 < r \leq 1 - p$. We easily obtain that $g(r) = r$ for $0 < r \leq 1 - p$. From (3.41) we find that

$$\frac{1}{2}(1-p)^2 = \int_0^{1-p} s ds + \int_{1-p}^1 g(s) ds \geq \frac{1}{2}(1-p)^2 + p(1-p),$$

which is obviously a contradiction since $0 < p < 1$. This illustrates that ESF is not a distortion risk measure. ∇

From (3.11) and the fact that $\text{TVaR}_p[X]$, $p \in (0, 1)$, corresponds to the distortion function given in (3.40), we find that $\text{CTE}_p[X]$, $p \in (0, 1)$, can be written in the form $\rho_g[X]$ with g given by

$$g(x) = \min\left(\frac{x}{1 - F_X(Q_p[X])}, 1\right), \quad 0 \leq x \leq 1. \quad (3.42)$$

This function g , however, depends on the cumulative distribution function of X ; hence we cannot infer that $\text{CTE}_p[\cdot]$ is a distortion risk measure. Actually, we can obtain the following result.

Remark 3.4.2 (CTE is not a distortion risk measure)

Using the same approach as in Remark 3.4.1, we assume by contradiction that the risk measure $\text{CTE}_p[X]$ can be expressed as (3.36) for some distortion function g . We first substitute to (3.36) a risk variable X with a uniform distribution on $(0, 1)$. Hence, for the given $p \in (0, 1)$, recalling (3.10), we find

$$p + \frac{1}{2}(1 - p) = \int_0^1 g(1 - x) dx.$$

Simplification on the above equation leads to

$$\int_0^1 g(x) dx = \frac{1}{2}(1 + p). \quad (3.43)$$

We then substitute to (3.36) another risk variable. To do this, we choose a Bernoulli variable Y with

$$\Pr(Y = 0) = 1 - \Pr(Y = 1) = 1 - r,$$

for an arbitrarily but fixed $0 < r \leq 1 - p$. Again by applying (3.10) we obtain that $\text{CTE}_p[Y] = 1$. Hence, by (3.36) it should hold that $g(r) = 1$. By virtue of the monotonicity of the distortion function g and the arbitrariness of $0 < r \leq 1 - p$ we conclude that $g(\cdot) \equiv 1$ on $(0, 1]$, which contradicts the equation (3.43). This illustrates that $\text{CTE}_p[X]$ is not a distortion risk measure. ∇

Example 3.4.1 (The Wang transform risk measure)

From (3.5), we see that the TVaR_p risk measure only uses the upper tail of the distribution. Hence, this risk measure does not create incentive for taking actions that increase the cumulative distribution function for outcomes smaller than Q_p . Also, from (3.9) we see that TVaR_p only accounts for the expected shortfall and hence, does not properly adjust for extreme low-frequency and high severity losses. The Wang Transform risk measure was introduced by Wang (2000) as an example of a risk measure that could give a solution to these problems. For any $0 < p < 1$, define the distortion function

$$g_p(x) = \Phi[\Phi^{-1}(x) + \Phi^{-1}(p)], \quad 0 \leq x \leq 1, \quad 0 < p < 1, \quad (3.44)$$

which is called the ‘Wang Transform at level p ’. The corresponding distortion risk measure is called the ‘Wang Transform risk measure’ and is denoted by $\text{WT}_p[X]$.

For a normally distributed random variable X , we find

$$1 - g_p(\overline{F}_X(x)) = \Phi\left[\frac{x - Q_p[X]}{\sigma}\right]$$

which implies that the ‘Wang Transform risk measure’ is identical to the quantile risk measure at the same probability level in case of a normal random variable:

$$\text{WT}_p[X] = Q_p[X]. \quad (3.45)$$

For a lognormal distributed random variable Y with parameters μ and σ^2 , we find

$$\text{WT}_p(Y) = Q_{\Phi[\Phi^{-1}(p) + \sigma/2]}[Y], \quad (3.46)$$

which is larger than $Q_p[Y]$.

Examples illustrating the fact that the WT risk measure uses the whole distribution and that it accounts for extreme low-frequency and high severity losses can be found in Wang (2001). ∇

It is easy to prove that any distortion risk measure ρ_g obeys the following properties, see also Wang (1996):

- **Additivity for comonotonic risks:**

For any distortion function g and all random variables X_i ,

$$\rho_g[X_1^c + X_2^c + \dots + X_n^c] = \sum_{i=1}^n \rho_g[X_i]. \quad (3.47)$$

- **Positive homogeneity:**

For any distortion function g , any random variable X and any non-negative constant a , we have

$$\rho_g[aX] = a\rho_g[X]. \quad (3.48)$$

- **Translation invariance:**

For any distortion function g , any random variable X and any constant b , we have

$$\rho_g[X + b] = \rho_g[X] + b. \quad (3.49)$$

- **Monotonicity:**

For any distortion function g and any two random variables X and Y where $X \leq Y$, we have

$$\rho_g[X] \leq \rho_g[Y] \quad (3.50)$$

The first property follows immediately from (3.37) and the additivity property of quantiles for comonotonic risks. The second and the third properties follow from (3.37) and Lemma 1.3.1. The fourth property follows from (3.37) and the fact that $X \leq Y$, implies that each quantile of Y exceeds the corresponding quantile of X . Note that in the literature the property of positive homogeneity

is often wrongly explained as ‘currency independence’. Take as an example the risk measure

$$\rho[X] = E[(X - d)_+], \quad (3.51)$$

where clearly X and d have to be expressed in the same monetary unit. This risk measure is not positive homogeneous but it is currency independent, see also Remark 3.5 in Goovaerts, Kaas, Dhaene & Tang (2003).

In the following theorem, stochastic dominance is characterized in terms of ordered distortion risk measures.

Theorem 3.4.1 (Stochastic dominance and distortion risk measures)

For any random pair (X, Y) we have that X is smaller than Y in stochastic dominance sense if and only if their respective distortion risk measures are ordered:

$$X \leq_{st} Y \Leftrightarrow \rho_g[X] \leq \rho_g[Y] \text{ for all distortion functions } g. \quad (3.52)$$

Proof. This follows immediately from (3.37) and Theorem 3.3.1. ■

3.4.2 Concave distortion risk measures

A subclass of distortion functions that is often considered in the literature is the class of concave distortion functions. A distortion function g is said to be concave if for each q in $(0, 1]$, there exist real numbers a_y and b_y and a line $l(x) = a_y x + b_y$, such that $l(q) = g(q)$ and $l(q) \geq g(q)$ for all q in $(0, 1]$. A concave distortion function is necessarily continuous in $(0, 1]$. For convenience, we will always tacitly assume that a concave distortion function is also continuous at 0. A risk measure with a concave distortion function is then called a ‘concave distortion risk measure’.

For any concave distortion function g , we have that $g(\overline{F}_X(x))$ is right-continuous, so that in this case the risk measure $\rho_g[X]$ can be interpreted as the expectation of X under a ‘distorted probability measure’. Note that the quantile risk measure is not a concave distortion risk measure whereas TVaR is a concave distortion risk measure.

In the following theorem, we show that stop-loss order can be characterized in terms of ordered concave distortion risk measures.

Theorem 3.4.2 (Stop-loss order and concave distortion risk measures)

For any random pair (X, Y) we have that $X \leq_{sl} Y$ if and only if their respective concave distortion risk measures are ordered:

$$X \leq_{sl} Y \Leftrightarrow \rho_g[X] \leq \rho_g[Y] \text{ for all concave distortion functions } g. \quad (3.53)$$

Proof. The “ $\Leftarrow$ ” implication follows immediately from Theorem 3.3.2. To complete the proof of Theorem 3.4.2, we first prove the “ $\Rightarrow$ ” implication for concave piecewise linear distortion functions g . Any such distortion function can be written in the form

$$g(x) = \sum_{i=1}^n a_i (\beta_i - \beta_{i+1}) \min(x/a_i, 1)$$

where $0 = a_0 < a_1 < \dots < a_{n-1} < a_n = 1$. Further, β_i is the derivative of g in the interval (a_{i-1}, a_i) and $\beta_{n+1} = 0$. Because of the concavity of g , we have that β_i is a decreasing function of i . From (3.37) and (3.40), it follows that the related risk measure $\rho_g[X]$ can be written as

$$\rho_g[X] = \sum_{i=1}^n a_i (\beta_i - \beta_{i+1}) \text{TVaR}_{1-a_i}[X].$$

In view of Theorem 3.3.2, we find that $X \leq_{sl} Y$ implies $\rho_g[X] \leq \rho_g[Y]$ for all concave piecewise linear distortion functions g .

Now we are able to prove the “ $\Rightarrow$ ” for general concave distortion functions g . If $\rho_g[Y] = \infty$, the result is obvious. Let us now assume that $\rho_g[Y] < \infty$. The concave distortion function g can be approximated from below by concave piecewise linear distortion functions g_n such that for any $x \in [0, 1]$, we have that $g_1(x) \leq g_2(x) \leq \dots \leq g_n(x) \leq \dots \leq g(x)$ and $\lim_{n \rightarrow \infty} g_n(x) = g(x)$. As we have just proven, the inequality $X \leq_{sl} Y$ implies $\rho_{g_n}[X] \leq \rho_{g_n}[Y]$ for all n . Further, $g_n(x) \leq g(x)$ implies $\rho_{g_n}[Y] \leq \rho_g[Y] < \infty$. From the monotone convergence theorem we find that $\lim_{n \rightarrow \infty} \rho_{g_n}[X] = \rho_g[X]$, so that we can conclude that $\rho_g[X] \leq \rho_g[Y]$. ■

Proofs for the theorem above can also be found in Yaari (1987), Wang & Young (1998) or Dhaene, Wang, Young & Goovaerts (2000).

Example 3.4.2 (The Beta distortion risk measure)

We will write $X <_{sl} Y$ if $X \leq_{sl} Y$ and $E[(X - d)_+] < E[(Y - d)_+]$ for at least one retention d . Wirth & Hardy (2000) give the following example that illustrates that $\text{TVaR}_{0.95}$ does not strongly preserve stop-loss order.

Let $\Pr[X = (0, 1, 2)] = (0.95, 0.025, 0.025)$ and $\Pr[Y = (1, 2)] = (0.975, 0.025)$. It is easy to verify that $X <_{sl} Y$ and $\text{TVaR}_{0.95}[X] = \text{TVaR}_{0.95}[Y] = 1.5$. This means that there exist random variables X and Y such that $X <_{sl} Y$ but $\text{TVaR}_{0.95}[X] = \text{TVaR}_{0.95}[Y]$.

More generally, they prove that any distortion risk measure derived from a distortion function g which is concave but not strictly concave (i.e. g has a linear part) does not strongly preserve stop-loss order. They also prove that for any distortion function g which is strictly concave, one has that $X <_{sl} Y$ implies $\rho_g[X] < \rho_g[Y]$. Wirth & Hardy (2000) start from the Beta cumulative distribution function

$$F_\beta(x) = \frac{1}{\beta(a, b)} \int_0^x t^{a-1} (1-t)^{b-1} dt, \quad 0 \leq x \leq 1, \quad (3.54)$$

where $\beta(a, b)$ is the Beta function with parameters $a > 0, b > 0$, i.e.

$$\beta(a, b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}, \quad (3.55)$$

to define the Beta distortion function

$$g(x) = F_\beta(x), \quad 0 \leq x \leq 1. \quad (3.56)$$

The Beta distortion function is strictly concave for any parameters $0 < a \leq 1$ and $b \geq 1$, provided a and b are not both equal to 1. This implies that the risk measure derived from the distortion function $F_\beta(q)$ strictly preserves stop-loss order. For random variables X and Y as defined above, and parameters $a = 0.1$ and $b = 1$, we find

$$\rho_{F_\beta}[X] = 1.4326 < \rho_{F_\beta}[Y] = 1.6915.$$

Note that the Beta distortion risk measure with the parameters $a = 0.1$ and $b = 1$ reduces to the PH-transform risk measure, which is considered in Wang (1995). ∇

Concave distortion risk measures are subadditive, which means that the risk measure for a sum of random variables is smaller than or equivalent to the sum of the risk measures.

• **Subadditivity:**

For any concave distortion function g , and any two random variables X and Y , we have

$$\rho_g[X + Y] \leq \rho_g[X] + \rho_g[Y]. \quad (3.57)$$

The proof follows immediately from Theorem 3.3.4, see also Wang & Dhaene (1998).

In Artzner (1999) and Artzner, Delbaen, Eber & Heath (1999) a risk measure satisfying the four axioms of subadditivity, monotonicity, positive homogeneity and translation invariance is called ‘coherent’. As we have proven, any concave distortion risk measure is coherent. As the quantile risk measure is not subadditive, it is not a ‘coherent’ risk measure. We believe the terminology ‘coherent’ is misleading as it gives the impression that any risk measure that is not ‘coherent’ is always wrong. We believe that the ‘coherent’ set of axioms might be appropriate in one situation but inadequate in another⁴, see Dhaene, Laeven, Vanduffel, Darkiewicz and Goovaerts (2004).

Note that the class of concave distortion risk measures is only a subset of the class of ‘coherent’ risk measures, as is shown by the following example.

Example 3.4.3 (The Dutch risk measure)

For any random variable X , consider the risk measure

$$\rho[X] = E[X] + \theta E[(X - \alpha E[X])_+], \quad \alpha \geq 1, \quad 0 \leq \theta \leq 1. \quad (3.58)$$

⁴ “There is no royal path which leads to geometry” - Euclid.

We will call this risk measure the ‘Dutch risk measure’, because for non-negative random variables, it is called the ‘Dutch premium principle’, see Kaas, van Heerwaarden & Goovaerts (1994).

In the sequel of this example we assume that the parameters α and θ are both equal to 1. In this case the Dutch risk measure is coherent. Indeed, the verifications of the properties of positive homogeneity, translation invariance and subadditivity are immediate. Finally, if $X \leq Y$, then $E[X] \leq E[Y]$, so that the property of monotonicity follows from

$$\rho[X] = E[\max(E[X], X)] \leq E[\max(E[Y], Y)] = \rho[Y].$$

Next, we will prove that the Dutch risk measure $\rho(\cdot)$ is in general not additive for comonotonic risks. Let (X_1^c, X_2^c) be a comonotonic random couple with Bernoulli marginal distributions: $\Pr[X_i = 1] = q_i$ with $0 < q_1 < q_2 < 1$ and $q_1 + q_2 > 1$. After some straightforward computations, we find

$$\rho[X_i] = q_i(2 - q_i), \quad i = 1, 2,$$

and

$$\rho[X_1^c + X_2^c] = 2q_1 + (1 - q_1)(q_1 + q_2),$$

from which we can conclude that the Dutch premium principle is in general not additive for comonotonic risks. Hence, the Dutch risk measure (with parameters equal to 1) is an example of a risk measure that is coherent, although it is not a distortion risk measure. The example also illustrates the fact that coherent risk measures are not necessarily additive for comonotonic risks. ∇

As we have seen, the quantile risk measure Q_p is not a concave distortion risk measure. The following theorem states that in the class of concave distortion risk measures, the one that leads to the minimal extra-capital compared to the quantile risk measure at probability level p is the TVaR risk measure at the same level p .

Theorem 3.4.3 (Characterization of TVaR) *For any $0 < p < 1$ and for any random variable X one has*

$$\text{TVaR}_p[X] = \min \{ \rho_g[X] \mid g \text{ is concave and } \rho_g \geq Q_p \}. \quad (3.59)$$

Proof. The distortion risk measure TVaR_p has a concave distortion function $\min\left(\frac{x}{1-p}, 1\right)$. Further, $\text{TVaR}_p \geq Q_p$. This implies that

$$\text{TVaR}_p(X) \geq \inf \{ \rho_g(X) \mid g \text{ is concave and } \rho_g \geq Q_p \}.$$

In order to prove the opposite inequality, consider a concave distortion function g such that $\rho_g(Y) \geq Q_p[Y]$ holds for all random variables Y . For any q with $1 - p < q < 1$, we define the Bernoulli random variable Y_q with

$$\Pr(Y_q = 1) = q.$$

It is easy to verify that $Q_p[Y_q] = 1$, and also $\rho_g(Y_q) = g(q)$. As $g(x) \leq 1$, we find that the condition $\rho_g(Y_q) \geq Q_p[Y_q]$ can be rewritten as $g(q) = 1$. This means that g is equal to 1 on the interval $(1 - p, 1]$. As g is concave, it follows immediately that

$$g(x) \geq \min\left(\frac{x}{1-p}, 1\right), \quad 0 < x < 1.$$

Hence,

$$\rho_g(X) \geq \text{TVaR}_p(X)$$

holds for all concave distortion risk measures g for which $\rho_g \geq Q_p$. This implies

$$\text{TVaR}_p(X) = \inf \{\rho_g(X) \mid g \text{ is concave and } \rho_g \geq Q_p\}.$$

which completes the proof. ■

A result with a taste similar to our Theorem 3.4.3 is Proposition 5.2 in Artzner, Delbaen, Eber & Heath (1999), which says that

$$\text{VaR}_p[X] = \inf \{\rho[X] \mid \rho \text{ coherent and } \rho \geq Q_p\}$$

holds for each risk variable X , see also Proposition 3.3 in Artzner (1999).

3.4.3 Risk measures for comonotonic sums

In this subsection, we will consider the problem of finding approximations for distorted expectations (such as quantiles and TVaR's) of a sum $S = \sum_{i=1}^n X_i$ of which the marginal distributions of the random variables X_i are given, but the dependency structure between the X_i is unknown or too cumbersome to work with. In view of Corollary 2.5.1, we propose to approximate (the c.d.f.) of S by (the c.d.f. of) $S^c = \sum_{i=1}^n F_{X_i}^{-1}(U)$ or (the c.d.f. of) $S^l = \sum_{i=1}^n E[X_i \mid \Lambda]$, and approximate $\rho_g[S]$ by $\rho_g[S^c]$ or by $\rho_g[S^l]$. Note that S^c is a comonotonic sum, hence from the additivity property for comonotonic risks we find

$$\rho_g[S^c] = \sum_{i=1}^n \rho_g[X_i]. \quad (3.60)$$

On the other hand, if the conditioning random variable Λ is such that all $E[X_i \mid \Lambda]$ are non-decreasing functions of Λ (or all are non-increasing functions of Λ), then S^l is a comonotonic sum too. Hence, in this case

$$\rho_g[S^l] = \sum_{i=1}^n \rho_g[E[X_i \mid \Lambda]]. \quad (3.61)$$

In case of a concave distortion function g , we find from Theorem 3.4.2 that $\rho_g[S^l]$ is a lower bound whereas $\rho_g[S^c]$ is an upper bound for $\rho_g[S]$:

$$\rho_g[S^l] \leq \rho_g[S] \leq \rho_g[S^c]. \quad (3.62)$$

In particular, we have that

$$\text{TVaR}_p[S^l] \leq \text{TVaR}_p[S] \leq \text{TVaR}_p[S^c]. \quad (3.63)$$

Note that the quantiles of S^l , S and S^c are not necessarily ordered in the same way.

Example 3.4.4 (Sums of lognormals)

Consider the sum

$$S = \sum_{i=1}^n \alpha_i e^{Z_i} \quad (3.64)$$

where the α_i are non-negative constants and the random vector $(Z_1, Z_2, \dots, Z_n)$ is assumed to have a multivariate normal distribution.

Let U be uniformly distributed on the unit interval. We find that the comonotonic upper bound $S^c = \sum_{i=1}^n F_{\alpha_i e^{Z_i}}^{-1}(U)$ of S is given by

$$S^c = \sum_{i=1}^n \alpha_i e^{\mathbb{E}[Z_i] + \sigma_{Z_i} \Phi^{-1}(U)}. \quad (3.65)$$

The upper bound S^c is obtained by replacing the original copula between the marginals of the sum S by the comonotonic copula, but keeping the marginal distributions unchanged.

From Theorem 3.3.3 and Example 3.2.2, we find the following expressions for the risk measures associated with S^c :

$$Q_p[S^c] = \sum_{i=1}^n \alpha_i e^{\mathbb{E}[Z_i] + \sigma_{Z_i} \Phi^{-1}(p)}, \quad (3.66)$$

$$\text{CTE}_p[S^c] = \sum_{i=1}^n \alpha_i e^{\mathbb{E}[Z_i] + \frac{1}{2}\sigma_{Z_i}^2} \frac{\Phi(\sigma_{Z_i} - \Phi^{-1}(p))}{1-p}, \quad p \in (0, 1) \quad (3.67)$$

where in deriving (3.67) we have used the fact that the CTE is additive for comonotonic risks with continuous marginal distributions, recall Remark 3.3.2 for details. From (3.12), Theorem 3.3.3 and Example 3.2.2 we also find

$$\text{CLTE}_p[S^c] = \sum_{i=1}^n \alpha_i e^{\mathbb{E}[Z_i] + \frac{1}{2}\sigma_{Z_i}^2} \frac{1 - \Phi(\sigma_{Z_i} - \Phi^{-1}(p))}{p}, \quad p \in (0, 1) \quad (3.68)$$

From (2.16) and (2.17) one can prove that S^c has a strictly increasing cumulative distribution function. This implies that in the expression for $\text{CLTE}_p[S^c]$, $Q_p^+[S^c]$ can be replaced by $Q_p[S^c]$. Finally, from (3.7) and (3.10) it follows that for $d > 0$,

$$\mathbb{E}[S^c - d]_+ = \sum_{i=1}^n \alpha_i e^{\mathbb{E}[Z_i] + \frac{1}{2}\sigma_{Z_i}^2} \Phi(\sigma_{Z_i} - \Phi^{-1}(p)) - d(1-p), \quad p \in (0, 1), \quad (3.69)$$

with p uniquely determined such that $Q_p[S^c] = d$. This holds true since S^c has a strictly increasing cumulative distribution function.

In order to define a stochastic lower bound for S , we choose a conditioning random variable Λ which is a linear combination of the Z_j :

$$\Lambda = \sum_{j=1}^n \gamma_j Z_j. \quad (3.70)$$

After some computations, we find that the lower bound $S^l = \sum_{i=0}^n \alpha_i E[e^{Z_i} | \Lambda]$ is given by

$$S^l = \sum_{i=1}^n \alpha_i e^{E[Z_i] + \frac{1}{2}(1-r_i^2)\sigma_{Z_i}^2 + r_i \sigma_{Z_i} \Phi^{-1}(U)}, \quad (3.71)$$

where the uniformly distributed random variable U follows from $\Phi^{-1}(U) \equiv \frac{\Lambda - E(\Lambda)}{\sigma_\Lambda}$, and r_i is the correlation between Z_i and Λ . The lower bound is obtained by changing both the copula and the marginals of the original sum. Intuitively, one can expect that an appropriate choice of the conditioning variable Λ will lead to much better approximations than the upper bound approximations.

If all r_i are positive, then S^l is a comonotonic sum, which means that quantiles and conditional tail expectations related to S^l can be computed by summing the associated risk measures for the marginal distributions involved. Assuming that all r_i are positive, we find the following expressions for the risk measures associated with S^l :

$$Q_p[S^l] = \sum_{i=1}^n \alpha_i e^{E[Z_i] + \frac{1}{2}(1-r_i^2)\sigma_{Z_i}^2 + r_i \sigma_{Z_i} \Phi^{-1}(p)}, \quad (3.72)$$

$$\text{CTE}_p[S^l] = \sum_{i=1}^n \alpha_i e^{E[Z_i] + \frac{1}{2}\sigma_{Z_i}^2} \frac{\Phi(r_i \sigma_{Z_i} - \Phi^{-1}(p))}{1-p}, \quad (3.73)$$

$$\text{CLTE}_p[S^l] = \sum_{i=1}^n \alpha_i e^{E[Z_i] + \frac{1}{2}\sigma_{Z_i}^2} \frac{1 - \Phi(r_i \sigma_{Z_i} - \Phi^{-1}(p))}{p}, \quad p \in (0, 1), \quad (3.74)$$

and also for $d > 0$

$$E[S^l - d]_+ = \sum_{i=1}^n \alpha_i e^{E[Z_i] + \frac{1}{2}\sigma_{Z_i}^2} \Phi(r_i \sigma_{Z_i} - \Phi^{-1}(p)) - d(1-p), \quad p \in (0, 1), \quad (3.75)$$

In the last expression p is determined uniquely by $Q_p[S^l] = d$. Indeed, it proves that S^l has a strictly increasing cumulative distribution function.

The correlation coefficients r_i follow from the correlations between the random variables Z_i . Let us now consider the important case that all Z_i are linear combinations of the random normal variables $Y_1, Y_2, \dots, Y_n$:

$$Z_i = \sum_{j=1}^n \lambda_{ij} Y_j. \quad (3.76)$$

The r.v. Λ can now also be defined as a linear combination of the yearly returns:

$$\Lambda = \sum_{j=1}^n \beta_j Y_j. \quad (3.77)$$

The relation between the β_j and the γ_k is given by

$$\beta_j = \sum_{k=1}^n \gamma_k \lambda_{kj}. \quad (3.78)$$

In the special case that all Y_i are i.i.d., we find

$$r_i = \frac{\sum_{j=1}^n \lambda_{ij} \beta_j}{\sqrt{\sum_{j=1}^n \lambda_{ij}^2} \sqrt{\sum_{j=1}^n \beta_j^2}}, \quad i = 1, 2, \dots, n. \quad (3.79)$$

The optimal choice for the coefficients γ_j (or β_j) and the performance of (the quantiles of) S^l and S^c as approximations for (the quantiles of) S are investigated in Chapter 4.

Finally, we also have the following ordering relations

$$\text{CTE}_p[S^l] \leq \text{CTE}_p[S] \leq \text{CTE}_p[S^c], \quad (3.80)$$

$$\mathbb{E}[S^l - d]_+ \leq \mathbb{E}[S - d]_+ \leq \mathbb{E}[S^c - d]_+, \quad (3.81)$$

$$\text{CLTE}_p[S^c] \leq \text{CLTE}_p[S] \leq \text{CLTE}_p[S^l], \quad p \in (0, 1), \quad (3.82)$$

Note however that such an ordering does not hold in general for $Q_p[S]$ and its approximations $Q_p[S^l]$ and $Q_p[S^c]$. One can prove however that the inequality $Q_p[S^l] \leq Q_p[S^c]$ holds, provided p is sufficiently high and the correlations r_i are positive. Indeed, in this case it is straightforward to verify that for all $p \in (0, 1)$, the condition

$$p \geq \Phi \left[\frac{1}{2} \max \{ (1 + r_i) \sigma_{Z_i} ; i = 1, \dots, n \} \right] \equiv p^+ \quad (3.83)$$

implies that $Q_p[S^l] \leq Q_p[S^c]$. Similarly, one can prove that the condition

$$p \leq \Phi \left[\frac{1}{2} \min \{ (1 + r_i) \sigma_{Z_i} ; i = 1, \dots, n \} \right] \equiv p^- \quad (3.84)$$

implies that $Q_p[S^c] \leq Q_p[S^l]$. Note that $p^- \geq 0.5$.

We can conclude that the cumulative distribution functions of S^l and S^c can only cross in the region where their cumulative distribution functions take a value that is contained in the interval $[p^-, p^+]$. In the region $p \leq p^-$, the cumulative distribution function of S^c will lay above the cumulative distribution function of S^l , while in the region $p \geq p^+$, the cumulative distribution function of S^c will lay under the cumulative distribution function of S^l . ∇

3.5 Theories of choice under risk

In Chapter 1 we embedded the actuarial ordering concepts in the expected utility framework. In this theory, a decision maker asserts a utility $u(x)$ to each possible wealth-level x and the real-valued function $u(\cdot)$ is called his utility function. As a rational decision-maker is assumed to prefer more to less, it is assumed that a utility function is non-decreasing. If the decision-maker, with initial wealth w , has to choose between random losses X and Y , then he compares $E[u(w - X)]$ with $E[u(w - Y)]$ and chooses the loss which gives rise to the highest expected utility. Hence, the decision-maker acts in order to maximize his expected utility.

Yaari (1987) presents a dual theory of choice under risk. In this dual theory, the decision-maker has a distortion function f . This ‘distortion function’ can be considered as the parallel to the concept of ‘utility function’ in utility theory. While in utility theory, choosing among risks is performed by comparing expected values of transformed wealth levels (utilities), in Yaari’s theory the quantities that are compared are the distorted expectations of wealth levels. Consider a decision-maker with initial wealth w , which has to choose between two random losses X and Y . The decision-maker acts in order to maximize his distorted expectation. Hence, he will prefer loss X over loss Y if and only if $\rho_f[w - X] \geq \rho_f[w - Y]$, where ρ_f is the distortion risk measure associated with the distortion function f . Comparing the expression

$$E[w - X] = \int_0^1 Q_{1-q}[w - X] dq \quad (3.85)$$

with

$$E[u(w - X)] = \int_0^1 u(Q_{1-q}[w - X]) dq \quad (3.86)$$

and

$$\rho_f[w - X] = \int_0^1 Q_{1-q}[w - X] df(q), \quad (3.87)$$

we see that both the expressions (3.86) and (3.87) transform the expected wealth level $E[w - X]$. Under the expected utility hypothesis, the possible levels-of-wealth are adjusted by a utility function, whereas under the distorted expectation hypothesis, the probabilities are adjusted. It is well-known that stochastic dominance and stop-loss order have a natural interpretation in terms of expected utility theory. The pairs of losses X and Y with $X \leq_{st} Y$ are exactly those pairs of losses about which all decision-makers (with a non-decreasing) utility function agree:

$$X \leq_{st} Y \Leftrightarrow E[u(w - X)] \geq E[u(w - Y)] \text{ for all utility functions } u. \quad (3.88)$$

In expected utility theory, a decision-maker is said to be risk averse if his utility function is concave. Stop-loss order represents the common preferences of all risk averse decision-makers:

$$X \leq_{sl} Y \Leftrightarrow E[u(w - X)] \geq E[u(w - Y)] \text{ for all concave utility functions } u. \quad (3.89)$$

For more details about actuarial applications of expected utility theory and its relation to ordering of random variables, see e.g. Kaas, Goovaerts, Dhaene & Denuit (2001).

On the other hand, one has that

$$X \leq_{st} Y \Leftrightarrow \rho_f[w - X] \geq \rho_f[w - Y] \text{ for all distortion functions } f. \quad (3.90)$$

Hence, stochastic dominance of loss Y over loss X holds if and only if all decision-makers in Yaari's dual theory of choice under risk prefer loss X over loss Y . This characterization for stochastic dominance follows from Theorem 3.4.1 by introducing the 'dual distortion function' $\bar{f}$ for each distortion function f :

$$\bar{f}(x) = 1 - f(1 - x), \quad x \in [0, 1]. \quad (3.91)$$

The dual distortion function is again a distortion function. It is clear that $\bar{\bar{f}} \equiv f$. Furthermore, we have that

$$\rho_f[-X] = -\rho_{\bar{f}}[X], \quad (3.92)$$

and therefore that

$$\rho_f[w - X] \geq \rho_f[w - Y] \Leftrightarrow \rho_{\bar{f}}[X] \leq \rho_{\bar{f}}[Y]. \quad (3.93)$$

The former relation (3.92) can easily be proven from the definition (3.36). Actually, we have

$$\begin{aligned} \rho_f[-X] &= - \int_{-\infty}^0 [1 - f(\Pr(-X > x))] dx + \int_0^{\infty} f(\Pr(-X > x)) dx \\ &= - \int_{-\infty}^0 \bar{f}(\Pr(-X \leq x)) dx + \int_0^{\infty} [1 - \bar{f}(\Pr(-X \leq x))] dx. \end{aligned}$$

Substituting $s = -x$ gives that

$$\begin{aligned} \rho_f[-X] &= - \int_0^{\infty} \bar{f}(\Pr(-X \leq -s)) ds + \int_{-\infty}^0 [1 - \bar{f}(\Pr(-X \leq -s))] ds \\ &= \int_{-\infty}^0 [1 - \bar{f}(\Pr(X \geq s))] ds - \int_0^{\infty} \bar{f}(\Pr(X \geq s)) ds \\ &= \int_{-\infty}^0 [1 - \bar{f}(\Pr(X > s))] ds - \int_0^{\infty} \bar{f}(\Pr(X > s)) ds, \end{aligned}$$

where in the last we used the fact that the Lebesgue measure of the set of all discontinuities of a monotone function is 0. This proves the stated result (3.92).

In Yaari's dual theory of choice under risk, a decision-maker is said to be risk averse if his distortion function is convex. Here we will tacitly assume that a convex distortion function is continuous on $[0, 1]$. This means that a risk averse decision-maker systematically underestimates his tail probabilities $g(\bar{F}_{w-X}(x))$

related to levels-of-wealth, which is a prudent attitude. One finds that stop-loss order of loss Y over loss X can be characterized as follows:

$$X \leq_{sl} Y \Leftrightarrow \rho_f[w - X] \geq \rho_f[w - Y] \text{ for all convex distortion functions } f. \quad (3.94)$$

Hence, also in Yaari's dual theory of choice under risk, stop-loss order represents the common preferences of all risk averse decision-makers. This characterization for stop-loss order follows from Theorem 3.4.2 by noting that a distortion function is convex if and only if its dual distortion function $\bar{f}$ is concave. A proof for these characterizations in case of non-negative random variables can be found e.g. in Wang & Young (1998), see also Dhaene, Wang, Young & Goovaerts (2000). Note that the relation between theories of choice under risk and distortion risk measures is also investigated in Denuit, Dhaene & Van Wouwe (1999) and Tsanakas & Desli (2003).

The zero utility risk measure $\rho(X)$ associated with a utility function u is the solution of the following indifference equation:

$$u(0) = E[u(\rho[X] - X)], \quad (3.95)$$

which has an intuitive interpretation in terms of utility theory, see e.g. Kaas, Goovaerts, Dhaene & Denuit (2001). An interesting risk measure arises when the utility function is of the exponential type:

$$u(x) = \frac{1}{\alpha} (1 - e^{-\alpha x}). \quad (3.96)$$

In this case we find

$$\rho(X) = \frac{1}{\alpha} \ln E[e^{\alpha X}]. \quad (3.97)$$

The class of distortion risk measures can be considered as Yaari's equivalent of the class of zero-utility risks measures in expected utility theory. Indeed, the solution of the indifference equation

$$\rho_f(0) = \rho_f[\rho[X] - X] \quad (3.98)$$

is given by

$$\rho[X] = \rho_{\bar{f}}[X]. \quad (3.99)$$

Hence, any (concave) distortion risk measure $\rho_g[X]$ can be considered as the solution of the indifference equation (3.98) of a (risk averse) decision-maker with distortion function $f(x) = \bar{g}(x)$.

Tsanakas & Desli (2003) introduce a class of risk measures which can be considered as the solutions of the indifference equations in 'Generalized Expected Utility Theory'. This theory combines both above mentioned theories of choice under risk, see Quiggin (1993).

3.6 Final remarks

In this chapter we examined and summarized properties of several well-known risk measures that can be used to facilitate decision-taking. Special attention was given to the class of (concave) distortion risk measures. We investigated the relationship between these risk measures and theories of choice under risk. We considered the problem of how to evaluate risk measures for sums of non-independent random variables. Approximations for these risk measures, based on the concept of comonotonicity, were proposed. Several examples were provided to illustrate properties or to prove that certain properties do not hold.

Several of the results presented in this chapter for (log)normal random variables can be generalized to the class of (log)elliptical distributions, see Valdez & Dhaene (2003).

Comparing approximations for sums of lognormals

4.1 Introduction

In this chapter we will consider and compare the performance of approximations for the c.d.f. and the risk measures related to a r.v. S given by

$$S = \sum_{i=1}^n \alpha_i e^{Z_i}. \quad (4.1)$$

Here, the α_i are non-negative real numbers and $(Z_1, Z_2, \dots, Z_n)$ is a multivariate normal distributed random vector.

The r.v. (4.1) plays a central role in the actuarial and financial theory. This is because the accumulated value at time n of a series of future deterministic saving amounts α_i can be written in the form (4.1), where Z_i denotes the random accumulation factor over the period $[i, n]$. Also the present value of a series of future deterministic payments α_i can be written in the form (4.1), where now Z_i denotes the random discount factor over the period $[0, i]$. Hence, the valuation of Asian or basket options in a Black & Scholes model and the setting of provisions and required capitals in an insurance context boils down to the evaluation of risk measures related to the cumulative distribution function of a random variable S as defined in (4.1) and this underlines the importance of the r.v. S . The application part of this work (Chapter 6-8), will also essentially deal with this r.v. We will therefore investigate how to (approximately) compute risk measures such as quantiles (Q) and conditional tail expectations (CTE) of the r.v. S . Notice that the quantile risk measure is often called the Value-at-Risk,

⁵ “Over the last few weeks I have become even more convinced than ever, about the remarkable accuracy of the comonotonic lower bound approximation” - Moshe Milevsky, Executive Director of the Individual Finance and Insurance Decisions Centre at the Fields Institute in Toronto (11/3/04).

whereas the conditional tail expectation coincides with the Tail-Value-at-Risk. The latter holds true because S is a continuous r.v., see (3.15).

The r.v. S defined in (4.1) will in general be a sum of non-independent lognormal r.v.'s. Its c.d.f. cannot be determined analytically and is too cumbersome to work with. In the literature, a variety of approximation techniques for this c.d.f. has been proposed.

Practitioners often use a moment matching *lognormal approximation* for the distribution of S . The lognormal approximation is chosen such that its first two moments are equal to the corresponding moments of S . The present value of a continuous perpetuity with lognormal return process has a reciprocal Gamma distribution, see for instance Milevsky (1997). This present value can be considered as the limiting case of a random variable S as defined above.

Motivated by this observation, Milevsky & Posner (1998) and Milevsky & Robinson (2000) propose a moment matching *reciprocal Gamma approximation* for the c.d.f. of S such that the first two moments match. They use this technique for deriving closed form approximations for the price of Asian and basket options.

In Example 3.4.4, see also Dhaene, Denuit, Goovaerts, Kaas & Vyncke (2002a,b) we derived *comonotonic upper bound and lower bound approximations* (in the convex order sense) for the c.d.f. of S . Especially the lower bound approximation, which is given by $E[S \mid \Lambda]$ for an appropriate choice of the conditioning r.v. Λ is extremely accurate, see for instance Vanduffel, Dhaene, Goovaerts & Kaas (2003).

Huang, Milevsky & Wang (2004) compare the performance of different approximations for the probability that a person outlives his money in case of a lifelong continuous consumption pattern. As a special case, they also consider approximations for such a probability when the consumption period is fixed.

This chapter is related to Huang, Milevsky & Wang (2004). However, we will only consider deterministic sums. Furthermore, instead of comparing the approximations of the ruin probabilities, we will evaluate the performance of the above mentioned techniques by comparing the approximated values of quantiles and conditional tail expectations of r.v.'s S as defined in (4.1).

The chapter is organized as follows. In Section 4.2, we recall from the previous chapter the comonotonic approximations. In this section we also focus on the optimal choice of the conditioning r.v. Λ of the comonotonic lower bound $E[S \mid \Lambda]$. We propose a conditioning r.v. which is likely to make the variance of the approximation 'as close as possible' to the exact variance, see also Example 2.6 of Chapter 1. In Section 4.3 we briefly recall the mathematical techniques behind the reciprocal Gamma and lognormal moment matching techniques. Finally, in Section 4.4 we compare the comonotonic approximations with the moment matching techniques, using an extensive Monte Carlo simulation as benchmark. This chapter is based on Vanduffel, Hoedemakers and Dhaene (2004).

4.2 Comonotonic approximations

4.2.1 General results

Let the r.v. S be given by (4.1), where the α_i are non-negative real numbers and the random vector $(Z_1, Z_2, \dots, Z_n)$ has a multivariate normal distribution. Consider the conditioning r.v. Λ , given by

$$\Lambda = \sum_{i=1}^n \gamma_i Z_i. \quad (4.2)$$

and also the r.v.'s S^l and S^c defined by

$$S^l = \sum_{i=1}^n \alpha_i e^{E[Z_i] + \frac{1}{2}(1-r_i^2)\sigma_{Z_i}^2 + r_i\sigma_{Z_i}\Phi^{-1}(U)} \quad (4.3)$$

and

$$S^c = \sum_{i=1}^n \alpha_i e^{E[Z_i] + \sigma_{Z_i}\Phi^{-1}(U)}, \quad (4.4)$$

respectively. Here U is a Uniform(0,1) r.v., Φ is the c.d.f. of the $N(0,1)$ distribution and r_i is the correlation between Z_i and Λ .

From Corollary 2.5.1 it follows that for the r.v.'s S , S^l and S^c , the following convex order relations hold:

$$S^l \leq_{cx} S \leq_{cx} S^c. \quad (4.5)$$

We have from Example 3.4.4 that

$$Q_p[S^c] = \sum_{i=1}^n \alpha_i e^{E[Z_i] + \sigma_{Z_i}\Phi^{-1}(p)}, \quad (4.6)$$

$$CTE_p[S^c] = \sum_{i=1}^n \alpha_i e^{E[Z_i] + \frac{1}{2}\sigma_{Z_i}^2} \frac{\Phi(\sigma_{Z_i} - \Phi^{-1}(p))}{1-p}, \quad p \in (0,1). \quad (4.7)$$

Provided all coefficients r_i are positive, we also find from the same Example 3.4.4 that for $p \in (0,1)$:

$$Q_p[S^l] = \sum_{i=1}^n \alpha_i e^{E[Z_i] + \frac{1}{2}(1-r_i^2)\sigma_{Z_i}^2 + r_i\sigma_{Z_i}\Phi^{-1}(p)}, \quad (4.8)$$

$$CTE_p[S^l] = \sum_{i=1}^n \alpha_i e^{E[Z_i] + \frac{1}{2}\sigma_{Z_i}^2} \frac{\Phi(r_i\sigma_{Z_i} - \Phi^{-1}(p))}{1-p}.. \quad (4.9)$$

Finally, notice that the expected values of the r.v.'s S , S^c and S^l are all equal:

$$E(S) = E(S^l) = E(S^c) = \sum_{i=1}^n \alpha_i e^{E[Z_i] + \frac{1}{2}\sigma_{Z_i}^2}, \quad (4.10)$$

whereas their variances are given by

$$Var(S) = \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j e^{E[Z_i] + E[Z_j] + \frac{1}{2}(\sigma_{Z_i}^2 + \sigma_{Z_j}^2)} (e^{cov(Z_i, Z_j)} - 1), \quad (4.11)$$

$$Var(S^l) = \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j e^{E[Z_i] + E[Z_j] + \frac{1}{2}(\sigma_{Z_i}^2 + \sigma_{Z_j}^2)} (e^{r_i r_j \sigma_{Z_i} \sigma_{Z_j}} - 1) \quad (4.12)$$

and

$$Var(S^c) = \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j e^{E[Z_i] + E[Z_j] + \frac{1}{2}(\sigma_{Z_i}^2 + \sigma_{Z_j}^2)} (e^{\sigma_{Z_i} \sigma_{Z_j}} - 1), \quad (4.13)$$

respectively.

4.2.2 The ‘maximal variance’ lower bound approach

From the comments made on (1.12) it follows that if $X \leq_{cx} Y$ whilst X and Y are not equal in distribution, then $Var[X] < Var[Y]$ must hold. An equality in variance would imply that $X \stackrel{d}{=} Y$. This indicates that if we want to replace S by the less convex S^l , the best approximations probably will be the ones where the variance of S^l is ‘as close as possible’ to the variance of S . In other words, we should try to choose the coefficients γ_i of the conditioning variable Λ defined in (4.2) such that the variance of S^l is maximized.

We will now prove that the first order approximation of the variance of S^l will be maximized for the following choice of the parameters γ_i :

$$\gamma_i = \alpha_i e^{E[Z_i] + \frac{1}{2}\sigma_{Z_i}^2}, \quad i = 1, \dots, n. \quad (4.14)$$

Indeed, from (4.12) we find that

$$\begin{aligned} Var(S^l) &\approx \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j e^{E[Z_i] + E[Z_j] + \frac{1}{2}(\sigma_{Z_i}^2 + \sigma_{Z_j}^2)} (r_i r_j \sigma_{Z_i} \sigma_{Z_j}) \\ &= \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j e^{E[Z_i] + E[Z_j] + \frac{1}{2}(\sigma_{Z_i}^2 + \sigma_{Z_j}^2)} \left(\frac{Cov[Z_i, \Lambda] Cov[Z_j, \Lambda]}{Var(\Lambda)} \right) \\ &= \frac{(Cov(\sum_{i=1}^n \alpha_i e^{E[Z_i] + \frac{1}{2}\sigma_{Z_i}^2} Z_i, \Lambda))^2}{Var(\Lambda)} \\ &= (Corr(\sum_{i=1}^n \alpha_i e^{E[Z_i] + \frac{1}{2}\sigma_{Z_i}^2} Z_i, \Lambda))^2 Var(\sum_{i=1}^n \alpha_i e^{E[Z_i] + \frac{1}{2}\sigma_{Z_i}^2} Z_i). \end{aligned} \quad (4.15)$$

Hence, the first order approximation of $Var(S^l)$ is maximized when Λ is given by

$$\Lambda = \sum_{i=1}^n \alpha_i e^{E[Z_i] + \frac{1}{2}\sigma_{Z_i}^2} Z_i. \quad (4.16)$$

Note that also in case the α_i are not all positive, the choice (4.14) will optimize the first order approximation of the variance of S^l . In the remainder of this chapter and also further in the second part of this work, we will always assume that the conditioning r.v. Λ is given by (4.16). Notice that this optimal choice for Λ is slightly different from the choice that was made for this r.v. in Dhaene, Denuit, Kaas, Goovaerts & Vyncke (2002b). Numerical comparisons reveal that the choice proposed here leads to more accurate approximations.

One can easily prove that the first order approximation for $Var(S^l)$ with Λ given by (3.30) is equal to the first order approximation of $Var(S)$. This observation gives an additional indication that this particular choice for Λ will provide a good fit. We emphasize that the conditioning r.v. Λ defined in (3.30) does not necessarily maximize the variance of S^l , but has to be understood as an approximation for the optimal Λ . Theoretically, one could use numerical procedures to determine the optimal Λ , but this would outweigh one of the main features of the convex bounds, namely that the quantiles and conditional tail expectations (and also other actuarial quantities such as stop-loss premiums) can easily be determined analytically. Having a ready-to-use approximation that can be implemented easily is important from a practical point of view.

4.3 Well-known moment matching approximations

It belongs to the toolkit of any actuary to approximate the c.d.f. of an unknown r.v. by a known c.d.f. in such a way that the first moments are preserved. In this section we will briefly describe the reciprocal Gamma and the lognormal moment matching approximations. These two methods are frequently used to approximate the c.d.f. of the r.v. S defined by (4.1).

4.3.1 The Reciprocal Gamma approximation

The r.v. X is said to be Gamma distributed when its probability density function (p.d.f.) is given by

$$f_X(x; \alpha, \beta) = \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta}, \quad x > 0, \quad (4.17)$$

where $\alpha > 0$, $\beta > 0$ and $\Gamma(\cdot)$ denotes the Gamma function:

$$\Gamma(\alpha) = \int_0^\infty u^{\alpha-1} e^{-u} du \quad (\alpha > 0). \quad (4.18)$$

Consider now the r.v. $Y = 1/X$. This r.v. is said to be reciprocal Gamma distributed. Its p.d.f. is given by

$$f_Y(y; \alpha, \beta) = f_X(1/y; \alpha, \beta)/y^2, \quad y > 0. \quad (4.19)$$

It is straightforward to prove that the quantiles and conditional tail expectations of Y are given by

$$Q_p[Y] = \frac{1}{F_X^{-1}(1-p; \alpha, \beta)}, \quad p \in (0, 1) \quad (4.20)$$

and

$$CTE_p[Y] = \frac{F_X(F_X^{-1}(1-p; \alpha, \beta); \alpha-1, \beta)}{(1-p)(\alpha-1)\beta}, \quad p \in (0, 1), \quad (4.21)$$

where $F_X(\cdot; \alpha, \beta)$ is the cumulative d.f. of the Gamma distribution with parameters α and β . Since the Gamma distribution is readily available in many statistical software packages, these risk measures can easily be determined.

The first two moments of the reciprocal Gamma distributed r.v. Y are given by

$$E[Y] = \frac{1}{\beta(\alpha-1)} \quad (4.22)$$

and

$$E[Y^2] = \frac{1}{\beta^2(\alpha-1)(\alpha-2)}. \quad (4.23)$$

Expressing the parameters α and β in terms of $E[Y]$ and $E[Y^2]$ gives

$$\alpha = \frac{2E[Y^2] - E[Y]^2}{E[Y^2] - E[Y]^2} \quad (4.24)$$

and

$$\beta = \frac{E[Y^2] - E[Y]^2}{E[Y]E[Y^2]}. \quad (4.25)$$

The c.d.f. of the r.v. defined in (4.1) is now approximated by a reciprocal Gamma distribution with first two moments that follow from (4.10) and (4.11), respectively. The coefficients α and β of the reciprocal Gamma approximation follow from (4.24) and (4.25). The reciprocal Gamma approximations for the quantiles and the conditional tail expectations are then given by (4.20) and (4.21). Dufresne (1990) proves that the present value of a continuous perpetuity with a Wiener logreturn process is reciprocal Gamma distributed, under suitable parameter restrictions. An elegant proof for this result can be found in Milevsky (1997).

Intuitively, one expects that the present value of a finite discrete annuity with a normal logreturn process with independent periodic returns, can be approximated by a reciprocal Gamma distribution, provided the time period involved is long enough. This idea was set forward and explored in Milevsky & Posner (1998), Milevsky & Robinson (2000) and Huang, Milevsky & Wang (2004).

4.3.2 The lognormal approximation

The r.v. X is said to be lognormal distributed if its p.d.f. is given by

$$f_X(x; \mu, \sigma^2) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{(\log x - \mu)^2}{2\sigma^2}}, \quad x > 0, \quad (4.26)$$

where $\sigma > 0$.

The quantiles and conditional tail expectations of X are given by

$$Q_p[X] = e^{\mu + \sigma\Phi^{-1}(p)}, \quad p \in (0, 1) \quad (4.27)$$

and

$$CTE_p[X] = e^{\mu + \frac{1}{2}\sigma^2} \frac{\Phi(\sigma - \Phi^{-1}(p))}{1 - p}, \quad p \in (0, 1). \quad (4.28)$$

The first two moments of X are given by

$$E[X] = e^{\mu + \frac{1}{2}\sigma^2} \quad (4.29)$$

and

$$E[X^2] = e^{2\mu + 2\sigma^2}. \quad (4.30)$$

Expressing the parameters μ and σ^2 of the lognormal distribution in terms of $E[X]$ and $E[X^2]$ leads to

$$\mu = \log \left(\frac{E[X]^2}{\sqrt{E[X^2]}} \right) \quad (4.31)$$

and

$$\sigma^2 = \log \left(\frac{E[X^2]}{E[X]^2} \right). \quad (4.32)$$

The same procedure as the one explained in the previous subsection can be followed in order to obtain a lognormal approximation for S , with the first two moments matched. Dufresne (2002) obtains a lognormal limit distribution for S as volatility σ tends to zero and this provides a theoretical justification for the use of the lognormal approximation.

4.4 Comparing the approximations

In order to compare the performance of the different approximations presented above, we consider the r.v. S_n , which is defined as the random present value of a series of n deterministic unit payment obligations due at times 1, 2, ..., n , respectively:

$$S_n = \sum_{i=1}^n e^{-Y_1 - Y_2 - \dots - Y_i} \stackrel{not}{=} \sum_{i=1}^n e^{Z_i}. \quad (4.33)$$

where the r.v. Y_i is the random return over the year $[i-1, i]$ and $e^{-(Y_1 + Y_2 + \dots + Y_i)}$ is the random discount factor over the period $[0, i]$.

We will assume that the yearly returns Y_i are i.i.d. normal distributed with mean $\left(0.075 - \frac{\sigma^2}{2}\right)$ and variance σ^2 . Notice that S_n is a r.v. of the general type defined in (4.1).

The provision or reserve to set up at time 0 for these future unit payment obligations can be determined as $Q_p[S_n]$ or $CTE_p[S_n]$, with p sufficiently large. A provision equal to $Q_{0.95}[S_n]$ for instance, will guarantee that all payments can be made with a probability of 0.95, see also Chapter 6 and 7.

As the time unit that we consider in practical situations is typically long (1 year), assuming a Gaussian model for the returns seems to be appropriate, at least approximately, by the Central Limit Theorem. In order to verify whether our theoretical setup can be approximately compared with the data generating mechanism of real situations, we refer to Cesari & Cremonini (2003). They investigate four well-known stock market indices in US dollars, from Morgan Stanley: MSCI World, North America, Europe and Pacific, covering all major stock markets in industrial as well as emerging countries. For the period 1997-1999, the authors conclude that weekly (and longer period) returns can be considered as normal and independent. Daily returns on the other hand are both non-normal and autocorrelated.

In order to compute the comonotonic approximations for quantiles and conditional tail expectations, notice that Z_i and its mean $E[Z_i]$ and variance $\sigma_{Z_i}^2$ are given by

$$Z_i = -Y_1 - Y_2 - \dots - Y_i \quad (4.34)$$

$$E[Z_i] = -i\left(0.075 - \frac{\sigma^2}{2}\right), \quad (4.35)$$

$$\sigma_{Z_i}^2 = i \sigma^2 \quad (4.36)$$

From (3.76), (3.78) and (3.79) it follows that the r_i are given by

$$r_i = \frac{\sum_{j=1}^i \sum_{k=j}^n \gamma_k}{\sqrt{i \sum_{j=1}^n \left(\sum_{k=j}^n \gamma_k\right)^2}}, \quad (4.37)$$

with γ_k given by

$$\gamma_k = e^{E[Z_k] + \frac{1}{2}\sigma_{Z_k}^2}, \quad k = 1, \dots, n.$$

Notice that the correlation coefficients r_i are positive, so that the formulae (4.8) and (4.9) can be applied.

We will now compare the performance of the different approximation methods that were presented in Sections 2 and 3: the comonotonic upper bound method (UB), the comonotonic ‘maximal variance’ lower bound method (LB), the reciprocal Gamma method (RG) and the lognormal method (LN). We will compare the different approximations for quantiles and conditional tail expectations with the values obtained by Monte-Carlo simulation. The results of the simulation are obtained using the generation of 500.000 random paths. The

n	σ	UB	LB	RECG	LN	MC ($\pm$ s.e.)
20	0.05	+3.24%	− 0.01%	+0.07%	−0.16%	12.1957 (0.04%)
	0.15	+8.02%	+ 0.02%	−0.15%	−0.06%	20.4592 (0.10%)
	0.25	+9.36%	+ 0.00%	−4.28%	+2.99%	41.5854 (0.25%)
	0.35	+7.50%	+ 0.35%	−14.27%	+9.04%	106.1389 (0.30%)
40	0.05	+4.39%	+ 0.00%	+0.06%	−0.23%	15.4733 (0.04%)
	0.15	+10.26%	− 0.06%	−0.55%	+0.58%	30.4033 (0.16%)
	0.25	+9.42%	+ 0.06%	−8.52%	+9.73%	87.7482 (0.32%)
	0.35	+1.47%	− 0.83%	−19.70%	+9.96%	427.0793 (0.49%)

Table 4.1: Approximations for the 0.95-quantile of S_n for different horizons and volatilities.

estimates obtained from this time-consuming simulation will serve as benchmark. The random paths are based on antithetic variables in order to reduce the variance of the Monte-Carlo estimates.

The tables that we will present display the Monte Carlo simulation result (MC) for the risk measure at hand, as well as the percentual deviations of the different approximation methods, relative to the Monte-Carlo result. For the quantiles and conditional tail expectations, these percentual deviations are defined as follows:

$$\frac{Q_p[S_n^{approx}] - Q_p[S_n^{MC}]}{Q_p[S_n^{MC}]} \times 100\%$$

and

$$\frac{CTE_p[S_n^{approx}] - CTE_p[S_n^{MC}]}{TVaR_p[S_n^{MC}]} \times 100\%,$$

where S_n^{approx} corresponds to one of the approximation methods and S_n^{MC} denotes the Monte Carlo simulation result. The figures displayed in bold in the tables correspond to the best approximations, this means the ones with the smallest percentual deviation compared to the Monte-Carlo results. In the tables, we also present the standard errors of the Monte Carlo estimates. Note that these standard errors are also expressed as a percentual deviation from the Monte Carlo estimate.

Table 4.1. summarizes the results for the 0.95-quantiles for different yearly volatilities σ and for a time horizon of $n = 20$ and $n = 40$, respectively. The ‘maximal variance’ lower bound approach (LB) turns out to fit the quantiles the best for all values of the parameters. Its approximated quantiles almost always fall in the confidence interval $Q_p[S_n^{MC}] \pm s.e.$. It appears that the performance of the reciprocal Gamma approximations is worse for higher levels of volatility and for longer time horizons. The latter result is surprising, given the convergence of the c.d.f. of S_n to the reciprocal Gamma distribution. The results indicate that this convergence occurs very slowly.

Table 4.2. compares the different approximations for some selected quantiles of S_{20} , with a fixed yearly volatility of 25% . The results are in line with the

p	UB	LB	RECG	LN	MC ($\pm$ s.e.)
0.999	+17.41%	-0.65%	+0.73%	-3.76%	84.0466 (0.51%)
0.90	+6.11%	+0.12%	-4.19%	+4.19%	32.0758 (0.25%)
0.75	+0.32%	-0.03%	-2.52%	+3.81%	21.2666 (0.12%)
0.5	-6.15%	-0.10%	+1.20%	+0.25%	13.8933 (0.04%)
0.25	-12.55%	+0.13%	+6.18%	-6.36%	9.3833 (0.09%)

Table 4.2: Approximations for several selected quantiles of S_{20} . The yearly volatility equals 0.25.

n	σ	UB	LB	RECG	LN	MC ($\pm$ s.e.)
20	0.05	+4.19%	-0.02%	+0.21%	-0.38%	12.8231 (1.04%)
	0.15	+10.98%	-0.14%	+1.18%	-1.88%	24.4591 (2.16%)
	0.25	+14.17%	-0.36%	-0.98%	-0.94%	59.6646 (2.90%)
	0.35	+12.98%	-0.59%	-15.41%	+4.56%	198.0164 (3.27%)
40	0.05	+5.86%	+0.09%	+0.28%	-0.48%	16.3994 (1.55%)
	0.15	+0.09%	-0.25%	+0.87%	-2.38%	38.2515 (2.61%)
	0.25	+0.28%	-0.59%	-7.49%	+4.18%	149.8569 (3.25%)
	0.35	-0.48%	-0.84%	-40.77%	+12.77%	1206.0858 (3.59%)

Table 4.3: Approximations for the 0.95-conditional tail expectation of S_n for different horizon and volatility levels.

p	UB	LB	RECG	LN	MC ($\pm$ s.e.)
0.25	+1.89%	-0.10%	-1.17%	+1.44%	21.0969 (1.96%)
0.50	+4.34%	-0.09%	-2.16%	+2.33%	25.8692 (2.23%)
0.75	+7.87%	-0.11%	-2.82%	+2.31%	34.6099 (2.52%)
0.90	+11.71%	-0.21%	-2.25%	+0.80%	47.9276 (2.77%)
0.995	+21.00%	-0.99%	+7.82%	-7.97%	111.5457 (3.24%)

Table 4.4: Approximations for $CTE_p[S_{20}]$. The yearly volatility equals 0.25.

previous ones. The lower bound approach outperforms all the others, for high as well as for low values of p .

Table 4.3. displays the approximated and simulated 95% conditional tail expectations for the same set of parameters as in Table 1. Again the lower bound approach approximates the exact conditional tail expectations extremely well.

The same conclusions can be drawn from the results in Table 4.4. This table reports the different approximations for $CTE_p[S_{20}]$ for different probability levels p and a fixed yearly volatility $\sigma = 0.25$.

From the 4 tables, one can observe that both moment matching techniques perform poorly for high levels of p and/or σ . The comonotonic lower bound approach, however, remains to produce accurate approximations, also in these extreme cases.

Finally, remark that in general the percentual deviation of the comonotonic upper bound compared to the MC-simulation is relatively high. From the Tables 1 and 3, however, we can conclude that for high volatility levels, the crude comonotonic upper bound approximation often performs better than the reciprocal Gamma approximation.

4.5 Final remarks

In this chapter, we compared some approximation methods for a standard actuarial and financial problem: the determination of quantiles and conditional tail expectations of the present value of a series of cash-flows, when discounting is done in a Gaussian setting. We tested the accuracy of the comonotonic lower and upper bound approximations and two moment matching approximations by comparing these approximations with the estimates obtained from extensive Monte Carlo simulations.

Overall, the comonotonic ‘maximal variance’ lower bound approach provides the best fit and leads to accurate approximations under varying parameter assumptions, which are in line with realistic market values.

The comonotonic approach has the additional advantage that it gives rise to easy computable approximations for any risk measure that is additive for comonotonic risks. Examples of such risk measures are the distortion risk measures that we considered in Chapter 3.

Constant continuous annuities

5.1 Introduction

⁶In Chapter 3 we considered the general problem of determining comonotonic approximations for risk measures of sums of random variables for which the dependency structure is unknown or too cumbersome to work with. In Example 3.4.4 we derived explicit expressions for several risk measures in case one deals with a sum of lognormals. Next, in Chapter 4, we compared these comonotonic approximations for determining the quantiles and the conditional tail expectations with popular moment-matching techniques. More in particular we analyzed in that chapter the present value of a series of cash-flows:

$$S_n = \sum_{i=1}^n \alpha_i e^{-Y_1 - Y_2 - \dots - Y_i}, \quad (5.1)$$

with $\alpha_1, \alpha_2, \dots, \alpha_n$ a series of non-negative deterministic payments due at times $1, 2, \dots, n$ and the random vector $(Y_1, Y_2, \dots, Y_n)$ being a multivariate normal random variable.

The results derived in Example 3.4.4 for the discrete case (sums of random variables) have a continuous counterpart (integrals of stochastic processes). We will not consider this in detail, but we will focus on some explicit results in case the stochastic process under consideration is a geometric Brownian motion which is a continuous equivalent of the Gaussian setting we discussed in Example 3.4.4 and explored further in Chapter 4.

Hence, in line with Dhaene, Denuit, Goovaerts, Kaas & Vyncke (2002b), we will now consider the continuous equivalent of (5.1) which is the continuous

⁶“One of the advantages of being disorderly is that one is constantly making exciting discoveries” - A.A. Milne

temporary annuity S_t defined by

$$S_t = \int_0^t \alpha(\tau) e^{-\delta\tau - \sigma^2 B(\tau)} d\tau, \quad (5.2)$$

where $\{B(\tau), \tau \geq 0\}$ represents a standard Brownian motion, i.e. the process has independent and stationary increments, $B(0) = 0$ and for any $\tau \geq 0$, the random variable $B(\tau)$ is normally distributed with mean 0 and variance τ . Furthermore, the drift δ and the volatility σ are positive real numbers. Finally, the payments are described by $\alpha(\tau)$ which is a non-negative and continuous function of τ .

Let $Y(\tau) = \delta\tau + \sigma B(\tau)$ and $X(\tau) = \exp\{-Y(\tau)\}$. Similar to Corollary 2.5.1 it can be proven that $S_t^l \leq_{cx} S_t \leq_{cx} S_t^c$, where the random variable S_t^c and S_t^l are defined by

$$S_t^c = \int_0^t F_{\alpha(\tau)X(\tau)}^{-1}(U) d\tau = \int_0^t \alpha(\tau) e^{-\delta\tau + \sigma \sqrt{\tau} \Phi^{-1}(U)} d\tau \quad (5.3)$$

and

$$S_t^l = E[S_t | \Lambda] = \int_0^t \alpha(\tau) e^{-\delta\tau + \frac{1}{2}\sigma^2\tau(1-r^2(\tau)) + r(\tau)\sigma\sqrt{\tau}\Phi^{-1}(V)} d\tau, \quad (5.4)$$

where U is a uniform(0,1) random variable, the conditioning variable Λ follows a normal distribution and $V = \Phi\left(\frac{\Lambda - E[\Lambda]}{\sigma_\Lambda}\right)$ is standard uniformly distributed. Furthermore, $r(\tau)$ is defined by

$$r(\tau) = \frac{\text{cov}[Y(\tau), \Lambda]}{\sigma_\Lambda \sigma \sqrt{\tau}}. \quad (5.5)$$

Since $B(\tau)$ is a Brownian motion process, the random variable $Y(\tau) | \Lambda = \lambda$ is normally distributed with mean

$$E[Y(\tau) | \Lambda = \lambda] = \delta\tau + r(\tau)\sigma\sqrt{\tau}\frac{\lambda}{\sigma_\Lambda} \quad (5.6)$$

and variance

$$\text{Var}[Y(\tau) | \Lambda = \lambda] = \sigma^2\tau(1 - r^2(\tau)). \quad (5.7)$$

We also define the quantity δ^* :

$$\delta^* = \left(\delta - \frac{1}{2}\sigma^2\right). \quad (5.8)$$

Throughout the remainder of the chapter we will assume that $\delta^* > 0$.

Since $\alpha(\tau)$ is non-negative, S_t^c will be an integral of comonotonic random variables. Hence, the quantiles, conditional tail expectations and stop-loss pre-

miums of S_t^c follow from

$$Q_p[S_t^c] = \int_0^t \alpha(\tau) e^{-\delta\tau + \sigma \sqrt{\tau} \Phi^{-1}(p)} d\tau, \quad (5.9)$$

$$CTE_p[S_t^c] = \int_0^t \alpha(\tau) e^{-\delta\tau + \sigma^2 \tau/2} \frac{\Phi[\sigma \sqrt{\tau} - \Phi^{-1}(p)]}{1-p} d\tau, \quad (5.10)$$

$$E[S_t^c - d]_+ = \int_0^t \alpha(\tau) e^{-\delta\tau + \sigma^2 \tau/2} \Phi[\sigma \sqrt{\tau} - \Phi^{-1}(p)] d\tau - d(1-p), \quad (5.11)$$

with $0 < p < 1$ and $d > 0$ determined as the unique root of $Q_p[S_t^c] = d$. The expressions (5.9), (5.10) and (5.11) are the continuous counterparts of (3.66), (3.67) and (3.69).

Likewise we find that S_t^l will be an integral of comonotonic random variables in case the function $f(\tau) = \text{cov}[Y(\tau), \Lambda]$ remains non-negative. Hence, we find the following continuous equivalents for the expressions (3.72), (3.73) and (3.75):

$$Q_p[S_t^l] = \int_0^t \alpha(\tau) e^{-\delta\tau + \frac{1}{2}\sigma^2 \tau(1-r^2(\tau)) + r(\tau)\sigma \sqrt{\tau} \Phi^{-1}(p)} d\tau, \quad (5.12)$$

$$CTE_p[S_t^l] = \int_0^t \alpha(\tau) e^{-\delta\tau + \sigma^2 \tau/2} \frac{\Phi[r(\tau)\sigma \sqrt{\tau} - \Phi^{-1}(p)]}{1-p} d\tau, \quad (5.13)$$

$$E[S_t^c - d]_+ = \int_0^t \alpha(\tau) e^{-\delta\tau + \sigma^2 \tau/2} \Phi[r(\tau)\sigma \sqrt{\tau} - \Phi^{-1}(p)] d\tau - d(1-p), \quad (5.14)$$

with again $0 < p < 1$ and $d > 0$ uniquely determined by $Q_p[S_t^l] = d$.

In the remainder of this chapter we will consider a constant unit annuity. Hence, we assume that $\alpha(\tau) \equiv 1$. We will show that, in case of the upper bound approximation, one obtains closed form results for the quantiles, conditional tail expectation and stop-loss premiums. We will also demonstrate that, for some specific choices of Λ , explicit results for these risk measures can be obtained in case one uses the lower bound approximation.

We agree that this is of course a rather theoretical exercise as in real life one deals with discrete sums and not with continuous integrals. However, many research is done using a continuous setting for the particular problem of interest. This is because this often leads to more tractable formulas and may therefore provide initial insights and as such pave the way for developing the results in real-life discrete settings. In case of constant perpetuities, for instance, it is known since Dufresne (1990) that in a continuous setting the cumulative distribution function of S_∞ can be calculated very easily since one can prove that S_∞^{-1} is Gamma distributed with parameters $\frac{2\delta}{\sigma^2}$ and $\frac{\sigma^2}{2}$. Using this insight it has been proposed by Huang, Milevsky and Wang (2004) to use the reciprocal Gamma distribution function as a suitable choice for approximating the risk measures of finite annuities. Therefore, a first reason to mention the explicit results is that it might be useful for future research. As a second reason, we

can compare our approximations with available explicit results. In the discrete setting we always need to compare with simulated results which always give rise to room of debate. In the continuous setting, however, there exist explicit closed form results for some interesting actuarial quantities and, hence, these can be compared with the results obtained by using the comonotonic approximations.

The remainder of the chapter is organized as follows: in Section 5.2, we derive explicit closed-form formulas for several risk measures related to continuous annuities. In Section 5.3 we compare these approximations with explicit results in case of constant continuous perpetuities. Section 5.4 provides the conclusions.

5.2 Closed-form comonotonic approximations

5.2.1 Upper bound approach

Since $\alpha(\tau) \equiv 1$ we have that S_t^c is an integral of comonotonic random variables. The quantiles of S_t^c follow from (5.9):

$$Q_p[S_t^c] = \int_0^t e^{-\delta\tau + \sigma \sqrt{\tau} \Phi^{-1}(p)} d\tau, \quad (0 < p < 1). \quad (5.15)$$

By substituting $y = \sqrt{\tau}$ and realizing that the resulting integral can be rewritten in terms of the standard-normal c.d.f. we find the following analytical expression:

$$\begin{aligned} Q_p[S_t^c] &= \frac{1}{\delta} \\ &\quad - \frac{1}{\delta} e^{-\delta t + \sigma \sqrt{t} \Phi^{-1}(p)} \\ &\quad + \frac{1}{\delta} \sqrt{2\pi a} e^{\frac{a^2}{2}} (\Phi[\sqrt{2\delta t} - a] - \Phi[-a]), \end{aligned} \quad (5.16)$$

with

$$a = \frac{\sigma \Phi^{-1}(p)}{\sqrt{2\delta}}, \quad (0 < p < 1). \quad (5.17)$$

From (5.10) we have that the conditional tail expectations are given by

$$CTE_p[S_t^c] = \int_0^t e^{-\delta\tau + \sigma^2\tau/2} \frac{\Phi[\sigma\sqrt{\tau} - \Phi^{-1}(p)]}{1-p} d\tau, \quad (0 < p < 1). \quad (5.18)$$

Using the same substitution $y = \sqrt{\tau}$ as in the case of the quantiles, we find after some computations the following closed form expression for the conditional tail

expectations:

$$\begin{aligned}
CTE_p[S_t^c] &= \frac{1}{\delta^*} \\
&\quad - \frac{1}{\delta^*} e^{-\delta^* t} \frac{\Phi[\sigma\sqrt{t} - \Phi^{-1}(p)]}{1-p} \\
&\quad + \frac{\sigma}{\delta^*} e^{\frac{a^2 - (\Phi^{-1}(p))^2}{2}} \sqrt{\frac{1}{2\delta}} \frac{\Phi[\sqrt{2\delta}t - a] - \Phi[-a]}{1-p}, \quad (0 < p < 1).
\end{aligned} \tag{5.19}$$

Finally, we obtain that the stop-loss premiums with retentions $d > 0$ are given by

$$\begin{aligned}
E[(S_t^c - d)_+] &= \frac{1}{\delta^*} (1-p) \\
&\quad - \frac{1}{\delta^*} e^{-\delta^* t} \Phi[\sigma\sqrt{t} - \Phi^{-1}(p)] \\
&\quad + \frac{\sigma}{\delta^*} e^{\frac{a^2 - (\Phi^{-1}(p))^2}{2}} \sqrt{\frac{1}{2\delta}} (\Phi[\sqrt{2\delta}t - a] - \Phi[-a]) \\
&\quad - d(1-p), \quad (0 < p < 1),
\end{aligned} \tag{5.20}$$

where p can be obtained by solving $Q_p[S_t^c] = d$. We remark that the expressions (5.19) and (5.20) are valid under the condition that $\delta^* > 0$.

5.2.2 Lower bound approaches

5.2.2.1 General results

In order to compute the risk measures of S_t^l , Dhaene, Denuit, Goovaerts, Kaas & Vyncke (2002b) proposed to use the conditioning random variable $\Lambda = \int_0^t e^{-\delta\tau} B(\tau) d\tau$ because this can be seen as a linear transformation of a kind of first order approximation of S_t . However, in the same way as in Chapter 4 one can prove that the alternative choice $\Lambda = \int_0^t e^{-\delta^*\tau} B(\tau) d\tau$ will maximize the first order Taylor approximation for the variance of S_t^l . The latter choice for Λ is therefore likely to provide better approximations for the risk measures of S_t . We have in this case that Λ is normally distributed with mean 0 and variance

$$\begin{aligned}
\sigma_\Lambda^2 = Var[\Lambda] &= \int_0^t \int_0^t e^{-\delta^*(\tau+\nu)} \min(\tau, \nu) d\tau d\nu \\
&= \frac{1}{2\delta^{*3}} + \frac{3 + 2\delta^*t - 4e^{\delta^*t}}{2\delta^{*3}e^{2\delta^*t}}.
\end{aligned} \tag{5.21}$$

$r(\tau)$ is given by

$$r(\tau) = \frac{cov[Y(\tau), \Lambda]}{\sigma_\Lambda \sigma \sqrt{\tau}} = \frac{1}{\sigma_\Lambda \sqrt{\tau}} \left[\frac{1 - e^{-\delta^*\tau}}{\delta^{*2}} - \frac{\tau e^{-\delta^*t}}{\delta^*} \right], \quad \tau \leq t. \tag{5.22}$$

Since the function $f(\tau) = \text{cov}[Y(\tau), \Lambda]$ is a non-negative function, S_t^l will be an integral of comonotonic random variables. Unfortunately, in case of finite annuities, there seems to be no closed form solutions for the quantiles, conditional tail expectations and stop-loss premiums of S_t^l , in case one uses any of the two discussed choices for Λ . We now propose 2 other choices for Λ , such that closed-form approximations for these risk measures can be obtained.

$$\mathbf{5.2.2.2} \quad \Lambda = \int_0^\infty e^{-\delta^* \tau} B(\tau) d\tau$$

The quantiles of S_t^l ($0 < p < 1$) follow from

$$Q_p[S_t^l] = \int_0^t e^{-\delta \tau + \frac{1}{2} \sigma^2 \tau (1 - r^2(\tau)) + r(\tau) \sigma \sqrt{\tau} \Phi^{-1}(p)} d\tau, \quad (5.23)$$

with

$$r(\tau) = \sqrt{\frac{2}{\delta^* \tau}} (1 - e^{-\delta^* \tau}). \quad (5.24)$$

Expression (5.23) can be rewritten as

$$Q_p[S_t^l] = \int_0^t e^{-\delta^* \tau - \frac{1}{2} c^2 (1 - e^{-\delta^* \tau})^2 + c \Phi^{-1}(p) (1 - e^{-\delta^* \tau})} d\tau, \quad (5.25)$$

with

$$c = \sigma \sqrt{\frac{2}{\delta^*}}. \quad (5.26)$$

By making the substitution $y = e^{-\delta^* \tau}$ one sees that the integral (5.25) can be rewritten in terms of the standard-normal c.d.f. Hence, we find the following analytical expression for the quantiles of S_t^l :

$$Q_p[S_t^l] = \frac{1}{c \delta^*} \sqrt{2\pi} e^{\frac{(\Phi^{-1}(p))^2}{2}} (\Phi(k_t) - 1 + p), \quad (5.27)$$

with

$$k_t = c(1 - e^{-\delta^* t}) - \Phi^{-1}(p). \quad (5.28)$$

The conditional tail expectations of S_t^l are now given by

$$CTE_p[S_t^l] = \int_0^t e^{-\delta^* \tau} \frac{\Phi[c(1 - e^{-\delta^* \tau}) - \Phi^{-1}(p)]}{1 - p} d\tau, \quad (5.29)$$

After some tedious algebra we find that

$$\begin{aligned} CTE_p[S_t^l] &= \frac{1}{\delta^* (1 - p)} \Phi(k_t) (1 - e^{-\delta^* t}) \\ &\quad + \frac{1}{\delta^* (1 - p)} \left(\frac{\Phi^{-1}(p)}{c} \right) (1 - p - \Phi(k_t)) \\ &\quad - \frac{1}{\sqrt{2\pi} c \delta^* (1 - p)} (e^{-\frac{1}{2} [\Phi^{-1}(p)]^2} - e^{-\frac{1}{2} k_t^2}), \end{aligned} \quad (5.30)$$

Finally, the stop-loss premiums of S_t^l with retentions $d > 0$ are now given by

$$\begin{aligned} E[(S_t^l - d)_+] &= \frac{1}{\delta^*} \Phi(k_t)(1 - e^{-\delta^* t}) \\ &\quad + \frac{1}{\delta^*} \left(\frac{\Phi^{-1}(p)}{c} \right) (1 - p - \Phi(k_t)) \\ &\quad - \frac{1}{\sqrt{2\pi c \delta^*}} (e^{-\frac{1}{2}[\Phi^{-1}(p)]^2} - e^{-\frac{1}{2}k_t^2}) \\ &\quad - d(1 - p), \end{aligned} \quad (5.31)$$

where p is the unique root of $Q_p[S_t^l] = d$.

We point out that this specific choice for $\Lambda = \int_0^\infty e^{-\delta^* \tau} B(\tau) d\tau$ cannot be expected to perform very well for finite annuities. However, when t reaches infinity, this choice for Λ leads to the continuous equivalent of the ‘maximal variance’ lower bound approach that we discussed in Chapter 4. Hence, our specific choice for Λ is likely to yield excellent results when t reaches infinity whilst allowing for finite t an analytical expression for the selected risk measures of S_t^l too.

5.2.2.3 $\Lambda = B(t)$

The quantiles of S_t^l ($0 < p < 1$) are still given by

$$Q_p[S_t^l] = \int_0^t e^{-\delta\tau + \frac{1}{2}\sigma^2\tau(1-r^2(\tau)) + r(\tau)\sigma\sqrt{\tau}\Phi^{-1}(p)} d\tau, \quad (5.32)$$

but now with $r(\tau)$ given by

$$r(\tau) = \sqrt{\frac{\tau}{t}}. \quad (5.33)$$

It proves easily that one obtains the following closed form expressions for the quantiles of S_t^l :

$$Q_p[S_t^l] = \frac{\sqrt{2\pi t}}{\sigma} e^{\frac{\gamma^2}{2}} (\Phi[\sigma\sqrt{t} - \gamma] - \Phi[-\gamma]), \quad (5.34)$$

with

$$\gamma = \Phi^{-1}(p) - \frac{\delta^* \sqrt{t}}{\sigma}. \quad (5.35)$$

The conditional tail expectations of S_t^l follow from

$$CTE_p[S_t^l] = \int_0^t e^{-\delta^* \tau} \frac{\Phi[\frac{\sigma\tau}{\sqrt{t}} - \Phi^{-1}(p)]}{1 - p} d\tau. \quad (5.36)$$

Again, after some computations we find that

$$CTE_p[S_t^l] = \frac{1}{\delta^*} - \frac{1}{\delta^*(1-p)} e^{-\delta^* t} (\Phi[\sigma\sqrt{t} - \Phi^{-1}(p)]) \quad (5.37)$$

$$+ \frac{1}{\delta^*(1-p)} e^{-(\frac{\delta^* \sqrt{t}}{\sigma})(\gamma + \frac{\delta^* \sqrt{t}}{2\sigma})} \times \left\{ \Phi[\sigma\sqrt{t} - \gamma] - \Phi[-\gamma] \right\}. \quad (5.38)$$

The stop-loss premiums of S_t^l with retentions $d > 0$ now follow from

$$E[(S_t^l - d)_+] = \int_0^t e^{-\delta^* \tau} \Phi\left[\frac{\sigma\tau}{\sqrt{t}} - \Phi^{-1}(p)\right] d\tau - d(1-p), \quad (5.39)$$

where p is the root of $Q_p[S_t^l] = d$. We find that

$$E[(S_t^l - d)_+] = +\frac{1}{\delta^*}(1-p) - \frac{1}{\delta^*} e^{-\delta^* t} (\Phi[\sigma\sqrt{t} - \Phi^{-1}(p)]) \quad (5.40)$$

$$+ \frac{1}{\delta^*} e^{-(\frac{\delta^* \sqrt{t}}{\sigma})(\gamma + \frac{\delta^* \sqrt{t}}{2\sigma})} \times \left\{ \Phi[\sigma\sqrt{t} - \gamma] - \Phi[-\gamma] \right\} - d(1-p). \quad (5.41)$$

5.3 Application on perpetuities

Consider the perpetuity S_∞ defined by

$$S_\infty = \int_0^\infty \exp[-\delta\tau - \sigma B(\tau)] d\tau. \quad (5.42)$$

For this annuity, the cumulative distribution function of the perpetuity S_∞ can be calculated very easily since one can prove that S_∞^{-1} is Gamma distributed with parameters $\frac{2\delta}{\sigma^2}$ and $\frac{\sigma^2}{2}$. This result can be found in Dufresne (1990) and Milevsky (1997). Hence, we can compare the cumulative distribution functions of the lower bound S_∞^l and the upper bound S_∞^c with the exact cumulative distribution function of S_∞ . We propose to use the ‘maximal variance’ lower bound approach that we discussed in Subsection 5.2.2.2, since this is likely to provide the best results in case of infinite annuities.

From (4.20), (5.16) and (5.27), with $t \rightarrow \infty$, we find for $0 < p < 1$ the following expressions for the quantiles of S_∞ , S_∞^c and S_∞^l respectively.

$$Q_p[S_\infty] = \frac{1}{F_X^{-1}(1-p; \frac{2\delta}{\sigma^2}, \frac{\sigma^2}{2})}, \quad (5.43)$$

$$Q_p[S_\infty^c] = \frac{1}{\delta}(1 + a\sqrt{2\pi}e^{\frac{a^2}{2}}\Phi[a]), \quad (5.44)$$

$$Q_p[S_\infty^l] = \frac{1}{c\delta^*}\sqrt{2\pi}e^{\frac{(\Phi^{-1}(p))^2}{2}}(\Phi[c - \Phi^{-1}(p)] - 1 + p),$$

with F_X the c.d.f. of the Gamma distribution and

$$a = \frac{\sigma\Phi^{-1}(p)}{\sqrt{2\delta}}, \quad (5.45)$$

$$c = \sigma\sqrt{\frac{2}{\delta^*}}. \quad (5.46)$$

From (4.20), (4.21), Theorem 3.2.1, (5.20) and (5.31), one also obtains closed-form results for the stop-loss premiums for S_∞ , S_∞^c and S_∞^l . This is left as an easy exercise for the interested reader.

Table 5.1 shows the quantiles of S_∞^l , S_∞^c and S_∞ in case $\delta = 0.07$ and $\sigma = 0.1$. These results can be compared with the results reported in Dhaene, Denuit, Goovaerts, Kaas & Vyncke (2002). The small differences we observe, can be explained as follows : Firstly, the authors computed the quantiles of S_∞^c and S_∞^l by numerical evaluation of the expressions (5.9) and (5.12) with $t \rightarrow \infty$ and $\alpha(\tau) \equiv 1$. Secondly, they used the conditioning variable $\Lambda = \int_0^\infty e^{-\delta\tau}B(\tau)d\tau$ whereas our explicit results rely on the ‘maximal variance’ lower bound approximation, involving $\Lambda = \int_0^\infty e^{-\delta^*\tau}B(\tau)d\tau$ as conditioning r.v.

Table 5.2. shows quantiles of S_∞ , S_∞^c and S_∞^l but now for $\delta = 0.07$ and $\sigma = 0.2$. This example is interesting because it proves that for suitable choices of Λ , the c.d.f.’s of S_∞^l and S_∞ do not necessarily cross only once.

Finally, Table 5.3 compares the stop-loss premiums for different retention values d . The same comments as for Table 5.1 can be made.

5.4 Final remarks

In this chapter we derived explicit comonotonic approximations for risk measures for constant continuous annuities, in case discounting is done using a Brownian motion process. We compared these approximations with available explicit results in case of perpetuities. The comparisons support the conclusions made in Chapter 4, namely that especially the comonotonic ‘maximal variance’ lower bound approximation provides an excellent fit.

p	$Q_p[S_\infty^l]$	$Q_p[S_\infty]$	$Q_p[S_\infty^c]$
0.95	23.62	23.63	25.90
0.975	26.09	26.13	29.34
0.99	29.37	29.49	34.08
0.995	31.90	32.10	37.86
0.999	38.00	38.49	47.38

Table 5.1: The table compares some selected exact quantiles of the constant perpetuity with the ‘maximal variance’ lower bound and upper bound approximations ($\delta=0.07$, $\sigma=0.1$).

p	$Q_p[S_\infty^l]$	$Q_p[S_\infty]$	$Q_p[S_\infty^c]$
0.25	11.13	11.07	9.34
0.50	15.74	15.76	14.29
0.75	23.51	23.50	23.11
0.95	46.30	46.14	51.84
0.99	79.64	80.71	100.45
0.999	98.35	101.09	130.77

Table 5.2: The table compares some selected exact quantiles of the constant perpetuity with the ‘maximal variance’ lower bound and upper bound approximations ($\delta=0.07$, $\sigma=0.2$).

d	$E[S_\infty^l - d]_+$	$E[S_\infty - d]_+$	$E[S_\infty^c - d]_+$
10	5.4430	5.4457	5.5554
15	1.8590	1.8626	2.2690
20	0.4917	0.4961	0.8337
25	0.1229	0.1270	0.3079
30	0.0316	0.0344	0.1192

Table 5.3: The table compares some selected exact stop-loss premiums of the constant perpetuity with the ‘maximal variance’ lower bound and upper bound approximations ($\delta=0.07$, $\sigma=0.1$).

Applications of Comonotonicity in Risk Management

Optimal portfolio selection problems

6.1 Introduction

⁷Strategic portfolio selection is the process used to identify the best allocation of wealth among a basket of securities for an investor with a given consumption/saving behavior over a given investment horizon. The basket of available securities will typically be a selection of risky assets such as stocks, bonds and real estate, and risk-free components such as cash and money market instruments. The individual investor or the asset manager chooses an initial asset mix and a particular tactical trading strategy within a given set of strategies, according to which he will buy and sell risky and risk-free assets, during the whole time period under consideration.

The simplest case is a static strategy called ‘buy and hold’: the investments are performed according to a given strategy and no rebalancing is performed during the investment period. Single-index benchmarking, e.g. replicating a single stock market index, is a buy-and-hold strategy in case the index weights are not changed over the investment period.

Other strategies are called dynamic in the sense that they imply a periodic rebalancing process of the assets. A ‘constant mix’ strategy implies keeping the initial proportions constant, as opposed to a ‘buy and hold’ strategy where the initial quantities are kept constant through time. As the prices of assets evolve randomly over time, a constant mix strategy requires buying and/or selling at each time instant in order to keep the fractions at the predetermined level. Such a strategy implies a ‘buy low and sell high’ rule in the sense that price and asset-purchase are counter-varying: if the price of a single asset goes up while the prices of the other assets remain constant, then the quantity of the single asset should be decreased and vice versa. Furthermore, if the stock market has

⁷ “No formula belongs to one man” - Jan Dhaene.

decreased, one has to buy stocks against the risk-free component, whereas if the stock market has increased, one has to sell stock and buy risk-free instruments. Multiple-index benchmarking, e.g. replicating a benchmark which consists of the average of a stock benchmark and a bond benchmark, implies a constant mix strategy in order to stick to the benchmark proportions.

In this chapter we will investigate multi-period optimal portfolio selection problems in a Black & Scholes (1973) lognormal setting. We will assume that the investor has to choose the optimal investment strategy for a given consumption or savings pattern, within the class of constant mix strategies. We will consider two general types of problems, which will be referred to as the terminal wealth problem and the reserving problem respectively.

In the *terminal wealth problem*, the decision-maker will invest a given series of positive saving amounts $\alpha_0, \alpha_1, \dots, \alpha_{n-1}$ at predetermined times $0, 1, \dots, n-1$ such that his terminal wealth at time n will reach or exceed a target capital K with a sufficiently large probability. As terminal wealth is a sum of dependent lognormal random variables, its c.d.f. cannot be determined exactly and is too cumbersome to work with. Therefore, we will use the comonotonic upper and lower bound approximations that we introduced in Chapter 2 and analyzed in detail in Chapter 4.

Recall that the ‘comonotonic upper bound’ is derived by keeping the marginal distributions exact, but approximating the copula that describes the dependency structure between the random accumulation factors involved, by the comonotonic copula. The ‘comonotonic lower bound’ approximation for the exact distribution is based on the technique of conditioning. In this approach, the marginal distributions are changed and the copula describing the dependency structure is replaced by the comonotonic copula. In Chapter 4 we showed that especially this lower bound will perform very accurately as an approximation to the exact distribution.

The approximations that we propose have the tremendous advantage that for any given investment strategy they provide an accurate and easy to compute approximation for any risk measure that is additive for comonotonic risks, such as distortion risk measures (VaR and TailVaR for instance). The comonotonic approximations reduce the multivariate randomness of the multiperiod problem to a univariate randomness.

The optimal investment mix could be defined as the one that requires the smallest constant amount α that has to be invested from period to period in order to reach a final wealth of at least K with a probability of at least $1 - \epsilon$. Or, one could define the optimal mix as the one that maximizes the probability of reaching terminal wealth of at least K for a given investment of α per period. The proposed methodology can be used to solve several personal finance problems.

A first problem is what one could call the ‘saving for retirement problem’. In this case, one wants to retire in n years with a ‘nest egg’ of K - in real terms, i.e. in today’s Euros. How much does one have to save monthly - in real terms - in

order to assure a $(1 - \epsilon)$ chance to reach the retirement financial goal? Clearly, the answer will depend on the investment mix. The theory on comonotonicity gives a quick, elegant and accurate answer to this question.

A second personal finance problem for which the methodology can be used is the situation in which an individual underwrites an n year loan of K at a yearly interest rate of i . The first way to pay back the loan is by a classical annuity, where the lender pays at the end of each of the coming n years a fixed amount of $\frac{K}{a_{\overline{n}|i}}$. From the viewpoint of the lender, an annuity loan is equivalent with a loan with yearly interest payments of Ki while a yearly amount of $\frac{K}{a_{\overline{n}|i}} - Ki$ is invested in an amortization fund with a fixed return of i . At the end of the n year period the amortization fund will have grown to the required amount of K . An alternative is to pay the yearly interest of Ki and in addition, invest a yearly amount α in an amortization fund with future stochastic returns. The amount α is chosen such that the probability that the value of the amortization fund at time n will exceed the amount K will be sufficiently large.

In the *reserving problem*, which is in some sense dual to the final wealth problem, the decision-maker targets a series of future consumptions $\alpha_1, \alpha_2, \dots, \alpha_n$ at times $1, 2, \dots, n$. He sets up an initial reserve R_0 and wants to invest this amount in such a way that the possibility of reaching his targets is maximized. For this type of investment problems, the optimal investment mix could be defined as the one that leads to the largest survival probability, given the initial reserve. Or, one could fix the required survival probability and determine the optimal investment strategy as the one that minimizes the required initial reserve.

One possible application is the problem of the decision-maker who is faced with a series of deterministic obligations or liabilities due at predetermined fixed points in time. He wants to determine the reserve (and/or total balance sheet capital requirement) and selects the optimal investment portfolio such that the possibility of ruin is minimized.

An application in the area of personal finance is the annuitization problem where an initial amount R_0 is invested and used to enable a series of future periodic payments.

A somewhat related problem is the so-called ‘after-retirement problem’. The difference with our setting is that in the after-retirement problem the time-horizon is random and equal to the remaining life time. This problem is considered in Milevsky, Ho & Robinson (1997) and Milevsky & Robinson (2000). These authors take the investment strategy as given and find the corresponding probability of lifetime ruin. Young (2004) on the other hand, finds the optimal dynamic investment strategy as the one that minimizes the probability of lifetime ruin. A related concept is the probability of financial regret. It has been investigated in Brockett, Cox & Witt (1984).

As the time unit that we consider is long (typically 1 year), assuming a Gaussian model seems to be appropriate, at least approximately, by the Central

Limit Theorem. In the insurance literature, the assumption of normal distributed investment returns can be traced back to Boyle (1976), see also Chapter 4.

Although our theoretical results find a natural application in the field of personal finance, they can also be applied for solving many general corporate investment decision problems, in particular in a (long term) insurance business context, see e.g. Dufresne (1990), Goovaerts, Dhaene & De Schepper (2001) or Dhaene, Goovaerts, Vanduffel & Vyncke (2002). In case of a sufficiently large portfolio of mutual independent risks, one could approximate the average yearly claim amounts by their mean. The investment risk is then investigated by considering the sum of the randomly discounted future average claim amounts. More generally, one could model the stochastic behavior of the yearly aggregate claims by a lognormal model. In this case the random variable of interest is the scalar product of two random vectors, one being the claims vector and the other the vector of discount factors. The latter approach is used for determining reserves for ‘Incurred But Not Reported claims’ (IBNR) for a non-life portfolio, see e.g. Goovaerts & Redant (1998) or Hoedemakers, Beirlant, Goovaerts & Dhaene (2003).

This chapter is organized as follows. In Section 6.2 we present some results concerning the Black & Scholes setting, constant mix portfolios and mean-variance analysis that will be used throughout this chapter. Next, in Section 6.3, we analyze the problem of finding optimal investment strategies in a general multivariate final wealth model with savings at discrete points in time. The dual problem of setting an initial provision and optimizing investments in a general model with consumptions at discrete points in time is considered in Section 6.4. Final conclusions and some possible generalizations are discussed in Section 6.5.

To the best of our knowledge, determining optimal investment strategies for terminal wealth and optimal provision problems by means of the comonotonic approach, as presented in Sections 6.3 and 6.4, is new. This chapter is taken from Dhaene, Vanduffel, Goovaerts, Kaas and Vyncke (2004).

6.2 Stochastic return processes

6.2.1 The Black & Scholes setting

Throughout this chapter, we will assume the classical continuous-time framework that was pioneered by Merton (1971) and is nowadays mostly referred to as the Black & Scholes (1973) setting. We suppose that there is a market in which $(m+1)$ securities (assets or investment accounts) are traded continuously. One of the assets is the risk-free asset. Let $P^0(0) = P^0 > 0$ be the current price, at time 0, of 1 unit of the risk-free asset, whereas $P^0(t)$ is its price at time t . This price is assumed to evolve according to the following ordinary differential

equation:

$$\frac{dP^0(t)}{P^0(t)} = r \, dt, \quad (6.1)$$

with $r > 0$. On the other hand, let $P^i(0) = P^i > 0$ be the current price, at time 0, of 1 unit of risky asset i , whereas $P^i(t)$ is the price at time t (including reinvestment of dividend income) of one unit of risky asset i . The price process $P^i(t)$ evolves according to a geometric Brownian motion stochastic process, represented by the following stochastic differential equation:

$$\frac{dP^i(t)}{P^i(t)} = \mu_i dt + \sum_{j=1}^d \bar{\sigma}_{ij} \, dW^j(t), \quad i = 1, \dots, m, \quad (6.2)$$

where $(W^1(\tau), W^2(\tau), \dots, W^d(\tau))$ is a d -dimensional standard Brownian motion process. The $W^i(\tau)$ are mutually independent standard Brownian motions.

The m -dimensional vector $\boldsymbol{\mu}^T = (\mu_1 \dots \mu_m)$ is called the drift vector of the risky assets. We will assume that $\boldsymbol{\mu} \neq r \mathbf{1}$, with $\mathbf{1}^T = (1 \ 1 \dots 1)$.

The $(m \times d)$ matrix $\bar{\boldsymbol{\Sigma}}$ defined by

$$\bar{\boldsymbol{\Sigma}} = \begin{pmatrix} \bar{\sigma}_{11} & \bar{\sigma}_{12} & \dots & \bar{\sigma}_{1d} \\ \bar{\sigma}_{21} & \bar{\sigma}_{22} & \dots & \bar{\sigma}_{2d} \\ \dots & \dots & \dots & \dots \\ \bar{\sigma}_{m1} & \bar{\sigma}_{m2} & \dots & \bar{\sigma}_{md} \end{pmatrix} \quad (6.3)$$

is called the diffusion matrix. Further, we define the $(m \times m)$ matrix $\boldsymbol{\Sigma}$ as

$$\boldsymbol{\Sigma} = \bar{\boldsymbol{\Sigma}} \cdot \bar{\boldsymbol{\Sigma}}^T = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \dots & \sigma_{1m} \\ \sigma_{21} & \sigma_2^2 & \dots & \sigma_{2m} \\ \dots & \dots & \dots & \dots \\ \sigma_{m1} & \sigma_{m2} & \dots & \sigma_m^2 \end{pmatrix}, \quad (6.4)$$

with coefficients σ_{ij} and σ_i^2 given by $\sigma_{ij} = \sum_{k=1}^d \bar{\sigma}_{ik} \bar{\sigma}_{jk}$ and $\sigma_i^2 = \sigma_{ii}$. We have that $\sigma_{ij} = \sigma_{ji}$. The matrix $\boldsymbol{\Sigma}$ is called the variance-covariance matrix. We will assume that $\boldsymbol{\Sigma}$ is positive definite. This means that for all non-zero vectors $\boldsymbol{\pi}^T = (\pi_1, \pi_2, \dots, \pi_m)$ we have that

$$\boldsymbol{\pi}^T \cdot \boldsymbol{\Sigma} \cdot \boldsymbol{\pi} > 0. \quad (6.5)$$

In particular, this assumption implies that all σ_i are strictly positive. Hence, all m risky assets are indeed risky. It also implies that $\boldsymbol{\Sigma}$ is non-singular, meaning that its determinant is strictly positive, and hence, $\boldsymbol{\Sigma}$ has a matrix inverse $\boldsymbol{\Sigma}^{-1}$. As we will see further on, the elements of the matrix $\boldsymbol{\Sigma}$ describe the covariances between the yearly returns of the different assets.

We define the process $B^i(\tau)$ by

$$B^i(\tau) = \frac{1}{\sigma_i} \sum_{j=1}^d \bar{\sigma}_{ij} W^j(\tau). \quad (6.6)$$

It is straightforward to verify that all $B^i(\tau)$ are (correlated) standard Brownian motions, with

$$\text{Cov}(B^i(t), B^j(t+s)) = \frac{\sigma_{ij}}{\sigma_i \sigma_j} t, \quad t, s \geq 0. \quad (6.7)$$

Rewriting equation (6.2), we find

$$\frac{dP^i(t)}{P^i(t)} = \mu_i dt + \sigma_i dB^i(t), \quad i = 1, \dots, m. \quad (6.8)$$

The solution to equation (6.8) is

$$P^i(t) = P^i \exp \left\{ \left(\mu_i - \frac{1}{2} \sigma_i^2 \right) t + \sigma_i B^i(t) \right\}, \quad (6.9)$$

which means that $\frac{P^i(t)}{P^i}$ is lognormally distributed with parameters $(\mu_i - \frac{1}{2} \sigma_i^2) t$ and $\sigma_i^2 t$, respectively. This implies that the expectation and standard deviation of the price of asset i at time t are given by

$$\begin{aligned} E[P^i(t)] &= P^i e^{\mu_i t}, \\ \sigma[P^i(t)] &= P^i e^{\mu_i t} \sqrt{e^{\sigma_i^2 t} - 1}. \end{aligned} \quad (6.10)$$

Let $k = 1, 2, \dots$. Investing an amount of 1 at time $k-1$ in asset i will grow to the random amount $e^{Y_k^i}$ at time k , where Y_k^i denotes the yearly return in year k of account i . One finds that

$$Y_k^i = \left(\mu_i - \frac{1}{2} \sigma_i^2 \right) + \sigma_i (B^i(k) - B^i(k-1)). \quad (6.11)$$

Hence, it follows that the random yearly returns Y_k^i of asset i are independent and have identical normal distributions with

$$\begin{aligned} E[Y_k^i] &= \mu_i - \frac{1}{2} \sigma_i^2, \\ \text{Var}[Y_k^i] &= \sigma_i^2, \\ \text{Cov}[Y_k^i, Y_l^j] &= \begin{cases} 0 & k \neq l \\ \sigma_{ij} & k = l \end{cases}. \end{aligned} \quad (6.12)$$

As stated earlier, the matrix Σ is the variance-covariance matrix of the yearly return vector $(Y_k^1, Y_k^2, \dots, Y_k^n)$. More details on the translation between the two formalisms (6.2) and (6.8) for describing the multivariate asset process can be found e.g. in Björk (1998).

6.2.2 Constant mix strategies

Assume one can invest wealth in one or more of the $m+1$ assets as defined above. Let $\boldsymbol{\pi}(\mathbf{t})^T = (\pi_1(t), \pi_2(t), \dots, \pi_m(t))$ be the vector describing the

portfolio process, i.e. $\pi_i(t)$ is the fraction of the wealth that is invested in risky asset i at time t . The residual, i.e. $1 - \sum_{i=1}^m \pi_i(t)$, is invested in the risk-free asset, or, if negative, finances the risky asset purchases. A negative proportion invested in the risk-free asset means borrowing (going short) on the risk-free asset.

We will restrict to constant portfolios $\boldsymbol{\pi}(t)^T = \boldsymbol{\pi}^T = (\pi_1, \pi_2, \dots, \pi_m)$, which means that the fractions invested in the different assets remain constant over time. Investing according to a constant portfolio process implies that one has to follow a dynamic trading strategy. Indeed, as the risky asset returns evolve randomly, one has to trade at each instant in order to keep the fractions invested in the different assets constant. Such investment strategies are known as constant mix strategies, or also as constant proportional investment strategies. The optimality of constant mix strategies in a utility theory setting is considered in Merton (1971).

Let us now consider one unit of a security that is constructed according to the continuously rebalanced investment strategy $(\pi_1, \pi_2, \dots, \pi_m)$, and let $P(t)$ be the price of that unit at time t , with $P(0) = P$. One can prove that the price process $P(t)$ evolves according to the dynamics

$$\begin{aligned} \frac{dP(t)}{P(t)} &= \sum_{i=1}^m \pi_i \frac{dP^i(t)}{P^i(t)} + \left(1 - \sum_{i=1}^m \pi_i\right) \frac{dP^0(t)}{P^0(t)} \\ &= \left(\sum_{i=1}^m \pi_i (\mu_i - r) + r \right) dt + \sum_{i=1}^m \pi_i \sigma_i dB^i(t). \end{aligned} \quad (6.13)$$

For a non-zero vector $\boldsymbol{\pi}$, define the process $B(\tau)$ by

$$B(\tau) = \frac{1}{\sqrt{\boldsymbol{\pi}^T \cdot \boldsymbol{\Sigma} \cdot \boldsymbol{\pi}}} \sum_{i=1}^m \pi_i \sigma_i B^i(\tau). \quad (6.14)$$

One can verify that $B(\tau)$ is a standard Brownian motion. Equation (6.13) can then be rewritten as

$$\frac{dP(t)}{P(t)} = \mu(\boldsymbol{\pi}) dt + \sigma(\boldsymbol{\pi}) dB(t) \quad (6.15)$$

with

$$\mu(\boldsymbol{\pi}) = r + \boldsymbol{\pi}^T \cdot (\boldsymbol{\mu} - r \mathbf{1}) \quad \text{and} \quad \sigma^2(\boldsymbol{\pi}) = \boldsymbol{\pi}^T \cdot \boldsymbol{\Sigma} \cdot \boldsymbol{\pi}, \quad (6.16)$$

where $\mathbf{1}$ is the m - vector $(1 \ 1 \dots 1)$. Hence, we find that when the portfolio is rebalanced in continuous time in order to keep the fractions constant, the portfolio return is also lognormal distributed. Recall that we assumed that the variance-covariance matrix $\boldsymbol{\Sigma}$ is positive definite. This means that any non-zero combination $\boldsymbol{\pi}$ of the risky assets is also risky in the sense that $\sigma^2(\boldsymbol{\pi}) > 0$. The solution to equation (6.15) is

$$P(t) = P \exp \left\{ \left(\mu(\boldsymbol{\pi}) - \frac{1}{2} \sigma^2(\boldsymbol{\pi}) \right) t + \sigma(\boldsymbol{\pi}) B(t) \right\}, \quad (6.17)$$

with expectation and standard deviation given by

$$\begin{aligned} E[P(t)] &= P e^{\mu(\pi)t}, \\ \sigma[P(t)] &= P e^{\mu(\pi)t} \sqrt{e^{\sigma^2(\pi)t} - 1}. \end{aligned} \quad (6.18)$$

The stochastic differential equation (6.15) was derived by Merton (1971, 1990), see also Rubinstein (1991). It can also be derived using elementary arguments by taking limits of lognormal sums, see Milevsky & Posner (1998).

Let k be a strictly positive integer. Investing according to investment strategy π , an amount of 1 at time $k-1$ will grow to the random amount $e^{Y_k(\pi)}$ at time k , where $Y_k(\pi)$ denotes the yearly return in year k of investment strategy π . One finds that

$$Y_k(\pi) = \left(\mu(\pi) - \frac{1}{2} \sigma^2(\pi) \right) + \sigma(\pi) (B(k) - B(k-1)). \quad (6.19)$$

Hence, the random yearly returns $Y_k(\pi)$ of the constantly rebalanced portfolio π are independent and identically distributed normal random variables with

$$\begin{aligned} E[Y_k(\pi)] &= \mu(\pi) - \frac{1}{2} \sigma^2(\pi), \\ Var[Y_k(\pi)] &= \sigma^2(\pi). \end{aligned} \quad (6.20)$$

Note that this observation about the yearly returns also holds in case π equals the zero vector. The price $P(k)$ can be written in terms of the yearly returns as follows:

$$P(k) = P \exp(Y_1(\pi) + Y_2(\pi) + \cdots + Y_k(\pi)). \quad (6.21)$$

6.2.3 Markowitz mean-variance analysis

In 1990, Harry M. Markowitz received the Nobel Prize in Economics (shared with William F. Sharpe and Merton H. Miller) for his theory on portfolio selection under uncertainty. The contribution for which he received the award was first published in Markowitz (1952) and more extensively in Markowitz (1959). As mentioned in the press release of the Royal Swedish Academy of Sciences, Markowitz's theory can be considered as the first approach to solving the problem that each investor faces, namely how to find the optimal trade-off between risk and return, i.e. how to find the optimal investment strategy under the two conflicting objectives of high expected return versus low risk of the investment portfolio. Markowitz proposed a way to reduce the complicated and multidimensional problem of finding the optimal portfolio with respect to a large number of different assets to a conceptual simple two-dimensional problem, known as mean-variance analysis. The Markowitz approach has become very popular due to the fact that it combines algebraic simplicity and suitability for practical applications. The mean-variance approach provides a fundamental basis for portfolio selection in a single period. A selected overview of the tremendous amount of research initiated by Markowitz's seminal work can be found in Steinbach (2001).

Several variants of the classical single-period mean-variance problem exist. Here, we will consider the formulation that we will need later on in the chapter. Among all constant mix portfolios $\boldsymbol{\pi}$ with a given portfolio volatility $\sigma(\boldsymbol{\pi}) = \sigma \geq 0$, we look for the one with the largest drift $\mu(\boldsymbol{\pi})$. Hence, for a given value of σ , we want to find the solution of the following problem:

$$\text{Max}_{\boldsymbol{\pi}} \mu(\boldsymbol{\pi}) \text{ subject to } \sigma(\boldsymbol{\pi}) = \sigma. \quad (6.22)$$

We will denote the portfolio that corresponds to the maximum in (6.22) by $\boldsymbol{\pi}^\sigma$. The assumption that $\boldsymbol{\mu} \neq r\mathbf{1}$, together with the assumptions that the variance-covariance matrix is positive definite and that short-selling is allowed, implies that there exists a unique local global minimum for problem (6.22). A Lagrange optimization yields:

$$\mu(\boldsymbol{\pi}^\sigma) = r + \sigma \sqrt{(\boldsymbol{\mu} - r\mathbf{1})^T \cdot \boldsymbol{\Sigma}^{-1} \cdot (\boldsymbol{\mu} - r\mathbf{1})}, \quad \sigma \geq 0 \quad (6.23)$$

and

$$\boldsymbol{\pi}^\sigma = \sigma \frac{\boldsymbol{\Sigma}^{-1} \cdot (\boldsymbol{\mu} - r\mathbf{1})}{\sqrt{(\boldsymbol{\mu} - r\mathbf{1})^T \cdot \boldsymbol{\Sigma}^{-1} \cdot (\boldsymbol{\mu} - r\mathbf{1})}}. \quad (6.24)$$

Note that $\mu(\boldsymbol{\pi}^\sigma)$ and $\boldsymbol{\pi}^\sigma$ are well-defined, because the inverse of a positive definite matrix is also positive definite.

The efficient frontier refers to the set of all solutions $\{(\sigma, \mu(\boldsymbol{\pi}^\sigma)) \mid \sigma \geq 0\}$ for the optimization problem (6.22). From (6.23) we see that the efficient frontier consists of a straight line in the (σ, μ) -plane. The portfolios $\boldsymbol{\pi}^\sigma$ belonging to the efficient frontier are called mean-variance efficient portfolios. The efficient frontier is often referred to as the ‘Capital Market Line’.

The ‘Capital Market Line’ can also be characterized as the set $\{(\sigma(\boldsymbol{\pi}^\mu), \mu) \mid \mu \geq r\}$ where $\boldsymbol{\pi}^\mu$ is the portfolio with the minimal variance, given that the drift equals μ . Hence, $\boldsymbol{\pi}^\mu$ is the solution of:

$$\text{Min}_{\boldsymbol{\pi}} \sigma(\boldsymbol{\pi}) \text{ subject to } \mu(\boldsymbol{\pi}) = \mu, \quad \mu \geq r. \quad (6.25)$$

In the following, we will call portfolios $\boldsymbol{\pi}$ that fulfill the condition $\mathbf{1}^T \times \boldsymbol{\pi} = 1$ risky-assets-only portfolios because such portfolios consist only of risky assets. It can be proven that if we only consider risky-assets-only portfolios, the efficient frontier corresponds to (the upper branch of) a hyperbola in the volatility - drift space (provided there are at least two risky assets with different drift). Now, consider the risky-assets-only global minimal variance portfolio $\boldsymbol{\pi}^{(m)}$, i.e. the portfolio that is the solution of the following problem:

$$\text{Min}_{\boldsymbol{\pi}} \sigma(\boldsymbol{\pi}) \text{ subject to } \mathbf{1}^T \cdot \boldsymbol{\pi} = 1. \quad (6.26)$$

This portfolio and its drift are given by

$$\begin{aligned} \boldsymbol{\pi}^{(m)} &= \frac{\boldsymbol{\Sigma}^{-1} \cdot \mathbf{1}}{\mathbf{1}^T \cdot \boldsymbol{\Sigma}^{-1} \cdot \mathbf{1}}, \\ \mu(\boldsymbol{\pi}^{(m)}) &= \frac{\mathbf{1}^T \cdot \boldsymbol{\Sigma}^{-1} \cdot \boldsymbol{\mu}}{\mathbf{1}^T \cdot \boldsymbol{\Sigma}^{-1} \cdot \mathbf{1}}. \end{aligned} \quad (6.27)$$

One can prove that under the condition

$$r < \mu(\boldsymbol{\pi}^{(m)}), \quad (6.28)$$

the Capital Market Line (6.23) is at a tangent to the upper branch of the hyperbola that corresponds to the efficient frontier of risky-asset-only portfolios. The portfolio that corresponds to the point of intersection between the Capital Market Line (6.23) and the upper branch of the hyperbola is called the ‘tangency portfolio’, and is denoted by $\boldsymbol{\pi}^{(t)}$. In the sequel, when we use $\boldsymbol{\pi}^{(t)}$, we will always assume that the condition (6.28) is fulfilled. One can easily verify that $\boldsymbol{\pi}^{(t)}$ is given by

$$\boldsymbol{\pi}^{(t)} = \frac{\boldsymbol{\Sigma}^{-1} \cdot (\boldsymbol{\mu} - r\mathbf{1})}{\mathbf{1}^T \cdot \boldsymbol{\Sigma}^{-1} \cdot (\boldsymbol{\mu} - r\mathbf{1})}. \quad (6.29)$$

Note that (6.24) can be rewritten as

$$\boldsymbol{\pi}^\sigma = \left(\frac{\mu(\boldsymbol{\pi}^\sigma) - r}{\mu(\boldsymbol{\pi}^{(t)}) - r} \right) \boldsymbol{\pi}^{(t)}. \quad (6.30)$$

This means that every mean-variance efficient portfolio $\boldsymbol{\pi}^\sigma$ consists of a fraction $\left(\frac{\mu(\boldsymbol{\pi}^\sigma) - r}{\mu(\boldsymbol{\pi}^{(t)}) - r} \right)$ invested in the risky-assets-only portfolio $\boldsymbol{\pi}^{(t)}$ and a fraction $\left(1 - \frac{\mu(\boldsymbol{\pi}^\sigma) - r}{\mu(\boldsymbol{\pi}^{(t)}) - r} \right)$ invested in the risk-free asset. Mean-variance optimizing investors only differ in terms of which fraction of their wealth they put in the tangency portfolio.

The result that all mean-variance investors will hold only two kinds of portfolios (or mutual funds), the exclusively risky portfolio $\boldsymbol{\pi}^{(t)}$ and the risk-free asset, is often called a ‘Mutual Fund Theorem’ or a ‘Two Fund Separation Theorem’.

Note that in case $r = \mu(\boldsymbol{\pi}^{(m)})$, there is no tangency portfolio, but any portfolio on the efficient frontier can still be constructed as a linear combination of two basic portfolios on the efficient frontier.

The Capital Market Line can be rewritten as

$$\mu(\boldsymbol{\pi}^\sigma) = r + \left(\frac{\mu(\boldsymbol{\pi}^{(t)}) - r}{\sigma(\boldsymbol{\pi}^{(t)})} \right) \sigma. \quad (6.31)$$

This equation describes the drift of the return as related to the volatility that the investor is willing to accept. The slope $\frac{\mu(\boldsymbol{\pi}^{(t)}) - r}{\sigma(\boldsymbol{\pi}^{(t)})}$ is referred to as the ‘Sharpe ratio’. It can be interpreted as the price of risk reduction: It shows how much the drift decreases if the volatility decreases 1 unit.

There are many papers that consider variants of the classical mean-variance portfolio selection criterion, where the variance is replaced by an alternative asymmetric risk measure that measures downside risk, in order to avoid penalization due to over-performance, see e.g. Emmer, Klüppelberg & Korn (2001)

or Li, Ng, Tan & Yang (2003). Note that Markowitz (1959) already introduced the idea to replace the variance by an alternative asymmetric risk measure in a more general mean - risk approach.

The single period mean-variance model was soon extended to multiperiod portfolio selection. Research on multiperiod portfolio selection has been dominated by the idea of maximizing expected utility functions of terminal wealth. Markowitz (1959) already considered long-term investment planning by considering multiperiod models based on attaching a utility to the levels of consumption of wealth over time.

6.3 Saving and terminal wealth

6.3.1 General problem description

In this section, we will consider the problem of how to invest periodic saving amounts in order to reach a target capital at a predetermined future time n . Let α_i be the positive amount that will be invested at time i , ($i = 0, 1, 2, \dots, n$). We assume that these amounts are invested according to a constant mix portfolio π as defined in Section 6.2. The choice of the constant portfolio mix has to be made at time 0. An amount of 1 unit invested at time i will grow to the random amount $e^{\sum_{j=i+1}^n Y_j(\pi)}$ at time n .

Let $W_j(\pi)$ be the wealth at time j , defined by the following recursive relation:

$$W_j(\pi) = W_{j-1}(\pi) e^{Y_j(\pi)} + \alpha_j, \quad j = 1, \dots, n, \quad (6.32)$$

with initial value $W_0(\pi) = \alpha_0$. Hence, $W_j(\pi)$ is the wealth that will be available at time j , including the savings amount α_j at time j . The realization of $W_j(\pi)$ will be known at time j . It depends on the investment returns (stochastic part) and on the savings (deterministic part) in the past. Note that the random variables $Y_j(\pi)$ are i.i.d. and normal distributed with parameters $\mu(\pi)$ and $\sigma(\pi)$ as defined in (6.16).

From the recursion (6.32) for the wealth process, we find the following explicit expression for terminal wealth $W_n(\pi)$:

$$W_n(\pi) = \sum_{i=0}^n \alpha_i e^{\sum_{j=i+1}^n Y_j(\pi)}. \quad (6.33)$$

where by convention, $\sum_{i=m}^n b_i$ is set equal to 0 if $m > n$.

In Chapter 3 we compared the utility theory framework of von Neumann & Morgenstern (1947) with Yaari's (1987) dual theory of choice under risk.

Within the expected utility theory framework of von Neumann & Morgenstern (1947), the investor could choose the investment strategy π that maximizes his expected utility of final wealth:

$$\max_{\pi} E[u(W_n(\pi))], \quad (6.34)$$

where u is the (non-decreasing) utility function he uses to appreciate the different levels of final wealth.

Within the framework of Yaari's (1987) dual theory of choice under risk, we choose the optimal investment strategy as the one that maximizes the distorted expectation of final wealth:

$$\max_{\pi} \rho_f [W_n(\pi)], \quad (6.35)$$

where f is the investor's distortion function (which means that $f(0) = 0$, $f(1) = 1$ and f is non-decreasing on the interval $[0, 1]$) and ρ_f is the 'distorted expectation', determined with $f(\Pr(W_n(\pi) > x))$:

$$\rho_f [W_n(\pi)] = \int_0^\infty f(\Pr(W_n(\pi) > x)) dx. \quad (6.36)$$

While in utility theory, choosing among risks is performed by comparing expected values of transformed wealth levels (utilities), in Yaari's theory the quantities that are compared are the 'distorted expectations' of wealth levels. The distorted expectation of final wealth $W_n(\pi)$ can be interpreted as an expectation of $W_n(\pi)$ evaluated with a 'distorted probability measure' in the sense of a Choquet-integral, see Denneberg (1994). The decision-maker acts in order to maximize the distorted expectation of final wealth.

For a distortion function f_p , $0 < p < 1$, given by

$$f_p(x) = \begin{cases} 0 & : 0 \leq x < p \\ 1 & : p \leq x \leq 1 \end{cases} \quad (6.37)$$

we find

$$\begin{aligned} \rho_{f_p} [W_n(\pi)] &= Q_{1-p}^+ [W_n(\pi)] \\ &= \sup \{x \in \mathbb{R} \mid \Pr(W_n(\pi) > x) \geq p\}. \end{aligned} \quad (6.38)$$

The optimization problem (6.35) with distortion function given by (6.37) determines the optimal investment strategy as the one that maximizes the largest amount that will be reached with a probability of at least p .

For the convex distortion function g_p , $0 < p < 1$, given by

$$g_p(x) = \begin{cases} 0 & : 0 \leq x < p \\ \frac{x-p}{1-p} & : p \leq x \leq 1 \end{cases} \quad (6.39)$$

we find

$$\rho_{g_p} [W_n(\pi)] = CLTE_{1-p} [W_n(\pi)]. \quad (6.40)$$

In Yaari's theory, a decision-maker is called risk averse if he has a convex distortion function. Hence, the optimization problem (6.35) with distortion function (6.39) can be interpreted as the problem to be solved by a risk averse decision-maker with distortion function g_p . The optimal investment strategy is the one

that maximizes the conditional expected value of final wealth, given that the p -target capital is not reached.

We have that

$$W_n(\boldsymbol{\pi}) \stackrel{d}{=} \sum_{i=0}^n \alpha_i e^{(n-i) [\mu(\boldsymbol{\pi}) - \frac{1}{2}\sigma^2(\boldsymbol{\pi})] + \sqrt{n-i} \sigma(\boldsymbol{\pi}) Z_i}, \quad (6.41)$$

with $(Z_0, Z_1, \dots, Z_n)$ a multivariate normal random vector with standard normal distributed marginals.

Now, consider two portfolios $\boldsymbol{\pi}_1$ and $\boldsymbol{\pi}_2$ with $\sigma(\boldsymbol{\pi}_1) = \sigma(\boldsymbol{\pi}_2)$ and $\mu(\boldsymbol{\pi}_1) < \mu(\boldsymbol{\pi}_2)$. Following a similar reasoning as in Ahcan, Darkiewicz, Dhaene, Goovaerts & Hoedemakers (2004), we find from (6.41) that

$$F_{W_n(\boldsymbol{\pi}_1)}(x) \geq F_{W_n(\boldsymbol{\pi}_2)}(x), \quad x \geq 0. \quad (6.42)$$

From the monotonicity property of distortion risk measures, we can conclude that

$$\rho_f[W_n(\boldsymbol{\pi}_1)] \leq \rho_f[W_n(\boldsymbol{\pi}_2)] \quad (6.43)$$

holds for any distortion risk measure $\rho_f[\cdot]$. This implies that the solution of the optimization problem (6.35) can be found on the Capital Market Line. Hence, we can replace (6.35) by the optimization problem

$$\max_{\boldsymbol{\sigma}} \rho_f[W_n(\boldsymbol{\pi}^{\boldsymbol{\sigma}})]. \quad (6.44)$$

Notice that the optimization (6.34) also reduces to a one-dimensional optimization over all portfolios on the Capital Market Line.

6.3.2 The case of a single investment

6.3.2.1 Time and portfolio diversification

Emmer, Klüppelberg & Korn (2001) remark that there seems to be common wisdom that long term stock investment leads to an almost sure gain over locally risk-free bond investments. In the long run stock indices are growing faster than risk-free rates, despite the repeated occurrence of stock market declines. The conventional perception therefore holds that the longer the investment horizon, the greater should be one's proportion invested in risky assets. In order to verify whether or not this common wisdom holds true, we will consider the terminal wealth problem with a single investment of 1 at time 0. Hence, $\alpha_0 = 1$ and $\alpha_1 = \alpha_2 = \dots = \alpha_n = 0$. We will assume that one can invest in one risk-free asset $P^0(t)$ and in m risky assets $P^i(t)$ as explained in section 3.1. We will also assume that $\mu(\boldsymbol{\pi}^{(m)}) > r$ holds, which implies that $\mu(\boldsymbol{\pi}^{(t)}) > r$.

The cumulative distribution function of final wealth $W_n(\boldsymbol{\pi})$ follows from

$$W_n(\boldsymbol{\pi}) \stackrel{d}{=} e^n [\mu(\boldsymbol{\pi}) - \frac{1}{2}\sigma^2(\boldsymbol{\pi})] + \sqrt{n} \sigma(\boldsymbol{\pi}) \Phi^{-1}(U). \quad (6.45)$$

Within the framework of expected utility theory, one determines the optimal constantly rebalanced portfolio as the one that maximizes the investor's expected utility of final wealth:

$$\max_{\pi} E[u(W_n(\pi))]. \quad (6.46)$$

For the logarithmic utility function $u(x) = \ln(x)$, one finds that the optimal portfolio π^* lies on the Capital Market Line and is given by

$$\pi^* = \left(\frac{\mu(\pi^t) - r}{\sigma^2(\pi^t)} \right) \pi^t. \quad (6.47)$$

Note that within the logarithmic utility framework, the optimal strategy (6.47) is independent of the investment horizon. For more details, see e.g. Merton (1990).

Within Yaari's dual theory of choice under risk, let us now consider the optimization problem

$$\max_{\pi} Q_{1-p}^+[W_n(\pi)] \quad (6.48)$$

with $\frac{1}{2} < p < 1$. The quantiles of final wealth $W_n(\pi)$ are given by

$$Q_{1-p}^+[W_n(\pi)] = Q_{1-p}[W_n(\pi)] = e^{n(\mu(\pi) - \frac{1}{2}\sigma^2(\pi)) - \sqrt{n}\sigma(\pi)\Phi^{-1}(p)}. \quad (6.49)$$

One can easily verify that for a fixed value of $\mu(\pi)$, the quantile is decreasing in $\sigma(\pi)$. On the other hand, for a fixed value of $\sigma(\pi)$, the quantile is increasing in $\mu(\pi)$, implying that the optimal portfolio of (6.48) will correspond to a point on the Capital Market Line and is given by

$$\pi^* = \left(\frac{\mu(\pi^t) - r}{\sigma^2(\pi^t)} - \frac{\Phi^{-1}(p)}{\sqrt{n}\sigma(\pi^t)} \right)_+ \pi^t, \quad \frac{1}{2} < p < 1, \quad (6.50)$$

where $(x)_+$ stands for $\max(0, x)$. From (6.50) it follows that in this setting, increasing the investment time horizon transforms the optimal investment strategy into a more risk-taking one. Hence, investors with a longer time horizon should have a larger exposure to stocks as compared to investors with a shorter time horizon. The optimal risky proportion converges to the optimal growth portfolio (6.47). We can conclude that in case the time horizon becomes infinitely large, the optimal growth portfolio will outperform any other portfolio, with respect to optimality criterion (6.48).

Next, consider the optimization problem of a risk averse decision-maker who determines the optimal investment strategy as the solution of the following maximization problem:

$$\max_{\pi} CLTE_{1-p}[W_n(\pi)]. \quad (6.51)$$

From (3.25) we find that

$$CLTE_{1-p}(W_n(\pi)) = e^{n\mu(\pi)} \frac{1 - \Phi(\sqrt{n}\sigma(\pi) + \Phi^{-1}(p))}{1 - p} \quad (6.52)$$

One can again verify that for a fixed value of $\mu(\boldsymbol{\pi})$, the $CLTE_{1-p}$ is decreasing in $\sigma(\boldsymbol{\pi})$, while for a fixed value of $\sigma(\boldsymbol{\pi})$, it is increasing in $\mu(\boldsymbol{\pi})$. This implies that the optimum of problem (6.51) will also correspond to a point on Capital Market Line.

Another way to look at the time diversification effect, is to consider the ‘Equity Shortfall Risk’ of the investment portfolio. Following Milevsky (2003), we define the Equity Shortfall Risk over a given investment period n and for a given investment strategy $\boldsymbol{\pi}$ by

$$\text{ESR}(\boldsymbol{\pi}, n) = \Pr(W_n(\boldsymbol{\pi}) \leq e^{nr}). \quad (6.53)$$

Hence, $\text{ESR}(\boldsymbol{\pi}, n)$ is the probability that the risky investment strategy $\boldsymbol{\pi}$ will underperform the risk-free investment strategy over a time horizon n . It can be interpreted as the probability of regretting the investment, where the investor regrets his choice $\boldsymbol{\pi}$ compared to the risk-free strategy, if this last strategy will have performed better than the risky investment strategy. The concept of Equity Shortfall Risk was introduced in the financial literature by Roy (1952), in a one-period discrete-time setting.

A straightforward calculation leads to

$$\text{ESR}(\boldsymbol{\pi}, n) = 1 - \Phi\left(\sqrt{n}\left(\frac{\mu(\boldsymbol{\pi}) - r}{\sigma(\boldsymbol{\pi})} - \frac{1}{2}\sigma(\boldsymbol{\pi})\right)\right). \quad (6.54)$$

From this expression, we see that, provided $\mu(\boldsymbol{\pi}) - \frac{1}{2}\sigma(\boldsymbol{\pi})^2 > r$, increasing the time horizon n will decrease $\text{ESR}(\boldsymbol{\pi}, n)$. Moreover, when the time horizon goes to infinity, the Equity Shortfall Risk disappears and the risky investment strategy $\boldsymbol{\pi}$ will outperform the risk-free investment strategy with probability 1. Important to note however, is that when the portfolio variance $\sigma(\boldsymbol{\pi})^2$ becomes large relative to $\mu(\boldsymbol{\pi}) - r$, $\text{ESR}(\boldsymbol{\pi}, n)$ increases with the time horizon and reaches level 1 at infinity. We can conclude that the general perception of time-diversification expressed in terms of decreasing Equity Shortfall Risk is in accordance with the theory, provided the expected yearly returns $E[Y_k(\boldsymbol{\pi})]$ exceed the risk-free yearly return r .

Comparing the optimal investment strategies (6.47) and (6.50), we can conclude that the time diversification benefit strongly depends on the optimality criterion that is considered. It has to be mentioned that the belief in time-diversification is not general and that the (non-) existence of a time-diversifying benefit is the subject of a heavy debate. The topic is considered in Samuelson (1989), Marshall (1994), Bodie (1995), Jagannathan & Kocherlakota (1996) and Milevsky (2003), amongst others.

Milevsky (2003) also considers the concept of space diversification, by which he means the diversification effect caused by increasing the number of risky assets in the investment portfolio. Provided $\mu(\boldsymbol{\pi}) > r$, we find that decreasing the portfolio volatility $\sigma(\boldsymbol{\pi})$ will, ceteris paribus, lead to an increase of the argument in the $\Phi(\cdot)$ function in (6.54). Hence, any increase of the number of risky assets which allows to reduce the portfolio volatility while keeping the portfolio drift constant (or increasing it), will decrease $\text{ESR}(\boldsymbol{\pi}, n)$. In other words, a better

p	$n = 1$	$n = 10$	$n = 20$	$n = 40$	$n = 100$
0.99	0	0	0	0.09	1.16
0.97	0	0	0	0.64	1.51
0.95	0	0	0.09	0.94	1.70
0.90	0	0	0.73	1.39	1.98

Table 6.1: Optimal proportion invested in $\pi^{(t)}$ in case of maximizing $Q_{1-p}[W_n(\pi)]$.

space-diversified portfolio implies a lower Equity Shortfall Risk.

This phenomenon can easily be illustrated in the case of a homogeneous market (i.e. all securities have equal drift and variance, all correlations are equal and positive) and applying the ‘naive’ constant mix strategy where all proportions are kept equal: $\pi_i = \frac{1}{m}$. In this particular case, it is straightforward to prove that increasing the number of securities m will keep the portfolio drift constant while it decreases the portfolio variance.

The interrelationship and trade-off between the two possible dimensions of diversification, i.e. the number of stocks in a portfolio and the number of periods over which an investment is held, is investigated in detail in Milevsky (2003).

6.3.2.2 Numerical illustration

In a Black & Scholes market there is a risk-free asset with a yearly return $r = 0.03$ and two risky assets with yearly drifts equal to $\mu_1 = 0.06$ and $\mu_2 = 0.10$ respectively. The volatilities of the risky assets are given by $\sigma_1 = 0.10$ and $\sigma_2 = 0.20$. Pearson’s correlation coefficient $\frac{\sigma_{12}}{\sigma_1\sigma_2}$ is given by 0.5. From (6.29) we find that the tangency portfolio is given by $\pi^{(t)} = (\frac{5}{9}, \frac{4}{9})$ with drift $\mu(\pi^{(t)}) = 7/90$ and volatility $\sigma(\pi^{(t)}) = \sqrt{\frac{43}{2700}}$.

We consider a single investment at time 0. In Table 6.1 and Table 6.2 we present the optimal proportions invested in the tangency portfolio, with respect to the optimization problems (6.48) and (6.51), for different values of the probability level p and the investment period n . For instance, for criterion (6.48) with $p = 0.99$ and $n = 40$ we find that the optimal investment strategy consists in investing 9% in the tangency portfolio and 91% in the risk-free asset.

The figures in Table 6.1 and Table 6.2 illustrate that increasing the time horizon leads to an increased optimal proportion invested in the tangency portfolio. Also, the lower the probability level p with which we want to reach the target, the higher the proportion to be invested in $\pi^{(t)}$. Finally, observe that the maximization of $CLTE_{1-p}[W_n(\pi)]$ leads to lower optimal risky proportions as compared to the maximization of $Q_{1-p}^+[W_n(\pi)]$.

p	$n = 1$	$n = 10$	$n = 20$	$n = 40$	$n = 100$
0.99	0	0	0	0	0.96
0.97	0	0	0	0.18	1.31
0.95	0	0	0	0.47	1.50
0.90	0	0	0	0.93	1.79

Table 6.2: Optimal proportion invested in $\pi^{(t)}$ in case of maximizing $CLTE_{1-p}[W_n(\pi)]$.

6.3.3 Comonotonic approximations for the general problem

Let us now consider the general terminal wealth problem as described in Section 6.3. From (6.33), we see that $W_n(\pi)$ is a sum of non-independent lognormal random variables. As it is impossible to determine the cumulative distribution function of $W_n(\pi)$ analytically, we will propose the convex order upper bound $W_n^c(\pi)$ and a convex order lower bound $W_n^l(\pi)$ for $W_n(\pi)$.

Rewriting $W_n(\pi)$ as

$$W_n(\pi) = \sum_{i=0}^n \alpha_i e^{Z_i}, \quad (6.55)$$

we see that we can apply the results of Example 3.4.4 with

$$\begin{aligned} Z_i &= Y_{i+1}(\pi) + Y_{i+2}(\pi) + \cdots + Y_n(\pi), \\ E[Z_i] &= (n-i) \left[\mu(\pi) - \frac{1}{2}\sigma^2(\pi) \right], \\ \sigma_{Z_i}^2 &= (n-i) \sigma^2(\pi). \end{aligned} \quad (6.56)$$

It follows that the comonotonic upper bound $W_n^c(\pi)$ for $W_n(\pi)$ is given by

$$W_n^c(\pi) = \sum_{i=0}^n \alpha_i e^{(n-i) [\mu(\pi) - \frac{1}{2}\sigma^2(\pi)] + \sqrt{n-i} \sigma(\pi) \Phi^{-1}(U)}. \quad (6.57)$$

For $p \in (0, 1)$, the quantiles of $W_n^c(\pi)$ are given by

$$Q_{1-p}^+[W_n^c(\pi)] = Q_{1-p}[W_n^c(\pi)] = \sum_{i=0}^n \alpha_i e^{(n-i) [\mu(\pi) - \frac{1}{2}\sigma^2(\pi)] - \sqrt{n-i} \sigma(\pi) \Phi^{-1}(p)}, \quad (6.58)$$

while $CLTE_p[W_n^c(\pi)]$ is given by

$$CLTE_{1-p}[W_n^c(\pi)] = \sum_{i=0}^n \alpha_i e^{(n-i) \mu(\pi)} \frac{1 - \Phi(\sqrt{n-i} \sigma(\pi) + \Phi^{-1}(p))}{1-p}. \quad (6.59)$$

In order to define the convex lower bound $W_n^l(\pi)$ for $W_n(\pi)$, we choose the conditioning random variable as follows:

$$\Lambda(\pi) = \sum_{j=0}^n \gamma_j(\pi) Z_j(\pi) \quad (6.60)$$

where the coefficients $\gamma_j(\pi)$ follow from (4.14). Notice that the lower bound approximation

$$W_n^l(\pi) = E[W_n(\pi) \mid \Lambda(\pi)] \quad (6.61)$$

is only determined up to a linear transformation of $\Lambda(\pi)$. Hence, we propose to choose the coefficients $\beta_j(\pi)$ as follows:

$$\gamma_j(\pi) = \alpha_j e^{-j \mu(\pi)} \quad (6.62)$$

For this choice of the parameters, the variance of the lower bound will be close to the variance of $W_n(\pi)$, provided $\sigma^2(\pi)$ is small enough, see Chapter 4. From Example 3.4.4, we find

$$W_n^l(\pi) = \sum_{i=0}^n \alpha_i e^{(n-i) \mu(\pi) - \frac{1}{2} r_i^2(\pi) (n-i) \sigma^2(\pi) + r_i(\pi) \sqrt{n-i} \sigma(\pi) \Phi^{-1}(U)} \quad (6.63)$$

where the coefficients $r_i(\pi)$ are given by

$$r_i(\pi) = \frac{\sum_{j=i+1}^n \sum_{k=0}^{j-1} \alpha_k e^{-k \mu(\pi)}}{\sqrt{n-i} \sqrt{\sum_{j=1}^n \left(\sum_{k=0}^{j-1} \alpha_k e^{-k \mu(\pi)} \right)^2}}. \quad (6.64)$$

Note that the correlation coefficients $r_i(\pi)$ are non-negative which implies that $W^l(\pi)$ is a comonotonic sum of lognormal random variables.

The following expression can be derived for the risk measure $Q_{1-p}^+[W_n^l(\pi)] = Q_{1-p}[W_n^l(\pi)]$, $p \in (0, 1)$:

$$Q_{1-p}^+[W_n^l(\pi)] = \sum_{i=0}^n \alpha_i e^{(n-i) \mu(\pi) - \frac{1}{2} r_i^2(\pi) \sigma^2(\pi) - r_i(\pi) \sqrt{n-i} \sigma(\pi) \Phi^{-1}(p)}, \quad (6.65)$$

while for $CLTE_{1-p}[W_n^l(\pi)]$ we find

$$CLTE_{1-p}[W_n^l(\pi)] = \sum_{i=0}^n \alpha_i e^{(n-i) \mu(\pi)} \frac{1 - \Phi(r_i(\pi) \sqrt{n-i} \sigma(\pi) + \Phi^{-1}(p))}{1-p}. \quad (6.66)$$

Still from Example 3.4.4 we have that

$$CLTE_{1-p}[W_n^c(\pi)] \leq CLTE_{1-p}[W_n(\pi)] \leq CLTE_{1-p}[W_n^l(\pi)], \quad 0 < p < 1. \quad (6.67)$$

Note however that the approximations $Q_{1-p}[W_n^l(\pi)]$ and $Q_{1-p}[W_n^c(\pi)]$ for the quantiles $Q_{1-p}[W_n(\pi)]$ are not necessarily ordered in the same way.

6.3.4 Determining the investment strategy that maximizes the target capital for a given probability level

6.3.4.1 The p - target capital

For a given probability level and a given investment strategy π , we define the p -target capital K as the $(1-p)$ -th order ‘+’-quantile of terminal wealth:

$$K = Q_{1-p}^+[W_n(\pi)]. \quad (6.68)$$

One immediately finds that

$$K = \sup \{x \in \mathbb{R} \mid \Pr[W_n(\pi) > x] \geq p\}. \quad (6.69)$$

Hence, the target capital at probability level p can be interpreted as the maximal amount that will be available at time n , with a probability of at least p .

Now assume that a probability level p is fixed and that the optimal investment strategy π^* is determined as the one that maximizes the p -target capital. Denoting the optimal target capital by K^* , we have

$$K^* = \max_{\pi} Q_{1-p}^+[W_n(\pi)]. \quad (6.70)$$

Note that from (6.35) and (6.38), it follows that this optimization problem can be interpreted in terms of Yaari’s dual theory of choice under risk.

Solving (6.70) is from a computational point of view a complicated problem because of the multi-dimensionality involved. Indeed, a ‘time-dimensionality’ occurs because $W_n(\pi)$ is a sum of n dependent accumulation factors. In the following sections we will show how to get rid of this ‘curse of dimensionality’.

At first sight, there is also a ‘portfolio-dimensionality’ involved, as the maximum has to be determined over all portfolios π . But from the reasoning that led to (6.44), we know that in (6.70), it suffices to consider the portfolios on the Capital Market Line. Hence, the optimal target capital K^* and the optimal investment strategy π^* follow from

$$K^* = \max_{\sigma} Q_{1-p}^+[W_n(\pi^{\sigma})], \quad (6.71)$$

where the efficient portfolios π^{σ} are given by (6.24).

6.3.4.2 Comonotonic bounds for $W_n(\pi)$

As it is impossible to determine $Q_{1-p}^+[W_n(\pi)]$ analytically, we first propose to approximate it by $Q_{1-p}^+[W_n^c(\pi)]$. The optimization problem (6.71) is then replaced by the following one:

$$K^c = \max_{\sigma} Q_{1-p}^+[W_n^c(\pi^{\sigma})], \quad (6.72)$$

where for each value of σ , the portfolios π^{σ} are given by (6.24), and where K^c is the approximated optimal target capital. The optimal portfolio π^* is

approximated by π^c , which is the portfolio that corresponds to the optimum in (6.72). Notice that $Q_{1-p}^+[W_n^c(\pi^\sigma)]$ follows from (6.58).

Next, we propose to approximate the optimal investment strategy π^* by π^l , where π^l is the investment strategy that maximizes $Q_{1-p}^+[W_n^l(\pi)]$ on the Capital Market Line. The p - target capital K^* is then approximated by K^l , which is given by

$$K^l = \max_{\sigma} Q_{1-p}^+[W_n^l(\pi^\sigma)]. \quad (6.73)$$

The quantiles $Q_{1-p}^+[W_n^l(\pi^\sigma)]$ follow from (6.65). As we will illustrate, the approximation defined in (6.73) will be extremely accurate, see also Chapter 4.

The approximated optimization problems (6.72) and (6.73) solve the curse of dimensionality. Introducing the comonotonic approximations $W_n^c(\pi^\sigma)$ and $W_n^l(\pi^\sigma)$ for $W_n(\pi^\sigma)$, reduce the multi-dimensionality caused by the n yearly returns involved to dimension 1.

6.3.4.3 Constant savings amounts

In this subsection, we consider the special case that the saving amounts are constant. For each investment strategy π we look for the required periodic saving amount $\alpha(\pi)$ that leads to a p -target capital equal to 1. From (6.68) we find that this saving amount $\alpha(\pi)$ is given by

$$\alpha(\pi) = \frac{1}{Q_{1-p}^+[\overline{W}_n(\pi)]}, \quad (6.74)$$

with $\overline{W}_n(\pi)$ given by

$$\overline{W}_n(\pi) = \sum_{i=0}^n e^{Y_{i+1}(\pi) + Y_{i+2}(\pi) + \dots + Y_n(\pi)}. \quad (6.75)$$

The optimal investment strategy is now defined as the one that minimizes the period savings. Denoting the minimal saving amount by α^* , we have

$$\alpha^* = \min_{\pi} \alpha(\pi). \quad (6.76)$$

Note that in the case of constant saving amounts, the investment strategy that maximizes the p - target capital K for given saving amounts α is identical to the investment strategy that minimizes the periodic savings α for a given target capital equal to 1.

Now we approximate $\overline{W}_n(\pi)$ by $\overline{W}_n^c(\pi)$ as explained in (6.57). The minimal periodic savings amount α^* is then approximated by α^c which is given by

$$\alpha^c = \min_{\sigma} \frac{1}{Q_{1-p}^+[\overline{W}_n^c(\pi^\sigma)]}. \quad (6.77)$$

Next, we propose to approximate $\overline{W}_n(\pi)$ by $\overline{W}_n^l(\pi)$ as explained in (6.63). The minimal periodic savings amount α^* is then approximated by α^l which is

determined by

$$\alpha^l = \min_{\sigma} \frac{1}{Q_{1-p}^+ \left[\overline{W}_n^l(\pi^\sigma) \right]}. \quad (6.78)$$

6.3.4.4 Numerical illustration

Consider the Black & Scholes type market with one risk-free and two risky assets as explained in Subsection 6.3.2.2. First, we assume saving amounts $\alpha_i = 1$ for $i = 0, 1, \dots, 39$, while $\alpha_{40} = 0$.

As we have seen, the solution of the problem (6.68) is to be found on the Capital Market Line. Because of the Two Fund Theorem, any portfolio on the Capital Market Line can be expressed as a linear combination of the risk-free portfolio and the tangency portfolio $\pi^{(t)}$. Hence, we can reduce the market to a market consisting of one risk-free asset with $r = 0.03$ and one risky asset with drift and volatility equal to the corresponding values of the tangency portfolio in the original setting.

Figure 6.1 shows the comonotonic lower bound approximation $Q_{0.05}[W_{40}^l(\pi^\sigma)]$ (solid line) and the comonotonic upper bound approximation $Q_{0.05}[W_{40}^c(\pi^\sigma)]$ (dashed line) for the 0.95 target capital $Q_{0.05}[W_{40}(\pi^\sigma)]$, for different values of the proportion invested in the tangency portfolio $\pi^{(t)}$, i.e. for the different portfolios on the mean-variance efficient set. We compare these quantiles with the simulated quantiles $Q_{0.05}[W_{40}^s(\pi^\sigma)]$ (dotted line). The simulation was performed by generating 20,000 paths using antithetic variables. We observe that the lower bound approximation is very close to the simulated results. Indeed, the maximum of the relative deviations $\left| \frac{Q_{0.05}[W_{40}^l(\pi^\sigma)] - Q_{0.05}[W_{40}^s(\pi^\sigma)]}{Q_{0.05}[W_{40}^s(\pi^\sigma)]} \right|$ is less than 0.5%. The maximum of the respective curves correspond to the (approximated or simulated) optimal portfolio.

For the comonotonic lower bound approximation we find that the optimal 0.95 - target capital K^l is given by $K^l = 89.78$. This target capital corresponds to a fraction 0.92 invested in the tangency portfolio.

The comonotonic upper bound approximation gives rise to an optimal fraction 0.51 invested in $\pi^{(t)}$, which corresponds to a maximal target capital $K^c = 82.25$. The simulated optimum is reached for a proportion equal to 0.92 and equals $K^s = 89.52$.

We can conclude that the lower bound approximation for the optimal investment strategy performs very well, compared to the simulated solution. From Figure 6.1 we also see that when the risky proportion increases, the target capital also increases until a certain level. Further increasing of the investment in the risky asset leads then to a decrease of the target capital. This observation is in accordance with intuition about optimal investment strategies.

In Figure 6.2, we assume constant saving amounts α at times $0, 1, \dots, 39$ and a target capital equal to 1 to be reached at time 40. We consider the investment strategy that minimizes the yearly savings amount for different probability levels p of the target capital. The computations were performed with the lower bound approximation $\overline{W}_{40}^l(\pi)$ for $\overline{W}_{40}(\pi)$.

Figure 6.1: The approximated target capitals $Q_{0.05}[W_{40}^l(\boldsymbol{\pi}^\sigma)]$ (solid line), $Q_{0.05}[W_{40}^c(\boldsymbol{\pi}^\sigma)]$ (dashed line) and the simulated target capital $Q_{0.05}[W_{40}^s(\boldsymbol{\pi}^\sigma)]$ (dotted line) as a function of the proportion invested in $\boldsymbol{\pi}^{(t)}$.

Figure 6.2: The minimal savings amount α^l (solid line - left scale) and the optimal proportion invested in $\pi^{(t)}$ (dashed line - right scale) as a function of p .

The solid line represents the (approximated) minimal savings amount α^l for different probability levels p of a target capital equal to 1 (left scale). As we see from the figure, when the required probability of reaching the target of 1, increases, the optimal savings amount increases as well. Note that the required savings amount in case of the risk-free investment, i.e. the one that corresponds to $p = 1$, is given by 0.0127.

The dashed line represents the (approximated) optimal risky proportion to be invested in the tangency portfolio $\pi^{(t)}$, for different probability levels p (right scale). As could be expected, when the probability of reaching the target capital increases, the optimal risky proportion in the portfolio will decrease.

6.3.4.5 Maximizing the CLTE

The optimal investment strategy π^* can also be defined as the one that maximizes the CLTE for a given value of p :

$$CLTE_{1-p}[W_n(\pi^*)] = \max_{\pi} CLTE_{1-p}[W_n(\pi)]. \quad (6.79)$$

Note that this optimization problem can be interpreted in terms of Yaari's dual theory of choice under risk. It is the problem faced by a risk averse decision-

maker with distortion function (6.40) who wants to optimize the distorted expectation of his final wealth. The optimization problem in this case can be expressed as follows: the decision-maker with target capital $Q_{1-p}^+[W_n(\boldsymbol{\pi})]$ maximizes the expected value of final wealth, given that the target capital is not reached. From (6.40) and (6.43) it follows that the n -dimensional maximization problem can again be reduced to a one-dimensional optimization problem over all portfolios on the Capital Market Line.

Approximating $W_n(\boldsymbol{\pi})$ by $W_n^l(\boldsymbol{\pi})$ or $W_n^c(\boldsymbol{\pi})$ and using the results of Example 3.4.4, lead to approximate solutions similar to the one derived above for problem (6.70). Indeed, the CLTE's of both approximations can be written as sums of CLTE's of lognormal random variables. Derivation of the results is left as an exercise to the reader.

6.3.5 Determining the investment strategy that maximizes the probability level for a given target capital

6.3.5.1 The probability of reaching the target

In this subsection, we will assume that the target capital $K > 0$ is given. For any investment strategy $\boldsymbol{\pi}$, the probability of reaching this target is given by

$$p = \overline{F}_{W_n(\boldsymbol{\pi})}(K), \quad (6.80)$$

where $\overline{F}_{W_n(\boldsymbol{\pi})}(x) = 1 - F_{W_n(\boldsymbol{\pi})}(x) = \Pr[W_n(\boldsymbol{\pi}) > x]$. The optimal investment strategy $\boldsymbol{\pi}^*$ is now defined as the one that maximizes the probability of reaching the target K . Denoting this optimal probability level by p^* , we have that

$$p^* = \max_{\boldsymbol{\pi}} \overline{F}_{W_n(\boldsymbol{\pi})}(K). \quad (6.81)$$

Consider two portfolios $\boldsymbol{\pi}_1$ and $\boldsymbol{\pi}_2$ with $\sigma(\boldsymbol{\pi}_1) = \sigma(\boldsymbol{\pi}_2)$ and $\mu(\boldsymbol{\pi}_1) < \mu(\boldsymbol{\pi}_2)$. From (6.41) we find that $\overline{F}_{W_n(\boldsymbol{\pi}_1)}(K) \leq \overline{F}_{W_n(\boldsymbol{\pi}_2)}(K)$. This implies that the solution of the optimization problem (6.81) is to be found on the Capital Market Line. Hence, we can replace (6.81) by

$$p^* = \max_{\sigma} \overline{F}_{W_n(\boldsymbol{\pi}^\sigma)}(K). \quad (6.82)$$

One possible choice for the target capital K is the final wealth that would arise if all savings were invested in the risk-free asset:

$$K^r = \sum_{i=0}^n \alpha_i e^{(n-i)r}. \quad (6.83)$$

Extending definition (6.53), the equity shortfall risk of a given investment strategy is now defined as

$$\text{ESR}(\boldsymbol{\pi}, n) = F_{W_n(\boldsymbol{\pi})}(K^r), \quad (6.84)$$

which is the probability that the investment strategy will underperform the risk-free investment strategy. Solving the maximization problem (6.81) with $K = K^r$ comes down to finding the investment strategy that minimizes the equity shortfall risk.

6.3.5.2 The comonotonic upper bound for $W_n(\pi)$

Neither the decumulative distribution function $\bar{F}_{W_n(\pi)}(x)$, nor its quantiles can be determined analytically. Therefore, we will introduce comonotonic approximations for $\bar{F}_{W_n(\pi)}(K)$.

The approximation π^c for the optimal investment strategy π^* of the problem (6.81) is defined as the investment strategy that maximizes $\bar{F}_{W_n^c(\pi)}(K)$ on the Capital Market Line. The approximation p^c for the optimal probability level p^* is then given by

$$p^c = \max_{\sigma} \bar{F}_{W_n^c(\pi^{\sigma})}(K). \quad (6.85)$$

In order to determine the probabilities $\bar{F}_{W_n^c(\pi^{\sigma})}(K)$, notice that for any $\sigma > 0$, it follows from (6.58) that $Q_p[W_n^c(\pi^{\sigma})]$ is a continuous and strictly increasing function of p , mapping $(0, 1)$ in $(0, \infty)$. This implies that $F_{W_n^c(\pi^{\sigma})}(x)$ is a strictly increasing and continuous function of x . Hence, for any K , we find that $F_{W_n^c(\pi^{\sigma})}(K)$ is the unique solution of the equation $Q_{F_{W_n^c(\pi^{\sigma})}(K)} = K$, which can be written as

$$\sum_{i=0}^n \alpha_i e^{(n-i) [\mu(\pi^{\sigma}) - \frac{1}{2}\sigma^2] - \sqrt{n-i} \sigma} \Phi^{-1}(\bar{F}_{W_n^c(\pi^{\sigma})}(K)) = K. \quad (6.86)$$

The approximation π^l for π^* is defined as the investment strategy that maximizes $\bar{F}_{W_n^l(\pi)}(K)$ on the Capital Market Line. The approximation p^l for the optimal probability level p^* is then given by

$$p^l = \max_{\sigma} \bar{F}_{W_n^l(\pi^{\sigma})}(K). \quad (6.87)$$

Similar as above, it can be proven that for any $\sigma > 0$, the probability $F_{W_n^l(\pi^{\sigma})}(K)$ follows from

$$\sum_{i=0}^n \alpha_i e^{(n-i) [\mu(\pi^{\sigma}) - \frac{1}{2}r_i^2(\pi^{\sigma})\sigma^2] - \sqrt{n-i} r_i^2(\pi^{\sigma}) \sigma} \Phi^{-1}(\bar{F}_{W_n^l(\pi^{\sigma})}(K)) = K. \quad (6.88)$$

Notice that the expressions (6.86) and (6.88) can be generalized to any investment strategy π with $\sigma(\pi) > 0$.

6.3.5.3 Numerical illustration

Assume the Black & Scholes market as described in Subsection 4.2.2. Consider saving amounts $\alpha_i = 1$ for $i = 0, 1, \dots, 39$, while $\alpha_{40} = 0$. We approximate the problem (6.81) by (6.87) for different values of the target capital K .

It suffices to consider portfolios that are linear combinations of the risk-free asset and the tangency portfolio $\pi^{(t)}$.

In Figure 6.3 we show the maximal probability p^l of reaching the target capital and the optimal proportion to be invested in $\pi^{(t)}$, as a function of the target capital K .

The solid line represents the (approximated) maximal probability levels p^l of

Figure 6.3: The maximal probability p^l of reaching the target capital (solid line - left scale) and the optimal proportion to be invested in $\pi^{(t)}$ (dashed line - right scale) as a function of the target capital K .

reaching a given target capital K (left scale), whereas the dashed line represents the (approximated) corresponding optimal proportion to be invested in $\pi^{(t)}$ (right scale).

The figure shows that increasing the level of the target capital leads to decreasing optimal probability levels and increasing risky proportions. Note that the risk-free investment corresponds to a target capital equal to 78.50 that is reached with a probability equal to 1.

6.4 Reserves for future obligations

6.4.1 General problem description

Consider a given set of deterministic obligations, i.e. a series of deterministic non-negative payments $\alpha_1, \alpha_2, \dots, \alpha_n$, that are due at times $1, 2, \dots, n$, respectively. Being able to meet these future obligations requires an appropriate funding, this means that appropriate assets have to be available to set up a reserve and/or a solvency margin. Following Atkinson & Dallas (2000), the reserve is defined as the amount of funds that have to be set aside as a liability in order to meet future obligations, whereas the solvency margin is the capital that regulators, rating agencies or the company itself deem necessary for the company to be able to withstand reasonable fluctuations in financial results.

At current time 0, assets of value R_0 are set up in order to be able to meet these future obligations. We will call R_0 the initial reserve. It can be interpreted as the reserve as defined above, or also as the total balance sheet requirement, i.e. reserve and solvency margin. We will assume that this reserve can be invested according to one of the constant mix portfolios π as defined in the Section 3.2. This investment strategy has to be chosen at time 0. Starting from the initial reserve R_0 and investing according to π , we define $R_j(R_0, \pi)$ at time j by the following recursion:

$$R_j(R_0, \pi) = R_{j-1}(R_0, \pi) e^{Y_j(\pi)} - \alpha_j, \quad j = 1, \dots, n, \quad (6.89)$$

with $R_0(R_0, \pi) = R_0$. Hence, $R_j(R_0, \pi)$ is the value of the assets that will be available at time j , after the payment of α_j , given that $R_0(R_0, \pi) = R_0$ is the initial reserve at time 0. The realization of $R_j(R_0, \pi)$ will be known at time j , and depends on the investment returns (stochastic part) and on the payments (deterministic part) in the past years. Often we will call $R_j(R_0, \pi)$ the (retrospective) reserve available at time j . Note that the random variables $Y_j(\pi)$ are i.i.d. and normal distributed with parameters $\mu(\pi)$ and $\sigma(\pi)$ as defined in (6.16). Solving the recursion (6.89), we find that the value of the assets available at time n is given by

$$R_n(R_0, \pi) = R_0 e^{\sum_{j=1}^n Y_j(\pi)} - \sum_{i=1}^n \alpha_i e^{\sum_{j=i+1}^n Y_j(\pi)}. \quad (6.90)$$

The random variable $S_j(\pi)$ is defined as the stochastically discounted value of all future payment obligations from time j on, given that the investment strategy is π :

$$S_j(\pi) = \sum_{i=j+1}^n \alpha_i e^{-(Y_{j+1}(\pi) + Y_2(\pi) + \dots + Y_i(\pi))}. \quad (6.91)$$

This random variable will be called the ‘stochastic future obligations’ at time j . The following relation holds between $S_j(\pi)$, $R_j(R_0, \pi)$ and $R_n(R_0, \pi)$:

$$R_n(R_0, \pi) = (R_j(R_0, \pi) - S_j(\pi)) e^{Y_{j+1}(\pi) + \dots + Y_n(\pi)}. \quad (6.92)$$

This relation implies that

$$R_n(R_0, \pi) \geq 0 \Leftrightarrow R_j(R_0, \pi) \geq S_j(\pi), \quad j = 0, \dots, n-1. \quad (6.93)$$

Hence,

$$\begin{aligned} \Pr[R_n(R_0, \pi) \geq 0] &= \Pr[R_j(R_0, \pi) \geq S_j(\pi), j = 0, \dots, n-1] \\ &= F_{S_0(\pi)}(R_0). \end{aligned} \quad (6.94)$$

Results similar to (6.94), applied in an obligation-based solvency framework for pension annuities, can be found in Olivieri & Pitacco (2003). Related results

in case of ‘deterministic future obligations’ can be found in Vanduffel, Dhaene, Goovaerts & Kaas (2003).

An investor will be interested in the probability $p = \Pr [R_n(R_0, \boldsymbol{\pi}) \geq 0] = F_{S_0(\boldsymbol{\pi})}(R_0)$ of ‘reaching the finish’, for different choices of the reserve R_0 and the investment strategy $\boldsymbol{\pi}$. In the next chapter we will deal with the situations where one does not only take into account the goal of ‘reaching the finish’, but also the conditions that year-to-year the provision R_j is larger than a given deterministic ‘hurdle’ value V_j with a sufficiently large probability.

Let us assume that the initial reserve for an investment strategy $\boldsymbol{\pi}$ is defined as $\rho_f[S_0(\boldsymbol{\pi})]$ where ρ_f is a distortion risk measure as defined in (6.36). The optimal investment strategy could then be defined as the one that corresponds to the following minimization problem:

$$\min_{\boldsymbol{\pi}} \rho_f[S_0(\boldsymbol{\pi})]. \quad (6.95)$$

Notice that

$$S_0(\boldsymbol{\pi}) \stackrel{d}{=} \sum_{i=1}^n \alpha_i e^{-i[\mu(\boldsymbol{\pi}) - \frac{1}{2}\sigma^2(\boldsymbol{\pi})] + \sqrt{i}\sigma(\boldsymbol{\pi})Z_i}, \quad (6.96)$$

where $(Z_1, \dots, Z_n)$ is a multivariate normal random vector with standard normal distributed marginals.

Now consider two portfolios $\boldsymbol{\pi}_1$ and $\boldsymbol{\pi}_2$ with $\sigma(\boldsymbol{\pi}_1) = \sigma(\boldsymbol{\pi}_2)$ and $\mu(\boldsymbol{\pi}_1) < \mu(\boldsymbol{\pi}_2)$. As in Ahcan, Darkiewicz, Dhaene, Goovaerts & Hoedemakers (2004), we find that

$$F_{S_0(\boldsymbol{\pi}_1)}(x) \leq F_{S_0(\boldsymbol{\pi}_2)}(x), \quad x \geq 0, \quad (6.97)$$

which implies

$$\rho_f[S_0(\boldsymbol{\pi}_1)] \geq \rho_f[S_0(\boldsymbol{\pi}_2)]. \quad (6.98)$$

Hence, the solution of the optimization problem (6.95) can be found on the Capital Market Line. The optimization problem (6.95) can be reduced to

$$\min_{\boldsymbol{\sigma}} \rho_f[S_0(\boldsymbol{\pi}^\sigma)]. \quad (6.99)$$

6.4.2 The case of a single obligation

In this subsection, we consider the special case of setting a reserve at time 0 for a single payment obligation at time n . Hence, $\alpha_1 = \alpha_2 = \dots = \alpha_{n-1} = 0$ and $\alpha_n = 1$. The cumulative distribution function of the stochastic provision $S_0(\boldsymbol{\pi})$ follows from

$$S_0(\boldsymbol{\pi}) \stackrel{d}{=} e^{-n[\mu(\boldsymbol{\pi}) - \frac{1}{2}\sigma^2(\boldsymbol{\pi})] + \sqrt{n}\sigma(\boldsymbol{\pi})\Phi^{-1}(U)}, \quad (6.100)$$

with U uniformly distributed on the unit interval. For a given investment strategy $\boldsymbol{\pi}$ and a given probability level p (with p sufficiently large), we determine the initial reserve R_0 as the p -quantile of $S(\boldsymbol{\pi})$:

$$R_0 = Q_p[S_0(\boldsymbol{\pi})]. \quad (6.101)$$

This reserving principle clearly makes sense, as one can easily prove that

$$Q_p[S_0(\boldsymbol{\pi})] = \inf \{x \mid \Pr[R_n(x, \boldsymbol{\pi}) \geq 0] \geq p\}. \quad (6.102)$$

This means that for a given investment strategy $\boldsymbol{\pi}$, the p -quantile reserving principle determines the initial reserve as the ‘smallest’ amount such that the probability of ‘reaching the finish’ is at least p . From (3.21) it follows that the quantile $Q_p[S_0(\boldsymbol{\pi})]$ is given by

$$Q_p[S_0(\boldsymbol{\pi})] = e^{-n(\mu(\boldsymbol{\pi}) - \frac{1}{2}\sigma^2(\boldsymbol{\pi})) + \sqrt{n}\sigma(\boldsymbol{\pi})\Phi^{-1}(p)}. \quad (6.103)$$

The optimal investment strategy $\boldsymbol{\pi}^*$ is defined as the one that minimizes $Q_p[S_0(\boldsymbol{\pi})]$. The initial provision R_0^* is then set equal to this minimal quantile:

$$R_0^* = \min_{\boldsymbol{\pi}} Q_p[S_0(\boldsymbol{\pi})]. \quad (6.104)$$

As

$$Q_p[S_0(\boldsymbol{\pi})] = \frac{1}{Q_{1-p}^+[W_n(\boldsymbol{\pi})]} \quad (6.105)$$

with $Q_{1-p}^+[W_n(\boldsymbol{\pi})]$ given by (6.49), we find that the optimal portfolio of problem (6.104) is identical to the optimal portfolio of problem (6.48). Hence, $\boldsymbol{\pi}^*$ is given by (6.50).

We can conclude that increasing the investment time horizon transforms the optimal investment strategy into a more risk-taking one. Investors with a longer time horizon should have a larger exposure to stocks relative to investors with a shorter time horizon. The optimal risky proportion converges to (6.47), which corresponds to the so-called optimal growth portfolio.

6.4.3 Comonotonic approximations for the general case

From (6.94) we see that in order to compute the ‘probability of reaching the finish’ i.e. $\Pr[R_n(R_0, \boldsymbol{\pi}) \geq 0]$, for a given pair $(R_0, \boldsymbol{\pi})$, we have to determine the c.d.f. of $S_0(\boldsymbol{\pi})$. However, the random variable $S_0(\boldsymbol{\pi})$ is a linear combination of dependent lognormal random variables. This implies that it is impossible to determine the cumulative distribution function of $S_0(\boldsymbol{\pi})$ analytically. Therefore, we will consider a convex order upper bound $S^c(\boldsymbol{\pi})$ and a convex order lower bound $S^l(\boldsymbol{\pi})$ for $S(\boldsymbol{\pi})$. Rewriting $S_0(\boldsymbol{\pi})$ as

$$S_0(\boldsymbol{\pi}) = \sum_{i=1}^n \alpha_i e^{Z_i}, \quad (6.106)$$

we can apply the results of Example 3.4.4 with

$$\begin{aligned} Z_i &= -Y_1(\boldsymbol{\pi}) - Y_2(\boldsymbol{\pi}) - \dots - Y_i(\boldsymbol{\pi}), \\ E[Z_i] &= -i \left[\mu(\boldsymbol{\pi}) - \frac{1}{2}\sigma^2(\boldsymbol{\pi}) \right], \\ \sigma_{Z_i}^2 &= i \sigma^2(\boldsymbol{\pi}). \end{aligned} \quad (6.107)$$

The comonotonic upper bound $S_0^c(\boldsymbol{\pi}) = \sum_{i=1}^n F_{\alpha_i}^{-1} e^{Z_i}(U)$ for $S_0(\boldsymbol{\pi})$ is given by

$$S_0^c(\boldsymbol{\pi}) = \sum_{i=1}^n \alpha_i \exp \left[-i \mu(\boldsymbol{\pi}) + \frac{1}{2} i \sigma^2(\boldsymbol{\pi}) + \sqrt{i} \sigma(\boldsymbol{\pi}) \Phi^{-1}(U) \right], \quad (6.108)$$

while for $p \in (0, 1)$, the quantiles and CTE's of $S_0^c(\boldsymbol{\pi})$ are given by

$$\begin{aligned} Q_p[S_0^c(\boldsymbol{\pi})] &= \sum_{i=1}^n \alpha_i e^{-i \mu(\boldsymbol{\pi}) + \frac{1}{2} i \sigma^2(\boldsymbol{\pi}) + \sqrt{i} \sigma(\boldsymbol{\pi}) \Phi^{-1}(p)}, \\ CTE_p[S_0^c(\boldsymbol{\pi})] &= \sum_{i=1}^n \alpha_i e^{-i \mu(\boldsymbol{\pi}) + i \sigma^2(\boldsymbol{\pi})} \frac{\Phi(\sqrt{i} \sigma(\boldsymbol{\pi}) - \Phi^{-1}(p))}{1-p}. \end{aligned} \quad (6.109)$$

In order to define a convex lower bound $S_0^l(\boldsymbol{\pi})$ for $S_0(\boldsymbol{\pi})$, we choose the conditioning random variables $\Lambda(\boldsymbol{\pi})$ as

$$\Lambda(\boldsymbol{\pi}) = \sum_{j=1}^n \gamma_j(\boldsymbol{\pi}) Z_j(\boldsymbol{\pi}), \quad (6.110)$$

with coefficients $\gamma_j(\boldsymbol{\pi})$ that follow from (4.14):

$$\gamma_j(\boldsymbol{\pi}) = -\alpha_j e^{-j(\mu(\boldsymbol{\pi}) - \sigma^2(\boldsymbol{\pi}))}.$$

For this choice of the parameters $\beta_j(\boldsymbol{\pi})$, the variance of the lower bound will be close to the variance of $S_0(\boldsymbol{\pi})$, provided $\sigma^2(\boldsymbol{\pi})$ is small enough. Hence, the lower bound

$$S_0^l(\boldsymbol{\pi}) = E[S_0(\boldsymbol{\pi}) | \Lambda(\boldsymbol{\pi})] \quad (6.111)$$

will have a cumulative distribution function that is close to the cumulative distribution function of $S_0(\boldsymbol{\pi})$, provided $\sigma(\boldsymbol{\pi})$ is small enough. We find from Example 3.4.4 that

$$S_0^l(\boldsymbol{\pi}) = \sum_{i=1}^n \alpha_i e^{-i \mu(\boldsymbol{\pi}) + (1 - \frac{1}{2} r_i^2(\boldsymbol{\pi})) i \sigma^2(\boldsymbol{\pi}) + r_i(\boldsymbol{\pi}) \sqrt{i} \sigma(\boldsymbol{\pi}) \Phi^{-1}(U)} \quad (6.112)$$

where the coefficients $r_i(\boldsymbol{\pi})$ are given by

$$r_i(\boldsymbol{\pi}) = \frac{\sum_{j=1}^i \sum_{k=j}^n \alpha_k e^{-k(\mu(\boldsymbol{\pi}) - \sigma^2(\boldsymbol{\pi}))}}{\sqrt{i} \sqrt{\sum_{j=1}^n \left(\sum_{k=j}^n \alpha_k e^{-k(\mu(\boldsymbol{\pi}) - \sigma^2(\boldsymbol{\pi}))} \right)^2}}. \quad (6.113)$$

Note that the correlation coefficients $r_i(\boldsymbol{\pi})$ are non-negative. This implies that $S^l(\boldsymbol{\pi})$ is a comonotonic sum of lognormal random variables. From Example

3.4.4, we find the following expressions for the risk measures $Q_p [S_0^l(\pi)]$ and $CTE_p [S_0^l(\pi)]$, $p \in (0, 1)$:

$$Q_p [S_0^l(\pi)] = \sum_{i=1}^n \alpha_i e^{-i \mu(\pi) + (1 - \frac{1}{2} r_i^2(\pi)) i \sigma^2(\pi) + r_i(\pi) \sqrt{i} \sigma(\pi)} \Phi^{-1}(p), \quad (6.114)$$

$$CTE_p [S_0^l(\pi)] = \sum_{i=1}^n \alpha_i e^{-i \mu(\pi) + i \sigma^2(\pi)} \frac{\Phi(r_i(\pi) \sqrt{i} \sigma(\pi) - \Phi^{-1}(p))}{1 - p}.$$

Finally, we have that

$$CTE_p [S_0^l(\pi)] \leq CTE_p [S_0(\pi)] \leq CTE_p [S_0^c(\pi)], \quad (6.115)$$

Note however that the approximations $Q_p [S_0^l(\pi)]$ and $Q_p [S_0^c(\pi)]$ for $Q_p [S_0(\pi)]$ are not necessarily ordered in the same way.

6.4.4 Determining the investment strategy that minimizes the p -quantile initial reserve for a given probability level

6.4.4.1 The p -quantile reserving principle

As in Section 6.3, we set the initial reserve for a given investment strategy π and a given probability level p , equal to the p -quantile of $S_0(\pi)$:

$$R_0 = Q_p [S_0(\pi)]. \quad (6.116)$$

As we noted in Section 5.2, the p -quantile reserving principle determines the initial reserve as the ‘smallest’ amount such that the probability of ‘reaching the finish’ is at least p , see (6.102).

For a given probability level p , we determine the optimal investment strategy π^* as the one that minimizes $Q_p [S_0(\pi)]$. The initial reserve R_0^* is then set equal to this minimal quantile:

$$R_0^* = \min_{\pi} Q_p [S_0(\pi)]. \quad (6.117)$$

Similar to the optimization problem (6.70), the optimization problem (6.117) suffers from a ‘curse of dimensionality’, as $S_0(\pi)$ is a sum of n non-independent random variables. The dimensionality problem will be solved by introducing comonotonic approximations for the stochastic provision $S_0(\pi)$. The portfolio dimensionality can be reduced because the solution of the minimization problem (6.117) will be an element of the Capital Market Line, see (6.95) and (6.99). Hence, we can transform (6.117) into

$$R_0^* = \min_{\sigma} Q_p [S_0(\pi^\sigma)]. \quad (6.118)$$

6.4.4.2 Comonotonic bounds for $S_0(\pi)$

First, we propose to approximate the optimal investment strategy π^* by π^c , which is the investment strategy that minimizes $Q_p[S_0^c(\pi)]$ on the Capital Market Line, and we approximate the initial provision R_0^* by R_0^c which follows from:

$$R_0^c = \min_{\sigma} Q_p[S_0^c(\pi^\sigma)], \quad (6.119)$$

where the quantiles $Q_p[S_0^c(\pi^\sigma)]$ are given by (6.109).

Secondly, we propose the approximation π^l for π^* , where π^l is the investment strategy that minimizes $Q_p[S_0^l(\pi)]$ on the efficient frontier, and we approximate the initial provision R_0^* by R_0^l which follows from

$$R_0^l = \min_{\sigma} Q_p[S_0^l(\pi^\sigma)]. \quad (6.120)$$

The quantiles $Q_p[S_0^l(\pi)]$ are given by (6.114).

The approximations (6.119) and (6.120) both reduce the time multidimensionality of problem (6.117), or equivalently (6.118), to dimension 1. As we will illustrate in the next subsection, it is especially the approximation (6.120) that will perform very well.

6.4.4.3 Numerical illustration

Consider the Black & Scholes market as considered in Section 4.2.2. consisting of one risk-free asset and two risky assets. We want to solve (approximately) problem (6.117) in case of a series of obligations α_i that are all equal to 1, for $i = 1, \dots, 40$. As the solutions of the problems (6.119) and (6.120) are to be found on the efficient frontier, we can again restrict to the case of a market consisting of the risk-free asset and a risky asset that corresponds to the tangency portfolio.

In Figure 6.4, the comonotonic lower bound approximation $Q_{0.95}[S_0^l(\pi^\sigma)]$ and the comonotonic upper bound approximation $Q_{0.95}[S_0^c(\pi^\sigma)]$ are given for different values of the proportion invested in the tangency portfolio $\pi^{(t)}$. The corresponding simulated quantiles $Q_{0.95}[S_0^s(\pi^\sigma)]$ are also given. The simulation is obtained by generating 20,000 paths using antithetic variables. We can conclude that the lower bound approximation performs extremely well. Indeed, the maximal value of the relative deviation $\left| \frac{Q_{0.95}[S_0^l(\pi^\sigma)] - Q_{0.95}[S_0^s(\pi^\sigma)]}{Q_{0.95}[S_0^s(\pi^\sigma)]} \right|$ was found to be as small as 0.20%.

The minimum of the respective curves corresponds to the (approximated or simulated) optimal portfolio. For the comonotonic lower bound approximation we find that the optimal proportion invested in the tangency portfolio is given by $\pi^l = 0.35$. This corresponds to an optimal initial reserve given by $R_0^l = 22.442$. The comonotonic upper bound approximation gives rise to an optimum $R_0^c = 22.945$ corresponding with an optimal proportion of 0.015.

The simulated optimal provision is given by $R_0^s = 22.444$, with a risky proportion of 0.345. Notice the high accuracy of the comonotonic lower bound approximation.

Figure 6.4: The approximated reserves $Q_{0.95}[S^l(\pi^\sigma)]$ (solid line) and $Q_{0.95}[S^c(\pi^\sigma)]$ (dashed line) and the simulated reserve $Q_{0.95}[S^s(\pi^\sigma)]$ (dotted line) as a function of the proportion to be invested in $\pi^{(t)}$.

Figure 6.5: The minimal Q_p -reserve R_0^l (solid line - left scale) and the optimal proportion to be invested in $\pi^{(t)}$ (dashed line - right scale) as a function of p .

In Figure 6.5, we consider the same obligations pattern and determine the investment strategy that minimizes the initial p -quantile reserve for different probability levels p . The computations were performed using the lower bound approximation $S_0^l(\pi)$ for $S_0(\pi)$. The dashed line represents the (approximated) optimal proportions to be invested in the tangency portfolio, for different probability levels p (right scale). The solid line represents the (approximated) initial reserve for different probability levels p (left scale). In accordance with intuition, we find that increasing the probability of reaching the finish will increase the optimal initial reserve and decrease the optimal risky proportion.

Note that if one sets the probability level p equal to 1, the investment strategy is completely risk-free. The initial reserve is in this case given by 22.946.

6.4.5 Determining the investment strategy that minimizes the CTE_p -quantile initial reserve for a given probability level

6.4.5.1 The CTE_p -quantile initial reserve

In this subsection we will set the initial reserve, for a given probability level p , and a given investment strategy π , equal to $CTE_p[S(\pi)]$:

$$R_0 = CTE_p[S_0(\pi)]. \quad (6.121)$$

One can prove that

$$CTE_p[S_0(\pi)] = Q_p[S_0(\pi)] + \left| E \left[R_n e^{-Y_1(\pi) - \dots - Y_n(\pi)} \mid R_n < 0 \right] \right|, \quad (6.122)$$

where R_n is given by

$$R_n = R_n(Q_p[S_0(\pi)], \pi). \quad (6.123)$$

The second term in (6.122) can be interpreted as the expected discounted shortfall, given that the initial reserve is set according to the p -quantile reserving principle and given that there is a shortfall. Hence, with the CTE_p reserving principle, the initial reserve is set equal to the sum of the p -quantile initial reserve and the expected discounted conditional shortfall of the p -quantile initial reserve.

Assuming that the probability level p is fixed, the optimal investment strategy π^* is now determined as the one that minimizes $CTE_p[S_0(\pi)]$. The initial provision R_0^* is set equal to this minimal conditional tail expectation:

$$R_0^* = \min_{\pi} CTE_p[S_0(\pi)]. \quad (6.124)$$

From the equivalence of (6.95) and (6.99), we find that the minimization problem (6.124) can be transformed into

$$R_0^* = \min_{\sigma} CTE_p[S_0(\pi^\sigma)]. \quad (6.125)$$

As the Conditional Tail Expectations cannot be determined analytically, we again propose comonotonic approximations for this minimization problem.

6.4.5.2 Comonotonic bounds for $S_0(\pi)$

First, we propose to approximate the optimal investment strategy π^* by the investment strategy π^c that minimizes $CTE_p[S_0^c(\pi)]$ on the efficient frontier. The initial reserve R_0^* is approximated by R_0^c :

$$R_0^c = \min_{\sigma} CTE_p[S_0^c(\pi^\sigma)]. \quad (6.126)$$

The quantities $CTE_p[S_0^c(\pi^\sigma)]$ can be determined from (6.109).

Next, we propose to approximate π^* by π^l , which is the investment strategy that minimizes $CTE_p[S_0^l(\pi)]$ on the efficient frontier. The initial reserve R_0^* is then approximated by

$$R_0^l = \min_{\sigma} CTE_p[S_0^l(\pi^\sigma)]. \quad (6.127)$$

The quantities $CTE_p[S_0^l(\pi)]$ are given by (6.114).

It is straightforward to prove that the following inequalities hold for any investment strategy π :

$$CTE_p[S_0^l(\pi)] \leq CTE_p[S_0(\pi)] \leq CTE_p[S_0^c(\pi)]. \quad (6.128)$$

These inequalities imply

$$R_0^l \leq R_0^* \leq R_0^c. \quad (6.129)$$

Figure 6.6: The minimal CTE_p -reserve R_0^l (solid line - left scale) and the optimal proportion invested in $\pi^{(t)}$ (dashed line - right scale) as a function of p .

6.4.5.3 Numerical illustration

Consider the Black & Scholes market as presented in Section 4.2.2. Now we want to solve (approximately) problem (6.124) in case of a series of obligations α_i that are all equal to 1, for $i = 1, \dots, 40$. As the solution of problem (6.124) is a portfolio on the efficient frontier, we can again restrict to the case of a market consisting of the risk-free asset and a risky asset that corresponds to the tangency portfolio.

In Figure 6.6, we consider the investment strategy that minimizes the initial CTE_p -reserve for different values of the probability level p . The computations were performed with the lower bound approximation $S^l(\pi)$ for $S(\pi)$.

The dashed line represents the (approximated) optimal risky proportion π^l to be invested in the tangency portfolio, for different probability levels p (right scale). As could be expected, increasing the required probability level of reaching the finish will decrease the optimal risky proportion in the portfolio.

The solid line represents the (approximated) CTE -reserve for different probability levels p (left scale). As we see from the figure, increasing the probability will increase the optimal reserve.

Comparing Figures 5 and 6, we see that for a given probability level p , the CTE provisioning principle will lead to a less risky investment strategy than the quantile provisioning principle.

6.4.6 Determining the investment strategy that maximizes the probability level for a given initial reserve

6.4.6.1 The probability of reaching the finish

From (6.94), we find that for a given investment strategy π and a given initial reserve R_0 , the probability p of ‘reaching the finish’ is given by

$$p = \Pr [R_n (R_0, \pi) \geq 0] = F_{S_0(\pi)}(R_0). \quad (6.130)$$

In this subsection, we will assume that the initial reserve $R_0 > 0$ is given. We propose to determine the optimal investment strategy π^* as the one that maximizes the probability of ‘reaching the finish’ in relation (6.130). Denoting this maximal probability by p^* , we find

$$p^* = \max_{\pi} F_{S_0(\pi)}(R_0). \quad (6.131)$$

Consider two portfolios π_1 and π_2 with $\sigma(\pi_1) = \sigma(\pi_2)$ and $\mu(\pi_1) < \mu(\pi_2)$. From (6.97) we find that the solution of the optimization problem (6.131) is to be found on the Capital Market Line. Hence, we can replace (6.131) by

$$p^* = \max_{\sigma} F_{S_0(\pi^\sigma)}(R_0).$$

Neither the probabilities $F_{S_0(\pi^\sigma)}(R_0)$, nor the quantiles $Q_p[S_0(\pi^\sigma)]$ can be determined analytically. Moreover, solving optimization problem (6.131) by performing a simulation is time-consuming, due to the multi-dimensionality in time. Therefore, we will again propose comonotonic approximations for π^* and p^* .

6.4.6.2 Comonotonic bounds for $S_0(\pi)$

A first approximation consists in approximating the optimal investment strategy π^* by the investment strategy π^c which is the one that maximizes $F_{S_0^c(\pi)}(R_0)$. The probability of ‘reaching the finish’ p^* is then approximated by p^c , which follows from

$$p^c = \max_{\sigma} F_{S_0^c(\pi^\sigma)}(R_0). \quad (6.132)$$

For any investment strategy π^σ , with $\sigma > 0$, it follows from (6.109) that $Q_p(S_0^c(\pi^\sigma))$ is a continuous and strictly increasing function of p , mapping $(0, 1)$ in $(0, \infty)$. This implies that $F_{S_0^c(\pi^\sigma)}(x)$ is a strictly increasing and continuous function of x . Hence, for any R_0 , we find that $F_{S_0^c(\pi^\sigma)}(R_0)$ is the unique solution of the equation $Q_{F_{S_0^c(\pi^\sigma)}(R_0)} = R_0$, or equivalently, $F_{S_0^c(\pi^\sigma)}(R_0)$ is the unique solution of

$$\sum_{i=1}^n \alpha_i e^{-i \mu(\pi^\sigma) + \frac{1}{2} i \sigma^2 + \sqrt{i} \sigma \Phi^{-1}(F_{S_0^c(\pi^\sigma)}(R_0))} = R_0. \quad (6.133)$$

The approximation π^l for π^* is the investment strategy that maximizes $F_{S_0^l(\pi)}(R_0)$. The probability of ‘reaching the finish’ p^* is then approximated by p^l , which is given by

$$p^l = \max_{\sigma} F_{S_0^l(\pi^\sigma)}(R_0). \quad (6.134)$$

Provided $\sigma > 0$, one finds that $F_{S_0^l(\pi^\sigma)}(R_0)$ is the unique solution of the following equation:

$$\sum_{i=1}^n \alpha_i e^{-i \mu(\pi^\sigma) + (1 - \frac{1}{2} r_i^2(\pi^\sigma)) i \sigma^2 + r_i(\pi^\sigma) \sqrt{i} \sigma \Phi^{-1}(F_{S_0^l(\pi^\sigma)}(R_0))} = R_0. \quad (6.135)$$

6.4.6.3 Numerical illustration

Consider again the Black & Scholes market as introduced in Section 4.2.2. We want to find the optimal investment strategy determined by (6.131) for a series of future obligations α_i that are all equal to 1, for $i = 1, \dots, 40$. We propose to approximate the exact solution by the solution of problem (6.134). In Figure 6.7 we show the maximal probability p^l of reaching the finish (solid line - left scale), as well as the optimal proportion to be invested in $\pi^{(t)}$ (dashed line - right scale), as a function of the initial reserve R_0 .

The figure shows that increasing the level of the initial reserve will lead to an increase of the maximal probability of reaching the finish. On the other hand, increasing the initial reserve will lead to an optimal investment strategy that is less risky.

6.5 Final remarks

In this chapter we considered the problem of how the available funds should be allocated among a basket of risk-free and risky assets, when the available investment strategies are to be chosen within the class of constant mix strategies. Two general asset allocation decision problems were distinguished. The *terminal wealth* problem considers the optimal investment mix in a situation where at regular points in time saving amounts are added to the available funds. The *reserving problem* describes situations where funds are set up and invested in order to be able to fulfill a future deterministic consumption pattern. For both problems several optimization criteria were considered.

These portfolio selection problems can be considered as multidimensional in two directions. First, they are multidimensional in terms of portfolio choice as the proportions related to the $m + 1$ assets have to be chosen. They are also multidimensional in terms of time as n yearly returns $Y_i(\pi)$ are involved. The portfolio-dimension is easily reduced to a single dimension, as it is straightforward to prove that optimal distorted expectations correspond to portfolios on the Markowitz mean-variance efficient frontier. On the other hand, the complexity caused by time-dimensionality and dependency cannot be solved in a straightforward way.

Figure 6.7: The maximal probability of reaching the finish p^l (solid line - left scale) and the optimal proportion invested in $\pi^{(t)}$ (dotted line - right scale) as a function of the initial reserve R_0 .

The classical way to solve the optimization problems we considered in this chapter is via Monte Carlo simulation. Take as an example the final wealth problem described in (6.70). Such a strategy starts with simulating series of outcomes $(u_1, u_2, \dots, u_n)$ of the multivariate random vector $(U_1, U_2, \dots, U_n)$ with mutually independent marginals that are uniformly distributed on the unit interval. From these simulated values, one determines series of yearly returns $(y_1(\boldsymbol{\pi}), y_2(\boldsymbol{\pi}), \dots, y_n(\boldsymbol{\pi}))$ for each possible portfolio $\boldsymbol{\pi}^T = (\pi_1, \pi_2, \dots, \pi_m)$ on the efficient frontier. This eventually leads to a simulated value of the quantiles $Q_{1-p}^+[W_n(\boldsymbol{\pi})]$ for all efficient portfolios. The optimal investment strategy is then the one that correspond to the largest quantile. Because of the multiple simulations and calculations involved, this method is time-consuming, with a trade-off to be made between speed and accuracy.

In this chapter we proposed a way to escape the ‘curse of time-dimensionality’. The comonotonicity approach reduces the time-dimensionality to one dimension, as the randomness of the n - dimensional vector of yearly returns is reduced to the randomness of one single uniform variable U . Moreover, the comonotonicity technique avoids simulation as analytical expressions for approximations of the quantiles are available. These expressions can be computed very quickly and are highly accurate at the same time.

Many of the results presented in this chapter can be generalized in several directions. A first immediate generalization consists in finding optimal invest-

ments mixes when restrictions are set on the proportions held in the different asset classes. E.g. there may be general linear constraints on the weights π_i , such as

$$l_j \leq a_0^j + \boldsymbol{\pi}^T \times (\mathbf{a}^j - a_0^j \mathbf{1}) \leq u_j, \quad j = 1, \dots, c, \quad (6.136)$$

with $\mathbf{a}^j = (a_1, a_2, \dots, a_m)$. These bounds might express restrictions on the individual weights, such as “the fraction of the portfolio in property must be between 0% and 15%” or “a fixed percentage of the fund should be available in cash” and “there shall be no short-selling of assets”. But also more general linear restrictions such as “the total fraction of the portfolio allocated to all international assets must not exceed 40%” can be expressed by a constraint of the form (6.136).

It is straightforward to prove that with constraints of this type, the approximations for the investor’s optimal portfolio choice of all problems considered in this chapter, will be found on the adjusted mean-variance efficient frontier $\{(\sigma, \mu(\boldsymbol{\pi}^\sigma))\}$ where for a given level of the volatility σ , the mean-variance efficient portfolio $\boldsymbol{\pi}^\sigma$ is the one that corresponds to the maximum in the following problem:

$$\text{Max}_{\boldsymbol{\pi} \in R} \mu(\boldsymbol{\pi}) \text{ subject to } \sigma(\boldsymbol{\pi}) = \sigma, \quad (6.137)$$

where R is the set of all portfolios $\boldsymbol{\pi}$ fulfilling the constraints (6.136). Note however that in this case, the efficient frontier cannot be expressed analytically anymore and the special structure of the efficient frontier is destroyed. However, powerful numerical methods are available to solve the mean-variance problem (6.137).

In this chapter, we restricted the optimization function to be a quantile, or a conditional (left) tail expectation. Many results can be generalized to the case where the optimization function is a distortion risk measure. Indeed, as is explained in Dhaene, Vanduffel, Tang, Goovaerts, Kaas & Vyncke (2004) for instance, any distortion risk measure of a sum of comonotonic random variables such as $W_n^l(\boldsymbol{\pi})$, $W_n^c(\boldsymbol{\pi})$, $S_0^l(\boldsymbol{\pi})$ and $S_0^c(\boldsymbol{\pi})$ can be expressed as the sum of the distortion risk measures of the (lognormal) random variables involved.

We considered the portfolio selection problem within the framework of a Black & Scholes market. In particular we assumed that the drift and volatility of the different asset classes are constant over time and that the yearly returns are lognormally distributed. The results can be generalized in a straightforward way to take into account the time-dependency of drifts and volatilities. Also many of the results presented here can be generalized to other than normal distributions for the yearly investment returns. In particular, many results can be generalized in a Lévy-type or elliptical-type world. Comonotonic approximations for sums of random variables with distributions of this type are considered in Valdez & Dhaene (2004) and Albrecher, Dhaene, Goovaerts & Schoutens (2004).

The hurdle-race problem

7.1 Introduction

⁸One of the optimal portfolio selection problems we considered in Chapter 6 was the so-called *reserving problem*: we determined an investment strategy that minimized the initial provision R_0 we had to set-up in order to meet, with a sufficiently high probability, deterministic (positive) obligations $\alpha_1, \alpha_2, \dots, \alpha_n$ that were due at times 1, 2, ..., n . Loosely speaking, we determined the initial provision such that it was, as small as possible, but still high enough in order to ‘reach the finish’ with probability $1 - \varepsilon_n$. A situation in which the first α_j are positive (hence payments) and moderate in absolute value, but the last ones are negative (hence incomes) and large in absolute value, this may lead to a situation where the available provision R_j in the first years becomes negative (or below a predefined level), which may be an undesirable situation.

In this chapter, we will present another way of determining the initial provision, which does not only take into account the goal of ‘reaching the finish’, but also the conditions that, year-to-year, the available provision R_j is larger than a predefined value V_j . These additional requirements V_j are the ‘hurdles’ that have to be taken with probabilities $1 - \varepsilon_1, 1 - \varepsilon_2, \dots, 1 - \varepsilon_n$. In case of an insurer establishing his reserve, they might be imposed by a supervisory authority or by internal policy. Determining the initial provision in this way allows one to make the probability of taking the hurdles time-dependent. In situations where year-to-year adjustments of the level of the reserve are possible, the required probabilities for taking the ‘hurdles’ in the first years could be chosen larger than the probabilities for the later years. Using time-dependent probabilities and/or considering cash-flows with mixed signs can lead to situations where the binding constraint is not ‘reaching the finish’ but ‘taking an intermediate hurdle’ and our method will reflect this in the required level for the provision.

We note that we work with ‘physical probabilities’ which implies that the

⁸ “He that would have the fruit must climb the tree” - Thomas Fuller.

required provision is different from the arbitrage free price of this cash-flow. Indeed, another way of determining the provision follows from noting that these future liabilities are equivalent to the liabilities associated with selling n zero coupon bonds with face values $\alpha_1, \alpha_2, \dots, \alpha_n$ respectively. The provision can then be determined as the price of this zero-coupon bond portfolio. Investing the provision in this sequence of zero-coupon bonds will exactly generate the amount α_i at times i ($i = 1, 2, \dots, n$), so that we will be able to meet our future obligations with certainty under this investment strategy.

In this chapter however, we will determine the level of the initial provision under a given risky investment strategy. We will assume that the provision will be invested such that it generates stochastic yearly returns $Y_1, Y_2, \dots, Y_n$ in the coming years. The provision will be determined such that the probability that we will be able to meet our future obligations at all intermediate times $i = 1, 2, \dots, n$ will be sufficiently large. Conversely, if the level of the provision is given, our methodology will enable us to compute the probability that we will be able to meet our future obligations under the given investment strategy.

An example of a situation in which our methodology is appropriate, is when an insurer wants to determine the current reserve or capital required to meet his future obligations. An example in the framework of personal finance is an individual who wants to invest today an amount that will provide him, with a certain percentage of certitude, a predetermined fixed income during the coming 20 years.

The agenda of this chapter is as follows. In Section 7.2 we describe how we can determine the initial provision R_0 . Next, in Section 7.3, we apply in a Gaussian setting the comonotonic upper and lower bounds to approximate R_0 . Finally, in the last Section, we illustrate our results by several numerical examples. This chapter is based on Vanduffel, Dhaene, Goovaerts & Kaas (2003).

7.2 Determining the provision

In line with (6.89), let R_0 be the provision at time 0, and consider the stochastic return process $(Y_1, \dots, Y_n)$, i.e. 1 unit invested at time 0 is assumed to grow to $e^{Y_1 + \dots + Y_j}$ at time j ($j = 1, 2, \dots, n$). Let R_j be defined recursively by

$$R_j(R_0) = R_{j-1}(R_0) e^{Y_j} - \alpha_j, \quad j = 1, \dots, n. \quad (7.1)$$

Hence, R_j is the (stochastic) provision that will be available at time j , after the payment of α_j , given that R_0 is the provision at time 0. The realization of R_j will be known at time j and depends on the investment returns (stochastic part) and on the payments (deterministic part) in the past years.

One could determine the initial provision as the minimal amount such that R_n will be non-negative with a probability of at least $1 - \varepsilon_n$, with ε_n sufficiently small. This means that we determine the initial provision in such a way that we will be able to ‘reach the finish’ with a predefined (high) probability:

$$\overline{R}_0 = \inf \{ R_0 \mid \Pr [R_n(R_0) \geq 0] \geq 1 - \varepsilon_n \}. \quad (7.2)$$

By solving the recursion (7.1), we find that the subsequent provisions R_j equal

$$R_j = R_0 e^{Y_1 + \dots + Y_j} - \sum_{i=1}^n \alpha_i e^{Y_{i+1} + \dots + Y_j}, \quad j = 1, 2, \dots, n. \quad (7.3)$$

We will always conventionally assume that $\sum_{i=m}^n b_i = 0$ if $m > n$. We find that

$$\Pr[R_n(R_0) \geq 0] = \Pr[S \leq R_0], \quad (7.4)$$

with S defined by

$$S = \sum_{i=1}^n \alpha_i e^{-(Y_1 + \dots + Y_i)}. \quad (7.5)$$

The random variable S is the stochastically discounted value of all future liabilities.

Hence, $\bar{R}_0$ is given by

$$\bar{R}_0 = \inf \{R_0 \mid \Pr[S \leq R_0] \geq 1 - \varepsilon_n\}, \quad (7.6)$$

which means that the initial reserve is determined in this case as the $(1 - \varepsilon_n)$ -quantile of S :

$$\bar{R}_0 = Q_{1-\varepsilon_n}[S]. \quad (7.7)$$

Remark that this is similar to expression (6.94). In general, it will be impossible to determine the cumulative distribution function and the quantiles of S analytically, because in any realistic model for the return process $(Y_1, \dots, Y_n)$ the random variable S will be a sum of strongly dependent random variables.

A drawback of the determination of the provision R_0 as a quantile of S is that the only goal that has to be achieved is ‘reaching the finish’. A situation in which the first α_j are positive (hence payments) and moderate in absolute value, but the last ones are negative (hence incomes) and large in absolute value, this may lead to a situation where the available provision R_j in the first years becomes negative (or below a predefined level), which may be an undesirable situation.

We will present now an (approximate) way of determining the initial provision, which does not only take into account the goal of ‘reaching the finish’, but also the condition that, year-to-year, the provision R_j is larger than a given deterministic value V_j with a sufficiently large probability. These additional requirements are the ‘hurdles’ that have to be taken. In case of an insurer establishing his reserve, these ‘hurdles’ might be imposed by a supervisory authority or by internal policy. Hence, in the sequel of this chapter the provision at time 0 is determined by

$$\bar{R}_0 = \inf \{R_0 \mid R_0 \geq V_0; \Pr[R_j(R_0) \geq V_j] \geq 1 - \varepsilon_j; j = 1, \dots, n\}, \quad (7.8)$$

for given hurdles $V_0, V_1, \dots, V_n$ and given (small) probabilities $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$. Determining the initial provision as in (7.8) allows one to make the probability

of taking the hurdles time-dependent. In situations where year-to-year adjustments of the level of the reserve are possible, the probabilities of taking the hurdles in the first years could be chosen larger than these probabilities in the later years.

Let us now define the random variables $S_{[0,j]}$, ($j = 1, \dots, n$) by

$$\begin{aligned} S_{[0,j]} &= \sum_{i=1}^{j-1} \alpha_i e^{-(Y_1 + \dots + Y_i)} + (V_j + \alpha_j) e^{-(Y_1 + \dots + Y_j)} \\ &= \sum_{i=1}^j \alpha_{i(j)} e^{-(Y_1 + \dots + Y_i)}, \quad j = 1, \dots, n, \end{aligned} \quad (7.9)$$

with the $\alpha_{i(j)}$ given by

$$\alpha_{i(j)} = \begin{cases} \alpha_i, & i \neq j \\ V_j + \alpha_j, & i = j \end{cases}. \quad (7.10)$$

The random variable $S_{[0,j]}$ can be interpreted as the stochastically discounted value of the future obligations in the restricted time period $[0, j]$.

Theorem 7.2.1 (Hurdle-dependent optimal provision) *The optimal initial provision $\bar{R}_0$ defined in (7.8) is given by*

$$\bar{R}_0 = \max \{V_0, Q_{1-\varepsilon_1}[S_{[0,1]}], Q_{1-\varepsilon_2}[S_{[0,2]}], \dots, Q_{1-\varepsilon_n}[S_{[0,n]}]\}$$

Proof. From (7.3), it is straightforward to verify that

$$\Pr [R_j(R_0) \geq V_j] = \Pr [S_{[0,j]} \leq R_0], \quad j = 1, \dots, n.$$

Hence, $\bar{R}_0$ follows from

$$\begin{aligned} \bar{R}_0 &= \inf \{R_0 \mid R_0 \geq V_0; \Pr [S_{[0,j]} \leq R_0] \geq 1 - \varepsilon_j, \quad j = 1, \dots, n\} \\ &= \inf \{R_0 \mid R_0 \geq V_0; R_0 \geq Q_{1-\varepsilon_j}[S_{[0,j]}], \quad j = 1, \dots, n\}. \end{aligned}$$

This proves the stated result. ■

Under certain constraints, the initial provision defined in (7.8) coincides with the initial reserve defined in (7.2). This follows from the following corollary.

Corollary 7.2.1 (Equivalence between provisions) *If $\alpha_j \geq 0$, $V_j = 0$ and $\varepsilon_j \geq \varepsilon_n$, ($j = 1, \dots, n$), then $\bar{R}_0$ is given by*

$$\bar{R}_0 = \max \{V_0, Q_{1-\varepsilon_n}[S_{[0,n]}]\}.$$

Proof. Under the constraints of the corollary, we have that

$$Q_{1-\varepsilon_j}[S_{[0,j]}] \leq Q_{1-\varepsilon_n}[S_{[0,j]}] \leq Q_{1-\varepsilon_n}[S_{[0,n]}],$$

where the last inequality follows from the fact that $F_{S_{[0,j]}}(x) \geq F_{S_{[0,n]}}(x)$ holds for all values of x . This proves the stated result. ■

If in addition to the conditions in the corollary, we also have that $V_0 = 0$, then the initial provisions defined in (7.2) and (7.8) respectively are identical. In practice, one will often choose the ε_i in the first years lower than the later ones, because the conditions in the immediate future have to be met with the highest probability. In this case, the conditions of the corollary will not be fulfilled, and $\bar{R}_0$ may be different from $Q_{1-\varepsilon_n}[S_{[0,n]}]$.

7.3 Approximations for $\bar{R}_0$

In order to find an accurate approximation for $\bar{R}_0$, we will use the theory on comonotonic sums that we discussed in Chapter 2. We will illustrate our methodology for lognormal returns. Hence, in the remainder of this chapter we will always assume that the Y_i are i.i.d. $N(\mu - \frac{\sigma^2}{2}, \sigma^2)$ distributed. We will also assume that all $\alpha_{i(j)}$ are positive.

In order to determine $\bar{R}_0$, we will need to compute quantiles of $S_{[0,j]}$, ($j = 1, \dots, n$), which cannot be done exactly. Following Corollary 2.5.1, we will approximate the c.d.f.'s of the random variables $S_{[0,j]}$, $j = 1, \dots, n$ by the respective cumulative distribution functions of the random variables $S_{[0,j]}^c$ and $S_{[0,j]}^l$.

Rewriting $S_{[0,j]}$, $j = 1, \dots, n$ as

$$S_{[0,j]} = \sum_{i=1}^j \alpha_{i(j)} e^{Z_i}, \quad (7.11)$$

we can apply the results of Example 3.4.4 with

$$\begin{aligned} Z_i &= -Y_1 - Y_2 - \dots - Y_i, \\ E[Z_i] &= -i \left[\mu - \frac{1}{2} \sigma^2 \right], \\ \sigma_{Z_i}^2 &= i \sigma^2. \end{aligned} \quad (7.12)$$

The comonotonic upper bound $S_{[0,j]}^c$ is given by

$$S_{[0,j]}^c = \sum_{i=1}^j \alpha_{i(j)} e^{-i(\mu - \frac{1}{2} \sigma^2) + \sqrt{i} \sigma \Phi^{-1}(U)}, \quad (7.13)$$

while for $p \in (0, 1)$, the quantiles of $S_{[0,j]}^c$ are given by

$$Q_p[S_{[0,j]}^c] = \sum_{i=1}^j \alpha_{i(j)} e^{-i(\mu - \frac{1}{2} \sigma^2) + \sqrt{i} \sigma \Phi^{-1}(p)}, \quad (7.14)$$

In order to define a convex lower bound $S_{[0,j]}^l$ for $S_{[0,j]}$, we choose the conditioning random variable as follows:

$$\Lambda(j) = \sum_{i=1}^j \gamma_{i(j)} Z_j, \quad (7.15)$$

with coefficients $\gamma_{i(j)}$ that follow from (4.14):

$$\gamma_{i(j)} = \alpha_{i(j)} e^{-i(\mu - \sigma^2)}. \quad (7.16)$$

For this choice of the parameters $\gamma_{i(j)}$, the variance of the lower bound $S_{[0,j]}^l$ will be close to the variance of $S_{[0,j]}$, provided σ^2 is small enough. Hence, the lower bound will have a cumulative distribution function that is close to the cumulative distribution function of $S(\boldsymbol{\pi})$, provided $\sigma(\boldsymbol{\pi})$ is small enough, see also Chapter 4.

We find from Example 3.4.4 that the coefficients $r_{i(j)}$ are given by

$$r_{i(j)} = \frac{\sum_{l=1}^i \sum_{k=l}^j \alpha_{k(j)} e^{-k(\mu - \sigma^2)}}{\sqrt{i} \sqrt{\sum_{l=1}^j \left(\sum_{k=l}^j \alpha_{k(j)} e^{-k(\mu - \sigma^2)} \right)^2}}. \quad (7.17)$$

Since the correlation coefficients $r_{i(j)}$ are non-negative, $S_{[0,j]}^l$ will be a comonotonic sum of lognormal random variables. Still from Example 3.4.4, we find the following expressions for $Q_p[S_{[0,j]}^l]$, $p \in (0, 1)$:

$$Q_p[S_{[0,j]}^l] = \sum_{i=1}^j \alpha_{i(j)} e^{-i\mu + (1 - \frac{1}{2}r_{i(j)}^2)i\sigma^2 + r_{i(j)}\sqrt{i}\sigma\Phi^{-1}(p)}, \quad (7.18)$$

Following Theorem 7.2.1 and the comonotonic approximations we just discussed, we propose the following approximations for $\bar{R}_0$:

$$\bar{R}_0^l = \max \left\{ V_0; Q_{1-\varepsilon_j}[S_{[0,j]}^l], j = 1, \dots, n \right\}, \quad (7.19)$$

$$\bar{R}_0^c = \max \left\{ V_0; Q_{1-\varepsilon_j}[S_{[0,j]}^c], j = 1, \dots, n \right\}. \quad (7.20)$$

In Section 7.4 we will numerically illustrate that $\bar{R}_0^l$ and $\bar{R}_0^c$ will often be good approximations for $\bar{R}_0$. Especially, the approximation $\bar{R}_0^l$ for $\bar{R}_0$ performs very well.

Note that we don't necessarily have that $\bar{R}_0^l \leq \bar{R}_0 \leq \bar{R}_0^c$. From (3.83) it follows that the inequality $\bar{R}_0^l \leq \bar{R}_0^c$ holds, provided the ε_j are sufficiently small. Indeed, if all ε_j are chosen such that

$$\varepsilon_j \leq 1 - p_{(j)}^+, \quad j = 1, \dots, n, \quad (7.21)$$

with

$$p_{(j)}^+ = \Phi \left[\frac{1}{2} \max \{ (1 + r_{i(j)}) \sigma_{Z_i}; i = 1, \dots, j \} \right], \quad j = 1, \dots, n, \quad (7.22)$$

then the inequality $\overline{R}_0^l \leq \overline{R}_0^c$ will hold.

From (7.12) it follows that sufficient conditions for (7.21) to hold are

$$\epsilon_j \leq 1 - \Phi \left[\sigma \sqrt{j} \right], \quad j = 1, \dots, n, \quad (7.23)$$

7.4 Numerical illustrations

7.4.1 Example 1

Consider a person who invests at time 0 an amount of 10 units in a fund with a yearly expected return $\mu = \ln 1.10$ and a yearly volatility $\sigma = 0.10$. Each of the coming 10 years, he wants to withdraw 0.8 units of this initial amount. In addition, at the end of the 10 year period he wants to recover his initially invested capital of 10. The investor believes it is very likely that his investment will meet his requirements because the investment period is long enough to eliminate the downside risk.

In terms of the notations introduced in Section 7.2, the investor has obligations $\alpha_i = 0.8$, $i = 1, \dots, 10$, with a final hurdle $V_{10} = 10$, while his initial provision equals $R_0 = 10$. Let us assume that his belief can be expressed as $\Pr[R_n \geq 10 \mid R_0 = 10] \geq 0.995$.

In order to judge the investor's belief, we determine

$$\Pr[R_n \geq 10 \mid R_0 = 10] = \Pr[S_{[0,10]} \leq 10].$$

From the expressions (7.14) and (7.18) one finds that the (approximate) probability that the investor will take the last hurdle successfully if his initial investment is 10, equals 65.28% if we perform the calculations with $S_{[0,10]}^l$ and 64.53% if we use $S_{[0,10]}^c$. In order to judge the quality of these approximations, we also determine $\Pr[S_{[0,10]} \leq 10]$ by simulating 100,000 values for the return vector $(Y_1, Y_2, \dots, Y_n)$. This simulation leads to a value of 65.35% for this probability, with a standard error equal to 0.077%. This indicates that especially the approximation $S_{[0,10]}^l$ performs very well. Moreover, the investor's belief of reaching the finish with a probability of at least 99.5% is certainly not true.

Let us now determine $\overline{R}_0$ which is the smallest R_0 such that the investor will take the final hurdle $V_{10} = 10$ with a probability of at least 99.5% :

$$\begin{aligned} \overline{R}_0 &= \inf \{ R_0 \mid \Pr[R_{10}(R_0) \geq 10] \geq 0.995 \} \\ &= \inf \{ R_0 \mid \Pr[S_{[0,10]} \leq 10] \geq 0.995 \} \\ &= Q_{0.995}[S_{[0,10]}]. \end{aligned}$$

Using $S_{[0,10]}^l$, we find $\bar{R}_0^l = Q_{0.995}[S_{[0,10]}^l] = 16.98$, while for $S_{[0,10]}^c$, we find $\bar{R}_0^c = Q_{0.995}(S_{[0,10]}^c) = 17.872$. The value for $\bar{R}_0$ obtained by the simulation of 100,000 paths for the returns equals 16.985, which is again very close to the approximation $S_{[0,10]}^l$.

Finally, we determine the maximal amount $\bar{V}_{10}$ that will be guaranteed at time 10 with a probability of at least 99.5% if the initial invested amount equals 10 :

$$\begin{aligned}\bar{V}_{10} &= \sup \{V_{10} \mid \Pr [R_{10}(R_0) \geq V_{10}] \geq 0.995\} \\ &= \sup \{V_{10} \mid \Pr [S_{[0,10]} \leq 10] \geq 0.995\} \\ &= \sup \{V_{10} \mid Q_{0.995}[S_{[0,10]}] \geq 10\}.\end{aligned}$$

This means that we want to determine the hurdle $\bar{V}_{10}$ as the maximal amount such that the 99.5% percentile of $S_{[0,10]}$ is at least equal to 10. Using the approximation $S_{[0,10]}^l$ we obtain $\bar{V}_{10}^l = 2.182$, while for $S_{[0,10]}^c$ we find $\bar{V}_{10}^c = 1.399$. The simulation of 100,000 paths leads to a value of 2.179 for the hurdle at time 10 with a standard error equal to 0.0929.

7.4.2 Example 2

Consider a person who invests an amount of 10 in the fund as described in Example 7.4.1, with a yearly expected return $\mu = \ln 1.10$ and a yearly volatility $\sigma = 0.10$. At the end of each of the coming 40 years he wants to withdraw 0.8 units of the fund. Moreover he would like to be sure that the value of the invested amount will never become lower than 10. He wonders if his requirements can be met by the given investment strategy.

In terms of the notations introduced in Section 7.2, the investor has obligations $\alpha_i = 0.8$, $i = 1, \dots, 40$ with an initial provision $R_0 = 10$. All hurdles V_i are equal to the initial invested amount of 10. We assume that the investor wants to take each of these hurdles with a probability of at least 99, 5%.

As before, we can use the approximations $Q_{0.995}[S_{[0,j]}^l]$ and $Q_{0.995}[S_{[0,j]}^c]$ for the percentiles $Q_{0.995}[S_{[0,j]}]$, $j = 1, \dots, 40$. Note that each $Q_{0.995}[S_{[0,j]}]$ represents the initial required amount needed to take the hurdle at time j . The quantiles $Q_{0.995}[S_{[0,j]}^l]$, $Q_{0.995}[S_{[0,j]}^c]$ and the corresponding quantiles ' $Q_{0.995}[S_{[0,j]}]$ ' obtained by simulating 50,000 paths are presented in Table 7.1 and Table 7.2.

If the investor wants to take all hurdles with the predetermined degree of certainty, the approximations (7.19) and (7.20) for the initial required amount $\bar{R}_0$ as defined in (7.8) are given by $\bar{R}_0^l = 17.55$ and $\bar{R}_0^c = 20.10$ respectively. The value for $\bar{R}_0$ obtained by simulating 50,000 paths equals 17.57.

In case of $\bar{R}_0^l$ the binding hurdles are the ones corresponding to years 21 to 23. For $\bar{R}_0^c$ however, the binding hurdle is the one at time 40. The binding hurdle for the simulation is the one after 26 years.

A probability of 90% for taking the hurdles would have resulted in an initial reserve equal to $\bar{R}_0^l = 12.36$ corresponding to the binding hurdle at time 12.

j	$Q_{0.995}[S_{[0,j]}^t]$	$Q_{0.995}[S_{[0,j]}]$	$Q_{0.995}[S_{[0,j]}^c]$
1	12.77	12.78	12.77
2	13.86	13.93	13.92
3	14.63	14.58	14.78
4	15.22	15.16	15.46
5	15.68	15.63	16.02
6	16.06	16.07	16.51
7	16.36	16.38	16.92
8	16.61	16.66	17.28
9	16.81	16.81	17.60
10	16.98	16.91	17.87
11	17.12	16.96	18.12
12	17.23	17.13	18.33
13	17.31	17.11	18.53
14	17.38	17.23	18.70
15	17.44	17.25	18.85
16	17.48	17.16	18.99
17	17.51	17.20	19.11
18	17.54	17.25	19.22
19	17.55	17.44	19.32
20	17.56	17.31	19.41

Table 7.1: Approximated and simulated values for the quantiles in Example 2.

j	$Q_{0.995}[S_{[0,j]}^t]$	$Q_{0.995}[S_{[0,j]}]$	$Q_{0.995}[S_{[0,j]}^c]$
21	17.57	17.33	19.49
22	17.57	17.45	19.56
23	17.57	17.45	19.62
24	17.56	17.49	19.68
25	17.56	17.50	19.73
26	17.55	17.57	19.78
27	17.54	17.54	19.82
28	17.53	17.54	19.86
29	17.52	17.51	19.89
30	17.51	17.56	19.92
31	17.50	17.52	19.95
32	17.49	17.45	19.97
33	17.48	17.54	19.99
34	17.48	17.47	20.01
35	17.46	17.42	20.03
36	17.46	17.49	20.05
37	17.45	17.44	20.06
38	17.44	17.39	20.08
39	17.43	17.39	20.09
40	17.43	17.41	20.10

Table 7.2: Approximated and simulated values for the quantiles in Example 2 (cont'd).

For $\overline{R}_0^c$, we find a value of 12.84 for the initial reserve and a binding hurdle at time 23, while by simulation we obtain an initial reserve of 12.39 and a binding hurdle at time 12.

The savings-retirement problem

8.1 Introduction

⁹The previous Chapters 6 and 7 underpinned the importance for several practical applications of the r.v. S defined as

$$S = \sum_{i=0}^n \alpha_i e^{Z_i}. \quad (8.1)$$

Here the α_i are real numbers and $(Z_0, Z_1, \dots, Z_n)$ is a multivariate normal random vector. The numerical illustrations that we provided in these chapters revealed that especially if one uses the ‘maximal variance’ lower bound $E[S \mid \Lambda]$, as proposed in Chapter 4, one obtains in general a very accurate and easy to compute approximate solution for the problem at hand. Recall also from Chapter 4 that the ‘maximal variance’ lower bound approximation outperformed some other popular and easy to handle approximation techniques for the c.d.f. of S , i.e. the moment matching *lognormal approximation* and the moment matching *reciprocal Gamma approximation* for the c.d.f. of S .

We point out that the lognormal and the reciprocal Gamma approximation have been proposed in the context of series of positive cash-flows, whereas the lower bound approximation can also be applied in case of changing signs for the α_i .

In this case however, and this in contrast to the situation that all α_i have equal signs, it is not possible to find a conditioning r.v. Λ that leads to an accurate approximation for the c.d.f. of S such that $E[S \mid \Lambda] = \sum_{i=0}^n \alpha_i E[e^{Z_i} \mid \Lambda]$ is a sum of non-decreasing functions of Λ . This implies that distortion risk

⁹ “The only real valuable thing is intuition” - Albert Einstein.

measures related to $E[S \mid \Lambda]$ cannot be obtained by simply summing the corresponding risk measures of the individual terms in the sum, as is the case when all α_i are positive.

Goovaerts, De Schepper, Hua, Darkiewicz & Vyncke (2003) propose an approach that uses convex bounds for the positive and negative sum separately. They connect the two r.v.'s involved by a copula of a particular family.

In this chapter we follow another path, based on the comonotonic bounds, to determine accurate and easy computable approximations for a sum S as defined in (8.1) in case one first has positive cash-flows (savings) followed by negative ones (withdrawals).

An important situation where we encounter this particular cash-flow pattern and, hence, where our results can be applied, is the ‘savings-retirement problem’. Take as an example the 20/65/95 retirement plan: A person of age 20 intends to save money on a savings account for 45 consecutive years (until retirement). After his retirement, he wants to withdraw money from his savings account on a regular basis and for a period of 30 years. Assume that his yearly savings are constant and equal to α , while his yearly withdrawal (pension) is constant and equal to 1. A relevant question to answer is, “for each Euro of pension income, what is the minimal yearly savings effort α such that he, with a probability of at least $(1 - \varepsilon)$, will not outlive his money during the 30 year withdrawal period?”

The chapter is further organized as follows. In Section 8.2 we mathematically describe the savings-retirement problem. In Section 8.3 we prove that under suitable conditions, the quantiles of final wealth can be obtained in a straightforward way. Finally, in Section 8.4 we numerically illustrate our results. This chapter is based on Vanduffel, Dhaene & Goovaerts (2004).

8.2 Problem description

Consider a set of deterministic amounts $\alpha_0, \alpha_2, \dots, \alpha_{n+m}$ with $n, m > 0$. The first n amounts $\alpha_0, \alpha_2, \dots, \alpha_{n-1}$ are strictly positive and correspond to amounts that are put on a savings account at respective times $0, 1, \dots, n-1$. The last $m+1$ amounts $\alpha_n, \alpha_{n+1}, \dots, \alpha_{n+m}$ are strictly negative and correspond to withdrawals from the savings account at times $n, n+1, \dots, n+m$, respectively.

We will assume that the return on the savings account is generated by a geometric Brownian motion process. An amount of 1 put on the account at time 0 is assumed to grow to the random amount $e^{(Y_1+Y_2+\dots+Y_i)}$ at time i , where for any i , the r.v. Y_i is the random return over the year $[i-1, i]$. Notice that the r.v.'s Y_i are i.i.d. and normal distributed, with parameters $\mu - \frac{\sigma^2}{2}$ and σ^2 , let's say.

Let V_j denote the value of the savings account at time j . By convention, the value at time j is to be understood as the value just after saving or withdrawal. Starting from the initial value $V_0 = a_0$, the value V_j present at time j is given

by the following recursive relation:

$$V_j = V_{j-1} e^{Y_j} + \alpha_j, \quad j = 1, \dots, n+m. \quad (8.2)$$

Note that for this cash-flow pattern we have that once the value becomes strictly negative, it will stay negative over the whole period and no recovery is possible anymore. This ‘ruin’ scenario can only occur from time n (retirement) on. Solving the recursion (8.2), we find that the final value V_{n+m} is given by:

$$V_{n+m} = \sum_{i=0}^{n+m} \alpha_i e^{Y_{i+1} + Y_{i+2} + \dots + Y_{n+m}}, \quad (8.3)$$

where, by convention, $\sum_{j=r}^s \alpha_j = 0$ if $r > s$. we have that V_{n+m} is a sum of non-independent lognormal random variables. As it is impossible to determine the cumulative distribution function of V_{n+m} analytically, we propose to approximate the c.d.f. of V_{n+m} by the c.d.f. of the lower bound approximation:

$$V_{n+m}^l = E[V_{n+m} \mid \Lambda], \quad (8.4)$$

for some suitable choice for the conditioning r.v. Λ .

Rewriting V_{n+m} as

$$V_{n+m} = \sum_{i=0}^{n+m} \alpha_i e^{Z_i}, \quad (8.5)$$

we can apply the results of Example 3.4.4 with

$$\begin{aligned} Z_i &= Y_{i+1} + Y_{i+2} + \dots + Y_{n+m}, \\ E[Z_i] &= (n+m-i) \left(\mu - \frac{\sigma^2}{2} \right), \\ \sigma_{Z_i}^2 &= (n+m-i)\sigma^2. \end{aligned} \quad (8.6)$$

where the conditioning random variable Λ is defined as a linear combination of the cumulative returns Z_i :

$$\Lambda = \sum_{j=0}^{n+m} \gamma_j Z_j. \quad (8.7)$$

The r.v. Λ can also be defined as a linear combination of the yearly returns:

$$\Lambda = \sum_{j=1}^{n+m} \beta_j Y_j \quad (8.8)$$

and from (3.78) we find that the relation between the β_j and the γ_j is given by

$$\beta_j = \sum_{k=0}^{j-1} \gamma_k. \quad (8.9)$$

We find that V_{n+m}^l is distributed as

$$V_{n+m}^l \stackrel{d}{=} \sum_{i=0}^{n+m} \alpha_i e^{E[Z_i] + \frac{1}{2}(1-r_i^2)\sigma_{Z_i}^2 + r_i \sigma_{Z_i} \Phi^{-1}(U)}, \quad (8.10)$$

with coefficients r_i given by

$$r_i = \frac{\text{cov}(Z_i, \Lambda)}{\sigma_{Z_i} \sigma_{\Lambda}} = \frac{\sum_{j=i+1}^{n+m} \beta_j}{\sqrt{n+m-i} \sqrt{\sum_{j=1}^{n+m} \beta_j^2}}, \quad i = 0, \dots, n+m-1 \quad (8.11)$$

and

$$r_{n+m} = 0. \quad (8.12)$$

The expectation of V_{n+m}^l is given by

$$E(V_{n+m}^l) \stackrel{d}{=} \sum_{i=0}^{n+m} \alpha_i E(e^{Z_i}). \quad (8.13)$$

Notice that the last term in the sum (8.10) reduces to the constant number a_{n+m} .

If all the terms in the sum (8.10) are non-decreasing functions of U (or all are non-increasing functions of U), then we have that V_{n+m}^l is a comonotonic sum. Hence, in this case we have that also the continuous function f defined by

$$f(p) = \sum_{i=0}^{n+m} \alpha_i e^{E[Z_i] + \frac{1}{2}(1-r_i^2)\sigma_{Z_i}^2 + r_i \sigma_{Z_i} \Phi^{-1}(p)}, \quad p \in (0, 1), \quad (8.14)$$

is non-decreasing. From Lemma 1.3.1 we find that the p -quantile of V_{n+m}^l is given by

$$Q_p[V_{n+m}^l] = f(p). \quad (8.15)$$

Moreover, in case V_{n+m}^l is a comonotonic sum, all distortion risk measures related to V_{n+m} are given by the sum of the risk measures related to the marginal terms in (8.10), see Chapter 3.

A conditioning r.v. Λ that makes f non-decreasing (and continuous) can always be found. Indeed, if we take all $\beta_j > 0$ for $j = 1, 2, \dots, n$, whereas $\beta_j = 0$ for $j = n+1, n+2, \dots, n+m$, then we will have that f is non-decreasing, so that (8.15) holds in this case. Of course, we cannot expect that such a choice of the parameters β_j will in general lead to an accurate approximation for the c.d.f. of V_{n+m} .

In general, and for appropriate choices of Λ , the r.v. V_{n+m}^l will not be a comonotonic sum of lognormal r.v.'s. Hence, the function $f(p)$ will not be non-decreasing on the whole interval $(0, 1)$, so that the quantiles of V_{n+m}^l cannot be determined easily because Lemma 1.3.1 cannot be applied in this case.

The cash-flow pattern that we consider (first saving, then withdrawing) implies that once the value V_j becomes negative, no recovery is possible in the

sense that all future values V_k will be negative too. Let us now assume that once the savings account reaches level 0, no further withdrawals from the savings account are allowed. If we define W_j as the r.v. that describes the wealth at time j , we find that $W_j = \max[V_j, 0]$. More in particular, we find that the final wealth W_{n+m} is given by

$$W_{n+m} = \max[V_{n+m}, 0]. \quad (8.16)$$

and we propose to approximate (the c.d.f. of) W_{n+m} by (the c.d.f. of) $W_{n+m}^l = \max[V_{n+m}^l, 0]$. From (8.10), we find that

$$W_{n+m}^l \stackrel{d}{=} \max[f(U), 0], \quad (8.17)$$

where the function f is defined by (8.14).

In the next section we will prove that under rather weak conditions the function $\max[f(p), 0]$ is non-decreasing and continuous. From Lemma 1.3.1 we find a straightforward way to compute the quantiles of W_{n+m}^l in this case.

8.3 Approximations for the quantiles of W_{n+m}

The following lemma gives sufficient conditions for the function $\max[f(p), 0]$ to be non-decreasing.

Lemma 8.3.1 (Conditions for $\max(f(p), 0)$ to be non-decreasing) *If all $\beta_j > 0$, for $j = 1, 2, \dots, n+m$, then the function $\max[f(p), 0]$ is non-decreasing on the interval $(0, 1)$.*

Proof. It suffices to prove that if $f(p) \geq 0$ for a value of p in the unit interval $(0, 1)$, then $f'(p) \geq 0$ must hold.

The standard deviation of Z_j is given by

$$\sigma_{Z_j} = \sigma \sqrt{n+m-i}.$$

Hence, from (8.11), one finds that all correlations $r_i > 0$ for $i = 0, 1, \dots, n+m-1$ and that the sequence $\{r_i \sigma_{Z_i}\}_{0 \leq i \leq n+m}$ is non-increasing and positive.

By application of the chain rule we have for $p \in (0, 1)$ that

$$f'(p) = \frac{1}{\Phi'[\Phi^{-1}(p)]} \sum_{i=0}^{n+m-1} \alpha_i e^{E[Z_i] + \frac{1}{2}(1-r_i^2)\sigma_{Z_i}^2 + r_i \sigma_{Z_i} \Phi^{-1}(p)} r_i \sigma_{Z_i}.$$

Now assume that $f(p) \geq 0$ for some value of p in the unit interval $(0, 1)$. Since $\frac{1}{\Phi'[\Phi^{-1}(p)]} > 0$ and $\alpha_{n+m} < 0$, we find that

$$\begin{aligned} f'(p) &> \frac{r_{n-1} \sigma_{Z_{n-1}}}{\Phi'[\Phi^{-1}(p)]} \sum_{i=0}^{n+m-1} \alpha_i e^{E[Z_i] + \frac{1}{2}(1-r_i^2)\sigma_{Z_i}^2 + r_i \sigma_{Z_i} \Phi^{-1}(p)} \\ &= \frac{r_{n-1} \sigma_{Z_{n-1}}}{\Phi'[\Phi^{-1}(p)]} (f(p) - \alpha_{n+m}) > 0, \end{aligned}$$

which ends the proof. ■

One can easily prove that when all $\beta_j > 0$, one has that

$$\lim_{p \rightarrow 0} f(p) = \alpha_{n+m} < 0 \quad (8.18)$$

and

$$\lim_{p \rightarrow 1} f(p) = +\infty. \quad (8.19)$$

From the proof of lemma 8.3.11 we also have that once f becomes non-negative, it is strictly increasing

$$f(p) \geq 0 \implies f'(p) > 0 \quad (8.20)$$

In the following theorem, we show that under the conditions of the lemma, the quantiles of W_{n+m}^l can easily be determined analytically.

Theorem 8.3.1 (Quantiles of W_{n+m}^l) *If all $\beta_j > 0$, $j = 1, 2, \dots, n+m$ then*

$$Q_p[W_{n+m}^l] = \max[f(p), 0].$$

Proof. As all β_j are assumed to be positive, we find from Lemma 1 that $\max[f(p), 0]$ is non-decreasing (and continuous) on the interval $(0, 1)$. The proof then follows from (8.17) and (1.3.1). ■

Under the conditions of the theorem, we have that the c.d.f. of W_{n+m}^l can be determined from

$$f(F_{W_{n+m}^l}(x)) = x, \quad x \geq 0, \quad (8.21)$$

Indeed, we have that for any $x \geq 0$,

$$\begin{aligned} F_{W_{n+m}^l}(x) &= \sup \left\{ p \in (0, 1) \mid p \leq F_{W_{n+m}^l}(x) \right\} \\ &= \sup \left\{ p \in (0, 1) \mid Q_p[W_{n+m}^l] \leq x \right\} \\ &= \sup \left\{ p \in (0, 1) \mid \max(f(p), 0) \leq x \right\} \\ &= \sup \left\{ p \in (0, 1) \mid f(p) \leq x \right\}. \end{aligned} \quad (8.22)$$

The relation (8.21) follows then from (8.18), (8.19), (8.20) and from the fact that f is continuous on $(0, 1)$.

In order to determine the most convenient choice for the coefficients β_j , we follow the procedure as explained in Chapter 4. Recall that the procedure we explained there also holds for general α_i . Therefore, we propose to use the following coefficients β_j , $j = 1, \dots, n+m$:

$$\beta_j = \sum_{i=0}^{j-1} \gamma_i = \sum_{i=0}^{j-1} \alpha_i e^{E[Z_i] + \frac{1}{2}\sigma_{Z_i}^2} = \sum_{i=0}^{j-1} \alpha_i E[e^{Z_i}]. \quad (8.23)$$

A sufficient condition for all coefficients β_j defined in (8.23) to be positive, and hence for Theorem 8.3.1 to hold, is that β_{n+m} is strictly positive. This main result is summarized in the following theorem.

Theorem 8.3.2 (Relation between β_{n+m} and quantiles of W_{n+m}^l) *If the β_j , $j = 1, 2, \dots, n+m$, are defined by (8.23), and if*

$$\beta_{n+m} = \sum_{i=0}^{n+m-1} \alpha_i E[e^{Z_i}] > 0, \quad (8.24)$$

then the quantiles of W_{n+m}^l are given by

$$Q_p[W_{n+m}^l] = \max[f(p), 0], \quad p \in (0, 1), \quad (8.25)$$

whereas the c.d.f. of W_{n+m}^l follows from

$$f(F_{W_{n+m}^l}(x)) = x, \quad x \geq 0, \quad (8.26)$$

with $f(p)$ defined by (8.14).

As we showed throughout this work, the choice (8.23) for the parameters β_j will in general lead to a very accurate approximation for the distribution of final wealth W_{n+m} . Moreover, it is easily verified that the condition (8.24) is fulfilled in case the saving-retirement plan is such that $E[V_{n+m}] \geq 0$. It is obvious that any reasonable saving-retirement plan should fulfill the condition that the average final wealth will be non-negative.

We will consider now the case that all saving amounts are equal to some $\alpha > 0$ whereas the unit amount will be withdrawn during the retirement period. The notation $\varepsilon^l(\alpha)$ is used to denote the approximation for the probability $\varepsilon(\alpha)$ of going bankrupt, using W_{n+m}^l . We will assume that the savings amount α is chosen such that $E[V_{n+m}] \geq 0$. From Theorem 8.3.2 and (8.14) we find that the probability $\varepsilon^l(\alpha)$ is determined as the unique root of the equation

$$\alpha = \frac{\sum_{i=n}^{n+m} e^{E[Z_i] + \frac{1}{2}(1-r_i^2)\sigma_{Z_i}^2 + r_i \sigma_{Z_i} \Phi^{-1}(\varepsilon^l(\alpha))}}{\sum_{i=0}^{n-1} e^{E[Z_i] + \frac{1}{2}(1-r_i^2)\sigma_{Z_i}^2 + r_i \sigma_{Z_i} \Phi^{-1}(\varepsilon^l(\alpha))}}. \quad (8.27)$$

Now, if α is such that $E[V_{n+m}] = 0$, then it follows from (8.13), (8.14) and the fact that the sequence $\{r_i \sigma_{Z_i}\}_{0 \leq i \leq n+m}$ will be non-increasing and positive, that

$$f(0.5) = \sum_{i=0}^{n-1} \alpha E[e^{Z_i}] e^{-\frac{1}{2}r_i^2 \sigma_{Z_i}^2} - \sum_{i=n}^{n+m-1} E[e^{Z_i}] e^{-\frac{1}{2}r_i^2 \sigma_{Z_i}^2} < 0, \quad (8.28)$$

so that $\varepsilon^l(\alpha) > 0.5$ in this case. On the other hand, $\varepsilon^l(\alpha)$ will converge to zero in case α reaches infinity. Hence, we have that for all α , chosen such that $E[V_{n+m}] \geq 0$, the approximated ruin probabilities $\varepsilon^l(\alpha)$ will be between 0 and t , for some $t > 0.5$. Hence, equation (8.27) can also be used to determine, for a given ruin probability $0 < \varepsilon < 0.5$, the approximate optimal savings amount denoted by $\alpha^l(\varepsilon)$.

p	$Q_p[W_{19}^l]$	$'Q_p[W_{19}]'$	$s.e.$
0.95	45.11	45.17	0.28
0.90	34.81	34.80	0.16
0.75	21.88	21.89	0.10
0.50	12.11	12.13	0.01
0.25	5.64	5.68	0.03
0.10	1.76	1.75	0.04
0.05	0.04	0	0.00
0.01	0	0	0.00

Table 8.1: Approximated and simulated values for the quantiles $Q_p[W_{19}]$.

α	$\varepsilon^l(\alpha)$
0.0320	71.29%
0.0500	55.38%
0.1000	23.22%
0.1500	9.89%
0.2500	2.24%
0.5000	0.14%

Table 8.2: Lower bound based approximations for the ruin probabilities.

8.4 Numerical illustration

In the numerical illustrations, we assume that the yearly returns Y_i have expectation and variance given by $\mu - \frac{\sigma^2}{2}$ and σ^2 , with $\mu = 0.075$ and $\sigma = 0.15$. We will also choose the parameters β_j as proposed in (8.23)

First, we consider a savings-retirement plan that consists in 10 constant payments $\alpha_0 = \alpha_2 = \dots = \alpha_9 = 1$, followed by 10 constant withdrawals $\alpha_{10} = \alpha_{11} = \dots = \alpha_{19} = -1$. In this case we find that β_{19} is given by $\beta_{19} = 17.02 > 0$, so that Theorem 8.3.2 can be used to approximate the quantiles of W_{19}^l in an elegant way. In Table 8.1 we compare the approximated quantiles

$$Q_p[W_{19}^l] = \max(f(p), 0)$$

with the simulated quantiles $'Q_p[W_{19}]'$. This table illustrates the accuracy of the lower bound based approximation W_{19}^l . From this table we find, for instance, that there is a 90% probability that the wealth at time 19 will exceed 1.75 (simulated value). The approximated value for this final wealth is 1.76. From (8.26) it follows that the (approximated) probability of outliving its money, given by $\Pr[W_{19}^l = 0]$ equals 4.83%.

Next, we consider the 20/65/95 savings-retirement problem we mentioned in the introduction. This consists in 45 constant payments $\alpha_0 = \alpha_2 = \dots = \alpha_{44} = \alpha$, the first one at the age of 20 and the last one at the age of 64. After retirement, yearly withdrawals equal to -1, will be made until the age of

95 has been reached, hence $\alpha_{45} = \alpha_{46} = \dots = \alpha_{75} = -1$. Table 8.2 contains the approximated probability $\varepsilon^l(\alpha)$ of outliving the money for different savings amounts α . These probabilities follow from (8.27). Notice that for the choice $\alpha = 0.0320$ we have that $E(V_{75}) = 0$, so that this choice for α is appropriate for an individual whose only concern is that on average he will not run bankrupt.

Again by applying (8.27), we find that the optimal savings amount $\alpha^l(\varepsilon)$ in case $\varepsilon = 5.00\%$ equals 0.1935, so that this yearly savings effort guarantees that the individual will not go bankrupt with a probability of 95%.

Bibliography

- [1] Ahcan, A., Darkiewicz, G., Dhaene, J., Goovaerts, M.J. & Hoedemakers, T. (2004). "Optimal portfolio selection: application in Insurance business and investments", Proceedings of the 8th International Congress on Insurance: Mathematics and Economics, Rome, June, 14-16.
- [2] Albrecher, H., Dhaene, J., Goovaerts, M.J. & Schoutens, W. (2003). "Static hedging of Asian options under Lévy models: the comonotonic approach", Research Report OR 0365, Department of Applied Economics, K.U.Leuven.
- [3] Artzner, P. (1999). "Application of coherent capital requirements", *North American Actuarial Journal*, 3(2), 11-25.
- [4] Artzner, P., Delbaen, F., Eber, J.M. & Heath, D. (1999). "Coherent measures of risk", *Mathematical Finance*, 9, 203-228.
- [5] Atkinson, D.B. & Dallas, J.W. (2000). "Life Insurance: Products and Finance", Society of Actuaries.
- [6] Björk, T. (1998). "Arbitrage Theory in Continuous Time", Oxford University Press, pp. 311.
- [7] Black, F. & Scholes, M. (1973). "The pricing of options and corporate liabilities", *Journal of Political Economy*, 81 (May-June), 637-659.
- [8] Bodie, Z. (1995). "On the risk of stocks in the long run", *Financial Analysts Journal*, 51(3), 18-22.
- [9] Boyle, P.P. (1976). "Rates of Return as Random Variables", *The Journal of Risk and Insurance*, 43(4), 693-713.
- [10] Brockett, P., Cox, S.H. & Witt, R.C. (1984). "Self-Insurance and the Probability of Financial Regret", *Journal of Risk and Insurance*, 4(4), 720-729.
- [11] Cesari, R. & Cremonini, D. (2003). "Benchmarking, portfolio insurance and technical analysis: a Monte Carlo comparison of dynamic strategies of asset allocation", *Journal of Economic Dynamics & Control*, 27, 987-1011.
- [12] Cox, J.C. & Leland, H.E. (1999). "On dynamic investment strategies", *Journal of Economic Dynamics & Control*, 24, 1859-1880.

- [13] Denneberg, D. (1994). "Non-additive measure and integral", Kluwer Academic Publishers, Boston.
- [14] Denuit, M. & Dhaene, J. (2003). "Simple Characterizations of Comonotonicity and Countermonotonicity by extremal correlation", *Belgian Actuarial Bulletin*, 3, 22-27.
- [15] Denuit, M., Dhaene, J. & Van Wouwe, M. (1999). "The economics of insurance: a review and some recent developments", *Mitteilungen der Schweiz. Aktuarvereinigung*, 1999(2), 137-175.
- [16] Dhaene, J., Denuit, M., Goovaerts, M.J., Kaas, R. & Vyncke, D., 2002(a). "The concept of comonotonicity in actuarial science and finance: Theory", *Insurance: Mathematics & Economics*, 31(1), 3-33.
- [17] Dhaene, J., Denuit, M., Goovaerts, M.J., Kaas, R. & Vyncke, D., 2002(b). "The concept of comonotonicity in actuarial science and finance: Applications", *Insurance: Mathematics & Economics*, 31(2), 133-161.
- [18] Dhaene, J. & Goovaerts, M.J. (1996). "Dependency of risks and stop-loss order", *ASTIN Bulletin* 26, 201-212.
- [19] Dhaene, J. & Goovaerts, M.J. (1997). "On the dependency of risks in the individual life model", *Insurance: Mathematics & Economics*, 19, 243-253.
- [20] Dhaene, J., Goovaerts, M.J. & Kaas, R. (2003). "Economic capital allocation derived from risk measures", *North American Actuarial Journal*, 7(2), 44-59.
- [21] Dhaene, J., Goovaerts, M.J., Vanduffel, S. & Vyncke, D. (2002). "How to determine the capital requirements for a portfolio of annuity liabilities", *Tijdschrift voor Economie en Management*, 46(4), 533-544.
- [22] Dhaene, J., Laeven, R.J.A., Vanduffel, S., Darkiewicz, G. & Goovaerts, M.J. (2004). "Can a coherent risk measure be too subadditive?", forthcoming.
- [23] Dhaene, J., Vanduffel, S., Goovaerts, M.J., Kaas, R. & Vyncke, D. (2004). "Comonotonic approximations for optimal portfolio selection problems", *Journal of Risk and Insurance*, to appear
- [24] Dhaene, J., Vanduffel, S., Goovaerts, M.J., Koch, R., Olieslagers, R. & Romijn, O. (2004). "The use of a stochastic loss given default in a credit default economic capital framework", forthcoming.
- [25] Dhaene, J., Vanduffel, S., Goovaerts, M.J., Olieslagers, R. & Koch, R. (2003). "On the computation of the capital multiplier in the Fortis Credit Economic Capital model", *Belgian Actuarial Bulletin*, 3, 50-57.

- [26] Dhaene, J., Vanduffel, S., Tang, Q., Goovaerts, M.J., Kaas, R. & Vyncke, D. (2004). "Solvency capital, risk measures and comonotonicity: a review", Research Report OR 0416, Department of Applied Economics, K.U.Leuven.
- [27] Dhaene, J., Wang, S., Young, V. & Goovaerts, M.J. (2000). "Comonotonicity and maximal stop-loss premiums", *Bulletin of the Swiss Association of Actuaries*, 2000(2), 99-113.
- [28] Dufresne, D. (1990). "The distribution of a perpetuity with applications to risk theory and pension funding", *Scandinavian Actuarial Journal*, 9, 39-79.
- [29] Dufresne, D. (2002). "Asian and Basket Asymptotics", Research Paper No. 100, Centre for Actuarial Studies, University of Melbourne.
- [30] Dybvig, P.H. (1988). "Inefficient Dynamic Portfolio Strategies or How to Throw Away a Million Dollars in the Stock Market", *The Review of Financial studies*, 1(1), 67-88.
- [31] Emmer, S., Klüppelberg, C. & Korn, R. (2001). "Optimal portfolios with bounded capital at risk", *Mathematical Finance*, 11 (4), 365-384.
- [32] Fréchet, M. (1951). "Sur les tableaux de corrélation dont les marges sont donnés", *Ann. Univ. Lyon Sect.A*, Series 3, 14, 53-77.
- [33] Gerber, H.U. & Shiu, E.S.W. (1996). "Actuarial bridges to dynamic hedging and option pricing", *Insurance: Mathematics & Economics*, 18, 183-218.
- [34] Goovaerts, M.J., De Vijlder, F. & Haezendonck, J. (1984). "Insurance Premiums", North-Holland, Amsterdam.
- [35] Goovaerts, M.J., De Schepper, A., Hua, Y., Darkiewicz, G. & Vyncke, D. (2003). "An investigation on the use of copulas when calculation general cash flow distributions", Proceedings of the 8th International Congress on Insurance: Mathematics and Economics, Rome, June, 14-16.
- [36] Goovaerts, M.J., Dhaene, J. & De Schepper, A. (2000). "Stochastic Upper Bounds for Present Value Functions", *Journal of Risk and Insurance Theory*, 67.1, 1-14.
- [37] Goovaerts, M.J. & Redant, R. (1999). "On the distribution of IBNR reserves", *Insurance: Mathematics & Economics*, 25, 1-9.
- [38] Harrison, J. & Kreps, D. (1979). "Martingales and arbitrage in multiperiod securities markets", *Journal of Economic Theory*, 20, 381-408.
- [39] Harrison, J. & Pliska, R. (1981). "Martingales and stochastic integrals in the theory of continuous trading", *Stochastic Processes and their Applications*, 11, 215-260.

- [40] Hoedemakers, T., Beirlant, J., Goovaerts, M.J. & Dhaene, J. (2003). "Confidence bounds for discounted loss reserves", *Insurance: Mathematics & Economics*, 33(2), 297-316.
- [41] Hoeffding, W. (1940). "Masstabinvariante Korrelationstheorie", *Schriften des mathematischen Instituts und des Instituts für angewandte Mathematik der Universität Berlin*, 5, 179-233.
- [42] Huang, H., Milevsky, M. & Wang, J. (2004). "Ruined Moments in Your Life: How Good Are the Approximations?", *Insurance: Mathematics & Economics*, 34(3), 421-447.
- [43] Jagannathan, R. & Kocherlakota, N. (1996). "Why should older people invest less in stocks than younger people?", *Federal Reserve Bank of Minneapolis Quarterly Review*, 20(3), 11-23.
- [44] Kaas, R., Dhaene, J. & Goovaerts, M. (2000). "Upper and lower bounds for sums of random variables", *Insurance: Mathematics and Economics*, 27, 151-168.
- [45] Kaas, R., Goovaerts, M.J., Dhaene, J. & Denuit, M. (2001). "Modern Actuarial Risk Theory", Kluwer Academic Publishers, pp. 328.
- [46] Kaas, R., van Heerwaarden, A.E. & Goovaerts, M.J. (1994). "Ordering of actuarial risks", Institute for Actuarial Science and Econometrics, University of Amsterdam, Amsterdam.
- [47] Li, Z.F., Ng, K.W., Tan, K.S. & Yang, H. (2003). "A closed-form solution to a dynamic portfolio optimization problem", forthcoming.
- [48] Markowitz, H. (1952). "Portfolio selection", *Journal of Finance*, 7, 77-91.
- [49] Markowitz, H. (1959). "Portfolio Selection - Efficient Diversification of Investments", Wiley, New York.
- [50] Marshall, J.F. (1994). "The role of the investment horizon in optimal portfolio sequencing", *Financial Review*, 29(4), 557-576.
- [51] Merton, R. (1971). "Optimum consumption and portfolio rules in a continuous-time model", *Journal of Economic Theory*, 3, 373-413.
- [52] Merton, R. (1990). "Continuous Time Finance", Cambridge, Blackwell.
- [53] Milevsky, M.A. (1997). "The present value of a stochastic perpetuity and the Gamma distribution", *Insurance: Mathematics and Economics*, 20(3), 243-250.
- [54] Milevsky, M.A. (2003). "Space-time diversification: Which dimension is better?", *Journal of Risk*, to appear.

- [55] Milevsky, M.A. , Ho, K. & Robinson, C. (1997). "Asset allocation via the conditional first exit time or how to avoid outliving your money", *Review of Quantitative Finance and Accounting*, 9(1), 53-70.
- [56] Milevsky, M.A. & Posner, S.E. (1998). "Asian Options, the Sum of Lognormals, and the Reciprocal Gamma Distribution", *Journal of financial and quantitative analysis*, 33(3), 409-422.
- [57] Milevsky, M.A. & Robinson C. (2000). "Self-Annuity and Ruin in Retirement", *North American Actuarial Journal*, 4(4), 112-124.
- [58] Müller, A. (1997). "Stop-loss order for portfolios of dependent risks", *Insurance: Mathematics & Economics*, 21(3), 219-223.
- [59] Olivieri, A. & Pitacco, E. (2003). "Solvency requirements for pension annuities", *Journal of Pension Economics and Finance*, 2(2), 127-157.
- [60] Overbeck, L. (2000). "Allocation of economic capital in loan portfolios", *Measuring risk in complex systems*, Franke, J., Haerdle, W., Stahl, G. (eds), Springer.
- [61] Panjer, H.H. (2002). "Measurement of risk, solvency requirements and allocation of capital within financial conglomerates", Institute of Insurance and Pension Research, University of Waterloo, Research Report 01-15.
- [62] Quiggin, J. (1993). "Generalised Expected Utility Theory: The Rank-Dependent Model", Kluwer Academic Publishers, Boston.
- [63] Rogers, L.C.G. & Shi, Z. (1995). "The value of an Asian option", *Journal of Applied Probability*, 32, 1077-1088.
- [64] Roy, A. (1952). "Safety first and the holding of assets", *Econometrica*, 20, 431-449.
- [65] Rubinstein, M. (1991). "Continuously rebalanced investment strategies", *The Journal of Portfolio Management*, Fall, 78-81.
- [66] Samuelson, P.A. (1989). "The judgement of economic science on rational portfolio management: Timing and long horizon effects", *The Journal of Portfolio Management*, 16(1), 4-12.
- [67] Shaked, M. & Shanthikumar, J.G. (1994). "Stochastic orders and their applications", Academic Press, pp. 545.
- [68] Steinbach (2001). "Markowitz revisited: mean-variance models in financial portfolio analysis", *SIAM Review*, 43(1), 31-85.
- [69] Tsanakas, A. & Desli E. (2003). "Risk measures and theories of choice", forthcoming.

- [70] Valdez, E. & Dhaene, J. (2004). "Bounds for sums of dependent log-elliptical risks", forthcoming.
- [71] Vanduffel, S., Dhaene, J., Goovaerts, M.J. & Kaas, R. (2003). "The hurdle-race problem", *Insurance: Mathematics & Economics*, 33(2), 405-413.
- [72] Vanduffel, S., Hoedemakers, T. & Dhaene, J. (2004). "Comparing approximations for sums of non-independent lognormal random variables", Research Report OR 0418, Department of Applied Economics, K.U.Leuven.
- [73] Vanduffel, S., Dhaene, J. & Goovaerts, M.J. (2004). "The savings-retirement problem", Proceedings of the 8th International Congress on Insurance: Mathematics and Economics, Rome, June, 14-16.
- [74] von Neumann, J. & Morgenstern, O. (1947). "Theory of games and economic behavior", Princeton University Press, Princeton.
- [75] Vyncke, D., Goovaerts, M.J. & Dhaene, J. (2001). "Convex upper and lower bounds for present value functions", *Applied Stochastic Models in Business and Industry*, 17(2), 149-164.
- [76] Wang, S. (1996). "Premium calculation by transforming the layer premium density", *ASTIN Bulletin* 26, 71-92.
- [77] Wang, S. (2000). "A class of distortion operators for pricing financial and insurance risks", *Journal of Risk and Insurance*, 67(1), 15-36.
- [78] Wang, S. (2001). "A risk measure that goes beyond coherence", Research Report 01-18, Institute of Insurance and Pension Research, University of Waterloo.
- [79] Wang, S. & Dhaene, J. (1998). "Comonotonicity, correlation order and stop-loss premiums", *Insurance: Mathematics & Economics*, 22, 235-243.
- [80] Wang, S. & Young, V. (1998). "Ordering risks: expected utility versus Yaari's dual theory of choice under risk", *Insurance: Mathematics & Economics*, 22, 145-162.
- [81] Wirth, J.L. & Hardy, M.R. (2000). "Ordering of risk measures for capital adequacy", Research Report 00-03, Institute of Insurance and Pension Research, University of Waterloo.
- [82] Yaari, M.E. (1987). "The dual theory of choice under risk", *Econometrica*, 55, 95-115.
- [83] Young, V.R. (2004). "Optimal investment strategy to minimize the probability of lifetime ruin", forthcoming.

Samenvatting

Alle informatie over een individuele kansvariabele X zit samengebond in de verdelingsfunctie F_X . Het lijkt daarom aangewezen dat het nemen van beslissingen met betrekking tot X steeds gebeurt op basis van F_X en het is de kerntaak van een actuaire om deze functie te bepalen. In een verzekeringscontext kan een dergelijke kansvariabele X bijvoorbeeld het risico van één welbepaalde polis vertegenwoordigen, maar ze kan ook het risico beschrijven van een gehele tak van activiteiten of zelfs de gehele verzekeringsportefeuille. Het risico X duikt ook op bij persoonlijke financiële planning: bij een reeks van toekomstige betalingen of bij het gevaar dat het pensioenplan onvoldoende zal zijn om de huidige levensstandaard ook later te behouden.

De actuaire zal wiskundige en statistische technieken gebruiken om de verdelingsfunctie F_X te bepalen, maar vaak zal hij geconfronteerd worden met kansvariabelen $S = \sum_{i=1}^n X_i$, waarbij hij de univariate verdelingsfunctie van elk van de X_i kent, maar niet de afhankelijkheidsstructuur van de kansvector $\underline{X} = (X_1, X_2, \dots, X_n)$. Het is ook mogelijk dat deze laatste te ingewikkeld is om ermee overweg te kunnen.

Vanuit wiskundig oogpunt is het handig te veronderstellen dat de kansvariabelen X_i onderling onafhankelijk zijn. In dit geval bestaat er immers een batterij van technieken zoals Panjer's recursie, convolutie,... die de verdelingsfunctie van S accuraat kunnen bepalen (of minstens benaderen). Men kan dan ook de Centrale Limiet Stelling aanwenden. Deze leert ons dat S 'ongeveer' normaal verdeeld zal zijn, indien het aantal risico's X_i voldoende hoog is. Men kan stellen dat de verzekeringsactiviteit haar bestaansrecht precies ontleent aan de hypothese van onafhankelijkheid tussen de risico's. Deze hypothese is inderdaad soms (ongeveer) in overeenstemming met de werkelijkheid. Meestal is het echter zo dat er een correlatie bestaat tussen de risico's. Men kan dan de cumulatieve verdelingsfunctie van S niet bepalen, wat het hele proces voor het nemen van beslissingen op losse schroeven zet. Zo een som S van afhankelijke risico's ontmoet men bijvoorbeeld, wanneer men de totale schade van een schadeverzekeringsportefeuille bekijkt. In dit geval hangen de verzekerde risico's in meer of mindere mate af van gemeenschappelijke factoren zoals klimaat, geografie of economie.

In deze thesis concentreren we ons vooral op het geval dat S een som van afhanke-

lijke lognormalen voorstelt:

$$S = \sum_{i=1}^n \alpha_i e^{Z_i}. \quad (1)$$

Hierbij zijn de α_i reële getallen en is $(Z_1, Z_2, \dots, Z_n)$ een multivariaat normaal verdeelde kansvector. De stochastische som (1) speelt een prominente rol in de actuariële en financiële theorie. Het is immers zo dat de opgerente waarde van een reeks toekomstige spaarbedragen precies in deze vorm kan geschreven worden, waarbij een dergelijke Z_i de stochastische oprentingsfactor over de periode $[i, n]$ voorstelt. Ook de huidige waarde van een reeks toekomstige betalingen α_i kan aan de hand van (1) uitgedrukt worden. In dit geval zullen de Z_i de stochastische verdisconteringsfactoren over de periode $[0, i]$ vertegenwoordigen.

Dit verklaart waarom de waardering van Aziatische of Basket opties in een Black & Scholes model, alsook de bepaling van provisies en vereiste kapitalen in een verzekeringscontext, zich precies herleidt tot het bepalen van de gepaste *risicomaat* voor kansvariabelen S zoals in (1).

We definiëren een risicomaat als een functie van de verzameling van kansvariabelen naar de reële getallen. Dergelijke risicomaten zijn een erg waardevol hulpmiddel in het hele beslissingsgebeuren. Ze vatten immers de informatie over het risico S samen in één enkel getal $\rho[S]$. Risicomaten duiken in het actuariaat ook op bij het bepalen van premieprincipes, maar in deze thesis bestuderen we haast uitsluitend risicomaten, die kunnen dienen om het niveau van reserves en solvabiliteitsmarges te bepalen.

The stochastische som (1) zal, als de Z_i stochastische verdisconterings- of oprentingsfactoren voorstellen, een som van sterk afhankelijke lognormalen zijn. Er is immers onvermijdelijk een sterke overlapping wanneer we verdisconteren (of oprenten) over de beschouwde tijdsperioden. Ook zullen de meeste risicomaten van S niet analytisch bepaald kunnen worden. In dergelijke gevallen is het zeker te overwegen om de afhankelijkheidsstructuur van de kansvector $(Z_1, Z_2, \dots, Z_n)$ zó te bepalen dat de som S ‘gevaarlijker’ wordt. Men kan immers verwachten dat als S ‘gevaarlijker’ wordt, dit aanleiding zal geven tot het nemen van meer conservatieve beslissingen. Vanuit het oogpunt van het risicobeheer is dat een goede zaak. We moeten dan uiteraard wel afspreken wat ‘gevaarlijker’ precies betekent, i.e. we moeten definiëren hoe we risico’s rangschikken in termen van ‘gevaarlijkheid’.

We zullen vooral werken binnen de klasse van convexe orde. Deze stemt immers overeen met de gemeenschappelijke voorkeuren van de risicomijdende beslissingsnemers, indien zij moeten kiezen tussen risico’s met dezelfde verwachtingswaarde, en dit geldt zowel in de klassieke nutsheorie als in de duale risictheorie van Yaari. De ‘gevaarlijkste’ som zal opduiken wanneer de kansvariabelen $\alpha_i e^{Z_i}$ comonotoon zijn. Dit wil zeggen dat ze monotone functies zijn van een gemeenschappelijke kansvariabele. Dit verklaart ook het woord comonotoniciteit (common monotonic). Het gebruiken van deze comonotone bovengrens, genoteerd met S^c , heeft eerst en vooral het erg belangrijke voordeel dat de meeste actuariële grootheden zoals de quantielen en stop-loss premies zeer eenvoudig kunnen bepaald worden en men dus geen beroep zal moeten doen op tijdrovende

simulaties. Bovendien leidt het gebruik van S^c in de regel ook tot het nemen van voorzichtige beslissingen. We zullen ook aantonen dat soms zin heeft om S te vervangen door de ‘optimistischer’ benedengrens S^l , die we bekomen door de conditionele verwachting $E[S \mid \Lambda] = \sum_{i=0}^n \alpha_i E[e^{Z_i} \mid \Lambda]$ met betrekking tot een conditioneringsvariabele Λ te bepalen. Indien de termen $E[e^{Z_i} \mid \Lambda]$ monotone functies zijn van Λ , bekomen we, net zoals in het geval van de comonotone bovengrens, een comonotone som en hiervan kunnen we de meeste risicomaten eenvoudig bepalen. Zich baseren op het minder gevaarlijke risico S^l in plaats van het meer gevaarlijke S^c , gaat in tegen de intuïtie van de risicomijdende beslissingsnemer, maar numerieke testen tonen aan dat de risicomaten van S^l , statistisch gesproken, haast niet onderscheiden kunnen worden van de gesimuleerde risicomaten van de originele variabele S . De actuaaris kan dan oordelen dat het criterium van ‘nauwkeurigheid’ zwaarder weegt dan het criterium van ‘voorzichtigheid’ en dat we dus toch de minder voorzichtige benedengrens S^l zullen nemen, zelfs al leidt deze tot het nemen van iets meer risico.

De veronderstelling dat de kansvector $(Z_1, Z_2, \dots, Z_n)$ een multivariate normale verdeling volgt, wordt in de literatuur vaak gemaakt en is in overeenstemming met empirische waarnemingen, tenminste voor wekelijkse rendementen of voor rendementen over langere periodes. Aangezien in onze toepassingen de tijdseenheid typisch één jaar bedraagt en ook de Centrale Limietstelling in ons voordeel speelt, is de veronderstelling van normaliteit zeker redelijk te noemen.

Deze thesis is als volgt ingedeeld.

In het eerste gedeelte leggen we de relatie uit tussen comonotoniciteit en het meten van risico. In het eerste hoofdstuk bespreken we bondig enkele belangrijke actuariële ordeningsconcepten. Vervolgens bediscussiëren we in het tweede hoofdstuk de theorie over comonotoniciteit en tonen we hoe de comonotone (convexe) bovengrens voor een stochastische som bepaald kan worden. In dit hoofdstuk tonen we ook hoe met het nemen van conditionele verwachtingswaarden een convexe benedengrens kan bekomen worden. Hoofdstuk 3 analyseert risicomaten en toont dat ze in het geval van comonotone sommen meestal eenvoudig te bepalen zijn. De risicomaat van een dergelijke som is immers meestal te schrijven als een som van risicomaten voor ieder van de termen in de som. Hoofdstuk 4 onderzoekt de precisie van de comonotone boven- en benedengrenzen in het geval dat S een som van afhankelijke lognormalen voorstelt. Een aantal (distortie-) risicomaten, berekend met deze grenzen, worden dan vergeleken met benaderingen gebaseerd op andere verdelingen met dezelfde verwachtingswaarde en variantie als de oorspronkelijke variabele S . Het hoofdstuk besluit dat de ‘maximale variantie’ benedengrens uitstekend presteert. In het laatste hoofdstuk van dit gedeelte bepalen we comonotone boven- en benedengrenzen voor constante continue annuïteiten en leiden we analytische uitdrukkingen voor verschillende risicomaten van deze benaderingen af.

In het tweede gedeelte van deze thesis bespreken we een aantal erg belangrijke

problemen uit de financiële en actuariële wereld, waar de stochastische som (1) centraal staat en waarvoor de comonotone resultaten uit het eerste gedeelte mooie en accurate resultaten geven.

In hoofdstuk 6, dat de kern van deze thesis is, gebruiken we de comonotone benaderingen om in een Black & Scholes setting, het meerperiode-portefeuille-selectieprobleem op te lossen. Strategische portefeuille selectie komt neer op het bepalen van de optimale asset mix en dit voor een gegeven risicoprofiel en een welbepaald consumptie- of spaarpatroon. De investeerder zal hierbij kunnen kiezen uit risicovolle activa of assets zoals aandelen, vastgoed en obligaties maar ook uit (quasi) risicoloze beleggingen zoals cash of money-market instrumenten. Hij bepaalt dan de initiële percentages die hij in elk van deze activa zal beleggen en een specifieke trading strategie volgens welke hij voortdurende activa zal kopen en verkopen. We zullen uitgaan van de hypothese dat de investeerder zo handelt dat hij de initiële percentages constant houdt en we bepalen dan, gegeven deze constante mix strategie, de voor hem ideale percentages, i.e. de optimale asset mix. We zullen twee problemen van naderbij bekijken, namelijk het spaarprobleem en het reserveringprobleem.

In het *spaarprobleem* gaan we uit van een investeerder die op regelmatige tijdstippen $0, 1, \dots, n-1$, vaste bedragen $\alpha_0, \alpha_1, \dots, \alpha_{n-1}$ zal sparen zodanig dat het bijeengespaard bedrag op einddatum n , met een voldoende grote kans, minstens even groot zal zijn als een welbepaald doelkapitaal K . Aangezien het bijeengespaard bedrag op einddatum n een som van lognormalen zal zijn, kan de verdelingsfunctie ervan niet exact bepaald worden en is deze som in het algemeen zeer moeilijk hanteerbaar. We zullen daarom comonotone boven- en benedengrenzen gebruiken. Deze benaderingen hebben immers het enorme voordeel dat ze voor iedere gekozen initiële mix, accurate en makkelijk uit te rekenen benaderingen geven voor iedere risicomaat die additief is voor comonotone risico's. Voorbeelden van dergelijke risicomaten zijn de distortie-risicomaten (bvb. VaR, TVaR). De comonotone benaderingen herleiden essentieel het multivariate karakter van het probleem tot een univariaat karakter. Stel nu dat we een kans $1 - \epsilon$ opleggen waarmee we het doelkapitaal zeker willen halen. De optimale mix kan dan bepaald worden als diegene waarvoor het periodiek vaste spaarbedrag om dit doel te bereiken zo klein mogelijk is. Andersom kunnen we voor een gegeven vast spaarbedrag de optimale mix bepalen zodanig dat de kans om een doelkapitaal te bereiken maximaal wordt. Het is zonneklaar dat dergelijke tool enorme toepassingen heeft, bijvoorbeeld in het domein van persoonlijke financiële planning. Een voorbeeld hiervan is pensioensparen. In dit geval wenst een persoon bij zijn pensioen te beschikken over een voldoende groot appeltje voor de dorst K . In welke asset mix moet hij nu beleggen, zodanig dat het maandelijks spaarbedrag zo klein mogelijk is, maar toch voldoende groot om dit doelkapitaal K bijeen te krijgen met een kans van minstens $1 - \epsilon$? De theorie over comonotoniciteit geeft een snel, elegant en nauwkeurig antwoord op deze vraag.

In het *reservingprobleem*, wordt een investeerder geconfronteerd met een serie van toekomstige verplichtingen of consumpties $\alpha_1, \alpha_1, \dots, \alpha_n$ waaraan moet

voldaan worden op tijdstippen $1, 2, \dots, n$.

Voor dit investeringsprobleem, zou men de initiële mix zo kunnen bepalen dat deze kans precies maximaal wordt. We zouden ook deze kans kunnen vastleggen en de initiële mix bepalen die ertoe leidt dat de beginreserve minimaal wordt.

In hoofdstuk 7, staat het *hordeloopprobleem* centraal. In dit hoofdstuk bepalen we op een andere manier de initiële reserve. We zullen immers niet allen rekening houden met de ‘eindstreep’ zoals we in het *reserveringprobleem* deden, maar ook met de ‘horden’. We zullen immers opleggen dat de aanwezige reserve R_j steeds groter moet zijn dan een vooraf vastgelegde waarde V_j . Deze bedragen zijn precies de horden die gehaald moeten worden met kansen $1 - \varepsilon_1, 1 - \varepsilon_2, \dots, 1 - \varepsilon_n$. In het geval van een verzekeraar die zijn reserve bepaalt, kan het management beslissen om dergelijke horden op te leggen, maar ze kunnen ook verplicht gesteld worden door een controlerende instantie. Deze invalshoek maakt het mogelijk om de kansen waarmee men zijn horden moet halen tijdsafhankelijk te maken iets wat soms aangewezen kan zijn. Een verzekeraar is in de eerste plaats geïnteresseerd om op de korte termijn zeker aan zijn verplichtingen te kunnen voldoen. Dit betekent dat de vereiste kansen om de eerste horden te nemen vaak groter moeten zijn dan de kansen om de laatste horden te halen. Dit kan uiteraard leiden tot situaties waarbij de bindend voorwaarde niet de eindstreep is, zoals in het reserveringprobleem, maar één van de vroegere horden en dit zal weer-spiegeld worden in een hogere initiële reserve.

De voorgaande hoofdstukken en de numerieke voorbeelden hierin, tonen aan dat vooral de ‘maximale variantie’ comonotone ondergrens, uitstekende en eenvoudig te berekenen oplossingen biedt voor het beschouwde probleem. Het is nu wel zo dat we in deze toepassingen enkel cash-flows α_i hebben beschouwd die allemaal positief (of allemaal negatief) zijn. Deze eigenschap van het cash-flow patroon is inderdaad cruciaal om de berekeningen eenvoudig (voor de benedengrens) en accuraat (voor de bovengrens) te houden. Indien in de cash-flows zowel negatieve als positieve bedragen zijn, zal het voor de comonotone benedengrens benadering niet mogelijk zijn een conditioneringsvariabele Λ te vinden die het mogelijk maakt dat de (distortie-) risicomaten eenvoudig én accuraat te berekenen zijn.

In het laatste hoofdstuk zullen we aantonen dat, indien de cash-flows eerst positief zijn (sparen) en dan pas negatief (consumptie), we toch de comonotone benedengrenzen kunnen gebruiken. Een belangrijke situatie waarin we dit cash-flow profiel ontmoeten, is het ‘spaar-consumptie’-probleem. Neem als voorbeeld het zogenaamde 20/65/95 pensioen plan: een persoon van 20 jaar wil maandelijks tot aan zijn pensioen vaste bedragen in zijn pensioenfonds storten. Vervolgens, als hij op rust is, wil hij iedere maand geld van deze rekening afhalen om zo zijn levensstandaard op peil te kunnen houden. We kunnen met onze comonotone benaderingen de volgende relevante vraag erg elegant oplossen: “Hoe hoog is de fractie α die hij gedurende 40 jaar iedere maand moet storten voor elke

euro die hij na zijn actieve loopbaan maandelijks en dit gedurende 30 jaar wil consumeren?”

In financiële en actuariële wetenschappen, botst men vaak op stochastische sommen $S = \sum_{i=1}^n X_i$, waarvan de multivariate structuur niet gekend of niet makkelijk hanteerbaar is. Comonotoniciteit herleidt het multivariate karakter tot een univariaat gegeven. In deze thesis tonen we hoe deze theorie ons helpt om een aantal (in de praktijk veel gebruikte) risicomaten van S te kunnen benaderen. Vervolgens tonen we, in een Black & Scholes wereld, dat de comonotone benaderingen ons in staat stellen om een aantal belangrijke risicobeheerproblemen efficiënt en elegant op te lossen.