

DECISION SUPPORT MODELING
VIA
MULTIVARIATE RISK MEASURES AND
STOCHASTIC OPTIMIZATION

BY JINWOOK LEE

A dissertation submitted to the
Graduate School—New Brunswick
Rutgers, The State University of New Jersey
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy
Graduate Program in Operations Research

Written under the direction of
András Prékopa
and approved by

New Brunswick, New Jersey

May, 2014

ABSTRACT OF THE DISSERTATION

Decision Support Modeling via Multivariate Risk Measures and Stochastic Optimization

by Jinwook Lee

Dissertation Director: András Prékopa

Whenever we have a decision to make, there is always some risk to take. From a mathematical perspective, risk is manifested by a random variable, and a risk measure simply characterizes the random variable in a more compact form. Risk, in general and in practice, is not be adequately described by a real valued random variable, but rather requires a random vector to capture the dimensions of the problem. To this end, multivariate risk measures are crucial ingredients for decision making processes, and stochastic optimization is a natural and superior skill to find a key to the optimal decision-making.

A recent paper by Prékopa (2012) presented results in connection with Multivariate Value-at-Risk (MVaR) that has been known for some time under the name of p -quantile or p -Level *Efficient Point* (pLEP) and introduced a new multivariate risk measure, called Multivariate Conditional Value-at-Risk (MCVaR). Lee and Prékopa (2013) studied new methods for numerical calculations and mathematical properties of these multivariate risk measures, presented in Chapter 2. Another new multivariate risk measure has been constructed and presented in Chapter 3. This is especially for

corporate mergers and acquisition (M&A) transactions, as the limited applicability of a coherent risk measure in the sense of Artzner et al (1999) for M&A transactions is already discussed in Kou et al (2013). A decision making scheme using that risk measure is introduced and surveyed, together with illustrative real-life numerical examples.

Insurance companies typically hold their money in bonds to pay out the random liabilities in the same periods. In Chapter 4, such bond portfolio construction problem is presented using various stochastic programming problem formulations. For a financial trading business, “price-bands” can be used as an indicator for successfully buying or short-selling shares of stock. Chapter 5 presents a mathematical model for the novel construction of price-bands using a stochastic programming formulation. Numerical examples using recent US stock market intraday data are presented.

Acknowledgements

I would like to acknowledge the inspirational instruction and the encouraging guidance of my adviser, Professor András Prékopa. It was my honor and pleasure to become his student and to work with him. His knowledge and insight of all aspects of relevant fields of study have truly amazed and inspired me. He encouraged both my academic and personal growth and development. I would like to express the highest and the most sincere gratitude for his academic and nonacademic influence on my life.

Dedication

To my wife, Yunkee

Table of Contents

Abstract	ii
Acknowledgements	iv
Dedication	v
List of Tables	viii
List of Figures	ix
1. Introduction	1
2. Properties and Calculation of Multivariate Risk Measures: MVaR and MCVaR	5
2.1. Introduction	5
2.2. The Notions of Multivariate Value-at-Risk (MVaR) and Multivariate Conditional Value-at-Risk (MCVaR)	9
2.3. Properties of MVaR and MCVaR	12
2.4. Calculation of MCVaR	25
2.4.1. The Case of a Continuous Distribution	25
2.4.2. The Case of a Discrete Distribution	33
Application of the Binomial Moment Bounding Scheme	33
Application of the Boolean Bounding Scheme	38
2.5. Concluding remarks	41
3. Decision-making from a Risk Assessment Perspective for Corporate Mergers and Acquisitions	42
3.1. Introduction	42

3.2. Risk assessment methods for corporate M&As from the mathematical perspective	44
3.3. Motivation to study a new risk measure for corporate M&As	45
3.4. Taking a Bottom-Up approach in the multidimensional case	47
3.5. Combined Value-at-Risk	49
3.6. Combined Value-at-Risk in the worst-case event	51
3.7. Decision-making via potential risk measure on corporate M&As	56
3.8. Numerical examples and discussion	58
3.9. Concluding remarks	65
4. Bond Portfolio Optimization for Insurance Companies	69
4.1. Introduction	69
4.2. Stochastic Programming Formulation	70
4.3. Numerical Example and Its Solutions	71
4.3.1. Random Liabilities	71
4.3.2. Bond Portfolio construction	74
4.4. Stochastic Programming Models	75
4.4.1. Probabilistic Constrained Model	76
4.4.2. Simple Recourse Model	77
4.4.3. Hybrid Model	80
5. Price-bands: A Technical Tool for Stock Trading	82
5.1. Introduction	82
5.2. Construction of Price-bands	83
5.3. Stochastic combinatorial optimization problem formulation	89
5.4. Binomial moment problem formulation	94
5.5. Numerical examples and discussion	101
5.6. Concluding remarks	108
6. Conclusion	110

List of Tables

2.1. Summary of MCVaR of different stochastic dependence relationship at various probability levels.	31
2.2. An insurance company's daily claims (in \$1,000) from 4 types of insurance.	35
2.3. $\text{MVar}_p(X)$ with $p = 0.9$, the unit claim size \$1,000.	36
3.1. Risk after M&A deals, i.e., risk of the new united firms	59
3.2. Sum of the risks of the companies before M&As	59
3.3. Risk before M&As for each company	60
3.4. Probability of the Worst Case Outcome	60
3.5. Risk after M&A deals, i.e., risk of the new united firms	64
3.6. Sum of the risks of the companies before M&As	64
3.7. Risk before M&As for each company	65
3.8. Probability of the Worst Case Outcome	65
4.1. Prices, coupon payments and maturities of 10 bonds	75
5.1. The Bollinger Bands	84
5.2. The Prékopa-Lee Bands	100
5.3. Comparison of the performances of Bollinger and Prékopa-Lee bands .	103

List of Figures

2.1. 2-D Illustration of the favorable set, and its complementary set where the Multivariate Conditional Value-at-Risk is defined in both types of a random vector – discrete and continuous.	11
2.2. 2-D Illustration of the first order stochastic dominance of random vectors: $X \preceq_{(1)} Y$	13
2.3. Illustration of the set $\{z \mid F(z) \geq p\}$	14
2.4. A counterexample of MCVaR Property (3): subadditivity, in the general case.	21
2.5. Illustration of $\text{MVaR}_{\mathbf{p}}(\mathbf{Z})$, where $\mathbf{Z} = \mathbf{X} + \mathbf{Y}$	22
2.6. Projection of $\text{MVaR}_{\mathbf{p}}((\mathbf{X}_1, \mathbf{X}_2)^{\mathbf{T}})$ from \mathbf{R}^2 onto the space of $\mathbf{X}_1, \mathbf{X}_2 \in \mathbf{R}$	24
2.7. Projection of $\text{MVaR}_{\mathbf{p}}((\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3)^{\mathbf{T}})$ from \mathbf{R}^3 onto the space of $(\mathbf{X}_1, \mathbf{X}_2)^{\mathbf{T}} \in \mathbf{R}^2$	24
2.8. Basic Chart of Fidelity Nasdaq ETF and Vanguard S&P 500 ETF.	27
2.9. Multivariate Value-at-Risk, the quantile function at the probability levels $p = 0.95$ and $p = 0.99$	28
2.10. Basic Chart of Deutsche Bank US Dollar Index ETF and Fidelity Nasdaq ETF.	29
2.11. Multivariate Value-at-Risk, the quantile function at the probability levels $p = 0.95$ and $p = 0.99$	30
2.12. Illustration of Compound Poisson Distributed Losses.	35
3.1. Favorable set and its complementary set in the case of a loss random variable	48
3.2. 2-D illustration of the favorable set and its complement.	50

3.3. 2-D Illustration of the worst-case unfavorable set	52
3.4. Description of projections of the set W	55
3.5. Comparison of risks among different M&A deals	67
3.6. Comparison of risks among different M&A deals	68
5.1. Past, present, future timeline	85
5.2. Description of Gaussian process	86
5.3. Description of data points and their corresponding probabilities	92
5.4. Description of data points and their corresponding probabilities	95
5.5. Prékopa-Lee bands on Apple Inc. (NASDAQ: AAPL) from May 1 to July 16, 2013	104
5.6. Prékopa-Lee bands on Verizon Communications Inc. (NYSE: VZ) from May 1 to July 16, 2013	105
5.7. Prékopa-Lee bands on Yahoo! Inc. (NASDAQ: YHOO) from May 1 to July 16, 2013	106
5.8. Prékopa-Lee bands on The Walt Disney Company (NYSE: DIS) from May 1 to July 16, 2013	107

Chapter 1

Introduction

When faced with any forward-looking decision, there will always be some risk of making an inappropriate or suboptimal decision. If there were no risk, then all ends are known and the optimal decision is clear to a clever individual – one might say there is hardly a decision to make. Given the prevalence of uncertainty in many problems of interest, risk analysis is a crucial aspect of an effective decision-making process.

The business environment is increasing not only in its complexity, but also in its intricacy. With the advent of smart systems and detailed data capture tracking every aspect of a business you can imagine, powerful datasets are now available and methods for their effective use are paramount for success in the competitive landscape. The explosion of this phenomenon – popularly known as “big data” – is transforming various industries from finance to healthcare through the use of analytics cleverly tracking key parameters to guide optimal decisions. Often, more often than not in fact, decisions must be made in the face of substantial risk and uncertainty, and methods for analytics under such risk are surprisingly incipient. Effective modeling and development of business analytics under suitable risk conditions will be a significant edge in the age of big data.

From a mathematical point of view, risk is a random variable itself, and a risk measure simply characterizes the random variable in a more compact form. Risk, in practice, may not be equal to a real valued random variable, rather, it is frequently represented by a finite collection of random variables, i.e., a random vector, to capture the dimensions of the problem. In order to characterize the diverse risk exposure of the entity, we therefore need to work with the joint probability distribution of the relevant random variables, which is often a significantly more complex issue. Hence, broadly

the discipline of risk management is fundamentally tied to the elements of random vectors, risk measures, and stochastic processes and their role in mitigating future negative outcomes. For effective risk management, companies must establish rigorous risk assessment processes via a “suitable risk measure.” However, a suitable risk measure is inherently related to a judgement of the outcome of a random experiment, and so, in this sense, a fluid understanding of the industry is decidedly synergistic.

A company typically has many different assets, portfolios, business sectors, exposed to different kinds of randomness, influencing the overall behavior of the company. In order to characterize it, from the point of view of risk exposure, we need to work with the joint probability distribution of the random variables involved. Thus, for a complex real world environment, I believe multivariate risk measures are crucial ingredients for the decision-making processes and stochastic optimization with relevant risk measures will play indispensable role in finding a key to optimal decision-making.

The organization of this dissertation is as follows. In Chapter 2, we have explored various properties of Multivariate Value at Risk, or MVaR and Multivariate Conditional Value at Risk, or MCVaR. We have shown that many properties enjoyed by VaR and CVaR, carry over to the multivariate risk measures. We also have proposed the numerical procedures to calculate or approximate MCVaR values. In Section 2.2 we recall the notions of Multivariate Value-at-Risk and Multivariate Conditional Value-at-Risk, following the guidelines of Prékopa (2012). In Section 2.3 basic properties of both risk measures are stated. While MVaR enjoys similar properties as the univariate counterpart, MCVaR does not have the convexity property. Explanation is supplied. In Section 2.4 we present numerical procedures, for both the continuous and discrete cases, to approximate MCVaR, by the use of bounding, based on the binomial moment method and the Boolean bounding scheme. The practical meaning of MCVaR is illustrated on two portfolios with different correlation structures. Finally, in Section 2.5 we present conclusions.

In Chapter 3, we construct a further multivariate risk measure: the worst case Combined Value-at-Risk (wCoVaR), where only one orthant of the space represent unfavorable set and its vertex is at the vector with components equal to the individual

VaR's. wCoVaR has a very strong connection with the univariate risk measure: Conditional Value-at-Risk (CVaR) – sum of individual CVaR equals wCoVaR if all random variables are independent. This makes it possible to compare values of sum of random variable and a random vector which has the random variables as components. And this idea is to help decision making in corporate mergers and acquisitions by comparison of risks before and after M&As or demergers activities before taking action on the deals. In Section 3.7 we show in what way it can be used in practice. In the numerical examples we look at one company which considers M&As with one or two of a few target candidates, we calculate which M&A deal is ideal in terms of risk and which M&A deals reduce risk. To do the above analysis we have introduced vector operations, where we put together risk vectors to create new risk vector with increased number of components, to describe M&As, and split a risk vector into parts, to describe demergers, i.e. restructuring of companies.

In Chapter 4, various stochastic problem formulations are presented for bond portfolio problem of insurance companies. Insurance companies typically hold their money in bonds to pay out the random liabilities in the same periods. Insurance claims are randomly occurring events, which is considered as liabilities to an insurance company. The probability of a number of events occurring in a fixed period of time can be expressed as a Poisson distribution. The problem is that how many of the different bonds should be purchased that minimizes the cost subject to the constraint that all liabilities can be payed out in the course of a given number of periods. Numerical examples are presented.

In Chapter 5, we present a mathematical model to construct “price-bands,” which are certainly helpful to deter investors from *entirely* following their feelings. This has been widely used in practice, especially for short term investment, to help people validate their investment decisions. As one variant, we construct new price-bands via binomial moment problem formulation under the assumption that stock prices follow a Gaussian process. Usage of conditional probability distributions is the key attribute

that differentiates our model. We presents a mathematical model for the novel construction of price-bands using a stochastic programming formulation. Numerical examples using recent US stock market intraday data are presented. This methodology for forecasting upper and lower bounds need not only apply in finance, but could also be applicable in many business management areas, e.g., supply chain management, production management, inventory control, reliability engineering, etc.

In any business practice, many different technical tools exist to guide decision makers through the swarm of information. Information is processed data, whereas data are plain facts – analytics are the link. Business analytics must provide timely information in order to play a material role in the decision making process. In Chapter 5, we consider a financial institution and, using intraday stock price data, present novel price-bands, one of the most widely used analytics in financial trading. Based on historical data, there are a huge number of possible cases on what will happen over the next day. All possible cases should be appropriately considered to make a reasonable decision, but in a systematic way allowing faster computation. To this end, we formulated the problem as a modified binomial moment problem, which effectively counts all possible cases without actual counting *per se* and renders the daunting problem solvable.

Through my doctoral residency with my adviser, Professor András Prékopa, I have found that a set theoretical approach is useful for many stochastic optimization problems, especially in a vector space, since set theory offers an intuitive way to represent data in a multi-dimensional space. With a suitable construction of sets, a mathematical programming formulation can be created using those sets as inputs for the objective function or constraints. For the calculation of multivariate risk measures in Chapter 2, the modified binomial moment and Boolean bounding schemes are used. For more complex problems, introduction of functions related to those sets was helpful, see, e.g., the equations and formulations in Chapter 5.

Chapter 2

Properties and Calculation of Multivariate Risk Measures: MVaR and MCVaR

2.1 Introduction

Value-at-Risk (VaR) has already existed in the statistical literature since the second half of the 19th century, under the name of quantile or percentile. The term Value-at-Risk was introduced at the beginning of the 1990s in the financial literature and became widely used in a short time. We refer the reader to Jorion (2006) and Saita (2007) for various topics of Value-at-Risk. Its multivariate counterpart turned up in the stochastic programming literature, primarily in the works of Prékopa (1970, 1973a, 1990, 1995, etc.). Based on this, Multivariate Conditional Value-at-Risk (MCVaR) was recently introduced by the same author (2012).

In stochastic programming one standard way to create a decision model out of one, where some of the parameters are random, is to prescribe a lower bound on the probability that the stochastic constraints are jointly satisfied. If, for example, a decision problem is an LP: $\min c^T x$ subject to $Tx \geq \xi$, $Ax = b$, $x \geq 0$, where ξ is a random vector, then we may formulate the problem: $\min c^T x$ subject to $P(Tx \geq \xi) \geq p$, $Ax = b$, $x \geq 0$, or $\min \{c^T x + \sum_{i=1}^r q_i E([\xi_i - T_i x]_+)\}$, subject to the same constraints, where T_i is the i th row of the $r \times n$ matrix T and the q_i , $i = 1, \dots, r$ are nonnegative constants. The practical application of this model goes in such a way that first we decide on the value of x and, after that, we observe the realized value of ξ . The probability p is chosen near 1 so that the inequality $Tx \geq \xi$ should be satisfied in most cases. If ξ has continuous distribution and its c.d.f. is $F(z) = P(\xi \leq z)$, $z \in R^r$, then the probabilistic constraint can be rewritten as: $Tx \geq z$, for at least one z such that $F(z) = p$. If ξ is discrete, then we may use the *p-Level Efficient Points* (pLEP's), or briefly *p-efficient*

points $z^{(1)}, \dots, z^{(N)}$, and reformulate the probabilistic constraint as $Tx \geq z^{(i)}$, for at least one $i = 1, \dots, N$. The above mentioned sets $\{z \mid F(z) = p\}$ and $\{z^{(1)}, \dots, z^{(N)}\}$ can be regarded as multivariate quantiles. In Prékopa (2012) the term Multivariate Value-at-Risk (MVaR) was introduced as an alternative name for the collection of p -efficient points. Methods to generate elements of MVaR in the case of a continuously distributed ξ and the entire MVaR, in the discrete case, has already been existed in the literature (see, e.g., Prékopa (1995) and the references therein; Prékopa, Vizvári, Badics (1998); Boros et al (2003); Dentcheva, Prékopa, Ruszczyński (2000), etc.).

The term Conditional Value-at-Risk (CVaR) was introduced by Rockafellar and Uryasev (2000). The same notion was named by Föllmer and Schied (2002) Average Value at Risk (AVaR). Earlier, in 1973, Prékopa has already used conditional expectation as a risk measure in stochastic programming. If the rows of the matrix T are T_1, \dots, T_r and the components of ξ are ξ_1, \dots, ξ_r , then the use of the constraints $E(\xi_i - T_i x \mid \xi_i - T_i x > 0) \leq d_i$, $i = 1, \dots, r$ was proposed as replacement of the computationally more complicated constraint: $P(Tx \geq \xi) \geq p$, or, as a supplement to it. One major advantage of the conditional expectation constraints is that if the components of ξ have continuous distributions with logconcave p.d.f.'s, then each of them is equivalent to a linear constraint (see Prékopa (1973a, 1995)).

Let F denote the probability distribution function of the random variable $X \in R$. Then the Value-at-Risk (VaR), for some fixed probability level p , is defined as the p -quantile of the probability distribution function F :

$$\text{VaR}_p(X) = F^{-1}(p), \quad (2.1)$$

where, by definition,

$$F^{-1}(p) = \min\{u \mid F(u) \geq p\}. \quad (2.2)$$

It can also be defined as the optimum value of the following two problems:

$$\begin{aligned} \min v \\ \text{subject to } P(X \leq v) \geq p, \end{aligned} \quad (2.3)$$

and

$$\begin{aligned} & \sup v \\ & \text{subject to } P(X \geq v) > 1 - p, \end{aligned} \tag{2.4}$$

where p is some fixed probability, $0 < p < 1$. The optimum values of problems (5.1) and (2.4) are equal. Value-at-Risk (VaR) has a property, considered undesirable by many authors for a risk measure: it is not convex, in general, and it measures the frequency, not the amount of losses beyond VaR (the predicted maximum amount of losses at a fixed probability level). This motivated the development of the notion of a coherent risk measure, equal to the conditional expectation of a random variable, given that it surpasses $\text{VaR}_p(X)$. Conditional Value at Risk, designated by $\text{CVaR}_p(X)$, where X is the random variable involved and p the probability, is defined as:

$$\text{CVaR}_p(X) = E(X \mid X \geq \text{VaR}_p(X)). \tag{2.5}$$

Uryasev and Rockafellar (2000) and Pflug (2000) have shown that

$$\text{CVaR}_p(X) = \min_a \left\{ a + \frac{1}{1-p} E([X - a]_+) \right\} \tag{2.6}$$

and that $\text{CVaR}_p(X) \geq \text{VaR}_p(X)$. Equation (2.6) can also be used as the definition of $\text{CVaR}_p(X)$.

An optimization problem, similar to that of (2.6) was introduced and applied to a “chance constrained problem” by Ben-Tal and Teboulle (1986). After reformulation of the problem, a random objective function is obtained, for which a new type of “certainty equivalent” is formulated. If X is a random variable and u is a (increasing and strictly concave) utility function, then it is equal to:

$$\sup_a \{a + E[u(X - a)]\}. \tag{2.7}$$

In the mentioned and in other papers (see, e.g., Ben-Tal and Ben-Israel (1991), Ben-Tal et al (1991), etc.) Ben-Tal, Ben-Israel and Teboulle expound a theory and application of the new certainty equivalent (2.7) and show that its negative enjoys the properties of a coherent risk measure in the sense of Artzner et al (1999).

Coherence, however, is not such a property of a risk measure that it would be imperative to rely on it, under all circumstances. VaR is not a coherent risk measure,

in general, but it is widely and successfully used in many applications, such as: testing statistical hypotheses, sequential analysis, decision theory, stochastic programming and others. If, for example, an unfavorable event may cause huge damage that has to be avoided, then VaR may be more important than CVaR. Another point is that CVaR takes average on rare events while small probabilities are multiplied by large numbers, making the estimation of the risk measure inaccurate and we need very long trial sequences to realize the benefit of the conditional expectation in practice, where the conditioning event has very low frequency. If the population is at hand at the same time, then CVaR may have reasonable practical interpretation.

We think that risk measures and axiomatic systems for risk measures should not be regarded in an exclusive manner. As in geometry, various axiomatic systems define various geometries out of which we may choose the one most suitable for a given application, the axiomatic systems for risk measures and the risk measures themselves should provide us only with a menu to choose one or more than one for our purpose.

Risk measures for multidimensional settings have previously been studied, and we refer the reader to the recent literature (e.g., see Dentcheva and Ruszczyński (2009), Noyan and Rudolf (2013), etc.). Multivariate risk measures, other than ours, exist in the literature, for example, see Cousin and Di Bernadino (2011) and the references therein. The results in connection with them, however, are little to do with ours, especially because those risk measures are mostly vectors while ours are numbers. On the other hand we have in mind applications in stochastic optimization which require convexity statements and algorithms to calculate the numerical values of the risk measures.

The organization of this paper is as follows. In Section 2.2 we recall the notions of Multivariate Value-at-Risk and Multivariate Conditional Value-at-Risk, following the guidelines of Prékopa (2012). In Section 2.3 basic properties of both risk measures are stated. While MVaR enjoys similar properties as the univariate counterpart, MCVaR does not have the convexity property. Explanation is supplied. In Section 2.4 we present numerical procedures, for both the continuous and discrete cases, to approximate MCVaR, by the use of bounding, based on the binomial moment method and the Boolean bounding scheme. The practical meaning of MCVaR is illustrated on two portfolios

with different correlation structures. Finally, in Section 2.5 we present conclusions.

2.2 The Notions of Multivariate Value-at-Risk (MVaR) and Multivariate Conditional Value-at-Risk (MCVaR)

While VaR and CVaR both have been around for some time, only VaR had a multivariate counterpart. However, the fact that we intend to take into account the stochastic dependence of the random variables involved, called for the introduction of the Multivariate Conditional Value at Risk or MCVaR. That was done in the recent paper by Prékopa (2012). For the sake of completeness below we recall the definitions of both MVaR and MCVaR (Definitions 2.2.1 and 2.2.2, respectively).

Definition 2.2.1. (*Prékopa 1990*) Let $X \in R^r$ be a random vector and F its c.d.f. A point $s \in R^r$ is said to be a p -Level Efficient Point, or briefly p -efficient point, of the probability distribution, or the distribution function F , if $F(s) \geq p$ and there is no y such that $y \leq s$, $y \neq s$, $F(y) \geq p$. $MVaR_p(X)$ is the set of all p -efficient points of the random vector X .

If X has discrete distribution on Z^r , then its support is finite or countably infinite. In both cases $MVaR_p(X)$ is a finite set by the following

Theorem 2.2.1. *If the components of the random vector ξ are integer-valued, then for any $p \in (0, 1)$ the set of p -level efficient points is nonempty and finite.*

Theorem 2.2.1 is an immediate consequence of Dickson's Lemma (3, Cor. 4.48). It was mentioned, in another context, by Vizvári (1987) and Dentcheva, Prékopa, Ruszczyński (2000), for p -efficient points. The assertion of Theorem 2.2.1 is not necessarily true if the support of the random vector is countable but not part of the integer lattice. In case of an integer valued Y , there exist $N \geq 1$ and $s^{(1)}, \dots, s^{(N)}$ such that

$$MVaR_p(Y) = \{s^{(1)}, \dots, s^{(N)}\},$$

where $s^{(i)} \in R^r$, $i = 1, \dots, N$. If X has continuous distribution, with strictly increasing c.d.f. ($F(s_1) > F(s_2)$ if $s_1 \geq s_2$, $s_1 \neq s_2$), then

$$MVaR_p(X) = \{s \mid F(s) = p\}. \quad (2.8)$$

The following concepts were introduced in Prékopa (2012). Suppose that the random vector X is related to losses, then the favorable event for X is defined as

$$X \in \bigcup_{s \in \text{MVar}_p(X)} (s + R_-^r). \quad (2.9)$$

The complementary of the event in (2.9) is the unfavorable event:

$$X \in \bigcap_{s \in \text{MVar}_p(X)} (s + R_-^r)^c. \quad (2.10)$$

Let us introduce the notation:

$$D_p = \bigcup_{s \in \text{MVar}_p(X)} (s + R_-^r). \quad (2.11)$$

The sets D_p, D_p^c are called favorable and unfavorable sets, respectively. As illustrated in the Figure 2.1, the unfavorable set D_p^c is the north east shaded region and the favorable set D_p is the south west unshaded region. Sometimes we write $D_p(X), D_p^c(X)$ to indicate the dependence on X .

Definition 2.2.2. (Prékopa (2012)) *The Multivariate Conditional Value-at-Risk, or MCVaR, of the random vector X , is defined as:*

$$\text{MCVaR}_p(X) = E(\psi(X) \mid X \in \overline{D_p^c}),$$

where ψ is some r -variate function such that $E(\psi(X))$ exists. The value $0 < p < 1$ (or $1 - p$) is called the level of MCVaR. The symbol $\overline{D_p^c}$ denotes the closure of D_p^c .

If the probability $P(X \in \text{MVar}_p(X))$ is negligible, then

$$E(\psi(X) \mid X \in \overline{D_p^c}) \approx E(\psi(X) \mid X \notin D_p).$$

Thus, we can write $\text{MCVaR}_p(X)$ as $E(\psi(X) \mid X \notin D_p)$ in case of $P(X \in \text{MVar}_p(X)) \approx 0$ (see Prékopa (2012)). Let us define the function $\psi(u)$ as

$$\psi(u) = \sum_{i=1}^r \lambda_i u_i, \quad (2.12)$$

where $\sum_{i=1}^r \lambda_i = 1$ and $\lambda_1, \dots, \lambda_r$ are nonnegative. If the components of X are losses in different portfolios, then λ_i weighs the loss in portfolio i , $i = 1, \dots, r$.

The following equation holds true:

$$E(\psi(X)) = E(\psi(X) \mid X \notin D_p)P(X \notin D_p) + E(\psi(X) \mid X \in D_p)P(X \in D_p), \quad (2.13)$$

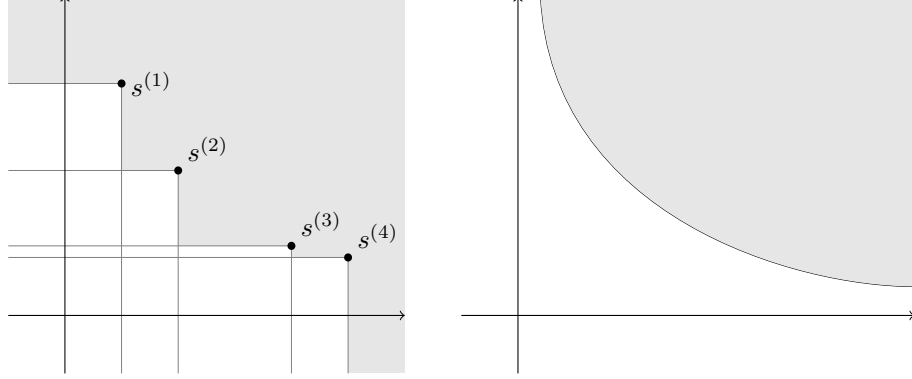


Figure 2.1: 2-D Illustration of the favorable set, and its complementary set where the Multivariate Conditional Value-at-Risk is defined in both types of a random vector – discrete and continuous.

X and Y are discrete and continuous random vectors, respectively. LHS: $\text{MVaR}_p(X) = \{s^{(1)}, \dots, s^{(4)}\}$, RHS: $\text{MVaR}_p(Y)$ is the boundary of the shaded region. MCVaR is defined in the shaded region (north east), i.e., the unfavorable set. The unshaded region (south west) is the favorable set.

from where we derive:

$$\text{MCVaR}_p(X) = E(\psi(X) \mid X \notin D_p) = \frac{1}{P(X \notin D_p)} \left(E(\psi(X)) - E(\psi(X) \mid X \in D_p)P(X \in D_p) \right). \quad (2.14)$$

Equation (3.8) can be written as:

$$\text{MCVaR}_p(X) = \frac{1}{1 - P(X \in D_p)} \left(\sum_{i=1}^r \lambda_i m_i - \sum_{i=1}^r \lambda_i E(X_i \mid X \in D_p)P(X \in D_p) \right), \quad (2.15)$$

where $m_i = E(X_i)$, $i = 1, \dots, r$.

While Definition 2.2.2 of MCVaR applies for the general case, a simpler definition can be given for the continuous case, as follows.

Definition 2.2.3. *The Multivariate Conditional Value-at-Risk, or MCVaR , of a continuous random vector $Z \in R^r$, is the value:*

$$\text{MCVaR}_p(Z) = E(\lambda^T Z \mid F_Z(Z) \geq p), \text{ where } F_Z(z) = P(Z \leq z).$$

Remark 1. *The Conditional Value-at-Risk (CVaR) measures the amount of losses beyond the Value-at-Risk (VaR). In the multivariate case, however, the elements of R^r are only partially ordered and MVaR is a set, not a single point, in general. On the other hand, we may not want to interpret the occurrence of an unfavorable event in terms*

of the sum of values of the r portfolios we are holding. In fact, a finance company typically creates a variety of portfolios, in an informative way, not only on their sum. The definition of MCVaR solves this problem in such a way that an unfavorable event is said to have occurred wherever X is larger than at least one element in MVaR, where “larger” means in the sense of partial order of vectors ($x > y$ iff $x \geq y, x \neq y$). The value of MCVaR is then the conditional expectation of the total loss, given that an unfavorable event occurs. In the total loss each asset has a multiplier $(\lambda_1, \dots, \lambda_r)$ to be able to value the different asset types on a common ground, by the use of a numeraire.

2.3 Properties of MVaR and MCVaR

The following definition and theorem about the multivariate stochastic ordering is well-known (see, e.g., Müller and Stoyan (2002)).

Definition 2.3.1. *Let X and Y be r -variate random vectors. Then we define*

- (SD1) *Stochastic dominance of order 1: $X \preceq_{(1)} Y$, if $Ef(X) \leq Ef(Y)$ for all bounded increasing functions $f : R^r \rightarrow R$.*
- (SD2) *Stochastic dominance of order 2: $X \preceq_{(2)} Y$, if $Ef(X) \leq Ef(Y)$ for all nondecreasing concave functions $f : R^r \rightarrow R$ such that the expectations exist.*

Theorem 2.3.1. *The following statements are equivalent.*

- (i) $X \preceq_{(1)} Y$,
- (ii) $P(X \in U) \leq P(Y \in U)$ for all upper sets U ,
- (iii) $P(X \in U) \leq P(Y \in U)$ for all closed upper sets U .

A set U is called an upper set if $x \in U$ implies $y \in U$ for every $y \geq x$. A set L is a lower set if $x \in L$ implies $y \in L$ for every $y \leq x$. Note that the Multivariate Conditional Value-at-Risk is defined on an upper set (both shaded sets in Figure 2.1).

The relationship between the Multivariate Value-at-Risk (MVaR) and multivariate stochastic dominance of order 1 is illustrated in Figure 2.2.

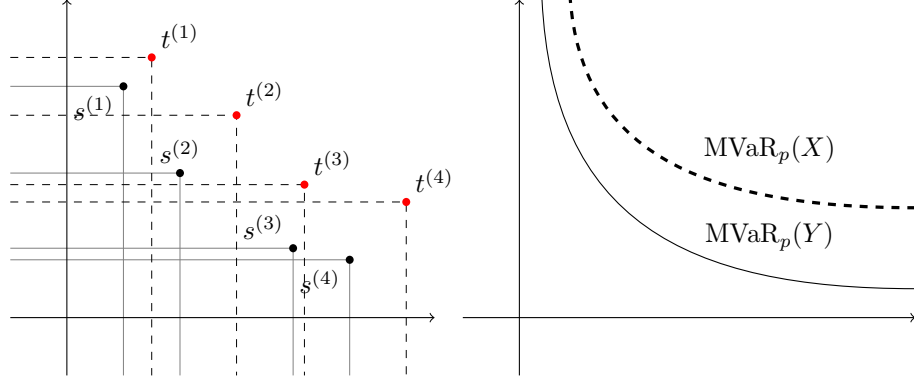


Figure 2.2: 2-D Illustration of the first order stochastic dominance of random vectors: $X \preceq_{(1)} Y$.

LHS: the case of discrete random vectors X and Y . $\text{MVaR}_p(X)$ is the collection of the p -efficient points $t^{(i)}$'s and $\text{MVaR}_p(Y)$ is the collection of the p -efficient points $s^{(i)}$'s. RHS: the case of continuous random vectors X and Y . $\text{MVaR}_p(X)$ is the hypersurface $\{t \mid F_X(t) = p\}$ and $\text{MVaR}_p(Y)$ is the hypersurface $\{s \mid F_Y(s) = p\}$.

We recall that the notions of logconcave p.d.f. and logconcave probability measure, designated by f and P , respectively, are defined by the inequalities

$$\begin{aligned} f(\lambda x + (1 - \lambda)y) &\geq [f(x)]^\lambda [f(y)]^{1-\lambda}, \\ P(\lambda A + (1 - \lambda)B) &\geq [P(A)]^\lambda [P(B)]^{1-\lambda}, \end{aligned} \tag{2.16}$$

where $x, y \in R^r$, $0 < \lambda < 1$, and A, B are convex subsets of R^r . Two basic theorems are as follows.

Theorem 2.3.2. (*Prékopa 1971, 1973b*) *If a probability measure is generated by a logconcave p.d.f., then it is a logconcave measure.*

Theorem 2.3.3. (*Prékopa 1973b*) *If the probability measure P is generated by a logconcave p.d.f. that is strictly logconcave in an open set $D \subset R^r$, then the c.d.f. is also strictly logconcave in D .*

It follows that if a multivariate p.d.f. is logconcave, then its c.d.f. is also logconcave for every $0 \leq p \leq 1$, and the set $\{z \mid F(z) \geq p\}$ is convex. If $0 < p < 1$, then the set $\{z \mid F(z) = p\}$ is an $r - 1$ dimensional hypersurface embedded in R^r which is illustrated by the boundary of the shaded set in Figure 2.3. If F is strictly logconcave, then it can also be described as a strictly concave function in which we take $r - 1$ (arbitrarily

chosen) variables as independent and one as dependent variable. For example, $x_r = x_r(x_1, \dots, x_{r-1})$. This idea is used for the calculation of MCVaR in case of continuous random vectors, presented in Section 5.41.

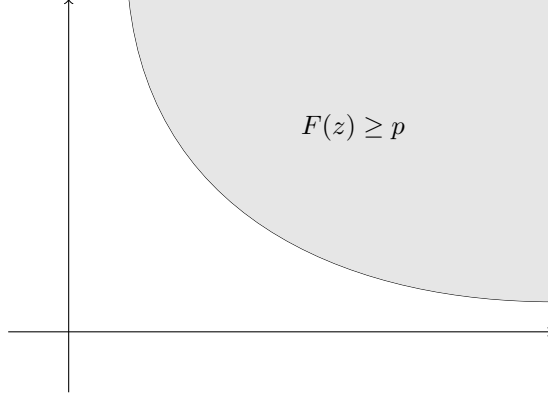


Figure 2.3: Illustration of the set $\{z \mid F(z) \geq p\}$.

The boundary of the shaded set is $\{z \mid F(z) = p\}$, i.e., an $r-1$ dimensional hypersurface embedded in R^r .

Consider the family of sets:

$$H(p) = \{z \mid F(z) - p \geq 0\}, \quad (2.17)$$

depending on the parameter p ($0 < p < 1$). For every fixed p , the set $H(p)$ is convex but now we want to consider $F(z) - p$ as a function of all variables in z and p . Since $F(z) - p$ is not a logconcave function of z, p , in general, we change the parameter and look at the family of sets:

$$K(u) = \{z \mid F(z) \geq e^u\}. \quad (2.18)$$

If $F(z) > 0, z \in R^r$, then $K(u) = \{z \mid \log F(z) - u \geq 0\}$. For any $-\infty < u < 0$, we have $K(u) \neq \emptyset$. We have the following

Theorem 2.3.4. *$K(u)$, $-\infty < u < 0$ is a concave family of sets, i.e., if u_1, u_2 are arbitrary negative numbers and $0 < \lambda < 1$, then $K(\lambda u_1 + (1 - \lambda)u_2) \supset \lambda K(u_1) + (1 - \lambda)K(u_2)$.*

Proof. Let $z_1 \in K(u_1)$, $z_2 \in K(u_2)$. Then

$$\begin{aligned} F(\lambda z_1 + (1 - \lambda)z_2) &\geq (F(z_1))^\lambda (F(z_2))^{1-\lambda} \\ &\geq (e^{u_1})^\lambda (e^{u_2})^{1-\lambda} = e^{\lambda u_1 + (1-\lambda)u_2}, \end{aligned}$$

which proves that

$$\lambda z_1 + (1 - \lambda)z_2 \in K(\lambda u_1 + (1 - \lambda)u_2). \quad \square$$

□

Corollary 2.3.5. *Let G be any convex subset of R^r and $K(u)$, $-\infty < u < 0$ the concave family of sets defined in (2.18) with a logconcave distribution function F . Then $K(u) \cap G$, $-\infty < u < 0$ is a concave family of sets.*

Theorem 2.3.2 implies

Theorem 2.3.6. *Let $f(z)$, $z \in R^r$ be any logconcave function. Then*

$$\int_{K(u)} f(z) dz$$

is a logconcave function of $u \in (-\infty, 0)$. In other words, the function

$$\int_{F(z) \geq p} f(z) dz$$

is logconcave in $\log p$.

Definition 2.3.2. *Two random variables X and Y defined on the same probability space (Ω, \mathcal{F}, P) are said to be comonotone, if for all $\omega_1, \omega_2 \in \Omega$,*

$$[X(\omega_1) - Y(\omega_1)][X(\omega_2) - Y(\omega_2)] \geq 0 \quad a.s.$$

We are now ready to state properties of $MVaR_p$ and $MCVaR_p$.

Theorem 2.3.7. *Let $X, Y \in R^r$ be random vectors on the same probability space. Then we have the following properties:*

- (1) *$MVaR_p$ is translation-equivariant: $MVaR_p(X+c) = MVaR_p(X)+c$, where $c \in R^r$.*
- (2) *$MVaR_p$ is positively homogeneous: $MVaR_p(cX) = cMVaR_p(X)$, where $c \in R_+$.*
- (3) *$\{z \mid P(-X \geq z) \geq p \text{ and it does not hold for any } y \geq z, y \neq z\} = -MVaR_p(X)$.*
- (4) *$MVaR_p$ is monotonic w.r.t. the first order stochastic dominance, i.e.:*

$$X \preceq_{(1)} Y \text{ implies } D_p(X) \supset D_p(Y)$$

- (5) If X_i and Y_i are comonotone, X, Y have independent components, continuous and increasing distribution functions, then

$$MVaR_p(X + Y) = \left\{ \left(\begin{array}{c} VaR_{\alpha_1}(X_1) \\ \vdots \\ VaR_{\alpha_r}(X_r) \end{array} \right) + \left(\begin{array}{c} VaR_{\alpha_1}(Y_1) \\ \vdots \\ VaR_{\alpha_r}(Y_r) \end{array} \right), \alpha_1 \dots \alpha_r = p, 0 < \alpha_i < 1, i = 1, \dots, r \right\}.$$

- (6) For any X and $0 < p < 1$, $MVaR_p(X)$ is bounded from below.

Proof. The proofs of (1), (2) and (4) are simple and therefore omitted.

(3)

$$\begin{aligned} & \{z \mid P(-X \geq z) \geq p \text{ and there is no } y \geq z, y \neq z \text{ such that } P(-X \geq y) \geq p\} \\ &= \{z \mid P(X \leq -z) \geq p \text{ and there is no } y \leq -z, y \neq -z \text{ such that } P(X \leq y) \geq p\} \\ &= -MVaR_p(X). \end{aligned} \tag{2.19}$$

It can also be written as

$$\begin{aligned} -MVaR_p(X) &= \{z \mid P(-X_i < z_i, \text{ for at least one } i = 1, \dots, r) < 1 - p \\ &\quad \text{and there is no } y \geq z, y \neq z \text{ such that} \\ &\quad P(-X_i < y_i, \text{ for at least one } i = 1, \dots, r) < 1 - p\}. \end{aligned} \tag{2.20}$$

- (5) If X has independent components and continuous and increasing distribution functions, then

$$MVaR_p(X) = \{u \mid u_i = VaR_{\alpha_i}(X_i), \alpha_1 \dots \alpha_r = p, 0 < \alpha_i < 1, i = 1, \dots, r\}.$$

We have the equation

$$MVaR_p(X + Y) = \{u \mid F_{X_i+Y_i}(u_i) = \alpha_i, \alpha_1 \dots \alpha_r = p, 0 < \alpha_i < 1, i = 1, \dots, r\}.$$

Since X_i and Y_i are comonotone, $i = 1, \dots, r$, it follows that

$$u_i = VaR_{\alpha_i}(X_i) + VaR_{\alpha_i}(Y_i), i = 1, \dots, r$$

which implies (5).

- (6) For every $z \in R^r$ we have the inequality $F_i(z_i) \geq F(z_1, \dots, z_r)$, hence $F(z_1, \dots, z_r) \geq p$ implies that $F_i(z_i) \geq p$. Since $F(VaR_p(X_1), \dots, VaR_p(X_r)) \geq p$, it follows that

$$F_i(VaR_p(X_i)) \geq p \text{ and } VaR_p(X_i) \geq F_i^{-1}(p), i = 1, \dots, r. \quad \square$$

□

Theorem 2.3.8. *Let $X, Y \in R^r$ be r -component random variables with finite expectations. Then $MCVaR_p$ exhibits the following properties:*

- (1) *$MCVaR_p$ is translation-equivariant: $MCVaR_p(X + c) = MCVaR_p(X) + c$.*
- (2) *$MCVaR_p$ is positively homogeneous: $MCVaR_p(cY) = cMCVaR_p(Y)$, $c \in R_+$.*
- (3) *$MCVaR_p$ is subadditive when X, Y are continuously distributed and all components in X and Y are independent, i.e., we have the inequality*

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

- (4) *If the components of $X = (X_1, \dots, X_r)$ are independent and have continuous distributions with logconcave p.d.f.'s, then $MCVaR_p(X)$ is logconcave in $\log p$, for $p \geq p_0$, where p_0 is a probability ($0 < p_0 < 1$) such that $VaR_p(X_i) \geq 0$, $i = 1, \dots, r$.*

Remark 2. *We can think, from Property (3) of $MCVaR$, about both cases of “good” and “bad” corporate M&A (Mergers and Acquisitions) deals. From a risk management perspective, Property (3) indicates that not all M&A deals would be successful, i.e., for some “bad” M&A deals, risk would not be reduced, since $MCVaR_p$ is not always subadditive. More detailed explanation is presented in Remark 5.*

Proof. Let us recall the following equation:

$$\begin{aligned} MCVaR_p(X) &= E(\psi(X) \mid X \notin D_p) \\ &= \frac{1}{1 - P(X \in D_p)} \left(\sum_{i=1}^r \lambda_i m_i - \sum_{i=1}^r \lambda_i E(X_i \mid X \in D_p) P(X \in D_p) \right), \end{aligned} \tag{2.21}$$

where $m_i = E(X_i)$ for $i = 1, \dots, r$, $\psi(X) = \sum_{i=1}^r \lambda_i X_i$ and $\sum_{i=1}^r \lambda_i = 1$, $\lambda_i \geq 0$ for $i = 1, \dots, r$.

- (1) Let $D'_p = \bigcup_{s' \in \text{MVaR}_p(X+c)} (s' + R_-^r)$, $D_p = \bigcup_{s \in \text{MVaR}_p(X)} (s + R_-^r)$. Then we have the equations:

$$\begin{aligned}
\text{MCVaR}_p(X+c) &= E(\psi(X+c) \mid X+c \notin D'_p) \\
&= E(\psi(X) \mid X+c \notin D'_p) + \sum_{i=1}^r \lambda_i c_i \\
&= E(\psi(X) \mid X \notin D_p) + \sum_{i=1}^r \lambda_i c_i \\
&= \text{MCVaR}_p(X) + c.
\end{aligned} \tag{2.22}$$

- (2) If we use the notations D_p, D'_p with $s = \frac{s'}{c}$, then we derive:

$$\begin{aligned}
\text{MCVaR}_p(cX) &= E(\psi(cX) \mid cX \notin D'_p) \\
&= E(c\psi(X) \mid cX \notin D'_p) \\
&= cE(\psi(X) \mid X \notin D_p) \\
&= c\text{MCVaR}_p(X).
\end{aligned} \tag{2.23}$$

The third equality holds since it can easily be seen that $cX \notin D'_p$ is equivalent to $X \notin D_p$, where D'_p and D_p are defined as above. If we use the second property of MVaR then we can obtain the following equations:

$$\begin{aligned}
cX \notin D'_p &= \bigcup_{s' \in \text{MVaR}_p(cX)} (s' + R_-^r) \\
\Leftrightarrow cX \notin D'_p &= \bigcup_{s' \in c\text{MVaR}_p(X)} (s' + R_-^r) \\
\Leftrightarrow X \notin \bigcup_{\frac{s'}{c} \in \text{MVaR}_p(X)} (\frac{s'}{c} + R_-^r) \\
\Leftrightarrow X \notin D_p &= \bigcup_{s \in \text{MVaR}_p(X)} (s + R_-^r), \quad s = \frac{s'}{c}.
\end{aligned} \tag{2.24}$$

- (3) First we remark that if $Z = (Z_1, \dots, Z_r)$ is a continuously distributed random vector and Z_1, \dots, Z_r are independent, then $P(Z \in D_p)$ is independent of the distribution of Z . In fact, let F_i be the c.d.f. of Z_i , $i = 1, \dots, r$. Then we have

$$\begin{aligned}
P(Z \in D_p) &= P(F_1(Z_1) \cdots F_r(Z_r) \geq p) = P(U_1 \cdots U_r \geq p) \\
&= P(-\log U_1 - \cdots - \log U_r \leq -\log p) = \int_0^{-\log p} \frac{z^{r-1} e^{-z}}{(r-1)!} dz,
\end{aligned} \tag{2.25}$$

where U_1, \dots, U_r are independent random variables, uniformly distributed in $(0,1)$. Hence, the numerator counts in the second equation of (2.15), only. We may disregard the linear terms and it is enough to look only at $E(X_1 \mid X \notin D_X), E(X_2 \mid X \notin D_p^X), E(Y_1 \mid Y \notin D_p^Y), E(Y_2 \mid Y \notin D_p^Y), E(X_1 + Y_1 \mid X + Y \notin D_p^{X+Y}), E(X_2 + Y_2 \mid X + Y \notin D_p^{X+Y})$, where

$$D_p^X = \bigcup_{s \in \text{MCVaR}_p(X)} (s + R_-^r), \quad D_p^Y = \bigcup_{t \in \text{MCVaR}_p(Y)} (t + R_-^r), \quad D_p^{X+Y} = \bigcup_{u \in \text{MCVaR}_p(X+Y)} (u + R_-^r).$$

It is enough to prove that

$$E(X_1 + Y_1 \mid X + Y \notin D_{X+Y}) \leq E(X_1 \mid X \notin D_X) + E(Y_1 \mid Y \notin D_Y), \quad (2.26)$$

$$E(X_2 + Y_2 \mid X + Y \notin D_{X+Y}) \leq E(X_2 \mid X \notin D_X) + E(Y_2 \mid Y \notin D_Y). \quad (2.27)$$

If we multiply the inequalities (2.26), (2.27) by -1 and add them, then by (2.15) and the fact that $P(X \notin D_X)$ is independent of the random variable, the convexity proof of MCVaR will be complete.

Proof. Proof of (2.26) (proof of (2.27) is the same):

$$E(X_1 + Y_1 \mid X + Y \notin D_{X+Y}) = E(X_1 + Y_1 \mid F_{X_1+Y_1}(X_1 + Y_1) F_{X_2+Y_2}(X_2 + Y_2) \geq p). \quad (2.28)$$

Note that $X_1 + Y_1$ and $X_2 + Y_2$ are independent

$$\begin{aligned} F_{X_1+Y_1}(X_1 + Y_1) &\sim U_1, \\ F_{X_2+Y_2}(X_2 + Y_2) &\sim U_2. \end{aligned} \quad (2.29)$$

It follows that (2.28) is further equal to

$$E(X_1 + Y_1 \mid F_{X_1+Y_1}(X_1 + Y_1) \geq \frac{p}{U_2}). \quad (2.30)$$

Incidentally we mention that if V is any random variable, then

$$E(V \mid F_V(V) \geq q) = \frac{\int_q^\infty [1 - G(v)] dv}{1 - G(q)} + q, \quad (2.31)$$

where G is c.d.f. of V , and this, as a function of q is increasing.

Let $\delta > 0$ and introduce the notations:

$$A = \int_{q+\delta}^{\infty} [1 - G(v)] dv, \quad B = 1 - G(q + \delta).$$

Then we have:

$$\begin{aligned}
& \frac{\int_{q+\delta}^{\infty} [1 - G(v)] dv}{1 - G(q + \delta)} + q + \delta - \frac{\int_q^{\infty} [1 - G(v)] dv}{1 - G(q)} - q \\
&= \delta + \frac{A(B + G(q + \delta) - G(q)) - \left(A + \int_q^{q+\delta} (1 - G(v)) dv \right) B}{(1 - G(q + \delta))(1 - G(q))} \\
&= \delta + \frac{A(G(q + \delta) - G(q)) - \int_q^{q+\delta} (1 - G(v)) dv}{(1 - G(q + \delta))(1 - G(q))} B \\
&= \delta + \frac{A}{1 - G(q + \delta)} - \frac{A}{1 - G(q)} - \frac{\int_q^{q+\delta} (1 - G(v)) dv}{1 - G(q)} \tag{2.32} \\
&\geq \delta - \frac{\int_q^{q+\delta} (1 - G(v)) dv}{1 - G(q)} \\
&\geq \delta - \frac{\delta(1 - G(q))}{1 - G(q)} \\
&= 0.
\end{aligned}$$

This implies that

$$E(X_1 + Y_1 \mid F_{X_1+Y_1}(X_1 + Y_1) \geq \frac{p}{U_2}) \leq E(X_1 + Y_1 \mid F_{X_1+Y_1}(X_1 + Y_1) \geq p). \tag{2.33}$$

□

In the same way, we have

$$E(X_2 + Y_2 \mid F_{X_2+Y_2}(X_2 + Y_2) \geq \frac{p}{U_1}) \leq E(X_2 + Y_2 \mid F_{X_2+Y_2}(X_2 + Y_2) \geq p) \tag{2.34}$$

and the assertion is proved.

A simple counterexample of Property (3) in the general case

Suppose that the random vectors $X, Y \in R^2$ have the following possible values with probability 0.25 for each point:

$$X = \{(1, 1), (4, 4)\}^T, (2, 1)^T, (2, 8)^T, (8, 4)^T\} \text{ and } Y = \{(1, 1)^T, (2, 2)^T, (3, 3)^T, (4, 4)^T\},$$

as depicted in Figure 2.4.

At $p = 0.75$, $\text{MVaR}_p(X) = \{(2, 8)\}$ and $\text{MVaR}_p(Y) = \{(3, 3)\}$. $E(X_1) = 3.275$, $E(X_2) = 4.35$ and $E(X_1 \mathbb{1}_{X \in D_p}) = 1.275$, $E(X_2 \mathbb{1}_{X \in D_p}) = 3.35$. Let $\lambda_1 = \lambda_2 = 0.5$. Plugging in those values into MCVaR formulation (2.15), $\text{MCVaR}_p(X) = 6$. For the random vector Y , $\text{MCVaR}_p(Y) = 4$ using $E(Y_1) = E(Y_2) = 2.5$, $E(Y_1 \mathbb{1}_{Y \in D_p}) = E(Y_2 \mathbb{1}_{Y \in D_p}) = 1.5$ and $\lambda_1 = \lambda_2 = 0.5$. Then $\text{MCVaR}_p(X) + \text{MCVaR}_p(Y) = 10$.

$$\text{Let } Z = X+Y. \text{ Then } Z_1 = \begin{cases} 2.1 & \text{with } p = 0.25^2 \\ 3 & \text{with } p = (0.5)(0.25) \\ 3.1 & \text{with } p = 0.25^2 \\ 4 & \text{with } p = (0.5)(0.25) \\ 4.1 & \text{with } p = 0.25^2 \\ 5 & \text{with } p = (0.5)(0.25) \\ 5.1 & \text{with } p = 0.25^2 \\ 6 & \text{with } p = (0.5)(0.25) \\ 9 & \text{with } p = 0.25^2 \\ 10 & \text{with } p = 0.25^2 \\ 11 & \text{with } p = 0.25^2 \\ 12 & \text{with } p = 0.25^2 \end{cases}, Z_2 = \begin{cases} 2 & \text{with } p = 0.25^2 \\ 3 & \text{with } p = 0.25^2 \\ 4 & \text{with } p = 0.25^2 \\ 5 & \text{with } p = 2 \times 0.25^2 \\ 5.4 & \text{with } p = 0.25^2 \\ 6 & \text{with } p = 0.25^2 \\ 6.4 & \text{with } p = 0.25^2 \\ 7 & \text{with } p = 0.25^2 \\ 7.4 & \text{with } p = 0.25^2 \\ 8 & \text{with } p = 0.25^2 \\ 8.4 & \text{with } p = 0.25^2 \\ 9 & \text{with } p = 0.25^2 \\ 10 & \text{with } p = 0.25^2 \\ 11 & \text{with } p = 0.25^2 \\ 12 & \text{with } p = 0.25^2 \end{cases}.$$

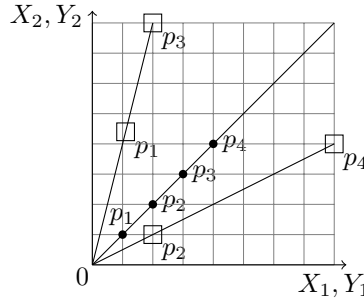


Figure 2.4: A counterexample of MCVaR Property (3): subadditivity, in the general case.

Each node of p_1, p_2, p_3, p_4 has probability 0.25. The length of each grid element is 1. The rectangle-nodes are the possible values of X and the circle-nodes are that of Y ; $X = \{(1, 1), (2, 1), (2, 8), (8, 4)\}^T$ and $Y = \{(1, 1), (2, 2), (3, 3), (4, 4)\}^T$. At the probability level $p = 0.75$, $\text{MCVaR}_p(X) = 6$ and $\text{MCVaR}_p(Y) = 4$.

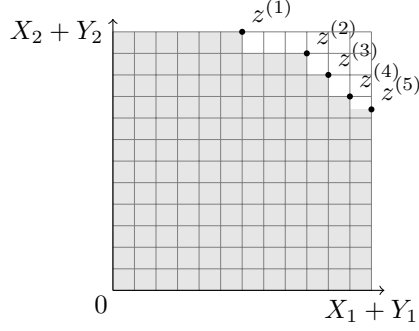


Figure 2.5: Illustration of $\text{MVaR}_p(\mathbf{Z})$, where $\mathbf{Z} = \mathbf{X} + \mathbf{Y}$.

The length of each grid element is 1. Let $Z = X + Y$, $Z \in R^2$. The points $z^{(i)}$, $i = 1, \dots, 5$ are the elements of $\text{MVaR}_p(Z)$, i.e., $\text{MVaR}_p(Z) = \{(6, 12), (9, 11), (10, 10), (11, 9), (12, 8.4)\}$. Under the probability level $p = 0.75$, $\text{MCVaR}_p(Z) = 11$.

At $p = 0.75$, we have $\text{MVaR}_p(Z) = \{(6, 12), (9, 11), (10, 10), (11, 9), (12, 8.4)\}$, since $F_Z(6, 12) = F_Z(12, 8.4) = 0.75$, $F_Z(9, 11) = F_Z(11, 9) = 0.7617875$ and $F_Z(10, 10) = 0.765625$ as described in Figure 2.5. From a simple calculation, we get $P(Z \in D_p) = 0.9609375$, $E(Z_1) = 5.775$, $E(Z_2) = 6.85$, $E(Z_1 \mathbb{1}_{Z \in D_p}) = 5.3453125$, $E(Z_2 \mathbb{1}_{Z \in D_p}) = 6.4203125$ and let $\lambda_1 = \lambda_2 = 0.5$. Plugging in those values into MCVaR formulation (2.15), we obtain $\text{MCVaR}_p(Z) = 11$. Thus, $\text{MCVaR}_p(Z) = 11 > 10 = \text{MCVaR}_p(X) + \text{MCVaR}_p(Y)$ and this is the counterexample of Property(3) of MCVaR in the general case.

(4) Since $P(X \in D_p)$ does not depend on the distribution of X , it is enough to prove that

$$\int_{F_1(z_1) \cdots F_r(z_r) \geq p} \psi(z) f(z) dz \quad (2.35)$$

is logconcave in $\log p$, for $p \geq p_0$. Since $\lambda_i > 0$, $i = 1, \dots, r$, the function $\psi(z)$ and also $\psi(z)f(z)$ is logconcave in $\{z \mid z \geq 0\}$. On the other hand, if $p \geq p_0$, then

$$\{z \mid F_1(z_1) \cdots F_r(z_r) \geq p\} \subset \{z \mid z \geq (\text{VaR}_p(X_1), \dots, \text{VaR}_p(X_r))\} \subset \{z \mid z \geq 0\}. \quad (2.36)$$

The rest of the proof is the same as the proof of Theorem 2.3.4. \square

\square

Remark 3. (*Relationship between VaR and MVaR*) Let us define

$$\left\{ \left\{ \begin{pmatrix} z_1 \\ \vdots \\ z_r \end{pmatrix} \mid P(\xi_1 \leq z_1, \dots, \xi_r \leq z_r) \geq p \right\} = \overline{D_p^c}(\xi) \subset R^r, \xi = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_r \end{pmatrix}, \right. \\ \left. \left\{ \begin{pmatrix} z_1 \\ \vdots \\ z_k \end{pmatrix} \mid P(\xi_1 \leq z_1, \dots, \xi_k \leq z_k) \geq p \right\} = \overline{D_p^c}(\eta) \subset R^k, \eta = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_k \end{pmatrix}, k < r. \right. \quad (2.37)$$

If we create a cylinder set out of $\overline{D_p^c}(\eta)$ in such a way that we take

$$\left\{ \left\{ \begin{pmatrix} z_1 \\ \vdots \\ z_r \end{pmatrix} \mid \begin{pmatrix} z_1 \\ \vdots \\ z_k \end{pmatrix} \in \overline{D_p^c}(\eta) \right\} = \overline{D_p^{rc}}(\eta), k < r, \right. \quad (2.38)$$

then we have the relation

$$\overline{D_p^c}(\xi) \subset \overline{D_p^{rc}}(\eta). \quad (2.39)$$

It is true that the projection of $MVaR_p(X)$ to a space of a smaller number of components of X , i.e., as before in (2.37), from R^r to R^k , then the lower bound of the projection in R^k of $MVaR_p(X)$ is equal to $MVaR_p((X_1, \dots, X_k)^T)$. The projection of $MVaR$ is illustrated in Figures 3.4 and 2.7. In Figure 3.4, for a random vector $X \in R^2$, the sets $\{z \mid z \geq VaR_p(X_1)\}$ and $\{z \mid z \geq VaR_p(X_2)\}$ are closures of the projections of $MVaR_p(X)$ onto the horizontal and vertical axes, respectively. The same sets are the closures of the projections of $D_p^c(X)$. In Figure 2.7, for a random vector $X \in R^3$, we illustrate a boundary of the set $\{(z_1, z_2, z_3)^T \mid (z_1, z_2)^T \geq MVaR_p(X_1, X_2)\}$. The set $\{z \in R^2 \mid z \geq MVaR_p(X_1, X_2)\}$ is the closure of the projection of $D_p^c(X)$ onto (X_1, X_2) -plane and also the closure of the projection of $MVaR_p(X)$ onto the same plane.

Let a random vector $X \in R^r$ denote losses from investment in a composite of portfolios, where each component X_i , $i = 1, \dots, r$ is just a single portfolio. Then we know that the unfavorable set $D_p^c(X)$ is only a part of $\{z \in R^r \mid z_i \geq VaR_p(X_i), i = 1, \dots, r\}$. It has practical meaning: if we have several portfolios put together in the random vector X , then a single portfolio may signal an unfavorable event, i.e., the realized value of X_i

is greater than $\text{VaR}_p(X_i)$, while X is still in the favorable set $D_p(X)$, i.e., there is no such signal for X , the composite of individual portfolios.

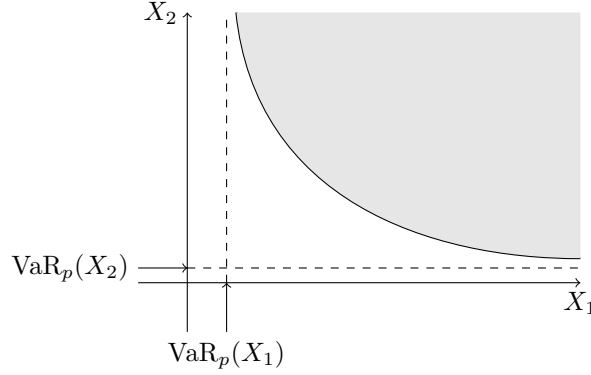


Figure 2.6: Projection of $M\text{VaR}_p((\mathbf{X}_1, \mathbf{X}_2)^T)$ from \mathbf{R}^2 onto the space of $\mathbf{X}_1, \mathbf{X}_2 \in \mathbf{R}$. The points $\text{VaR}_p(X_1)$ and $\text{VaR}_p(X_2)$ are the lower bounds of the projection of $M\text{VaR}_p(X)$, $X = (X_1, X_2)^T \in \mathbf{R}^2$, onto the space of $X_1, X_2 \in \mathbf{R}$ respectively.

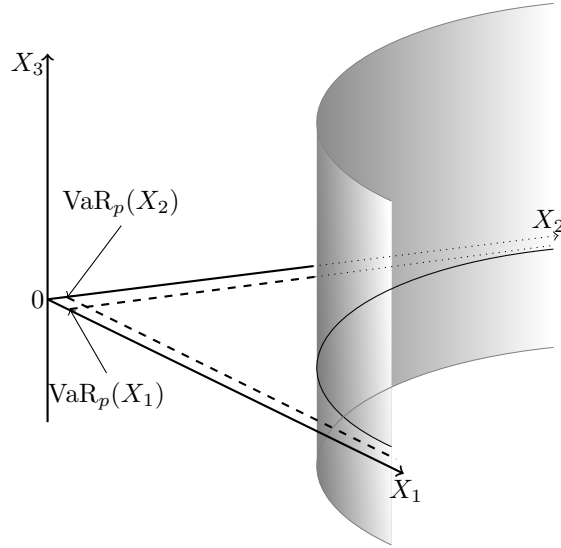


Figure 2.7: Projection of $M\text{VaR}_p((\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3)^T)$ from \mathbf{R}^3 onto the space of $(\mathbf{X}_1, \mathbf{X}_2)^T \in \mathbf{R}^2$.

The shaded surface is the $M\text{VaR}_p((X_1, X_2)^T)$, which is a cylinder set of the lower bound of the projection of $M\text{VaR}_p(X)$, $X = (X_1, X_2, X_3)^T \in \mathbf{R}^3$, onto the space of $(X_1, X_2)^T \in \mathbf{R}^2$.

2.4 Calculation of MCVaR

2.4.1 The Case of a Continuous Distribution

Assume that a random vector Z has continuous distribution, its p.d.f. and c.d.f. are $f_Z(z)$ and $F_Z(z)$, respectively, where $Z = (z_1, \dots, z_r)^T$. Then

$$\begin{aligned} \text{MCVaR}_p(Z) &= E(\lambda^T Z \mid F_Z(Z) \geq p) \\ &= \frac{\int \dots \int_{D_p^c} (\lambda_1 z_1 + \dots + \lambda_r z_r) f_Z(z_1, \dots, z_r) dz_1 \dots dz_r}{\int \dots \int_{D_p^c} f_Z(z_1, \dots, z_r) dz_1 \dots dz_r}, \end{aligned} \quad (2.40)$$

where $D_p^c = \{z \mid F_Z(z) \geq p\}$.

If $f_Z(z)$ is a log concave function, then so is $F_Z(z)$ (see Prékopa (1995)) and the set D_p^c is convex. Its boundary is a convex surface that can be represented in the form of a function provided that no coordinate axis is a supporting line of D_p^c . In this case we can take any $r - 1$ variables out of z_1, \dots, z_r , the remaining variable will be a function of them and this function uniquely describes the surface of D_p^c (see Busemann (2008)). The set D_p^c can then be represented as

$$D_p^c = \{(z_1, \dots, z_r) : \text{VaR}_p(Z_1) < z_1, l_1(z_1) < z_2, l_2(z_1, z_2) < z_3, \dots, l_{r-1}(z_1, \dots, z_{r-1}) < z_r\}, \quad (2.41)$$

where $l_k(z_1, \dots, z_k)$'s are k -variate lower bound functions, $k = 1, \dots, r - 1$ and $l_1(z_1)$ is continuous on the domain $\{\text{VaR}_p(Z_1) < z_1\}$, $l_2(z_1, z_2)$ is continuous on the 2-dimensional domain $\{\text{VaR}_p(Z_1) < z_1, l_1(z_1) < z_2\}$, \dots , $l_{r-1}(z_1, z_2, \dots, z_{r-1})$ is continuous on the " $r-1$ "-dimensional domain $\{\text{VaR}_p(Z_1) < z_1, l_1(z_1) < z_2, l_2(z_1, z_2) < z_3, \dots, l_{r-2}(z_1, z_2, \dots, z_{r-2}) < z_{r-1}\}$. Then (2.40), together with (2.41) can be written as

$$\text{MCVaR}_p(Z) = \frac{\int_{\text{VaR}_p(Z_1)}^\infty \int_{l_1}^\infty \dots \int_{l_{r-2}}^\infty \int_{l_{r-1}}^\infty (\lambda_1 z_1 + \dots + \lambda_r z_r) f_Z(z_1, \dots, z_r) dz_r dz_{r-1} \dots dz_2 dz_1}{\int_{\text{VaR}_p(Z_1)}^\infty \int_{l_1}^\infty \dots \int_{l_{r-2}}^\infty \int_{l_{r-1}}^\infty f_Z(z_1, \dots, z_r) dz_r dz_{r-1} \dots dz_2 dz_1}. \quad (2.42)$$

Each lower bound of the integrals in (2.42) represents MVaR_p on its corresponding multidimensional space, i.e.,

$$\begin{aligned} \{(z_1, \dots, z_r)^T \mid z_r = l_{r-1}(z_1, \dots, z_{r-1})\} &= \text{MVaR}_p((Z_1, \dots, Z_r)^T), \\ \{(z_1, \dots, z_{r-1})^T \mid z_{r-1} = l_{r-2}(z_1, \dots, z_{r-2})\} &= \text{MVaR}_p((Z_1, \dots, Z_{r-1})^T), \\ &\vdots \\ \{(z_1, z_2)^T \mid z_2 = l_1(z_1)\} &= \text{MVaR}_p((Z_1, Z_2)^T). \end{aligned} \quad (2.43)$$

Generally, there is no closed form of the quantile function for a multivariate distribution. Thus, the functions of lower bounds in the set D_p^c of (2.41), i.e., the lower limits of the integrals of (2.42) can be constructed by some numerical methods, e.g., multivariate nonlinear approximation. For various methods of multivariate function fitting, we refer the readers to related books and literature (see, e.g., (Atkinson, 1988), (Gasca and Sauer, 2000), (Sauerbrei et al, 2006), (Strang, 2007), etc.).

We generate a multidimensional grid of equally spaced points in the following set:

$$\{(z_1, \dots, z_r)^T : z_k \in I_k \text{ for } k = 1, \dots, r\}, \text{ where } I_k = [\text{VaR}_p(Z_k), x \text{ such that } F_{Z_k}(x) \approx 1]. \quad (2.44)$$

Then we generate a collection of the closest points to the $\text{MVaR}_p(Z) = \{z \mid F_Z(z) = p\}$. By a nonlinear approximation based on the collection of such points, functions of lower bounds in the set D_p^c of (2.41) can be constructed. This is followed by (2.42), the calculation of MCVaR.

Two numerical examples of recent finance market data are presented. The type of financial securities is exchange-traded funds (ETFs), which can be regarded as mutual funds that can be bought and sold just like common stocks, i.e. exchange-traded products. We use 6 months of time period from February 15, 2012 to August 14, 2012 for the calculation of MCVaR with probability levels $p = 0.80$, $p = 0.90$, $p = 0.95$ and $p = 0.99$. From Yahoo Finance, online finance portal, we download the data of daily closing prices for the time period of 6 months from February 15, 2012 to August 14, 2012.

We want to show, by Examples 1 and 2, how MCVaR works on a set of “stochastically dependent” random variables. In Example 1 we have two “positively” correlated funds: Fidelity Nasdaq Composite Index Tracking (ONEQ) ETF and Vanguard S&P 500 (VOO) ETF that closely resemble two of the US major indices, Nasdaq and S&P 500, respectively. For Example 2, we select Deutsche Bank US Dollar Index Bullish (UUP) ETF which tracks the performance of the Deutsche Bank Long US Dollar Futures index, while keeping Fidelity Nasdaq Composite Index Tracking (ONEQ) ETF, in order to see how MCVaR measures a risk on the set of index funds “negatively” correlated with each other.

Example 1 (Positively correlated two ETFs). We choose two ETFs: Fidelity Nasdaq Composite Index Tracking (ONEQ) ETF and Vanguard S&P 500 (VOO) ETF, “positively” correlated, as in Figure 2.8.



Figure 2.8: Basic Chart of Fidelity Nasdaq ETF and Vanguard S&P 500 ETF. Over 6 months from February 15 2012 to August 14 2012, it is a basic chart of Fidelity Nasdaq Composite Index Tracking (ONEQ) ETF and Vanguard S&P 500 (VOO) ETF from Yahoo Finance.

Initial prices per share on February 15, 2012 are \$115.02 for ONEQ, \$61.56 for VOO. Assume that we have total available amount of \$1,000,000 for the investment which is intended for the equal investment in each kind, i.e. \$500,000 each. However, since there is no fractional shares for those securities in real financial market, the initial investment is \$499,991.94 and \$499,990.32 for ONEQ and VOO, respectively. The corresponding number of shares is 4,347 for ONEQ and 8,122 for VOO. We manipulate the data of daily closing prices into daily losses from the following equation:

$$\text{Loss} = \frac{\text{initial investment} - (\text{number of shares} \times \text{price per share})}{\text{initial investment}}. \quad (2.45)$$

Let X_1 , X_2 denote the random variables of losses from an investment in Fidelity Nasdaq Composite Index Tracking (ONEQ) ETF and Vanguard S&P 500 (VOO) ETF, respectively. Each is assumed normally distributed. The random vector $X = (X_1, X_2)^T$ has a bivariate normal distribution with parameters $E(X_1) = \mu_{x_1} = -0.01185$, $\sigma_{x_1} =$

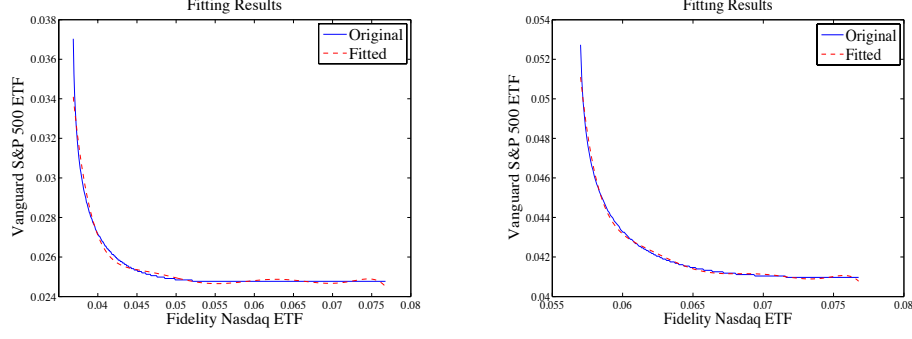


Figure 2.9: Multivariate Value-at-Risk, the quantile function at the probability levels $p = 0.95$ and $p = 0.99$.

At probability level $p = 0.95$ for LHS and $p = 0.99$ for RHS, over 6 months of time period, $MVaR_p(X)$ is well approximated by polynomial fitting in the sense of the least-squares (dotted curve). The random vector has components of losses from the investment in Fidelity Nasdaq Composite Index Tracking (ONEQ) ETF and Vanguard S&P 500 (VOO) ETF from Yahoo Finance. The piecewise linear line (looks like a curve) is the set of line segments between the points (x_1, x_2) such that $F_X(x_1, x_2) = p$.

0.02956, $E(X_2) = \mu_{x_2} = -0.01439$, $\sigma_{x_2} = 0.02477$ and $\rho = 0.95139$. We calculate

$$\begin{aligned}
 MCVaR_p(X) &= E(\lambda^T X \mid F_X(X) \geq 0.95), \text{ where } F_X(x) = P(X \leq x) \\
 &\approx \frac{\int_{VaR_p(X_1)}^{\infty} \int_{l(x_1)}^{\infty} (\lambda_1 x_1 + \lambda_2 x_2) f_X(x_1, x_2) dx_2 dx_1}{\int_{VaR_p(X_1)}^{\infty} \int_{l(x_1)}^{\infty} f_X(x_1, x_2) dx_2 dx_1} \\
 &\approx \frac{\int_{VaR_p(X_1)}^{\mu_{x_1} + 7\sigma_{x_1}} \int_{l(x_1)}^{\mu_{x_2} + 7\sigma_{x_2}} (\lambda_1 x_1 + \lambda_2 x_2) f_X(x_1, x_2) dx_2 dx_1}{\int_{VaR_p(X_1)}^{\mu_{x_1} + 7\sigma_{x_1}} \int_{l(x_1)}^{\mu_{x_2} + 7\sigma_{x_2}} f_X(x_1, x_2) dx_2 dx_1},
 \end{aligned} \tag{2.46}$$

where $f_X(x)$ and $F_X(x)$ denote bivariate normal p.d.f. and c.d.f., respectively; $\lambda_1 = \lambda_2 = 1/2$ and

$$\begin{aligned}
 l(x_1) &= -12407756274.6875x_1^7 + 5103841979.8231x_1^6 - 893105499.6990x_1^5 + \\
 &\quad 86172769.9900x_1^4 - 4951126.6304x_1^3 + 169402.1044x_1^2 - 3196.2936x_1 + 25.6869,
 \end{aligned} \tag{2.47}$$

with domain of $\{VaR_p(X_1) \leq x_1 \leq \mu_{x_1} + 7\sigma_{x_1}\}$, which is simply constructed by the use of Matlab “polyfit” function fitting the polynomial in the sense of the least squares.

We used the same upper bounds $\mu_{x_i} + 7\sigma_{x_i}$, $i = 1, 2$ for the integrals in (2.46) since $P(X_i \geq E(X_i) + 7\sigma_{x_i}) = 0.000000019$ which is small enough. At $p = 0.95$ and $p = 0.99$,

the p -quantile set $\{(x_1, x_2)^T \mid x_2 = l(x_1)\} \approx MVaR_p((X_1, X_2)^T)$ is illustrated in Figure 2.9.

From (2.46), we obtain $MCVaR_p(X) = 0.00048388$, and it means that \$483.88 is the expected loss amount beyond the $MVaR_p(X)$ at probability level $p = 0.95$. In other words, that amount of loss is expected to exceed at least one element in $MVaR$ with $p = 0.95$. With more critical probability level $p = 0.99$, we approximate $MVaR_p(X)$ as in Figure 2.9 and calculate the value of $MCVaR_p(X) = 0.00082815$, i.e., we can expect \$828.15 of loss from the investment at probability level $p = 0.99$. Note that $MCVaR_p(X)$ at probability level $p = 0.99$ is clearly larger than that at $p = 0.95$.

Example 2 (Negatively correlated two ETFs). While keeping Fidelity Nasdaq Composite Index Tracking (ONEQ) ETF, we replace S & P 500 ETF with Deutsche Bank US Dollar Index Bullish (UUP) ETF, which is “negatively” correlated to Nasdaq ETF. As we observe in Figure 2.10, the two index funds do not move in the same directions.



Figure 2.10: Basic Chart of Deutsche Bank US Dollar Index ETF and Fidelity Nasdaq ETF.

Basic chart over 6 months of Deutsche Bank US Dollar Index Bullish (UUP) ETF and Fidelity Nasdaq Composite Index Tracking (ONEQ) ETF from Yahoo Finance.

Initial prices per share on February 15, 2012 are \$115.02 for ONEQ, \$22.21 for UUP. Like the previous example, we assume that we have total available amount of

\$1,000,000 for the investment which is intended for the equal investment in each kind, i.e. \$500,000 each. And due to no fractional shares for those securities in real financial market, the initial investment is \$499,991.94 and \$499,991.52 for ONEQ and UUP, respectively. The corresponding number of shares is 4,347 for ONEQ and 22,512 for UUP. Again, we manipulate the data of daily closing prices into daily losses using the equation (2.45). Let Y_1, Y_2 denote the random variables of losses from an investment in Fidelity Nasdaq Composite Index Tracking (ONEQ) ETF and Deutsche Bank US Dollar Index Bullish (UUP) ETF, respectively. Each is assumed normally distributed. The random vector $Y = (Y_1, Y_2)^T$ have a bivariate normal distribution with parameters $\mu_{y_1} = -0.01185, \mu_{y_2} = -0.00875, \sigma_{y_1} = 0.02956, \sigma_{y_2} = 0.01705$ and $\rho = -0.70933$.

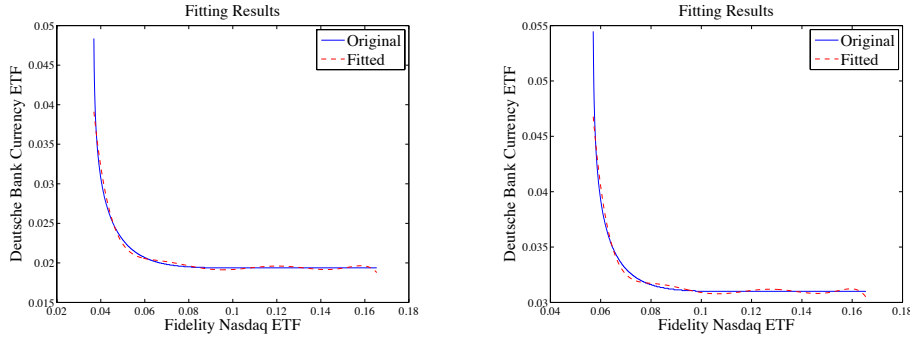


Figure 2.11: Multivariate Value-at-Risk, the quantile function at the probability levels $p = 0.95$ and $p = 0.99$.

At probability levels $p = 0.95$ for LHS and $p = 0.99$ for RHS, over 6 months of time period, $MVaR_p(Y)$ is well approximated by polynomial fitting in the sense of the least-squares (dotted curve). The random vector has components of the losses from the investment in Deutsche Bank US Dollar Index Bullish (UUP) ETF and Fidelity Nasdaq Composite Index Tracking (ONEQ) ETF from Yahoo Finance. The piecewise linear line (looks like a curve) is the set of line segments between the points (y_1, y_2) such that $F_Y(y_1, y_2) = p$.

With $MVaR_p(Y)$ of probability level $p = 0.95$, depicted in LHS of Figure 2.11, we calculate the following:

$$MCVaR_p(Y) \approx \frac{\int_{VaR_p(Y_1)}^{\mu_{y_1} + 7\sigma_{y_1}} \int_{l(y_1)}^{\mu_{y_2} + 7\sigma_{y_2}} (\lambda_1 y_1 + \lambda_2 y_2) f_Y(y_1, y_2) dy_2 dy_1}{\int_{VaR_p(Y_1)}^{\mu_{y_1} + 7\sigma_{y_1}} \int_{l(y_1)}^{\mu_{y_2} + 7\sigma_{y_2}} f_Y(y_1, y_2) dy_2 dy_1}, \quad (2.48)$$

where $f_Y(y)$ denote bivariate normal p.d.f., $\lambda_1 = \lambda_2 = 1/2$ and

$$\begin{aligned} l(y_1) = & -7207383.7882y_1^7 + 5413617.7299y_1^6 - 1703609.1254y_1^5 + \\ & 290811.0271y_1^4 - 29061.2545y_1^3 + 1700.3294y_1^2 - 54.01701y_1 + 0.74111, \end{aligned} \quad (2.49)$$

with domain of $\{VaR_p(Y_1) \leq y_1 \leq \mu_{y_1} + 7\sigma_{y_1}\}$, constructed by the use of Matlab “polyfit” function. In Figure 2.11, for both probability levels $p = 0.95$ and $p = 0.99$, $\{(y_1, y_2)^T \mid y_2 = l(y_1)\} \approx MVaR_p((Y_1, Y_2)^T)$ is illustrated.

The result is $MCVaR_p(Y) = 0.00040766$, which means \$407.66 is the expected amount of loss beyond the $MVaR_p(Y)$ with probability level $p = 0.95$. At probability level $p = 0.99$, we obtain $MCVaR_p(Y) = 0.00060555$, i.e. \$605.55 is the expected amount of loss beyond $MVaR_p(Y)$. We have conducted both Examples 1 and 2 with probability levels: $p = 0.8, 0.9, 0.95, 0.99$ and summarized in Table 3.1 for an easy and quick comparison.

Table 2.1: Summary of MCVaR of different stochastic dependence relationship at various probability levels.

	Nasdaq and S&P 500 ($\rho = 0.9510393$)		Nasdaq and US currency ($\rho = -0.7093342$)	
p -levels	MCVaR	$P(X \in D_p^c)$	MCVaR	$P(Y \in D_p^c)$
$p = 0.8$	0.00020031396148	0.15378401420764	0.00010897751748	0.00010152096654
$p = 0.9$	0.00033556530624	0.07132564927678	0.00020802596277	0.00000159763343
$p = 0.95$	0.00048387995078	0.03287958274500	0.00040766191359	0.00000002176898
$p = 0.99$	0.00082775999651	0.00477476826419	0.00060555466408	0.00000000000078

As in Table 3.1, the set of assets with negative correlation has a lower level of risk at each probability level. We also observe that, for the set of negatively correlated assets, the chance being beyond the $MVaR$ is much smaller than the case of positive correlation among assets. That is the power of low-correlation investment. Stochastic dependence structure among assets must be taken into account and that is one of the great features of desirable multivariate risk measures. It is shown that MCVaR can be used as a risk measure on correlated assets.

Remark 4. In the multivariate case it is reasonable to choose the value of p smaller than what we choose in the univariate case and it may depend on the number of components of a random vector.

Remark 5. (*Corporate M&A (Mergers and Acquisitions)*) Adequate measure of potential risk in a corporate M&A is essential. This is one of the most important factors in analyzing M&A deals from a risk management perspective. Examples 1 and 2 can also be considered as risk evaluation processes for different M&A deals. Depending on the categorization, there are many types of risk evaluation useful for M&A deals including strategic risk, compliance risk, operational risk, financial risk, reputation risk, etc. Additionally, each type of risk needs to be evaluated by a suitable risk measurement. Let us present a simple example of operational risk evaluation: suppose there are two companies of the same size (asset-based), involved in an M&A deal, and each board of directors wants to gauge its potential risk. Company A has 5 business sectors: smart phone, tablet PC, laptop computer, TV and digital camera, each with a total asset of \$1 billion. Let X_i , $i = 1, \dots, 5$ denote the random variable of operational loss from business sectors: smart phone, tablet PC, laptop computer, TV and digital camera, respectively. Company B has only 2 business sectors: display (LCD, LED panels) and real estate with total asset \$4 billion and \$1 billion, respectively. Let Y_j , $j = 1, 2$ denote the random variable of operational loss from these two business sectors, respectively. Let us assume Y_1 and X_i for $i = 1, 2, 3, 4, 5$ are highly correlated and Y_2 has a low correlation with others. With given asset value of each business sector, we have weight vectors $\lambda^X = (1/5, 1/5, 1/5, 1/5, 1/5)^T$ and $\lambda^Y = (4/5, 1/5)^T$ for company A and B, respectively.

For the potential operational risk evaluation of this M&A deal, we calculate

$$MCVaR_p(X_1, \dots, X_5, Y_1, Y_2) = E \left(\sum_{i=1}^5 \lambda_i^X X_i + \sum_{i=1}^2 \lambda_i^Y Y_i \mid H(X_1, \dots, X_5, Y_1, Y_2) \geq p \right), \quad (2.50)$$

where H is the c.d.f. of the random vector $(X_1, X_2, X_3, X_4, X_5, Y_1, Y_2)$. If the value of (2.50) is less than the sum of (2.51) and (2.52), risk measures of Company A and Company B, respectively:

$$MCVaR_p(X_1, \dots, X_5) = E \left(\sum_{i=1}^5 \lambda_i^X X_i \mid G(X_1, \dots, X_5) \geq p \right), \quad (2.51)$$

where G is the c.d.f. of the random vector $(X_1, X_2, X_3, X_4, X_5)$,

$$MCVaR_p(Y_1, Y_2) = E \left(\sum_{i=1}^2 \lambda_i^Y Y_i \mid F(Y_1, Y_2) \geq p \right), \quad (2.52)$$

where F is the c.d.f. of the random vector (Y_1, Y_2) , then it may be considered as a signal that this M&A deal is desirable from the point of view on managing risk. Comparison of (2.50) and the sum of (2.51) and (2.52) may serve the process of risk management in the M&A decision-making and it is advisable to do it for several p values to have an overview before the final decision. In the real world, the estimation of operational risk is complex and necessitates input from subject matter experts. However, MCVaR captures a stochastically dependent structure among correlated business sectors in M&A deals. In this respect, MCVaR will be able to play an important role as a risk measure in M&A analyses as well.

2.4.2 The Case of a Discrete Distribution

We use bounding schemes to obtain sharp lower and upper bounds for the probability $P(X \in D_p)$ as well as for the expectations $E(X_i \mathbb{1}_{X \in D_p}) = E(X_i \mid X \in D_p)P(X \in D_p)$ (see Prékopa (1988, 1990a,b, 1995, 2003)). If the lower and upper bounds are close to each other, we can use them for approximation.

Application of the Binomial Moment Bounding Scheme

The binomial moment problem for the probability of the union of events was introduced in Prékopa (1988). If A_1, \dots, A_N are arbitrary events and

$$S_k = \sum_{1 \leq i_1 < \dots < i_k \leq N} P(A_{i_1} \dots A_{i_k}), \quad k = 1, \dots, m,$$

then we solve the LP's:

$$\begin{aligned} & \min(\max) \sum_{i=1}^N p_i \\ & \text{subject to} \\ & \sum_{i=1}^N \binom{i}{k} p_i = S_k, \quad k = 1, \dots, m \\ & p_i \geq 0, \quad i = 1, \dots, N, \end{aligned} \quad (2.53)$$

and where m is a fixed integer, in practice $m \ll N$. Let W_{\min} , W_{\max} designate the optimum values, respectively. Then we have the sharp bounds, for the probability of the union,

$$W_{\min} \leq P\left(\bigcup_{i=1}^N A_i\right) \leq \min(W_{\max}, 1).$$

If the bounds are close to each other then they can be used to approximate the probability of the union. In the above formulation we are bounding probability but the method can be applied, in a straightforward manner, for subsets of an arbitrary set with finite measure.

In order to obtain lower and upper bounds for MCVaR, we take the sets $\{u \mid u \leq s^{(i)}\}$, $i = 1, \dots, N$, the union of which is D_p and define the measure on the Borel sets of D_p , generated by functions of the type: $u_i f(u)$, $u \in R^r$. Finally we construct lower and upper bounds for MCVaR, by the use of the obtained bounds.

Let $f(u)$ denote the p.d.f. of the random vector $X \in R^r$. Then

$$P(X \in D_p) = \int_{X \in D_p} f(u) du, \quad (2.54)$$

$$E(X_i \mathbb{1}_{D_p}) = E(X_i \mid X \in D_p) P(X \in D_p) = \int_{X \in D_p} u_i f(u) du, \quad (2.55)$$

where $D_p = \bigcup_{s \in \text{MVaR}_p(X)} (s + R_-^r)$.

The set $A_i = s^{(i)} + R_-^r$, $i = 1, \dots, N$ are the orthants in R^r , so are their intersections.

The vertex of $A_{i_1} \cdots A_{i_k}$ is

$$\left(\min(s_1^{(i_1)}, \dots, s_1^{(i_k)}), \dots, \min(s_r^{(i_1)}, \dots, s_r^{(i_k)}) \right).$$

Let us define k^{th} "binomial moment" S_k as follows.

$$S_k = \sum_{i_1 < \dots < i_k} \int_{A_{i_1} \cdots A_{i_k}} g(y) dy, \quad k = 1, \dots, N. \quad (2.56)$$

Example 3 (Compound Poisson Processes, Insurance Claims). *Suppose that various insurance claims occur according to independent, homogeneous Poisson processes. For simplicity the claims are assumed to be integer valued and independent of each other*

within each claim process and of the claims in the other claim processes. In the numerical example $M = 4$, and the types are: auto, health, home, life. The time period in which the claims are observed is one day. Information of daily claims to an insurance company is summarized in Table 2.2 and depicted in Figure 2.12.

Table 2.2: An insurance company's daily claims (in \$1,000) from 4 types of insurance.

ξ_1 :	Auto:	$N_1 \sim \text{Poisson}(0.55)$,	$Z_{N_1} \sim U(1, 2)$; the average of each claim= \$1,500
ξ_2 :	Health:	$N_2 \sim \text{Poisson}(0.12)$,	$Z_{N_2} \sim U(1, 3)$; the average of each claim= \$2,000
ξ_3 :	Home:	$N_3 \sim \text{Poisson}(0.08)$,	$Z_{N_3} \sim U(1, 5)$; the average of each claim= \$3,000
ξ_4 :	Life:	$N_4 \sim \text{Poisson}(0.01)$,	$Z_{N_4} \sim U(1, 5)$; the average of each claim= \$3,000

Let $N_i(t)$ and $X_i(t)$ designate the number of events and the total claim up to time t in the i th process, respectively. Then

$$P(N_i(t) = x) = \frac{(\lambda_i t)^x}{x!} e^{-\lambda_i t}, \quad x = 0, 1, \dots; \quad i = 1, \dots, M \quad (2.57)$$

$$X_i(t) = Z_{i1} + Z_{i2} + \dots + Z_{iN_i(t)}, \quad i = 1, \dots, M,$$

where Z_{ij} is the j th claim amount in claim process i .

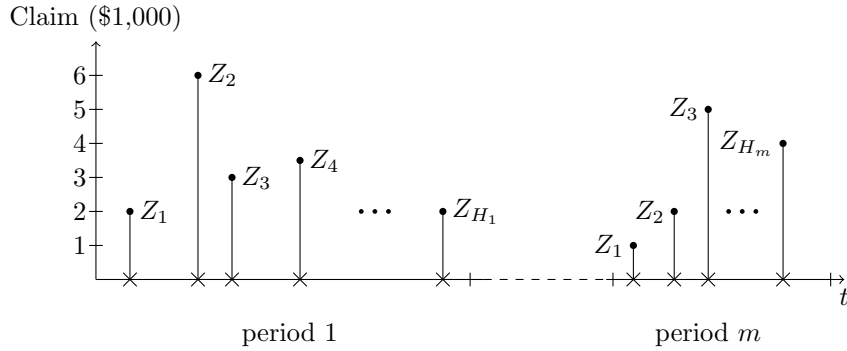


Figure 2.12: Illustration of Compound Poisson Distributed Losses. H_i is the number of events incurred over the period i . The unit claim size is \$1,000.

The company is concerned about the loss at probability level $p = 0.9$. Let $f_i(x) = P(X_i(t) = x)$. Then, Panjer's formula (see Bowers et al (1997)) provides us with recursions to calculate the probabilities $f_i(x)$, $x = 1, 2, \dots$, $i = 1, 2, 3, 4$:

$$f_i(x) = \frac{\lambda_i}{x} \sum_{j=1}^x j p_i(j) f_i(x-j), \quad x = 1, 2, \dots, \quad (2.58)$$

$$f_i(0) = e^{-\lambda_i}, \quad i = 1, 2, 3, 4.$$

Table 2.3: $MVaR_p(X)$ with $p = 0.9$, the unit claim size \$1,000.

ξ_1	ξ_2	ξ_3	ξ_4
3	4	6	1
3	7	5	6
4	3	5	3
4	3	6	1
4	4	3	1
4	6	2	6
4	7	2	5
5	2	6	1
5	3	3	4
5	3	4	1
5	4	1	1
6	2	5	5
6	3	3	1
7	2	5	3

For $MVaR$, we obtained the following 14 p -efficient points, presented in Table 2.3.

The binomial moment bounding scheme, to obtain lower and upper bounds for $E(X_i | X \in D_p)P(X \in D_p)$ is:

$$\begin{aligned}
& \min(\max) \sum_{i=1}^{14} p_i \\
& \text{subject to} \\
& \sum_{i=1}^{14} \binom{i}{k} p_i = S_k, \quad k = 1, \dots, m \\
& p_i \geq 0, \quad i = 1, \dots, 14,
\end{aligned} \tag{2.59}$$

where

$$\begin{aligned}
S_k &= \sum_{i_1 < \dots < i_k} \int_{A_{i_1} \dots A_{i_k}} g(y) dy, \quad k = 1, \dots, m, \\
g(y) &= y_i f(y), \text{ where } y = \xi \in R^4, \\
A_i &= \{s^{(i)} + R_-^4, \quad s^{(i)} \in R^4\}, i = 1, \dots, 14
\end{aligned} \tag{2.60}$$

and f is the p.d.f. of the random vector $X \in R^4$.

With $m = 14$, the constraints in (2.59) uniquely determine the unknowns and we obtain the values:

$$\begin{aligned}
E(X_1 | X \in D_p)P(X \in D_p) &= 0.82115806, \\
E(X_2 | X \in D_p)P(X \in D_p) &= 0.23898585, \\
E(X_3 | X \in D_p)P(X \in D_p) &= 0.22968684, \\
E(X_4 | X \in D_p)P(X \in D_p) &= 0.02975693.
\end{aligned} \tag{2.61}$$

For the bounds of $P(X \in D_p)$, we solve the LP's with $m = 14$ and integrand $g(y) = f(y)$:

$$\begin{aligned} & \min(\max) \sum_{i=1}^{14} p_i \\ & \text{subject to} \\ & \sum_{i=1}^{14} \binom{i}{k} p_i = S_k, \quad k = 1, \dots, m \\ & p_i \geq 0, \quad i = 1, \dots, 14, \end{aligned} \tag{2.62}$$

where

$$\begin{aligned} S_k &= \sum_{i_1 < \dots < i_k} \int_{A_{i_1} \dots A_{i_k}} f(y) dy, \quad k = 1, \dots, m, \\ A_i &= \{s^{(i)} + R_-^4, \quad s^{(i)} \in R^4\}, \quad i = 1, \dots, 14 \end{aligned} \tag{2.63}$$

and f is the p.d.f. of the random vector $X \in R^4$.

The optimum values coincide up to 8 digits and the resulting number is accepted as approximation of $P(X \in D_p)$:

$$P(X \in D_p) = 0.99832959. \tag{2.64}$$

Let $\lambda_i = (\text{the amount of premium in type } i) / (\text{total premium to all types})$ and assume that $\lambda_i = \frac{1}{4}$ for $i = 1, \dots, 4$. Then we have

$$\sum_{i=1}^4 \lambda_i E(X_i | X \in D_p) P(X \in D_p) = 0.32989692. \tag{2.65}$$

Simple calculation gives:

$$\begin{aligned} m_1 &= E(X_1) = E(N_1)E(Z_1) = \lambda_1 E(Z_1) = 0.55 \times 1.5, \\ m_2 &= E(X_2) = E(N_2)E(Z_2) = \lambda_2 E(Z_2) = 0.12 \times 2, \\ m_3 &= E(X_3) = E(N_3)E(Z_3) = \lambda_3 E(Z_3) = 0.08 \times 3, \\ m_4 &= E(X_4) = E(N_4)E(Z_4) = \lambda_4 E(Z_4) = 0.01 \times 3, \end{aligned} \tag{2.66}$$

and

$$\sum_{i=1}^4 \lambda_i m_i = 0.33375. \tag{2.67}$$

Plugging in the values of (2.64), (2.65) and (2.67) into (2.15), we obtain a loss amount of 2.30666989 with the unit claim size of \$1,000. The insurance company will expect

the daily loss, i.e., the amount of total claims per day, of \$2,306.67 with probability level 90%. Thus, we conclude that under that probability level, the insurance company would like to collect the amount of premium at least \$2,306.67 per day for accepting the risk, in order to make some underwriting profit.

Application of the Boolean Bounding Scheme

The Boolean bounding scheme can also be used for bounding measures of the union of sets. Again, for the details of the Boolean bounding scheme we refer the readers to Prékopa (2003). Let us present a LP formulation for bounding the probability of the union, i.e., $P(X \in D_p)$, the probability of the union of favorable domain. In order to formulate the problem we introduce the notations:

$$\begin{aligned} a_{IJ} &= \begin{cases} 1, & \text{if } I \subset J \\ 0, & \text{if } I \not\subset J, I, J \subset \{1, \dots, n\} \end{cases}, \\ x_J &= P\left(\left(\bigcap_{j \in J} A_j\right) \cap \left(\bigcap_{j \notin J} \bar{A}_j\right)\right), \\ p_I &= P\left(\bigcap_{j \in I} A_j\right), I, J \subset \{1, \dots, n\}. \end{aligned}$$

The probability p_I means that all events A_j , $j \in K$ occur and the probability x_J means that all events A_j , $j \in I$ occur but the other do not occur. The Boolean probability bounding problem, or scheme for the probability of the union is the following:

$$\begin{aligned} &\min(\max) \sum_{\emptyset \neq J \subset \{1, \dots, n\}} x_J \\ &\text{subject to} \\ &\sum_{J \subset \{1, \dots, n\}} a_{IJ} x_J = p_I, I \subset \{1, \dots, n\}, |I| \leq m \\ &x_J \geq 0, J \subset \{1, \dots, n\}. \end{aligned} \tag{2.68}$$

Problem (2.68) has $1 + \sum_{i=1}^m \binom{n}{i}$ equality constraints and 2^n variables. If we remove x_0 and the equality constraint containing x_0 (meaning that the sum of the variables is equal to 1), then we obtain an equivalent Boolean problem, for bounding the probability

of the union:

$$\begin{aligned}
& \min(\max) \sum_{\emptyset \neq J \subset \{1, \dots, n\}} x_J \\
& \text{subject to} \\
& \sum_{\emptyset \neq J \subset \{1, \dots, n\}} a_{IJ} x_J = p_I, \quad \emptyset \neq I \subset \{1, \dots, n\}, |I| \leq m \\
& x_J \geq 0, \quad \emptyset \neq J \subset \{1, \dots, n\}.
\end{aligned} \tag{2.69}$$

We also need to introduce the notations x_{J_i} and E_{I_i} as the following:

$$x_{J_i} = E \left(X_i \mathbb{1}_{\left(\bigcap_{j \in J} A_j \right) \cap \left(\bigcap_{j \notin J} \bar{A}_j \right)} \right), \tag{2.70}$$

$$E_{I_i} = E \left(X_i \mathbb{1}_{\bigcap_{j \in I} A_j} \right), \quad I, J \subset \{1, \dots, n\}. \tag{2.71}$$

For example, let $I, J \subset \{1, 2, 3\}$, $|I| \leq 2$. Then E_{I_i} can be described as follows.

$$E_{I_i} = \begin{pmatrix} \int_{A_1} x_i f(x) dx \\ \int_{A_2} x_i f(x) dx \\ \int_{A_3} x_i f(x) dx \\ \int_{A_1 A_2} x_i f(x) dx \\ \int_{A_1 A_3} x_i f(x) dx \\ \int_{A_2 A_3} x_i f(x) dx \end{pmatrix}, \tag{2.72}$$

where $f(x)$ is the p.d.f. of the random vector $X \in R^r$. Now we are ready to apply Boolean bounding scheme to the same numerical data of the examples in the previous section.

Example 4 (using the same data of Example 3 with the Boolean bounding scheme).

The following LP formulation (2.73) with $m = 4$, $n = 14$ is for $P(X \in D_p)$:

$$\begin{aligned}
& \min(\max) \sum_{\emptyset \neq J \subset \{1, \dots, n\}} x_J \\
& \text{subject to} \\
& \sum_{\emptyset \neq J \subset \{1, \dots, n\}} a_{IJ} x_J = p_I, \quad \emptyset \neq I \subset \{1, \dots, n\}, |I| \leq m \\
& x_J \geq 0, \quad \emptyset \neq J \subset \{1, \dots, n\},
\end{aligned} \tag{2.73}$$

where $x_J = P\left(\left(\bigcap_{j \in J} A_j\right) \cap \left(\bigcap_{j \notin J} \bar{A}_j\right)\right)$ and $p_I = P\left(\bigcap_{j \in I} A_j\right)$, $I, J \subset \{1, \dots, n\}$. The corresponding Boolean matrix is very large – the size of the matrix is 1480×16383 because $\sum_{i=1}^4 \binom{14}{i} = 1470$ and $2^{14} - 1 = 16383$.

The LP formulation (2.73) provides us with the following result:

$$0.99832959468795 \leq P(X \in D_p) \leq 0.99832959471865. \quad (2.74)$$

Since the difference between lower and upper bounds in (2.74) is very small, let us present the bounds as one number of 8 decimal places: $P(X \in D_p) \approx 0.99832959$ which is the same as (2.64). In order to calculate $E(X_i | X \in D_p)P(X \in D_p)$, we solve the following Boolean bounding problem again with $m = 4$, $n = 14$:

$$\begin{aligned} & \min(\max) \sum_{\emptyset \neq J \subset \{1, \dots, n\}} x_{J_i} \\ & \text{subject to} \\ & \sum_{\emptyset \neq J \subset \{1, \dots, n\}} a_{IJ} x_{J_i} = E_{I_i}, \quad \emptyset \neq I \subset \{1, \dots, n\}, |I| \leq m \\ & x_J \geq 0, \quad \emptyset \neq J \subset \{1, \dots, n\}, \end{aligned} \quad (2.75)$$

where $E_{I_i} = E\left(X_i \mathbb{1}_{\bigcap_{j \in I} A_j}\right)$, $I, J \subset \{1, \dots, n\}$, and $x_{J_i} = E\left(X_i \mathbb{1}_{\left(\bigcap_{j \in J} A_j\right) \cap \left(\bigcap_{j \notin J} \bar{A}_j\right)}\right)$, where $A_i = \{s^{(i)} + R_+^4, s^{(i)} \in R^4\}$, which is an orthant with vertex $(s_1^{(i)}, s_2^{(i)}, \dots, s_r^{(i)})$ for $i = 1, \dots, 14$ since we have 14 p-Level Efficient Points, as those are enumerated in Table 2.3.

From the above LP formulation the results are obtained as follows.

$$\begin{aligned} 0.821158068554149 & \leq E(X_1 | X \in D_p)P(X \in D_p) \leq 0.821158068554182, \\ 0.238985857588695 & \leq E(X_2 | X \in D_p)P(X \in D_p) \leq 0.238985857631699, \\ 0.229686849617116 & \leq E(X_3 | X \in D_p)P(X \in D_p) \leq 0.229686849618277, \\ 0.029756933777791 & \leq E(X_4 | X \in D_p)P(X \in D_p) \leq 0.029756933782136. \end{aligned} \quad (2.76)$$

Again, let us use numbers of 8 decimal places. Then all expectations are represented as one number, which are the same as (2.61) in Example 3. Since all the inputs are the same as one with binomial moment scheme, we obtain the same result that MCVaR for an insurance company is the loss of \$2,306.67. Thus, under the probability level

$p = 0.9$, the insurance company would like to collect the amount of premium at least \$2,306.67 per day for accepting the risk.

2.5 Concluding remarks

We have explored various properties of Multivariate Value at Risk, or MVaR and Multivariate Conditional Value at Risk, or MCVaR. We have shown that many properties enjoyed by VaR and CVaR, carry over to the multivariate risk measures. In addition we have derived some properties of MVaR and MCVaR, based on multivariate logconcave theory, that do not have univariate counterpart or it is trivial. As regards the convexity of MVaR and MCVaR, none of them has that property, in general, but we have proved the convexity of MCVaR under the assumption that the components of the random vector are independent. We have proposed the numerical procedures to calculate or approximate MCVaR values. In case of a continuously distributed random vector we have used approximation and numerical integration. In case of a discrete random vector we have used the recently developed binomial moment and Boolean bounding schemes to approximate MCVaR. The results are illustrated on real life data and it is shown how MCVaR depends on the probability level and the correlation between the components of the random vector, representing different portfolios.

Chapter 3

Decision-making from a Risk Assessment Perspective for Corporate Mergers and Acquisitions

3.1 Introduction

Mergers and acquisitions (M&A) and corporate restructuring (e.g., combining divisions, demergers, etc.) have a significant impact on financial markets. Investment bankers on Wall Street and in financial centers worldwide arrange M&A transactions daily, and deals can be worth hundreds of millions, and in many of the largest cases, billions of dollars. Total M&A deals are in excess of tens and hundreds of billions annually. We hear of the many deals frequently, and indeed they happen almost continuously – often hundreds per quarter (see, e.g., Ernst&Young (2013), etc.).

Granted, M&A deals make headlines, but what does this all mean to decision makers in the involved companies? Presumably, they want to achieve a synergetic effect as a consequence of the M&A deal. In other words, decision makers considering M&A deals seek robust, sustainable profitability from their future business model over the long term. To this end, corporate M&As must be analyzed in multiple ways. For details about analyses of corporate M&As, we refer the reader to the literature (e.g., see Ernst&Young (2013), Rose and Frame (2011), Jaruzelski et al (2009), Gregoriou and Renneboog (2007), Shimizu et al (2004), etc.). Prediction of cash flow, especially in the near future, may be the most important matter for the united firm in terms of the business' sustainability. However, future cash flow prediction is very difficult, especially in the early phases after M&As, because so many factors are uncertain in the post-M&A process. In this respect, reasonable risk assessment methods should play a key role in realistically gauging the outcomes of M&A transactions.

From a risk management perspective, assessment of the risk *before and after* proposed M&A deals is instrumental in the decision-making process. This is often because decision makers are interested in determining if the M&A deals will result in a *risk reduction*, allowing for more robust operations and, in many cases, revenues. This begs the question: How can we systematically and effectively evaluate the complex risk profiles of M&A deals? Would deals necessarily reduce risk? Can we think of corporate M&A as analogous to asset diversification? If this is the case, then would it also be cost-effective? Most research indicates that M&A transactions from 1995 to 2005 have an overall success rate of about 50% (see, e.g., Rose and Frame (2011).), indicating that the answers to these questions are as often “yes” as “no,” and demonstrating the need for more sophisticated technical tools for assessing the risk/return profile.

In a broad sense, corporate M&As can be regarded as an addition of assets, since each company has many business divisions, which, in some sense, can be viewed as “assets.” The question arises: Is it then logically equivalent to the sum of distinct portfolios? We would contend that it is not. First of all, for the portfolio construction, from the standpoint of investors, there is no resistance on adding (or subtracting) assets to (or from) the existing portfolios. In case of M&As, on the other hand, restructuring of the business divisions will be a very difficult and demanding task, and takes quite a long time until the proper and satisfactory functioning of the reorganized business units.

Moreover, business divisions within a company are very strongly correlated in a steady and sustainable way; they are communicating all the time and try to create a more efficient network. Each business division plays its own role, and the values (or, profit/cost profile) generated by each division have unique patterns of net profit and loss. Assets in a portfolio are also correlated with each other, however, not as strongly and consistently as the business units within a firm. Clearly, business divisions operate in a timely and organized manner, i.e., responding immediately to the actions of other divisions.

3.2 Risk assessment methods for corporate M&As from the mathematical perspective

Suppose that there is an ongoing analysis for a corporate M&A deal between Companies X and Y , and both companies run many different businesses. Assume that there are n and m business units in Companies X and Y , respectively. Let the random variables X_i , $i = 1, \dots, n$ and Y_j , $j = 1, \dots, m$ denote losses of the associated business sectors of Companies X and Y , respectively. Then the random vectors $X = (X_1, \dots, X_n) \in R^n$ and $Y = (Y_1, \dots, Y_m) \in R^m$ mean losses of the associated companies. Vectors are written in row form but if they appear in matrix operations then they are written in column form. Let ρ denote a risk measure. Then $\rho(X)$ and $\rho(Y)$ are the risk measures of Company X and Y , respectively.

For a risk evaluation process on the corporate M&A deal for companies X and Y , we suggest the use of the random vector $(X, Y) \in R^{n+m}$, instead of the sum of random variables, i.e., $Z = \sum_{i=1}^n X_i + \sum_{j=1}^m Y_j$. Then a potential risk measurement of the united firm can be written as

$$\rho((X, Y)), \quad (3.1)$$

where the random vector $(X, Y) \in R^{n+m}$ means the losses of the new merged entity, and all components of the random vector (X, Y) are loss random variables of corresponding business sectors of a new firm after the M&A deal of Company X and Y .

The reason behind this mathematical expression in (3.1) is that after an M&A deal the new company becomes a multi-plant firm and each business sector is still producing the same goods or services as it did for the former individual companies. The component of the loss random vector $(X, Y) \in R^{n+m}$ can be thought of losses of corresponding business units of a multi-plant firm, immediately following an M&A deal. It is paramount to include all remaining business units of both companies, as each unit would maintain its operations in the critical initial period following the deal. Although a restructuring strategy for a united firm can begin in this initial period, the implementation necessarily takes significant time to establish integrated operations, and the early phases of the integration process are sensitive and critical to the success of

the M&A deal.

We know that timing is crucial in a decision-making process. A risk measurement for a more critical period would be exceedingly valuable in setting up a suitable risk management plan. Thus, for risk evaluation for decision-making on M&As, it would be best to simultaneously consider all the business units individually as they were before M&As and all operating units as a whole. This implies that the risk measure on a random vector (X, Y) well-suited for a decision-making processes on M&As.

As we mentioned earlier in this section, $\rho(X) + \rho(Y)$ and $\rho((X, Y))$ represent the risk measurements *before and after* the M&A transaction, respectively, and their comparison could be an important component of analysis for (in)validation of execution on the deal. Question arises: What kind of risk measure ρ should be used here for comparison of the values? To calculate and compare the magnitudes of risks *before and after* M&As, a convex risk measure might not be suitable for the risk evaluation processes in M&A events.

The reason for this is that a convex risk measure always indicates that the merged company will become less risky after the event because of its subadditivity property, no matter how badly the M&A fits the existing business models of the target and acquiring firms. This suggests that a convex risk measure will play a limited role in the risk evaluation processes for the decision-making on M&As. The problem of the applicability of a subadditive risk measure for M&A transactions has already been discussed in Kou et al (2013). In this respect, an appropriate risk evaluation method necessitates the use of a reasonable risk measure. Many examples of bad M&A deals strongly motivate our research on the risk evaluation processes and investigation of suitable risk measures for corporate M&A deals.

3.3 Motivation to study a new risk measure for corporate M&As

Value-at-Risk and Conditional (or Average) Value-at-Risk are widely accepted and used by both academics and practitioners. Value-at-Risk (VaR) has already existed in the statistical literature since the second half of the 19th century, under the name of

quantile or percentile. The term Value-at-Risk was introduced at the beginning of the 1990s in the financial literature. For various topics of Value-at-Risk see Jorion (2006), Saita (2007), etc. Its multivariate counterpart turned up in the stochastic programming literature, primarily in the works of Prékopa (1990, 1995), etc.

The term Conditional Value-at-Risk (CVaR) was introduced in Uryasev and Rockafellar (2000). The same notion was named Average Value at Risk (AVaR) in Föllmer and Schied (2002). This is also called *Expected Shortfall*, or *Tail Value at Risk*. However, it had already been presented in the earlier literature in Prékopa (1973a) and Ben-Tal and Teboulle (1986). CVaR is a coherent risk measure in the sense of Artzner et al (1999) while VaR is generally not (see, e.g., Pflug (2000).). For more about coherent (or convex) risk measures, the reader is referred to Acerbi and Tasche (2001), Szegö (2002), Frittelli and Gianin (2002), Jarrow and Purnanandam (2005), Föllmer and Penner (2006), Ben-Tal and Teboulle (2007), Föllmer and Schied (2010), etc.

In the case of multiple correlated assets, however, it is not imperative to require convexity for a reasonable risk measure, and indeed, it is generally misguided. There are reasons for that, as we mentioned earlier in this section, such as bad corporate M&A deals and poorly constructed portfolios, which may have undesirable risk-return characteristics. For decision-making on M&As, especially from a risk management perspective, we believe that comparison of the risks before and after the M&A deals would be useful, since a positive decision can be made if the M&A is expected to reduce risk, i.e., the risk of a merged firm is less than the sum of the risks of the separate acquiring and target companies.

M&As are very complex since a number of things are involved and correlated to each other, and thus the use of a risk measure capable of handling a multidimensional situation may be a useful tool for the decision-making in M&As. Multidimensional settings can be managed in large part by capturing the dependence structure among key elements involved in the M&As. Risk measures for multidimensional settings have previously been studied, and we refer the reader to the recent literature (see, e.g., Prékopa (2012), Dentcheva and Ruszczyński (2009), Noyan and Rudolf (2013), Lee and Prékopa

(2013), etc.) In order to deal with dependency structures among multiple correlated objects, copula approach has been developed and used for various practical applications, including finance, risk management, etc. For the theoretical and applicable aspects of copula, we refer the reader to Joe (1997), Embrechts et al (1999), Embrechts et al (2001), Ané and Kharoubi (2003), Luciano et al (2004), Junker and May (2005), etc. Although copula itself is not supported by realistic modelling procedure, a suitable application of copula might be useful for the risk management of M&As.

However, we are especially interested in the decision-making on M&As and a decision should be based on a suitable decision analysis process before taking action on the deal. We believe that, from a risk management perspective, comparisons of the expected losses from before and after M&As should be one of the decision criteria. In other words, we not only need to handle the dependence structure of the components of (X, Y) , but also simultaneously quantify the expected losses from before and after the M&As in order to see if the M&As would result in a risk reduction. For this reason, application of copula would not be suitable for our goal, although it may play a role in the risk management in post-M&A processes.

In a Bottom-Up way of thinking, we came up with a new multivariate risk measure which quantifies the conditional expected loss of multiple correlated assets in some unfavorable situations, and also simultaneously incorporates stochastic dependency structures among the objects. In Section 3.6 we introduce such risk measure under the name of the worst-case Combined Value-at-Risk (CoVaR), which is developed through Sections 3.4 and 3.5, by a Bottom-Up approach. For decision-making on corporate M&As, especially from a risk management perspective, an appropriate methodology is presented in Section 3.7. Numerical examples about corporate M&As are presented and discussed in Section 3.8.

3.4 Taking a Bottom-Up approach in the multidimensional case

Let $X \in R$ be a random variable, interpreted as loss, the probability distribution function of X : $F(z) = P(X \leq z)$, $-\infty < z < \infty$. Let Q denote the p -quantile,

equal to $\text{VaR}_p(X) = F^{-1}(p)$, where, by definition, $F^{-1}(p) = \min\{u \mid F(u) \geq p\}$. The event $X \leq Q$ (or $X \leq \text{VaR}_p(X)$) is preferred for large p to the event $X > Q$ (or $X > \text{VaR}_p(X)$). In other words, the set $\{x \mid x \leq Q\}$ is favorable and its complement $\{x \mid x > Q\}$ is unfavorable. If X means profit, then this holds with reversed inequalities, i.e., $\text{Loss} = -\text{Profit}$. For a single asset, its Value-at-Risk can be represented as a single point on the real line as demonstrated in Figure 5.3.

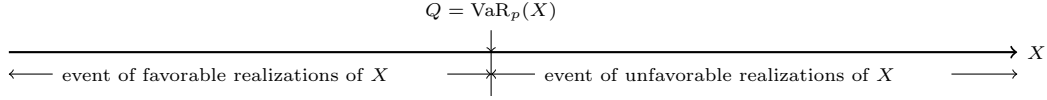


Figure 3.1: Favorable set and its complementary set in the case of a loss random variable

If we are dealing with multiple assets jointly, then the level of loss of some of the assets can be expressed as a set rather than a single value. For a set of multiple assets, the level of *loss* is called Multivariate Value-at-Risk (MVaR) that has been known for some time as *p*-quantile or *p-Level Efficient Point* (pLEP), or briefly *p*-efficient point. The latter concept was introduced in Prékopa (1990) and further studied in Prékopa (1995), Prékopa et al (1998), Prékopa (2012), Lee and Prékopa (2013), Boros et al (2003). Multivariate Value-at-Risk (MVaR), the multivariate counterpart of VaR, is a set of points, rather than a single point as it is in the univariate case.

For the construction of portfolios, however, analysis of individual financial assets is essential, since every individual asset has its own attributes in various aspects – for example, categorization of stocks in the market can be done by business sector (health-care, technology, services, etc.), capitalization (large, mid, or small Cap), style (growth or value) and many other different ways. Clearly, assets in different countries have different characteristics, even though the assets are of the same type, since there are various types of country risks, already reflected in the rating of assets. Furthermore, assets can also be handled in various ways by their classes; real estates, bonds, commodities, etc. Examining the specific assets, followed by analysis of a set of these assets, can be called a Bottom-Up Approach for investment.

The same reasoning applies to corporate M&A (Mergers and Acquisitions) activities. Before taking action for the M&A integration, a clear identification of the target company is essential. Furthermore, every relevant business sector should be closely examined for any potential effects of the event. In this respect, the detailed M&A plan can be made based on the Bottom-Up analysis approach to facilitate decision making. Note that M&A is a complex process so the risk measurement is only one of the key resources in a successful M&A integration.

Let us assume that we have a set of n different assets (or n different business sectors in a company). Let X_k be the loss random variable of asset k , $k = 1, \dots, n$ and p a given probability level for asset k , $k = 1, \dots, n$. With the probability distribution functions F_{X_k} the p -quantile points are $Q_k = F_{X_k}^{-1}(p) = \text{VaR}_p(X_k)$, $k = 1, \dots, n$. Let Q be the p -quantile vector in the following sense:

$$Q = (F_{X_1}^{-1}(p), \dots, F_{X_n}^{-1}(p))^T = (\text{VaR}_p(X_1), \dots, \text{VaR}_p(X_n))^T. \quad (3.2)$$

Let B denote the most favorable outcome of the loss random vector $X \in R^n$, i.e.,

$$B = \{x \in R^n \mid x_1 \leq \text{VaR}_p(X_1), x_2 \leq \text{VaR}_p(X_2), \dots, x_n \leq \text{VaR}_p(X_n)\}. \quad (3.3)$$

Also let $A_j = \{X_j \leq \text{VaR}_p(X_j)\}$, for $j = 1, \dots, n$. Then we can write the favorable set $B = \bigcap_{j=1}^n A_j$, and its complementary is an unfavorable set $B^c = \bigcup_{j=1}^n A_j^c$. Both are illustrated in Figure 3.2.

3.5 Combined Value-at-Risk

In the definition of CVaR (or AVaR), we take the expectation of X given that $X > \text{VaR}_p(X)$ (unfavorable outcome) if X means loss, i.e., $\text{CVaR}_p(X) = E(X \mid X > \text{VaR}_p(X))$. For the multivariate case (a set of multiple assets), we propose the following

Definition 3.5.1. *The Combined Value-at-Risk, or CoVaR, of the loss random vector $X \in R^n$ is designated and defined, with a fixed individual probability level p , as:*

$$\text{CoVaR}_p(X) = E(\psi(X) \mid X \notin B), \quad (3.4)$$

where ψ is some n -variate function, B is the most favorable outcome as defined in (3.3).

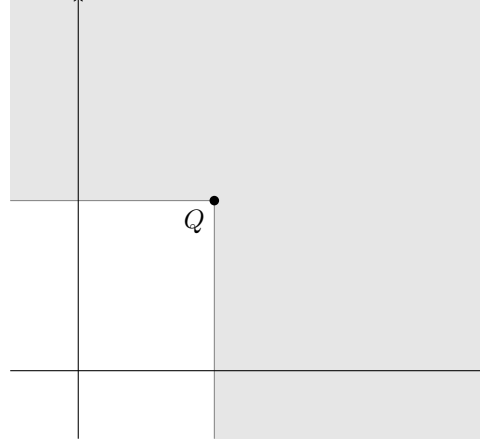


Figure 3.2: 2-D illustration of the favorable set and its complement. The unshaded region represents the most favorable set B , and the shaded region describes its complement. Q is the p -quantile point in the sense that $p = P(X_1 \leq Q_1, \dots, X_n \leq Q_n)$, $Q_k = \text{VaR}_p(X_k)$, $k = 1, \dots, n$.

CoVaR in (3.4) can also be rewritten as

$$\text{CoVaR}_p(X) = E \left(\psi(X) \mid X \in \bigcup_{j=1}^n A_j^c \right), \quad (3.5)$$

where $A_j^c = \{X_j > \text{VaR}_p(X_j)\}$ for $j = 1, \dots, n$. Note that the event $\bigcup_{j=1}^n A_j^c$ allows for X the entire space excluding the single orthant $\{x \in R^n \mid x_j \leq \text{VaR}_p(X_j), j = 1, \dots, n\}$.

Let us define the function $\psi(u)$, $u = (u_1, \dots, u_n)^T$ in the following way:

$$\psi(u) = \sum_{i=1}^n \lambda_i u_i, \quad (3.6)$$

where λ_i , $i = 1, \dots, n$ can be chosen in a suitable way depending on how random variables are defined. For a risk evaluation process on corporate M&As, random variables may be defined as losses of business units of the companies involved in the M&As. In this case we may have $\lambda_i = 1$, $i = 1, \dots, n$ in order to count each business unit once to quantify risk of the united firm after the M&As. For measuring risk on stock portfolios, random variables can be designated as losses of stocks. Then it would be suitable to have $\lambda_1, \dots, \lambda_n$ to be integer-valued as their interpretation is the number of shares of corresponding stocks. We may also allow negative λ values, meaning short selling. If we want to assign weights on investment in Assets $1, \dots, n$, then $\lambda_1, \dots, \lambda_n$ are nonnegative constants satisfying $\sum_{i=1}^n \lambda_i = 1$. Depending on the meaning of random variables, a function $\psi(u)$, $u \in R^n$ can be specialized in an appropriate way.

The calculation of CoVaR in (3.5) is not simple because it is defined in the space of the union of sets A_1^c, \dots, A_n^c (shaded region in Figure 3.2). To calculate, we can use the following equation:

$$E(\psi(X)) = E(\psi(X) \mid X \notin B)P(X \notin B) + E(\psi(X) \mid X \in B)P(X \in B), \quad (3.7)$$

from which we derive:

$$\begin{aligned} \text{CoVaR}_p(X) &= E(\psi(X) \mid X \notin B) \\ &= \frac{1}{P(X \notin B)} \left(E(\psi(X)) - E(\psi(X) \mid X \in B)P(X \in B) \right). \end{aligned} \quad (3.8)$$

Equation (3.8) can be written as:

$$\text{CoVaR}_p(X) = \frac{1}{1 - P(X \in B)} \left(\sum_{i=1}^n E(X_i) - \sum_{i=1}^n E(X_i \mid X \in B)P(X \in B) \right). \quad (3.9)$$

Since the set B is only a single-orthant in the n dimensional space (unshaded region in Figure 3.2), the formulation of (3.9) can be calculated. The set B represents the event of the best-case realizations of the random vector $X \in R^n$ as it is illustrated as an unshaded region in Figure 3.2. Combined Value-at-Risk (CoVaR) gauges the expected loss amount in the set B^c , where the unfavorable events occur. B^c represents the whole space excluding the best-case scenarios and therefore B^c is the least risky of unfavorable events.

Let us now turn our attention to the worst-case event, the riskiest event among all possible outcomes under a set of individual probability levels. The expected loss amount in the worst-case event is clearly the largest, and so would cover any other risky situation. For this reason, it may be used for the calculation of minimum (but safe) required reserve for financial institutions, and this is our motivation in the next section.

3.6 Combined Value-at-Risk in the worst-case event

The worst-case event should be considered in practice for various purposes: trading operations, asset management, or any business where a short-term catastrophe could result in complete collapse of the entity. The worst-case event will focus on a possible

loss given that no favorable event occurs. Let W denote the worst-case event. If a random vector $X = (X_1, \dots, X_n)^T \in R^n$ means losses, then the event W can be written as

$$\begin{aligned} W &= \{x \in R^n \mid x_i > \text{VaR}_p(X_i), i = 1, \dots, n\} \\ &= \{x \in R^n \mid x > Q\}, \text{ where } Q = (\text{VaR}_p(X_1), \dots, \text{VaR}_p(X_n))^T. \end{aligned} \quad (3.10)$$

Definition 3.6.1. *The worst-case CoVaR of the loss random vector $X \in R^n$ is designated and defined as:*

$$w\text{CoVaR}_p(X) = E(\psi(X) \mid X \in W), \quad (3.11)$$

where ψ is some n -variate function, W denotes the worst-case set as in (3.10). Let us define the function $\psi(u)$ as in (3.6) with $\lambda_i = 1$, $i = 1, \dots, n$. Note that the function ψ can be specialized in various ways, depending on the characteristics of the business, as mentioned earlier in Section 3.5.

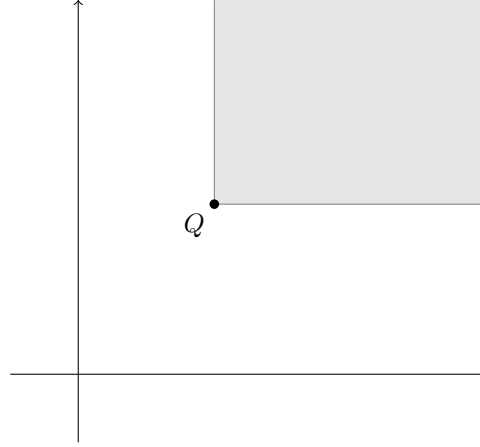


Figure 3.3: 2-D Illustration of the worst-case unfavorable set
2-D Illustration of the worst-case unfavorable set $W = \{x \in R^n \mid x > Q\}$ is the shaded region, where $Q = (\text{VaR}_p(X_1), \dots, \text{VaR}_p(X_n))^T$.

The worst-case CoVaR (wCoVaR) can easily be calculated directly from the definition because the set W is a single-orthant in n -dimensional space, as described in Figure 3.3. If a random loss vector $X \in R^n$ has a continuous distribution, then the

worst-case CoVaR (wCoVaR) can be formulated as:

$$\begin{aligned}
\text{wCoVaR}_p(X) &= E(\psi(X) \mid X \in W) = E(\psi(X) \mid X > Q) \\
&= E(X_1 + \dots + X_n \mid X_1 > \text{VaR}_p(X_1), \dots, X_n > \text{VaR}_p(X_n)) \\
&= \frac{\int_{\text{VaR}_p(X_n)}^{\infty} \dots \int_{\text{VaR}_p(X_1)}^{\infty} (t_1 + \dots + t_n) f_X(\mathbf{t}) dt_1 \dots dt_n}{\int_{\text{VaR}_p(X_n)}^{\infty} \dots \int_{\text{VaR}_p(X_1)}^{\infty} f_X(\mathbf{t}) dt_1 \dots dt_n},
\end{aligned} \tag{3.12}$$

where f_X is the probability density function of a random vector $X \in R^n$.

Theorem 3.6.1. *For a loss random vector $X = (X_1, \dots, X_n)$ with independent components, we have the equation,*

$$wCoVaR_p(X) = CVaR_p(X_1) + \dots + CVaR_p(X_n), \tag{3.13}$$

where we assume that $E(X_i)$, $i = 1, \dots, n$ exist.

Proof. We have the equations

$$\begin{aligned}
\text{wCoVaR}_p(X) &= E\left(X_1 + \dots + X_n \mid \bigcap_{j=1}^n A_j^c\right) \\
&= \sum_{i=1}^n E(X_i \mid A_1^c \dots A_n^c) \\
&= \sum_{i=1}^n E(X_i \mid A_i^c) \\
&= \sum_{i=1}^n E(X_i \mid X_i \geq \text{VaR}_p(X_i)) \\
&= \sum_{i=1}^n CVaR_p(X_i).
\end{aligned} \tag{3.14}$$

□

Note that for a univariate random variable X , $\text{wCoVaR}_p(X)$ is the same as $CVaR_p(X)$.

Theorem 3.6.2. *Let $X, Y \in R^n$ be random vectors with finite expectations. Then the worst-case CoVaR ($wCoVaR$) exhibits the following properties:*

(1) $wCoVaR_p$ is translation-invariant:

$$wCoVaR_p(X + c) = wCoVaR_p(X) + c, \quad c \in R$$

(2) $wCoVaR_p$ is positively homogeneous:

$$wCoVaR_p(cX) = c \times wCoVaR_p(X), \quad c \in R_+.$$

(3) $wCoVaR_p$ is subadditive: $wCoVaR_p(X + Y) \leq wCoVaR_p(X) + wCoVaR_p(Y)$,

when all $2n$ components in X and Y are independent.

(4) $wCoVaR_p$ is monotonic with respect to the second order stochastic dominance:

$$X \prec_{SD(2)} Y \text{ implies } wCoVaR_p(X) \leq wCoVaR_p(Y),$$

when all $2n$ components of X and Y are independent.

(5) $wCoVaR_p$ is additive in the sense that

$$wCoVaR_p((X, Y)^T) = wCoVaR_p(X) + wCoVaR_p(Y),$$

when all $n + m$ components in $X \in R^n$ and $Y \in R^m$ are independent.

Proof. The proofs of (1) and (2) are simple and therefore omitted.

(3) By Theorem 3.6.1, if all the components of the random vectors are independent, then $wCoVaR_p(X) = \sum_{i=1}^n CVaR_p(X_i)$ and $wCoVaR_p(Y) = \sum_{i=1}^n CVaR_p(Y_i)$. Conditional (or Average) Value-at-Risk satisfies subadditivity, one of the coherence axioms of Artzner et al (1999), which implies that $\sum_{i=1}^n CVaR_p(X_i + Y_i) \leq \sum_{i=1}^n CVaR_p(X_i) + \sum_{i=1}^n CVaR_p(Y_i)$. Thus it can be written as follows.

$$\begin{aligned} wCoVaR_p(X + Y) &= \sum_{i=1}^n CVaR_p(X_i + Y_i) \\ &\leq \sum_{i=1}^n CVaR_p(X_i) + \sum_{i=1}^n CVaR_p(Y_i) \\ &= wCoVaR_p(X) + wCoVaR_p(Y). \end{aligned}$$

(4) About multivariate stochastic orders we refer the reader to the literature, e.g., Müller and Stoyan (2002), etc. If X is second order stochastically dominated by Y , that is, $X \prec_{SD(2)} Y$, together with the independence assumption on the all components of X and Y , then $X_i \prec_{SD(2)} Y_i$, $i = 1, \dots, n$, which implies $CVaR_p(X_i) \leq CVaR_p(Y_i)$, $i = 1, \dots, n$ (see Pflug (2000)). By Theorem 3.6.1, we have $wCoVaR_p(X) \leq wCoVaR_p(Y)$ if $CVaR_p(X_i) \leq CVaR_p(Y_i)$, $i = 1, \dots, n$.

- (5) It is a simple consequence of Theorem 3.6.1. When all components of the random vectors $X \in R^n$ and $Y \in R^m$ are independent of each other, we have

$$\begin{aligned}
\text{wCoVaR}_p((X, Y)^T) &= \text{wCoVaR}_p((X_1, \dots, X_n, Y_1, \dots, Y_m)^T) \\
&= \sum_{i=1}^n \text{CVaR}_p(X_i) + \sum_{i=1}^m \text{CVaR}_p(Y_i) \\
&= \text{wCoVaR}_p((X_1, \dots, X_n)^T) + \text{wCoVaR}_p((Y_1, \dots, Y_m)^T) \\
&= \text{wCoVaR}_p(X) + \text{wCoVaR}_p(Y).
\end{aligned}$$

□

Remark 6 (Geometrical relationships between wCoVaR and VaR). *For a random vector $X \in R^n$, if the worst-case set W in (3.10) is projected onto spaces of $X_j \in R$, $j = 1, \dots, n$, then the lower bound of the projection onto the space of $X_j \in R$ is $\text{VaR}_p(X_j)$, $j = 1, \dots, n$. The projection of W is illustrated in Figure 3.4 for the 2-dimensional case. In Figure 3.4, the sets $\{z \mid z > \text{VaR}_p(X_1)\}$ and $\{z \mid z > \text{VaR}_p(X_2)\}$ are the projections of $W = \{x \in R^2 \mid x_1 > \text{VaR}_p(X_1), x_2 > \text{VaR}_p(X_2)\}$ onto the horizontal and vertical axes, respectively. If the random vector $X \in R^2$ has independent components X_1 and X_2 , then, by Theorem 3.6.1, the following equation holds true: $\text{wCoVaR}_p(X) = \text{CVaR}_p(X_1) + \text{CVaR}_p(X_2)$. Figure 3.4 can be used to illustrate this.*

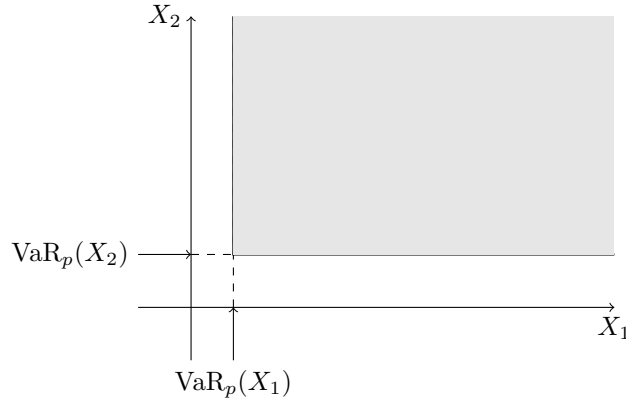


Figure 3.4: Description of projections of the set W
Projections of the set W , i.e., the worst case outcome as in (3.10) of the loss random vector $X = (X_1, X_2)^T$, from \mathbf{R}^2 onto the space of $\mathbf{X}_1, \mathbf{X}_2 \in \mathbf{R}$. The points $\text{VaR}_p(X_1)$ and $\text{VaR}_p(X_2)$ are the lower bounds of the projection of the set $W \in R^2$, onto the space of $X_1, X_2 \in R$, respectively.

Remark 7 (Individual Probability levels of wCoVaR). *The probability level for a set of multiple assets is determined by the probability levels for individual assets, as $P(X \in W)$, where $W = \{x \mid x_1 \geq \text{VaR}_p(X_1), \dots, x_n \geq \text{VaR}_p(X_n)\}$. If $P(X \in W)$ is too small (i.e., individual probability levels are large), then the wCoVaR associated with the corresponding individual p -level would not be a useful value in decision-making. This is because if the probability of the worst-case scenario is close to zero, then wCoVaR may become way too large. From a practical point of view, if the wCoVaR is used for the calculation for the reserve requirement of a financial institution, then it will result in an excessive reserve amount, and as a consequence, a negative effect on cash flow may occur. With this in mind, if we are dealing with a “high” dimensional case then “small” individual probability levels should be used to obtain a reasonable risk measurement.*

3.7 Decision-making via potential risk measure on corporate M&As

For decision-making on M&A deals, comparison of risk measurements *before and after* the M&A deal would be useful. Let us consider the simplest case: a univariate random variable corresponding to loss for each company, and a bivariate random vector meaning loss for a united firm. Let the random variables $X \in R$ and $Y \in R$ denote losses of Companies X and Y , respectively. Then $\text{CVaR}_p(X)$ and $\text{CVaR}_p(Y)$ indicate the amounts of losses of the respective Companies X and Y (i.e., the expected magnitude of losses beyond the $\text{VaR}_p(X)$, $\text{VaR}_p(Y)$, respectively). Suppose that M&A activity occurs between these companies and a new merged company is formed. Now let X and Y be random variables corresponding to losses of the business sectors X and Y of the new firm, respectively. If these business sectors are operating totally independently, then the risk associated with the new company should be $\text{CVaR}_p(X) + \text{CVaR}_p(Y)$: a simple sum of the risks, by Theorem 3.6.1.

However, if these are correlated, then the risk *after* the M&A deal is not the same, in general, as the simple sum of risks of the companies. In case of correlated losses from the business sectors of X and Y of the new company, risk measurement *before and after*

the deal would be of either

$$\begin{aligned} \text{superadditivity:} \quad & \text{wCoVaR}_p((X, Y)^T) \geq \text{CVaR}_p(X) + \text{CVaR}_p(Y), \text{ or} \\ \text{subadditivity:} \quad & \text{wCoVaR}_p((X, Y)^T) \leq \text{CVaR}_p(X) + \text{CVaR}_p(Y). \end{aligned} \tag{3.15}$$

Between the two cases in (3.15), the case of subadditivity would be desirable for M&A deals, while the case of superadditivity is ideal for demerger (split-up) activities. This is because the LHS of (3.15) indicates a risk measurement of a united firm, and each term of the RHS of (3.15) represents that of a single company before an M&A event (or after a demerger activity). The same way of reasoning applies to an M&A event for multi-plant firms. It is advisable to require for a decision that the new risk measure should take better value than a combination of individual risk measures (e.g., sum of two or more) for every $p \geq p_0$, where p_0 can be chosen based on the problem.

It is important to note that the risk evaluation process must be very detailed. For example, all business sectors that belong to the companies should be included and analyzed in the evaluation process, especially for their correlation structure. Furthermore, there are many types of risk evaluations useful for M&A deals, including strategic risk, compliance risk, operational risk, financial risk, reputation risk, etc., and each type of risk needs to be evaluated by a suitable risk measurement.

In the real world, the estimation of operational risk is complex and necessitates input from subject matter experts. Indeed, risk measurement is just one of the various key factors to analyze for in M&A decision-making processes. There are many other key factors to be considered as well, such as acculturation, human resource issues, post-marketing, and taxation to name just a few. For more about corporate M&A activities from a general point of view, we refer the reader to the literature (e.g., see Gregoriou and Renneboog (2007), Shimizu et al (2004), Nahavandi and Malekzadeh (1988), Walsh (1988), Berger et al (1998), Erickson (1998), Hagedoorn and Duysters (2002), Wulf and Singh (2011), etc.).

In section 3.8 numerical examples of corporate M&A deals with the comparisons of the risk measurements as in (3.15) are presented and discussed.

3.8 Numerical examples and discussion

Adequate measure of potential risk in a corporate M&A (Mergers and Acquisitions) is essential. In this section we want to show, with illustrative examples of corporate M&A deals, how the worst-case Multivariate Individual Value-at-Risk (wCoVaR) plays a role in risk evaluation processes for various cases in terms of the stochastic dependence structure. Examples 5 and 6 present the cases of corporate M&A deals.

Example 5. *Company A, a local newspaper company, wants to expand its business scope through a good M&A deal before the fourth quarter (Q4) of the year. Suppose that there are six candidates for the deal labeled by 1, 2, 3, 4, 5 and 6, local business competitors of Company A. The risk management arm of Company A wants to gauge and compare the risks to select the best case, in terms of risk, for the decision-making. The expected Q4 operational profits for the companies are: 4% for Company A, -3% for Company 1, -4% for Company 2, 0.5% for Company 3, 2% for Company 4, 3% for Company 5 and 5% for Company 6.*

Let X denote the Q4 operational loss of Company A, a normal random variable with mean -4% and variance 1, i.e., $X \sim N(-0.04, 1)$. Let Y_i , $i = 1, 2, 3, 4, 5, 6$ be the normal random variables with

$$\begin{aligned} Y_1 &\sim N(0.03, 0.7^2), & Y_2 &\sim N(0.04, 1.7^2), & Y_3 &\sim N(-0.005, 2^2), \\ Y_4 &\sim N(-0.02, 1.6^2), & Y_5 &\sim N(-0.03, 1.8^2), & Y_6 &\sim N(-0.05, 1.5^2), \end{aligned} \quad (3.16)$$

meaning the Q4 operational losses of Companies 1, 2, 3, 4, 5 and 6, respectively.

Then the random vectors $(X, Y_i)^T \in R^2$, $i = 1, 2, 3, 4, 5, 6$ are assumed to have bivariate normal distributions, interpreted as Q4 operational losses of business units of the newly united firms. Suppose that the pairs (X, Y_i) , $i = 1, 2, 3, 4, 5, 6$ have correlation coefficients: $\rho = 0.9, -0.6, 0, 0.5, -0.3, 0.7$, respectively. Then we have

$$\begin{aligned} \text{Cov}(X, Y_1) &= (0.9)(0.7) = 0.63, & \text{Cov}(X, Y_2) &= (-0.6)(1.7) = -1.02, \\ \text{Cov}(X, Y_3) &= 0, & \text{Cov}(X, Y_4) &= (0.5)(1.6) = 0.8, \\ \text{Cov}(X, Y_5) &= (-0.3)(1.8) = -0.54, & \text{Cov}(X, Y_6) &= (0.7)(1.5) = 1.05. \end{aligned} \quad (3.17)$$

To make a decision regarding all the possible M&A events, especially in a risk evaluation process, two types of comparisons could be useful:

(a) Which M&A deal is ideal in terms of the risk?

We need to choose the smallest value among $\text{wCoVaR}_p((X, Y_i)^T)$, $i = 1, \dots, 6$, i.e., compare the risk measurements *after* the M&A deals for all cases.

(b) Which M&A deals reduce risk?

We need to check which case will reduce the risk through the M&A deal, or equivalently, check if $\text{wCoVaR}_p((X, Y_i)^T) < \text{CVaR}_p(X) + \text{CVaR}_p(Y_i)$, $i = 1, \dots, 6$.

This is the comparison of the risk measurements *before and after* for the M&A deal.

Let us use the same individual probability levels for each company. Then for each pair of merged companies the worst-case Multivariate Individual Value-at-Risk (wCoVaR) can be calculated by:

$$\begin{aligned} \text{wCoVaR}_p((X, Y_i)^T) &= E(X + Y_i \mid X > \text{VaR}_p(X), Y_i > \text{VaR}_p(Y_i)) \\ &= \frac{\int_{\text{VaR}_p(X)}^{\infty} \int_{\text{VaR}_p(Y_i)}^{\infty} (x + y) f(x, y) \, dx dy}{\int_{\text{VaR}_p(X)}^{\infty} \int_{\text{VaR}_p(Y_i)}^{\infty} f(x, y) \, dx dy}. \end{aligned} \quad (3.18)$$

In (3.18), f is the bivariate standard normal p.d.f. of (X, Y_i) , $i = 1, \dots, 6$ with expectations in (3.16) and covariances (3.17).

Table 3.1: Risk after M&A deals, i.e., risk of the new united firms

The worst-case CoVaR (wCoVaR), i.e., risk after M&As					
Individual p -levels:	$p = 0.60$	$p = 0.70$	$p = 0.80$	$p = 0.90$	$p = 0.95$
$\text{wCoVaR}_p(X, Y_1)$	1.79025356	2.12712366	2.54391028	3.15397777	3.67959583
$\text{wCoVaR}_p(X, Y_2)$	1.89906848	2.42685181	3.10301480	4.11400437	4.99177280
$\text{wCoVaR}_p(X, Y_3)$	2.85256302	3.43191854	4.15441809	5.21993051	6.14310303
$\text{wCoVaR}_p(X, Y_4)$	2.72411371	3.23586774	3.86782594	4.79169631	5.58741982
$\text{wCoVaR}_p(X, Y_5)$	2.33154399	2.87028090	3.54970143	4.55937126	5.43700554
$\text{wCoVaR}_p(X, Y_6)$	2.61648455	3.11191386	3.72329776	4.61625343	5.38458160

Table 3.2: Sum of the risks of the companies before M&As

Risk of companies before the merger					
p -levels:	$p = 0.60$	$p = 0.7$	$p = 0.8$	$p = 0.9$	$p = 0.95$
$\text{CVaR}_p(X) + \text{CVaR}_p(Y_1)$	1.63195576	1.96025814	2.36967632	2.97347163	3.49661176
$\text{CVaR}_p(X) + \text{CVaR}_p(Y_2)$	2.60781203	3.12923346	3.77948583	4.73845478	5.56932425
$\text{CVaR}_p(X) + \text{CVaR}_p(Y_3)$	2.85256302	3.43191853	4.15441808	5.21993051	6.14310303
$\text{CVaR}_p(X) + \text{CVaR}_p(Y_4)$	2.45122645	2.95333598	3.57950495	4.50295660	5.30305326
$\text{CVaR}_p(X) + \text{CVaR}_p(Y_5)$	3.11719163	3.75444932	4.54913547	5.72102543	6.73620487
$\text{CVaR}_p(X) + \text{CVaR}_p(Y_6)$	2.32464083	2.80743845	3.40952400	4.29745828	5.06678200

Table 3.3: Risk before M&As for each company

Risk of each company					
p -levels:	$p = 0.60$	$p = 0.7$	$p = 0.8$	$p = 0.9$	$p = 0.95$
Company \mathcal{A} : $\text{CVaR}_p(X)$	0.92585633	1.11897538	1.35980960	1.71498331	2.02271280
Company 1: $\text{CVaR}_p(Y_1)$	0.70609943	0.84128276	1.00986672	1.25848832	1.47389896
Company 2: $\text{CVaR}_p(Y_2)$	1.68195571	2.01025808	2.41967623	3.02347147	3.54661145
Company 3: $\text{CVaR}_p(Y_3)$	1.92670669	2.31294315	2.79460848	3.50494720	4.12039023
Company 4: $\text{CVaR}_p(Y_4)$	1.52537012	1.83436060	2.21969535	2.78797329	3.28034046
Company 5: $\text{CVaR}_p(Y_5)$	2.19133530	2.635473941	3.18932587	4.00604212	4.71349207
Company 6: $\text{CVaR}_p(Y_6)$	1.39878450	1.68846307	2.04971440	2.58247497	3.04406920

Table 3.4: Probability of the Worst Case Outcome

1- probability of the worst case outcome					
Individual p -levels:	$p = 0.60$	$p = 0.70$	$p = 0.80$	$p = 0.90$	$p = 0.95$
Case of (X, Y_1)	0.66947743	0.76241307	0.85006754	0.93113505	0.96813223
Case of (X, Y_2)	0.93272725	0.97723374	0.99550175	0.99976102	0.99998881
Case of (X, Y_3)	0.84000011	0.91000008	0.96000005	0.99000002	0.99750001
Case of (X, Y_4)	0.76087275	0.84323267	0.91284943	0.96759847	0.98781057
Case of (X, Y_5)	0.88492698	0.94497326	0.98094402	0.99700071	0.99954134
Case of (X, Y_6)	0.72237245	0.80948141	0.88709824	0.95322102	0.98040069

We calculate wCoVaR at the same individual probabilities (two decimal places) from 0.50 to 0.99, and plot the results on the subfigures of Figure 3.5. Note that the wCoVaR is a function of the individual probability levels. For clarification, risk *after* M&As, i.e., wCoVaR at $p = 0.6, 0.7, 0.8, 0.9, 0.95$ are summarized in Table 3.1. In Table 3.2, risks *before* M&As, i.e., $\text{CVaR}_p(X) + \text{CVaR}_p(Y_i), i = 1, \dots, 6$ are also presented. We have obtained the following risk measurements for *before and after* the M&A deals between Companies \mathcal{A} and its six target companies. For the M&As with Companies 1, 4 and 6, we have *superadditive* relationships:

$$\text{wCoVaR}_p((X, Y_i)^T) > \text{CVaR}_p(X) + \text{CVaR}_p(Y_i), \quad i = 1, 4, 6, \quad (3.19)$$

for M&A activity with Company 3, we have additive relationship:

$$\text{wCoVaR}_p((X, Y_3)^T) = \text{CVaR}_p(X) + \text{CVaR}_p(Y_3), \quad (3.20)$$

and for the M&As with Companies 2 and 5, we have *subadditive* relationships:

$$\text{wCoVaR}_p((X, Y_i)^T) < \text{CVaR}_p(X) + \text{CVaR}_p(Y_i), \quad i = 2, 5. \quad (3.21)$$

From a risk management perspective, cases of (3.21) are desirable since the mergers result in a risk reduction. These ideal cases can also be found from subfigures (b) and (e) of Figure 3.5. If decision will be made solely on the risk measurements, then (b) will be preferred to (e) since $\text{wCoVaR}_p(X, Y_2) < \text{wCoVaR}_p(X, Y_5)$.

Note the case of the M&A between Companies \mathcal{X} and 1 could also be beneficial, since $w\text{CoVaR}_p(X, Y_1)$ has the smallest value among all cases. In this respect, a thorough case study over the pairs of (X, Y_1) and (X, Y_2) will help decision makers to find a better deal and take action for the successful integration. In this example, we can see a connection between $w\text{CoVaR}$ and CVaR . From Tables 3.1 and 3.2 and equation (3.20), we observe that the values of $w\text{CoVaR}$ in the case of $\rho = 0$ are just a sum of CVaR of each company (Zero correlation does not imply independence, in general.). By the additive relationship (3.20), Theorem 3.6.1 can be checked numerically.

In a setting of more than two variables, dependency structure would be more complex as we observe from Example 6.

Example 6. (Example 1 Continued) *Suppose that Company \mathcal{A} , the local newspaper company in the previous example, were able to grow throughout the M&A deal. This company is now considering acquiring the other media companies running other types of business, in order to become a media group. The target companies are as follows. Company 1 runs a magazine business, Company 2 is a radio broadcasting company, Company 3 has web-based technology and Company 4 does digital TV broadcasting service.*

Company \mathcal{A} wants to acquire two of these companies before the second half of the year. The half-year business forecasts for Company \mathcal{A} and Companies k , $k = 1, 2, 3, 4$ are 2% profit with standard deviation 0.8, -3% profit with standard deviation 0.6, 1.5% profit with standard deviation 1.2, 7% profit with standard deviation 2.3, and 9% profit with standard deviation 2.5, respectively. The risk management center of Company \mathcal{A} wants to compare the risk measurements before and after the all possible M&A activities, in order to use the result as one of the key factors of the decision-making on the event.

Let X , Z_k , $k = 1, 2, 3, 4$ denote normally distributed random variables, meaning the half-year operational losses of Companies \mathcal{A} and k , $k = 1, 2, 3, 4$, respectively:

$$\begin{aligned} X &\sim N(-0.02, 0.8^2), \\ Z_1 &\sim N(0.03, 0.6^2), Z_2 \sim N(-0.015, 1.2^2), Z_3 \sim N(-0.07, 2.3^2), Z_4 \sim N(-0.09, 1.9^2). \end{aligned} \quad (3.22)$$

Let us further assume for the correlation coefficients between the companies that $\rho_{X, Z_1} = -0.7$, $\rho_{X, Z_2} = 0.2$, $\rho_{X, Z_3} = 0.5$, $\rho_{X, Z_4} = -0.6$, $\rho_{Z_1, Z_2} = 0.3$, $\rho_{Z_1, Z_3} = -0.2$, $\rho_{Z_1, Z_4} =$

$0.6, \rho_{Z_2, Z_3} = -0.5, \rho_{Z_2, Z_4} = -0.2, \rho_{Z_3, Z_4} = -0.4$. Then the random vectors $(X, Z_i, Z_j)^T \in R^3$, $i \neq j$, $i, j = 1, 2, 3, 4$ mean the set of operating losses from the corresponding business sectors of the united firm after the M&As.

To make a decision regarding all the possible M&A events, especially in a risk evaluation process, two different types of comparison may be useful:

(a) Which M&A deal is ideal in terms of the risk?

We need to choose the smallest value among $wCoVaR_p((X, Z_i, Z_j)^T)$, $i \neq j$, $i = 1, 2, 3, 4$, i.e., compare the risk measurements *after* the M&A deals for all cases.

(b) Which M&A deals reduce risk?

We need to check if

$wCoVaR_p((X, Z_i, Z_j)^T) < CVaR_p(X) + CVaR_p(Z_i) + CVaR_p(Z_j)$, $i \neq j$, $i, j = 1, 2, 3, 4$. This is the comparison of the risk measurements *before and after* for the M&A deal.

Let us use the same probability levels for individual companies. Then for each case (X, Z_i, Z_j) , $i \neq j$, $i = 1, 2, 3, 4$, the worst-case Multivariate Individual Value-at-Risk ($wCoVaR$) can be calculated by:

$$\begin{aligned} wCoVaR_p((X, Z_i, Z_j)^T) &= E(X + Z_i + Z_j \mid X > VaR_p(X), Z_i > VaR_p(Z_i), Z_j > VaR_p(Z_j)) \\ &= \frac{\int \int \int_W (x + y + z) f(x, y, z) \, dx dy dz}{\int \int \int_W f(x, y, z) \, dx dy dz}, \end{aligned} \quad (3.23)$$

where $W = \{X > VaR_p(X), Z_i > VaR_p(Z_i), Z_j > VaR_p(Z_j)\}$.

In (3.23), f is the trivariate standard normal p.d.f. with expectation vector $\boldsymbol{\mu}$ and covariance matrix C_k , $k = 1, \dots, 6$ in (3.24) according to the random vectors (X, Z_i, Z_j) , $i \neq j$, $i, j = 1, 2, 3, 4$ as follows.

$$\begin{aligned}
(X, Z_1, Z_2)^T \text{ has } \boldsymbol{\mu}_1 &= \begin{pmatrix} -0.020 \\ 0.030 \\ -0.015 \end{pmatrix} \text{ and } C_1 = \begin{pmatrix} 0.640 & -0.336 & 0.192 \\ -0.336 & 0.360 & 0.216 \\ 0.192 & 0.216 & 1.440 \end{pmatrix}, \\
(X, Z_1, Z_3)^T \text{ has } \boldsymbol{\mu}_2 &= \begin{pmatrix} -0.020 \\ 0.030 \\ -0.070 \end{pmatrix} \text{ and } C_2 = \begin{pmatrix} 0.640 & -0.336 & 0.920 \\ -0.336 & 0.360 & -0.276 \\ 0.920 & -0.276 & 5.290 \end{pmatrix}, \\
(X, Z_1, Z_4)^T \text{ has } \boldsymbol{\mu}_3 &= \begin{pmatrix} -0.020 \\ 0.030 \\ -0.090 \end{pmatrix} \text{ and } C_3 = \begin{pmatrix} 0.640 & -0.336 & -1.200 \\ -0.336 & 0.360 & 0.900 \\ -1.200 & 0.900 & 3.610 \end{pmatrix}, \\
(X, Z_2, Z_3)^T \text{ has } \boldsymbol{\mu}_4 &= \begin{pmatrix} -0.020 \\ -0.015 \\ -0.070 \end{pmatrix} \text{ and } C_4 = \begin{pmatrix} 0.640 & 0.192 & 0.920 \\ 0.192 & 1.440 & -1.380 \\ 0.920 & -1.380 & 5.290 \end{pmatrix}, \\
(X, Z_2, Z_4)^T \text{ has } \boldsymbol{\mu}_5 &= \begin{pmatrix} -0.020 \\ -0.015 \\ -0.090 \end{pmatrix} \text{ and } C_5 = \begin{pmatrix} 0.640 & 0.192 & -1.200 \\ 0.192 & 1.440 & -0.600 \\ -1.200 & -0.600 & 3.610 \end{pmatrix}, \\
\text{and } (X, Z_3, Z_4)^T \text{ has } \boldsymbol{\mu}_6 &= \begin{pmatrix} -0.020 \\ -0.070 \\ -0.090 \end{pmatrix} \text{ and } C_6 = \begin{pmatrix} 0.640 & 0.920 & -1.200 \\ 0.920 & 5.290 & -2.300 \\ -1.200 & -2.300 & 3.610 \end{pmatrix}.
\end{aligned} \tag{3.24}$$

The risk measurements in the cases of different dependency structures as in (3.24) are calculated at the same individual probabilities (two decimal places) from 0.50 to 0.99. Due to the same individual p -levels they can be plotted in 2-D as in the subfigures of Figure 5.5. We observe some interesting results from Figure 5.5 – the cases in the subfigures (a), (b), (d) are neither subadditive nor superadditive. So we can expect that the stochastic dependence structure will be complicated in the risk evaluation processes in practice, as we observe that some cases are neither superadditive nor subadditive. Depending on stochastic dependence structures, these properties coexist and are separated by a certain level of probability.

Numerical results at individual probability levels $p = 0.6, 0.7, 0.8, 0.9, 0.95$ are summarized in Tables 3.5, 3.6, 3.7 and 3.8. At these individual probability levels, we have relationships of superadditivity (3.25) and subadditivity (3.26) – comparisons of risks *before* and *after* the M&As. As illustrated in the subfigures (a),(b),(d) of Figure 5.5

and presented in Tables 3.5 and 3.6, at high individual probability levels we have *superadditive* relationships:

$$\begin{aligned}
\text{wCoVaR}_p((X, Z_1, Z_2)^T) &> \text{CVaR}_p(X) + \text{CVaR}_p(Z_1) + \text{CVaR}_p(Z_2), p > 0.7 \\
\text{wCoVaR}_p((X, Z_1, Z_3)^T) &> \text{CVaR}_p(X) + \text{CVaR}_p(Z_1) + \text{CVaR}_p(Z_3), p > 0.8 \\
\text{wCoVaR}_p((X, Z_2, Z_3)^T) &> \text{CVaR}_p(X) + \text{CVaR}_p(Z_2) + \text{CVaR}_p(Z_3), p > 0.95
\end{aligned} \tag{3.25}$$

We have *subadditive* relationships for all other cases:

$$\text{wCoVaR}_p((X, Z_i, Z_j)^T) < \text{CVaR}_p(X) + \text{CVaR}_p(Z_i) + \text{CVaR}_p(Z_j), \quad i \neq j, \quad i, j = 1, 2, 3, 4. \tag{3.26}$$

Table 3.5: Risk after M&A deals, i.e., risk of the new united firms

The worst-case CoVaR (wCoVaR), i.e., risk after M&As					
Individual p -levels:	$p = 0.60$	$p = 0.70$	$p = 0.80$	$p = 0.90$	$p = 0.95$
$\text{wCoVaR}_p(X, Z_1, Z_2)$	2.49329287	3.11460792	3.91970243	5.14037649	6.21496657
$\text{wCoVaR}_p(X, Z_1, Z_3)$	3.42086954	4.22224287	5.22684633	6.69666789	7.94759881
$\text{wCoVaR}_p(X, Z_1, Z_4)$	1.98392938	2.64570195	3.49924961	4.77204115	5.87065544
$\text{wCoVaR}_p(X, Z_2, Z_3)$	3.70888752	4.56629510	5.66226146	7.30957330	8.75154331
$\text{wCoVaR}_p(X, Z_2, Z_4)$	2.42305069	3.20355632	4.20574281	5.69781483	6.98608444
$\text{wCoVaR}_p(X, Z_3, Z_4)$	2.21334102	3.25538668	4.62309213	6.65213752	8.38476756

Table 3.6: Sum of the risks of the companies before M&As

Risk of companies before the merger					
p -levels:	$p = 0.60$	$p = 0.7$	$p = 0.8$	$p = 0.9$	$p = 0.95$
$\text{CVaR}_p(X) + \text{CVaR}_p(Z_1) + \text{CVaR}_p(Z_2)$	2.50622646	3.00833599	3.63450496	4.55795663	5.35805322
$\text{CVaR}_p(X) + \text{CVaR}_p(Z_1) + \text{CVaR}_p(Z_3)$	3.51354385	4.22805169	5.11907625	6.43304858	7.57134263
$\text{CVaR}_p(X) + \text{CVaR}_p(Z_1) + \text{CVaR}_p(Z_4)$	3.10732475	3.74461726	4.53936957	5.71144107	6.72694516
$\text{CVaR}_p(X) + \text{CVaR}_p(Z_2) + \text{CVaR}_p(Z_3)$	4.04805773	4.87843701	5.91396195	7.44103858	8.76397032
$\text{CVaR}_p(X) + \text{CVaR}_p(Z_2) + \text{CVaR}_p(Z_4)$	3.64183854	4.39500261	5.33425538	6.71943107	7.91957284
$\text{CVaR}_p(X) + \text{CVaR}_p(Z_3) + \text{CVaR}_p(Z_4)$	4.64915592	5.61471818	6.81882662	8.59452302	10.13286217

From a risk management perspective, the subadditive cases of (3.26) are ideal since the M&As result in a risk reduction. As we mentioned in Remark 3.6, at “high” individual probability levels we will have unrealistic (way too large) magnitude of risk since probability level for the worst-case event will be too high. In Table 3.8 such situation can be confirmed – individual probability levels $p = 0.6$ or 0.7 are only reasonable to be used. All the cases at lower individual p -levels are subadditive and so would be beneficial in terms of risk reduction. The last case depicted in subfigure (f) of Figure 5.5 would be the best case concerning efficiency of risk reduction throughout the M&As. However, a proper case study is necessary to find out logical and reasonable descriptions of why and how this results in risk reduction, in order to make an appropriate final decision on the deal.

Table 3.7: Risk before M&As for each company

Risk of each company					
p -levels:	$p = 0.60$	$p = 0.7$	$p = 0.8$	$p = 0.9$	$p = 0.95$
Company \mathcal{A} : $\text{CVaR}_p(X)$	0.75268507	0.90718030	1.09984768	1.38398666	1.63017025
Company 1: $\text{CVaR}_p(Z_1)$	0.60951380	0.72538523	0.86988576	1.08298999	1.26762768
Company 2: $\text{CVaR}_p(Z_2)$	1.14402760	1.37577046	1.66477152	2.09097998	2.46025537
Company 3: $\text{CVaR}_p(Z_3)$	2.15134501	2.59548615	3.14934281	3.96607194	4.67354471
Company 4: $\text{CVaR}_p(Z_4)$	1.74512586	2.11205173	2.56963612	3.24446443	3.82914722

Table 3.8: Probability of the Worst Case Outcome

1- probability of the worst case outcome					
Individual p -levels:	$p = 0.60$	$p = 0.70$	$p = 0.80$	$p = 0.90$	$p = 0.95$
Case of (X, Z_1, Z_2)	0.95860665	0.98867999	0.99845649	0.99996197	0.99999925
Case of (X, Z_1, Z_3)	0.96670203	0.99153104	0.99893355	0.99997599	0.99999955
Case of (X, Z_1, Z_4)	0.98374077	0.99775031	0.99989708	0.99999968	0.99999999
Case of (X, Z_2, Z_3)	0.92911062	0.97178483	0.99268452	0.99934013	0.99994519
Case of (X, Z_2, Z_4)	0.98899152	0.99868821	0.99995276	0.99999991	0.99999999
Case of (X, Z_3, Z_4)	0.99886459	0.99998574	0.99999999	0.99999999	0.99999999

Remark 8. *Note that M&As and demerger activities are like two sides of the same coin. Thus, these examples are also considered as the cases about corporate demerger events – corporate restructuring processes. For general corporate demerger activities, there could be several candidates in terms of how to split business sectors of the company. For example, if there are m different ways of restructuring, then we may need to choose the least risky case. By comparison of risks among different cases, we can see what would be the most desirable case in a demerger from a risk management perspective, and eventually it will be used to make a decision on the demerger event. For demerger events in the real world, analyses are required on each pertinent business sector and on the correlation among all the sectors. Then this should be followed by some suitable partition options for the inputs of the risk evaluation processes. Furthermore, a decision should not be made based only on risk measurement since demerger is a very complex process with various reasons behind it. For successful M&As and demergers, detailed post-event plans should be set up before taking action on the deals.*

3.9 Concluding remarks

Risk, in practice, may not be equal to a real valued random variable, rather, it is frequently represented by a finite collection of random variables, i.e., a random vector. A company typically has many different assets, portfolios, business sectors, exposed

to different kinds of randomness, influencing the overall behavior of the company. In order to characterize it, from the point of view of risk exposure, we need to work with the joint probability distribution of the random variables involved. Multivariate risk measures have already been introduced in the literature and the starting point of our investigation is the paper by Prékopa (2012), where the concepts of Multivariate-Value-at-Risk (MVaR) and Multivariate-Conditional-Value-at-Risk (MCVaR) are introduced and explored. See also Lee and Prékopa (2013), where new methods for numerical calculations of these concepts are presented.

In this paper we construct a further multivariate risk measure: the worst case Combined Value-at-Risk (wCoVaR), where only one orthant of the space represent unfavorable set and its vertex is at the vector with components equal to the individual VaR's. Properties of this risk measure are derived but they fail to satisfy the convexity inequality required by one of the risk measure axioms of Artzner et al (1999). In our opinion a risk measure should signal the advantages or disadvantages of corporate M&As and demergers rather than to always satisfy an axiom, however, attractive it is from a purely mathematical point of view; wCoVaR is constructed to serve this objective. In Section 3.7 we show in what way it can be used in practice. In the numerical examples we look at one company which considers M&As with one or two of a few target candidates, we calculate which M&A deal is ideal in terms of risk and which M&A deals reduce risk. To do the above analysis we have introduced vector operations, where we put together risk vectors to create new risk vector with increased number of components, to describe M&As, and split a risk vector into parts, to describe demergers.

Comparison of risks before and after M&As or demergers activities will be useful for decision-making before taking action on the deals. It is the job of the decision makers to find out reasonable justification for an increase or decrease in risks throughout the M&A deals, in order to make the best decision. We hope that these aspects of wCoVaR will attract interest for future research.

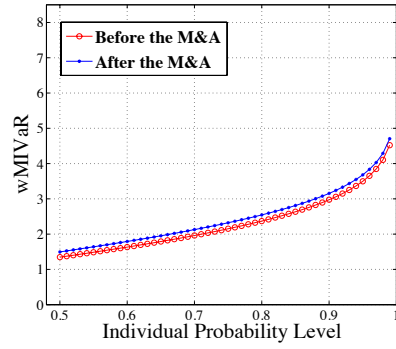
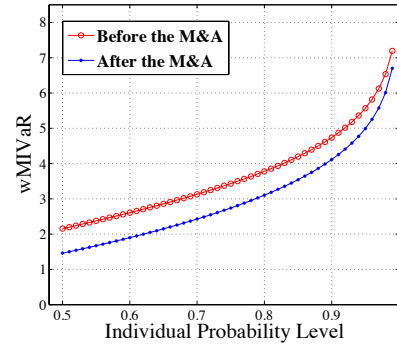
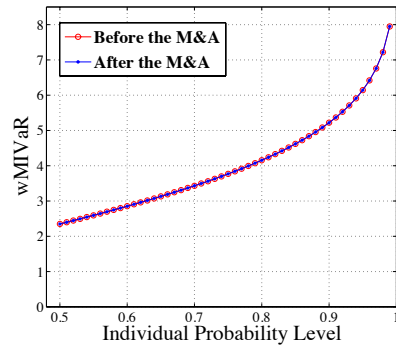
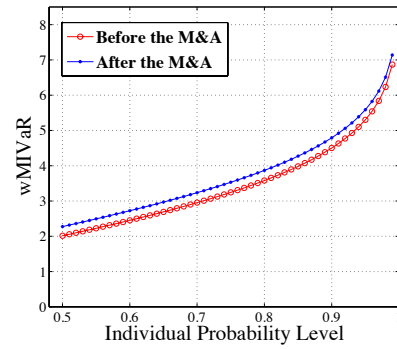
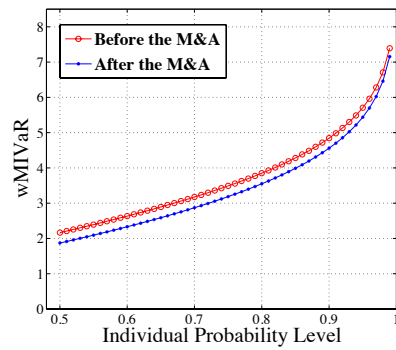
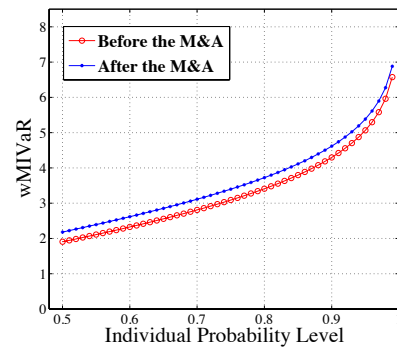
(a) Risk measurement of (X, Y_1) (b) Risk measurement of (X, Y_2) (c) Risk measurement of (X, Y_3) (d) Risk measurement of (X, Y_4) (e) Risk measurement of (X, Y_5) (f) Risk measurement of (X, Y_6)

Figure 3.5: Comparison of risks among different M&A deals

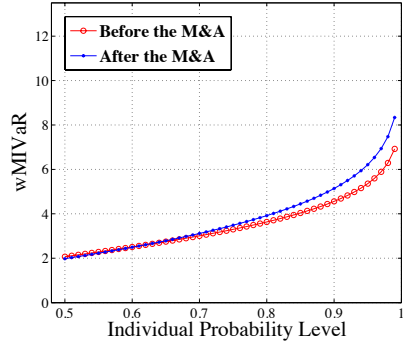
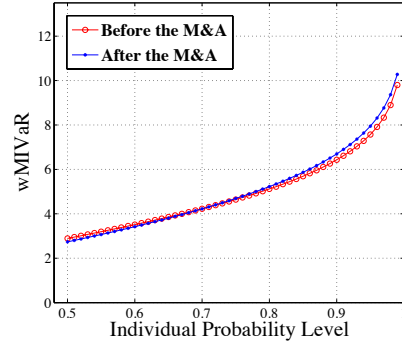
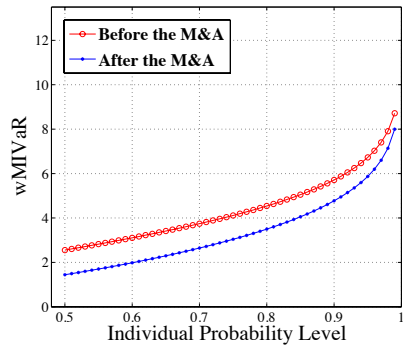
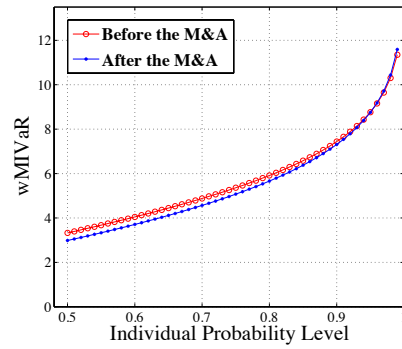
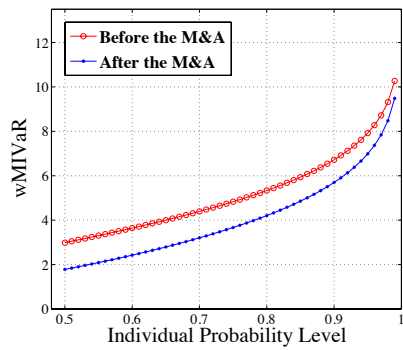
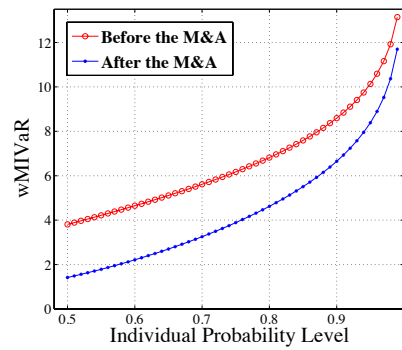
(a) Risk measurement of (X, Z_1, Z_2) (b) Risk measurement of (X, Z_1, Z_3) (c) Risk measurement of (X, Z_1, Z_4) (d) Risk measurement of (X, Z_2, Z_3) (e) Risk measurement of (X, Z_2, Z_4) (f) Risk measurement of (X, Z_3, Z_4)

Figure 3.6: Comparison of risks among different M&A deals

Chapter 4

Bond Portfolio Optimization for Insurance Companies

4.1 Introduction

The investment objectives vary by different financial entities. The objective of the insurance company is to earn a higher rate than that offered on the policy, that is, higher return than its cost. In this paper, the total amount of claims from random events is the only considered one as the financial cost. We also assume that the only possible payoff source for the claims is the cash flows (fixed income) from the investment in U.S treasury securities. The problem is that how many of the different bonds should be purchased for minimization of the cost. With the cash flows generated by the bonds invested, all the claims need to be payed out in the course of a given number of periods. Since the claims are uncertain, a suitable stochastic programming formulation is needed. Three stochastic programming models and their numerical solutions are presented in this paper.

Bonds are financial instruments that pay the owner fixed amounts of money, in subsequent periods. Insurance companies typically hold their money in bonds to pay out the random liabilities in the same periods. Insurance claims are randomly occurring events, which is considered as liabilities to an insurance company. The probability of a number of events occurring in a fixed period of time can be expressed as a Poisson distribution. Let X denote the number of events in the interval. If we let λ denote the expected number of events in the period, then we can write the probability that x events occur in the interval as:

$$P(X = x) = \frac{\lambda^x}{x!} e^{-\lambda}, \quad x = 0, 1, \dots$$

Suppose that N denote the number of events in the interval and let λ be its expected

number of events in a given period. And suppose that X_1, X_2, \dots, X_N denote the claim at each event $1, 2, \dots, N$, respectively. Then we can introduce the random sum $S = X_1 + X_2 + \dots + X_N$ which represents the aggregate claims of an insurance company in the interval. Then S has a compound Poisson distribution with the Poisson parameter λ . To introduce the suitable probability density function of S we need to assume:

- The positive claim amounts are positive integers,
- the claim size distribution is denoted by $p(i)$, $i = 1, 2, \dots$

If we have m periods, then there are m compound Poisson distributions. Simply depicted below is compound Poisson distributed liabilities in the course of different length of periods.

Now we need to compute the probability of S , the total of the incurred claims. Let $f(x) = P(S = x)$, $x = 0, 1, \dots$. Since a compound sum of integer claims has a recursive relationship, the recursive formula to compute the probability of S is De Pril (1986a):

$$f(x) = \frac{\lambda}{x} \sum_{i=1}^x i p(i) f(x-i), \quad x = 1, 2, \dots \quad (4.1)$$

$$f(0) = P(N = 0) = e^{-\lambda},$$

where $N \sim \text{Poisson}(\lambda)$, i.e., $P(N = k) = \frac{\lambda^k}{k!} e^{-\lambda}$.

4.2 Stochastic Programming Formulation

The problem is that how many of the different bonds should be purchased that minimizes the cost subject to the constraint that all liabilities can be payed out in the course of a given number of periods. Let us introduce the notations:

If the liabilities were deterministic values then our optimal bond portfolio model

n	number of bond types which are candidates for inclusion into the portfolio
m	number of periods
a_{ik}	cash flow of a bond of type k in period $i, k = 1, \dots, n$ and $i = 1, \dots, m$
p_k	unit price of bond of type k
r_k	remaining maturity of bond of type k
ξ_i	random liability value in period $i, i = 1, \dots, m$
x_k	decision variable, number of bonds of type k to include into the portfolio

would be the following Hodges and M. (1977)

$$\begin{aligned}
 & \min \left\{ \sum_{k=1}^n p_k x_k \right\} \\
 & \text{subject to} \\
 & \sum_{k=1}^n a_{ik} x_k \geq \xi_i, \quad i = 1, \dots, m \\
 & x_k \geq 0, \quad k = 1, \dots, n,
 \end{aligned} \tag{4.2}$$

where the positivity of the variables means no short-selling allowed.

The probabilistic constrained variant of it can be formulated as

$$\begin{aligned}
 & \min \left\{ \sum_{k=1}^n p_k x_k \right\} \\
 & \text{subject to} \\
 & P \left(\sum_{k=1}^n a_{ik} x_k \geq \xi_i, \quad i = 1, \dots, m \right) \geq p \\
 & x_k \geq 0, \quad k = 1, \dots, n
 \end{aligned} \tag{4.3}$$

where p is a safety(reliability) level chosen by ourselves, e.g., $p = 0.8, 0.9, 0.95$ etc.

4.3 Numerical Example and Its Solutions

4.3.1 Random Liabilities

We assume that the length of a period is 1 month and the expected number of events of incurred claims is 6, i.e., the Poisson parameter $\lambda = 6$. For simplicity, we also assume

that there are only two kinds of the claim size, which are 1 and 2 with corresponding probabilities $p(1) = 0.8$, $p(2) = 0.2$. Then the recursive formula to compute the probability of the total incurred claims, i.e., $P(S = x) = f(x)$ would be:

$$f(x) = \frac{6}{x} \sum_{i=1}^x ip(i)f(x-i), \quad x = 1, 2, \dots \quad (4.4)$$

$$f(0) = P(N = 0) = e^{-6},$$

where $N \sim \text{Poisson}(6)$, i.e., $P(N = k) = \frac{6^k}{k!}e^{-6}$. Equivalently,

$$\begin{aligned} f(0) &= e^{-6} \\ f(1) &= 6p(1)f(0) = 6(0.8)e^{-6} \\ f(2) &= \frac{6}{2}(p(1)f(1) + 2p(2)f(0)) = 3(0.8f(1) + 0.4e^{-6}) \\ f(3) &= \frac{6}{3}(p(1)f(2) + 2p(2)f(1) + 3p(3)f(0)) = 2(0.8f(2) + 2(0.2)f(1)) \\ &\vdots \end{aligned} \quad (4.5)$$

For finite number of possible values, computing $f(0), f(1), f(2), \dots$ up to a certain point where the probability already small, that is the point s such that $F(s) \approx 1$ gives us:

$$\begin{aligned} F(0) &= 0.0025 & F(10) &= 0.8559 \\ F(1) &= 0.0144 & F(11) &= 0.9090 \\ F(2) &= 0.0459 & F(12) &= 0.9451 \\ F(3) &= 0.1059 & F(13) &= 0.9682 \\ F(4) &= 0.1967 & F(14) &= 0.9823 \\ F(5) &= 0.3128 & F(15) &= 0.9905 \\ F(6) &= 0.4419 & F(16) &= 0.9951 \\ F(7) &= 0.5703 & F(17) &= 0.9976 \\ F(8) &= 0.6861 & F(18) &= 0.9988 \\ F(9) &= 0.7820 & F(19) &= 0.9994 \\ & & F(20) &= 0.9997 \end{aligned} \quad (4.6)$$

Since $F(20) = 0.9997 \approx 1$, we can choose $s = 20$ and we have set up the *p.d.f.* of ξ : $f(x), x = 0, 1, \dots, 20$.

Consider three periods. Suppose that all three periods are of the same length. Let ξ_i = total insurance claims in period i , $i = 1, 2, 3$. And assume that ξ_1 , ξ_2 and ξ_3 are independent. Since it is a total incurred claim amount in the interval, the possible values of ξ_i are $0, 1, \dots, 20$, $i = 1, 2, 3$. If we let \$1,000 be the unit claim size, then, for example, the probability that \$1,000 claim occurs in the interval is 1.19%, i.e.,

$$P(S = 1) = f(1) = 0.0119 = 1.19\%.$$

Thus, in this simple example, the possible claim dollar amount in one interval is $0, \$1,000, \dots, \$20,000$ with probability $f(0), f(1), \dots, f(20)$, respectively.

An insurance company needs to generate a cash flow to pay off such random claims. Let z_i = total cash flow in period i . Then the insurance company want to make sure about the following relationship:

$$P(\xi_1 \leq z_1, \xi_2 \leq z_2, \xi_3 \leq z_3) = p \approx 1,$$

where p is a safety(reliability) level chosen by ourselves, e.g., $p = 0.8, 0.9, 0.95$ etc.

If we let $f^i(x)$, $F^i(x)$ designate the *p.d.f* and *c.d.f.* of ξ for period $i = 1, 2, 3$, respectively, then we can write the following equations:

$$\begin{aligned} P(\xi_1 \leq z_1, \xi_2 \leq z_2, \xi_3 \leq z_3) &= \sum_{i \leq z_1, j \leq z_2, k \leq z_3} f^1(i) f^2(j) f^3(k) \\ &= F^1(z_1) F^2(z_2) F^3(z_3) \end{aligned} \quad (4.7)$$

Since the set of p -level efficient points serves as the p -quantile of the probability distribution determined by F . If we choose the safety(reliability) level $p = 0.9$, then we can generate the p -level efficient points (or PLEP) with $p = 0.9$.

Recall

Definition 4.3.1. *Prékopa (1990) A point $s \in R^r$ is said to be a p -level efficient point of the probability distribution F , if $F(s) \geq p$ and there is no y such that $y \leq s$, $y \neq s$, $F(y) \geq p$.*

and recall the following

Theorem 4.3.1. *Dentcheva et al (2000) If the components of the random vector ξ are integer-valued, then for any $p \in (0,1)$ the set of p -level efficient points is nonempty and finite. The set of p -level efficient points serves as the p -quantile of the probability distribution determined by F .*

Thus, the set of p -level efficient points of our example is nonempty and finite. Computation of PLEP's for this example ($p=0.9$) is as follows.

$$\begin{array}{ccccccccc}
 (11, 15, 20) & (11, 16, 16) & (11, 20, 15) & (12, 13, 15) & (12, 14, 14) & & & & \\
 (12, 15, 13) & (13, 12, 15) & (13, 13, 13) & (13, 15, 12) & (14, 12, 14) & & & & \\
 (14, 14, 12) & (15, 11, 20) & (15, 12, 13) & (15, 13, 12) & (15, 20, 11) & & & & \\
 (16, 11, 16) & (16, 16, 11) & (20, 11, 15) & (20, 15, 11) & & & & &
 \end{array} \tag{4.8}$$

4.3.2 Bond Portfolio construction

In the previous section, we enumerate 19 p -efficient points with $p = 0.9$. We have $s^{(i)} \in R^3$, $i = 1, \dots, 19$ such that $F(s) = F^1(s_1)F^2(s_2)F^3(s_3) \geq 0.9$ and there is no y such that $y \leq s$, $y \neq s$, $F(y) \geq 0.9$. The following relationship:

$$P(\xi_1 \leq z_1, \xi_2 \leq z_2, \xi_3 \leq z_3) \geq 0.9, \quad z_i = \text{total cash flow in period } i$$

needs to be satisfied for the insurance company.

Suppose that the company wants to construct a bond portfolio to generate the cash flows (z_i denote total amount of cash flows in period i) to pay off the random liabilities. For the sake of simplicity, we assume that bonds are offered with monthly coupon payments. Suppose that there are only 10 kinds of bonds to invest in as:

Based on the notations and the table above, the matrix for the cash flows can be written as:

$$A = \begin{pmatrix}
 5 & 10 & 15 & 20 & 25 & 30 & 35 & 1040 & 1045 & 1050 \\
 5 & 10 & 15 & 20 & 1025 & 1030 & 1035 & 0 & 0 & 0 \\
 5 & 10 & 1015 & 1020 & 0 & 0 & 0 & 0 & 0 & 0
 \end{pmatrix}$$

Note that the full principal amount of the bond of type 8,9,10 is included in the cash flow of period 1; that of the bond of type 5,6,7 is included in the cash flow of period 2;

Table 4.1: Prices, coupon payments and maturities of 10 bonds

	Price	Coupon Payment	Remaining Maturity
x_1	\$970	\$5	5 month
x_2	\$980	\$10	4 month
x_3	\$990	\$15	3 month
x_4	\$1,000	\$20	3 month
x_5	\$1,010	\$25	2 month
x_6	\$1,020	\$30	2 month
x_7	\$1,030	\$35	2 month
x_8	\$1,040	\$40	1 month
x_9	\$1,050	\$45	1 month
x_{10}	\$1,060	\$50	1 month

that of the bond of type 3,4 is included in the cash flow of period 3.

For the sake of consistency, we need to scale down the components of the matrix A as the same as the unit size of the random liability, which is \$1000. Multiplying A by $\frac{1}{1000}$ gives us:

$$T = \begin{pmatrix} .005 & .010 & .015 & .020 & .025 & .030 & .035 & 1.040 & 1.045 & 1.050 \\ .005 & .010 & .015 & .020 & 1.025 & 1.030 & 1.035 & 0 & 0 & 0 \\ .005 & .010 & 1.015 & 1.020 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad (4.9)$$

4.4 Stochastic Programming Models

The problem is how to construct bond portfolio with 10 different bonds (as in Section 3.2) that minimizes the cost at initial time subject to the constraint that all liabilities (as in Section 3.1) can be paid out in the course of three periods of the same length. We assume that the length of each period is 1 month and the expected number of events of incurred claims is 6, i.e., the Poisson parameter $\lambda = 6$. For simplicity, we also assume that there are only two kinds of the claim size, which are 1 and 2 with corresponding probabilities $p(1) = 0.8$, $p(2) = 0.2$, where the unit size is \$1,000. Our example can be described as follows.

It is solved in three ways; Probabilistic Constrained Model; Simple Recourse Model; Hybrid Model. The corresponding numerical solutions are presented in the following

three sections.

4.4.1 Probabilistic Constrained Model

The probabilistic constrained model is formulated as follows.

$$\begin{aligned}
 & \min \{c^T x\} \\
 & \text{subject to} \\
 & Tx \geq s^{(i)}, \text{ for at least one } i = 1, \dots, 19 \\
 & x \geq 0,
 \end{aligned} \tag{4.10}$$

where $x = (x_1, \dots, x_{10})$ the number of each type of 10 bonds to be invested; $c = (970, 980, 990, 1000, 1010, 1020, 1030, 1040, 1050, 1060, 1070)^T$, which designates the current price of each bond as in Table 3.1; the matrix T is as in (3.6); PLEP's, $s^{(i)}$, $i = 1, \dots, 19$ are as in (3.5).

The next form of the problem (3.7) is a relaxation of problem in which we take the convex hull of the p -efficient points and write up the first constraint in (3.7) by the use of the convex combination of $s^{(1)}, \dots, s^{(19)}$. The relaxed problem can be formulated as the following:

$$\begin{aligned}
 & \min \{c^T x\} \\
 & \text{subject to} \\
 & Tx \geq \sum_{i=1}^{19} \delta_i s^{(i)}, \text{ where } s^{(i)} = (s_1^{(i)}, s_2^{(i)}, s_3^{(i)})^T \\
 & \sum_{i=1}^{19} \delta_i = 1 \\
 & x \geq 0, \delta_i \geq 0, i = 1, \dots, 19.
 \end{aligned} \tag{4.11}$$

In the relaxed problem (3.8) the set $\left\{ \sum_{i=1}^{19} \delta_i s^{(i)} \mid \sum_{i=1}^{19} \delta_i = 1, \delta_i \geq 0 \right\}$ is the convex hull of the set of p -efficient points $\{s^{(1)}, \dots, s^{(19)}\}$.

As we solve the relaxed problem (3.8) using Matlab with the code attached as an

appendix, the optimal solution, i.e., the optimal bond portfolio comes up as:

$$x^* = (0, 0, 12.8079, 0, 12.4955, 0, 0, 12.0149, 0, 0),$$

which can be interpreted that the investment in bond of type 3, 5, 8 as the optimal solution x^* is for both the cost minimization and the liability payoff.

4.4.2 Simple Recourse Model

$$\begin{aligned} & \min \{c^T x + \sum_{i=1}^3 q_i^+ E([\xi_i - T_i x]_+)\} \\ & \text{subject to} \\ & x \geq 0, \end{aligned} \tag{4.12}$$

where $c = (970, 980, 990, 1000, 1010, 1020, 1030, 1040, 1050, 1060, 1070)^T$ and the second term in the objective function is the penalty. Let $q_1^+ = 1000$, $q_2^+ = 200$, $q_3^+ = 100$. Then this means that we pay 1,000, 200 and 100 dollars for each unit deviation of ξ_i from $T_i x$ such that $\xi_i > T_i x$, $i = 1, 2, 3$, respectively.

Theorem 4.4.1. *Prékopa (1995) Let $r = 1$. If the random vector (T, ξ) , where T is a $r \times n$ random matrix and ξ is an r -component random vector, has a discrete probability distribution with a finite number of possible values, then the function*

$$E([\xi_i - T_i x]_+),$$

is piecewise linear and convex in R^n .

Since we have the relation

$$E([\xi_i - T_i x]_+) = \int_{T_i x}^{\infty} [1 - F_i(z)] dz$$

where F_i is the c.d.f. of ξ_i , $i = 1, \dots, r$ and for any real number a we have that $[a]_+ - [-a]_+ = a$, it follows that

$$\begin{aligned} & q_i^+ E([\xi_i - T_i x]_+) + q_i^- ([T_i x - \xi_i]_+) \\ &= q_i^+ E([T_i x - \xi_i]_+) + q_i^+ (E(\xi_i) - T_i x) + q_i^- E([T_i x - \xi_i]_+) \\ &= ((q_i^+ + q_i^-) E([T_i x - \xi_i]_+) + q_i^+ (E(\xi_i) - T_i x)) \end{aligned} \tag{4.13}$$

We may omit the term $q_i^+ E(\xi)$ because it is constant. The new form of our problem is:

$$\begin{aligned}
 & \text{minimize} \quad \{c^T x + \sum_{i=1}^3 (q_i^+ \int_{-\infty}^{T_i x} F_i(z) dz - q_i^+ T_i x)\} \\
 & \text{subject to} \\
 & \quad x \geq 0,
 \end{aligned} \tag{4.14}$$

In our problem, the possible values of ξ_i are $0, \dots, 20$. and pick two values -50 and 50 with zero probability. Each ξ_i , $i = 1, 2, 3$ is set to have the 23 values of $-50, 0, 1, \dots, 20, 50$, i.e., $j = 0, \dots, 22$. If we replace $y_i = T_i x$ in problem (3.11) then we get

$$\begin{aligned}
 & \min \{c^T x + \sum_{i=1}^3 (-q_i^+ y_i + q_i^+ \int_{-\infty}^{y_i} F_i(z) dz)\} \\
 & \text{subject to} \\
 & \quad T_i x = y_i, \quad i = 1, \dots, 3 \\
 & \quad x \geq 0.
 \end{aligned} \tag{4.15}$$

Let

$$f_i(y_i) = -q_i^+ y_i + q_i^+ \int_{-\infty}^{y_i} F_i(z) dz \tag{4.16}$$

This function is piecewise linear, continuous and convex in the interval $[-50, 50]$. Hence we can represent $f_i(y_i)$ as the optimum value of the LP:

$$\begin{aligned}
 & \min \sum_{j=0}^{22} f_i(z_{ij}) \lambda_{ij} \\
 & \text{subject to} \\
 & \quad \sum_{j=0}^{22} z_{ij} \lambda_{ij} = y_i, \quad i = 1, \dots, 3 \\
 & \quad \sum_{j=0}^{22} \lambda_{ij} = 1, \lambda_{ij} \geq 0 \text{ for all } i, j.
 \end{aligned} \tag{4.17}$$

The next step is to combine (3.12) and (3.14) to get the LP:

$$\begin{aligned}
& \min \{c^T x + \sum_{i=1}^3 \sum_{j=0}^{22} f_i(z_{ij}) \lambda_{ij}\} \\
& \text{subject to} \\
& T_i x = y_i, \quad i = 1, \dots, 3 \\
& \sum_{j=0}^{22} z_{ij} \lambda_{ij} = y_i, \quad i = 1, \dots, 3 \\
& \sum_{j=0}^{22} \lambda_{ij} = 1 \\
& x \geq 0, \lambda_{ij} \geq 0 \text{ for all } i, j.
\end{aligned} \tag{4.18}$$

The final form of the problem is obtained if we remove the superfluous variables y_1, \dots, y_3 from (3.15) and introduce the notation: $c_{ij} = f_i(z_{ij})$. Then we get

$$\begin{aligned}
& \min \{c^T x + \sum_{i=1}^3 \sum_{j=0}^{22} c_{ij} \lambda_{ij}\} \\
& \text{subject to} \\
& T_i x - \sum_{j=0}^{22} z_{ij} \lambda_{ij} = 0, \quad i = 1, \dots, 3 \\
& \sum_{j=0}^{22} \lambda_{ij} = 1, \quad i = 1, \dots, 3 \\
& x_1, \dots, x_{10} \geq 0, \quad \lambda_{ij} \geq 0 \text{ for all } i, j,
\end{aligned} \tag{4.19}$$

Thus, the simple recourse problem with discrete ξ , where each component has finite support, can be reformulated as a specially structured LP.

When solved with the penalty constant $q_1^+ = 1000$, $q_2^+ = 200$, $q_3^+ = 100$, a trivial solution comes up, however.

The solution is $x^* = (0, \dots, 0)^T$.

It is because each ξ has a value of zero and the corresponding probability is: $f(0) = 0.0025$ as in (3.3). As we run the model, their corresponding λ coefficients, i.e., in $\lambda_{i1} f(0)$, are set to one and everything else is set to zero, i.e., $\lambda_{i1} = 1$ for all $i = 1, 2, 3$ and $\lambda_{ij} = 0$, for all $i = 1, 2, 3$ and $j \neq 1$. To satisfy the equality constraint, the model forces to let x be zero.

Let us consider about a huge penalty. With a huge penalty, the company gets the business shut down, if they cannot afford to pay off the claims from their clients. Thus, they need to do generate enough cash flows to pay off the claims with huge penalties. With $q_1^+ = 15000$, $q_2^+ = 10000$ and $q_3^+ = 5000$, the optimal solution becomes $x^* = (0, 0, 4.9262, 0, 4.8010, 0, 0, 4.6211, 0, 0)^T$.

4.4.3 Hybrid Model

We take the objective function from the penalty model, the probabilistic constraint from the probabilistic constrained model and then our model is:

$$\begin{aligned}
& \min \{c^T x + \sum_{i=1}^3 \sum_{j=0}^{22} f_i(z_{ij}) \lambda_{ij}\} \\
& \text{subject to} \\
& Tx \geq \sum_{i=1}^{19} \delta_i s^{(i)}, \text{ where } s^{(i)} = (s_1^{(i)}, s_2^{(i)}, s_3^{(i)})^T \\
& \sum_{i=1}^{19} \delta_i = 1 \\
& T_i x - \sum_{j=0}^{22} z_{ij} \lambda_{ij} = 0, \quad i = 1, 2, 3 \\
& \sum_{j=0}^{22} \lambda_{ij} = 1 \\
& x_1, \dots, x_{10} \geq 0, \quad \lambda_{ij} \geq 0 \text{ for all } i, j.
\end{aligned} \tag{4.20}$$

As in the section 3.3.1, with a safety (reliability) level $p = 0.9$, we have 19 p -level efficient points and the set $\left\{ \sum_{i=1}^{19} \delta_i s^{(i)} \mid \sum_{i=1}^{19} \delta_i = 1, \delta_i \geq 0 \right\}$ is the convex hull of the set of p -efficient points $\{s^{(1)}, \dots, s^{(19)}\}$.

As in the section 3.3.2, we penalize with $q_1^+ = 1000$, $q_2^+ = 200$, $q_3^+ = 100$. And, again, this means that we pay 1,000, 200 and 100 dollars for each unit deviation of ξ_i from $T_i x$ such that $\xi_i > T_i x$, $i = 1, 2, 3$, respectively.

The numerical solution of this model is:

$$\begin{aligned}
x^* &= (0, 0, 14.7783, 0, 12.4666, 0, 0, 11.0256, 0, 0) \\
\delta_4 &= 1 \\
\lambda_{1,9} &= 0.9048 \\
\lambda_{1,22} &= 0.952 \\
\lambda_{2,9} &= 0.8810 \\
\lambda_{2,22} &= 0.1190 \\
\lambda_{3,9} &= 0.8333 \\
\lambda_{3,22} &= 0.1667
\end{aligned} \tag{4.21}$$

Interestingly, the optimal solution $x^* = (0, 0, 14.7783, 0, 12.4666, 0, 0, 11.0256, 0, 0)$ is the same when the penalty of $q_1^+ = 15000$, $q_2^+ = 10000$, $q_3^+ = 5000$.

Chapter 5

Price-bands: A Technical Tool for Stock Trading

5.1 Introduction

Stock trading can be approached in a multitude of ways, and the one fact practitioners agree upon is that there is no clear and easy way to outperform the market consistently. It is intuitively appealing that a company should have an “intrinsic value,” and in the long run the stock price would converge to this value. This is the conventional wisdom of so-called “fundamental analysis,” and the valuation of the company amounts to estimating this unknown “intrinsic value.” This process is concerned mostly with the economic climate, interest rates, products, earnings, management, etc. There is a large number of valuation studies, and many of them are actually being used in financial practice.

On the other hand, technical (or quantitative) analysis is an evaluation process of securities based entirely on charting patterns, statistical approaches, and/or mathematical formulae. The technical analysis approach is particularly suited for short-term investing. In a certain sense, this amounts to an analysis of crowd psychology and behavior as well as investor philosophy, since it is believed that short-term patterns and trends result primarily from decisions by human investors. The tacit assumption underlying technical analysis is that a future price can be predicted by quantitative analysis of the past price movement. However, the efficient-market hypothesis (e.g., see Fama (1970), Damodaran (2012), etc.) asserts that no one can consistently outperform the market, since the market incorporates all information instantaneously. On the other hand, behavioral economists criticize the efficiency of the market for many reasons, such as irrationality of investors, information asymmetry, etc. For more details, we refer the reader to the literature, e.g., see Kahneman and Tversky (1979), Shleifer (2000), etc.

In this paper we suppose that stock prices are at least *partially* predictable based on recent market trends. Price patterns can be elusive, and the difficulty is amplified by the sheer complexity of the financial market, as well as market participants' philosophical and psychological states. We believe that price movement discloses investors' expectations in light of these (and many more) factors, and in this way accounts for them. There are some clear patterns in the stock market; for example, when a stock is in an up-trend with increasing volume, it is regarded as a sign of an up-market trend. Such patterns can be found by technical analysis, and this motivates the study of such methodologies. In what follows, we present a mathematical model for constructing new and reliable trading-bands (or price-bands) for price forecasting.

5.2 Construction of Price-bands

In the stock trading business, many different technical tools exist to guide traders through the swarm of information, e.g., trading bands, envelopes, channels, etc. Bollinger Bands is probably one of the most popular and successful models. Many traders use it daily as a tool for pattern recognition, augmenting technical trading strategies, and so on (see, e.g., Bollinger (2002), Grimes (2012), etc.). It is simple to see that Bollinger Bands is a collection of individual confidence intervals of future stock prices. Stock prices are expected to remain within the bands with a certain probability for each future time, depending on the width of the bands. Observation of stock price outside of the bands is considered a sign for “buying” or “(short) selling.”

Let us consider the following simple example: Suppose we happen to observe a stock's price moving below the lower band. Then our expectation would be that the price will go up, moving back into the bands, *ceteris paribus*. If we make our investment decision solely on the basis of price-bands, then we would choose to “buy” while the price is below the lower band and subsequently “sell” within the bands. Consider, on the other hand, the case that a stock price is observed above the upper band. As one might expect, the proper action on the stock would be to “short sell” above the upper band and then “buy” within the bands.

Table 5.1: The Bollinger Bands

Upper band = n -day moving average + $k\sigma$
Middle band = n -day moving average
Lower band = n -day moving average - $k\sigma$

There are several ways to calculate moving average, e.g., simple moving average, front-weighted moving average, exponential moving average, and so on. The band width is determined by the multiplier k and standard deviation $\sigma = \sqrt{\frac{\sum (x_i - \bar{x})^2}{N-1}}$, where x_i is the data point, μ the average, and N the number of points. The multiplier k can be chosen depending on the time periods n . Recommended (by Bollinger) width parameters with time periods are $k = 1.9$ if $n = 10$, $k = 2.0$ if $n = 20$, $k = 2.1$ if $n = 50$, etc.

For completeness we present the Bollinger Bands formulae in Table 5.1. Bollinger uses the mean and standard deviation to create price-bands as in Table 5.1, where the mean can be thought of as a central tendency and standard deviation as its volatility, thus determining the width of the bands. Like Bollinger, we assume the stock price in a time period to be a Gaussian process. However, rather than using simple mean and variance, we use conditional mean and conditional variance. By conditioning on a recent historical stock price data, we construct price-bands that are more sensitive to recent market information than Bollinger Bands.

Let stochastic process $X(s)$ be the stock price at time s , and let $I_\tau = \{X(\tau_1), X(\tau_2), \dots, X(\tau_N)\}$, $\tau_1 < \tau_2 < \dots < \tau_N$ be a sequence of past (τ_1, \dots, τ_N) stock prices (see Figure 5.1.). Then given the information set I_τ , the probability p of the future stock prices running within $[a_1, b_1], \dots, [a_l, b_l]$ at times $t_1 < \dots < t_l$, is the following:

$$p = P(a_i \leq X(t_i) \leq b_i, i = 1, \dots, l \mid I_\tau). \quad (5.1)$$

If we use the past N data points (equally spaced), say, in order to predict the future price changes for n time points, then (5.1) has the following practical meaning. With probability p , we can expect that the future stock price is likely to run within the lower and upper bounds $a_i, b_i, i = 1, \dots, l$. These bounds are paramount in trading strategy, and in this paper we present an efficient method of computing these values.

Let us assume that the stochastic process $\{X(t), t \geq 0\}$ is Gaussian. Let $\mu = (\mu_1, \dots, \mu_n)^T$ denote the expectation of the random vector $X = (X(t_1), \dots, X(t_n))^T$.

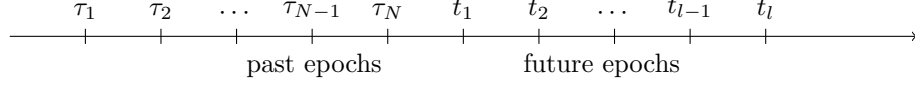


Figure 5.1: Past, present, future timeline

Then the covariance matrix $\Sigma = (\Sigma_{ij})$ is defined by

$$\Sigma_{ij} = E((X(t_i) - \mu_i)(X(t_j) - \mu_j)). \quad (5.2)$$

Let us recall that we have the closed form of p.d.f. of Gaussian models. Let random vector $X = X_1, \dots, X_p$, Σ its covariance matrix. Suppose that we are interested in the conditional distribution of $X_A = (X_1, \dots, X_k)$ given $X_B = (X_{k+1}, \dots, X_p)$. Let μ_A and μ_B denote the corresponding expectation vectors of X_A and X_B , respectively. We partition the covariance matrix Σ into

$$\Sigma = \begin{pmatrix} \Sigma_{AA} & \Sigma_{AB} \\ \Sigma_{BA} & \Sigma_{BB} \end{pmatrix}.$$

Then we have

$$X_A | X_B \sim N_k(\mu_A + \Sigma_{AB} \Sigma_{BB}^{-1} (X_B - \mu_B), \Sigma_{AA} - \Sigma_{AB} \Sigma_{BB}^{-1} \Sigma_{BA}),$$

where k denotes the number of dimension of the distribution function.

For a price-bands construction, let $X(\tau_j)$ denote the random stock prices on day j , $j = 1, \dots, m$, and let \mathbf{P} denote “Past.” We form a random vector:

$$\mathbf{X}^{\mathbf{P}} = \begin{pmatrix} X^P(\tau_1) \\ \vdots \\ X^P(\tau_m) \end{pmatrix}. \quad (5.3)$$

Let e_j denote the expected stock price on day i , $i = 1, \dots, m$. Then the expectation vector of (5.3), i.e., an m by 1 vector, on day j , $j = 1, \dots, m$ can be written as

$$\mathbf{e}^{\mathbf{P}} = E(\mathbf{X}^{\mathbf{P}}) = \begin{pmatrix} e_1^P \\ \vdots \\ e_m^P \end{pmatrix}. \quad (5.4)$$

For the future random variable, let us use F for “Future.” Then we form a future random vector:

$$\mathbf{X}^F = \begin{pmatrix} X^F(t_1) \\ \vdots \\ X^F(t_n) \end{pmatrix} \quad (5.5)$$

and expectation vector:

$$\mathbf{e}^F = E(\mathbf{X}^F) = \begin{pmatrix} e_1^F \\ \vdots \\ e_n^F \end{pmatrix}. \quad (5.6)$$

Using the notations (5.3), (5.4), (5.5) and (5.6), let us define components of covariance matrix of the random variables $X^P(\tau_j)$, $j = 1, \dots, m$, $X^F(t_i)$, $i = 1, \dots, n$ by

$$\begin{aligned} S &= E[(\mathbf{X}^F - \mathbf{e}^F)(\mathbf{X}^F - \mathbf{e}^F)^T] \\ U &= E[(\mathbf{X}^F - \mathbf{e}^F)(\mathbf{X}^P - \mathbf{e}^P)^T] \\ T &= E[(\mathbf{X}^P - \mathbf{e}^P)(\mathbf{X}^P - \mathbf{e}^P)^T]. \end{aligned} \quad (5.7)$$

Then the covariance matrix C can be written as

$$C = \begin{pmatrix} S & U \\ U^T & T \end{pmatrix}. \quad (5.8)$$

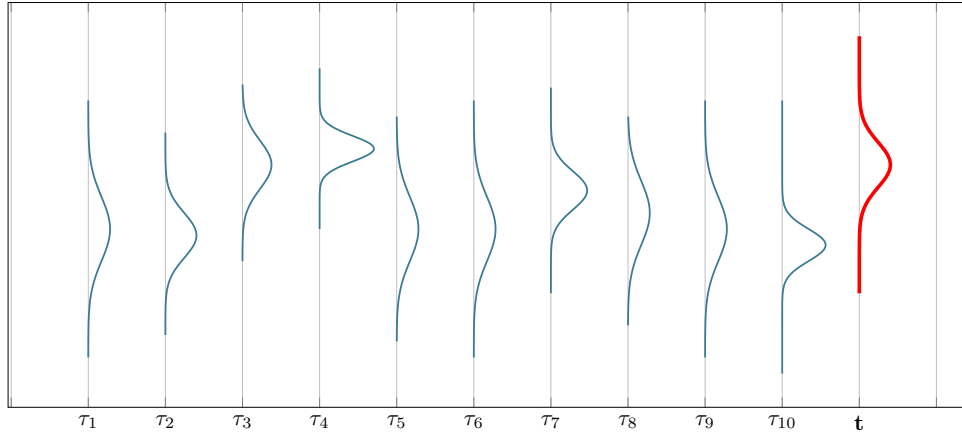


Figure 5.2: Description of Gaussian process

$X(\tau_1), X(\tau_2), \dots, X(\tau_{10})$, $\tau_1 < \tau_2 < \dots < \tau_{10}$ be a sequence of random variables of past stock prices on day τ_1, \dots, τ_{10} . We want to find reasonable lower and upper bounds of $X(\mathbf{t})$.

In our setting, given that $\mathbf{X}^P = \mathbf{x}^P$, \mathbf{X}^F has a normal distribution with “conditional” expectation vector

$$e^C = \mathbf{e}^F + UT^{-1}(\mathbf{x}^P - \mathbf{e}^P), \quad (5.9)$$

and covariance matrix

$$S - UT^{-1}U^T. \quad (5.10)$$

The conditional probability density of \mathbf{X}^F , given $\mathbf{X}^P = \mathbf{x}^P$, can be written up as

$$\begin{aligned} f(x^F \mid \mathbf{x}^P) &= \left[\frac{|(S - UT^{-1}U^T)^{-1}|}{(2\pi)^N} \right]^{1/2} \\ &\times \exp \left\{ -\frac{1}{2}(x^F - e^C)^T (S - UT^{-1}U^T)^{-1} (x^F - e^C) \right\}. \end{aligned} \quad (5.11)$$

We want to predict upper and lower bounds of a stock price for the next business day, based on historical data over a certain time period. Since we are interested in trading on the next day, \mathbf{X}^F is a random variable, not a random vector. As S is the variance of \mathbf{X}^F , it follows that S in the covariance matrix (5.8) is a number. Hence the conditional covariance matrix $S - UT^{-1}U^T$ is a number. Given $\mathbf{X}^P = \mathbf{x}^P$, let us denote the standard deviation of \mathbf{X}^F by

$$\sigma^C = \sqrt{S - UT^{-1}U^T}. \quad (5.12)$$

Then after the way of Bollinger, we can construct our trading bands, but with inputs of the conditional variance and expectation (5.12) and (5.9), respectively.

Definition 5.2.1 (Prékopa-Lee Bands). *Based on n intraday data points from the past m days, we construct next-day trading bands by predicting volatility $\sigma^C = \sqrt{S - UT^{-1}U^T}$ and average price $e^C = e^F + UT^{-1}(x^P - e^P)$ as follows:*

$$\begin{aligned} \text{Upper band} &= e^C + k\sigma^C \\ \text{Middle band} &= e^C \\ \text{Lower band} &= e^C - k\sigma^C, \end{aligned} \quad (5.13)$$

where k is a multiplier that can be chosen based on investors’ preference (or risk tolerance level), i.e., the larger k , the wider bands. Together with σ^C , k determines the width of the bands. The central tendency (i.e. e^C) determines the moving direction of

the bands.

To construct the bands, we need to find the values of $\sigma^C = \sqrt{S - UT^{-1}U^T}$ and $e^C = e^F + UT^{-1}(x^P - e^P)$. Since the stock price distribution of the next business day is unknown, there is no way to find the exact values of σ^C and e^C . However, given $X^P = \mathbf{x}^P$, the calculation of reasonable upper and lower bounds for the volatility σ^C is possible.

Among the terms in the expression for σ^C , the “Future” random variable X^F appears in $S = E[(\mathbf{X}^F - \mathbf{e}^F)(\mathbf{X}^F - \mathbf{e}^F)^T]$ and $U = E[(\mathbf{X}^F - \mathbf{e}^F)(\mathbf{X}^P - \mathbf{e}^P)^T]$. For the next-day variance S , we want to use the sample variance (as in Bollinger (2002)):

$$S = \frac{\sum_{i=1}^N (x_i^P - \bar{x})^2}{N - 1}, \quad (5.14)$$

where x_i^P , $i = 1, \dots, N$, are all the given data points (n intraday data points on each of the past m days, i.e., the total number of data points is mn), \bar{x} their mean, and N the number of data points, i.e., $N = mn$. In a short period of time, say up to 30 days, estimation of S by (5.14) is known to be acceptable to use in practice. More importantly, the success of Bollinger bands in practice proves (5.14) works just fine with a proper multiplier k .

Cross covariance U represents the relationship between past and future stock prices. In the short term (of, in this case, at most ten days), a myriad of factors can materially contribute to wildly varying degrees. For example, investor sentiment can, and often does, shift in dramatic fashion as herd psychology sweeps the market. When patterns do exist, distinguishing them in an actionable fashion is virtually impossible from historical time series data alone.

However, the cross-covariance matrix U is typically estimated from historical time series via linear regression. This approach effectively averages out the myriad of factors alluded to above that contribute to short-term fluctuations; indeed, this is part of the value of the approach in long-term forecasting. For the purpose of short-term technical trading, however, a more detailed approach is required. For this reason, effective estimation of $U = E[(\mathbf{X}^F - \mathbf{e}^F)(\mathbf{X}^P - \mathbf{e}^P)^T]$ is paramount to obtaining the

upper and lower bounds of the conditional volatility σ^C .

To this end, suppose we could choose the cross-covariance to maximize or minimize the conditional variance (next-day volatility). Given n intraday data points z_i , measured at equally spaced time points, then we have $n!$ possible permutations on their ordering. That is, we have the possible cases

$$\left\{ \begin{array}{l} z_1 < z_2 < \cdots < z_{n-1} < z_n \\ z_1 < z_2 < \cdots < z_n < z_{n-1} \\ \vdots \\ z_n < z_{n-1} < \cdots < z_2 < z_1. \end{array} \right. \quad (5.15)$$

As n gets arbitrarily large, the complexity of bounding σ^C therefore increases exponentially. Thus, in order to efficiently obtain the upper and lower bounds of σ^C , we propose an efficient stochastic programming formulation.

5.3 Stochastic combinatorial optimization problem formulation

We want to find reasonable lower and upper bounds of the next day's stock price. The unit time period could be a day, a week, or a month, depending on the investment strategy in terms of realization of profit in a preferred length of time. Generally, price-bands are used for day trading, short-term investments, etc. For a short-term trading we use intraday stock price data.

Suppose that for the next day stock trading, we use n intraday data points (per day) with historical data points for the past m days. Let us define "Past Data Matrix" by:

$$Y^T = \begin{pmatrix} y_{11} & y_{21} & \cdots & y_{m1} \\ y_{12} & y_{22} & \cdots & y_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ y_{1n} & y_{2n} & \cdots & y_{mn} \end{pmatrix} = \begin{pmatrix} x_{11}^P - e_1^P & x_{21}^P - e_2^P & \cdots & x_{m1}^P - e_m^P \\ x_{12}^P - e_1^P & x_{22}^P - e_2^P & \cdots & x_{m2}^P - e_m^P \\ \vdots & \vdots & \ddots & \vdots \\ x_{1n}^P - e_1^P & x_{2n}^P - e_2^P & \cdots & x_{mn}^P - e_m^P \end{pmatrix}, \quad (5.16)$$

where x_{ji}^P denotes the past stock price at i th time point on day j ; e_j^P denotes the estimated average stock price on day j , $j = 1, \dots, m$ (i.e., $e_j^P = \frac{1}{n} \sum_{i=1}^n x_{ji}$).

Since we are interested in "unknown" stock prices on the next business day, let us

define “Future Data Vector” by:

$$\mathbf{z} = (z_1, \dots, z_n) = (X_1^F - e^F, X_2^F - e^F, \dots, X_n^F - e^F), \quad (5.17)$$

where random variables X_i^F , $i = 1, \dots, n$ are normally distributed with mean e^F and variance S , and denote stock price at i th time point of the next day. In (5.16) and (5.17), Y^T is an $n \times m$ matrix, and \mathbf{z} is an $1 \times n$ row vector.

Components of the covariance matrix

$$C = \begin{pmatrix} S & U \\ U^T & T \end{pmatrix}. \quad (5.18)$$

need to be represented by the use of “Past Data Matrix” and “Future Data Vector” as in (5.16) and (5.17), respectively.

Using the “Past Data Matrix” Y^T , $T = E[(\mathbf{X}^P - \mathbf{e}^P)(\mathbf{X}^P - \mathbf{e}^P)^T]$ can be estimated by

$$T = \frac{1}{n-1} Y Y^T, \quad (5.19)$$

which is an $m \times m$ matrix.

Now we turn our attention to the other components of C : S and U , involving the “Future” random variable X^F . As in Bollinger (2002), we estimate next day’s variance S by (5.14). The future is unpredictable, especially short term trading, so we cannot predict the volatility of next day based on existing patterns or market trends (if there is any) in the past couple of days. Thus the bridge between past and future for should be closely examined, and we have to keep the randomness for the best selection.

The “unknown” cross covariance $U = E[(\mathbf{X}^F - \mathbf{e}^F)(\mathbf{X}^P - \mathbf{e}^P)^T]$ can be written up as

$$\frac{1}{n-1} \sum_{i=1}^n z_i [Y^T]_i, \quad (5.20)$$

where z_i , $i = 1, \dots, n$ denote i th component of “Future Data Vector” \mathbf{z} in (5.17), and $[Y^T]_i$ $i = 1, \dots, n$ denote i th row of the Past Data Matrix Y^T in (5.16). Note that this is a $1 \times m$ row vector.

By (5.14), (5.19) and (5.20), the conditional variance $S - UT^{-1}U^T$ (i.e., the variance

of $\mathbf{X}^F | \mathbf{X}^P = \mathbf{x}^P$) can be written up as:

$$\frac{\sum_{i=1}^N (x_i^P - \bar{x})^2}{N-1} - \left(\frac{\sum_{i=1}^n z_i [Y^T]_i}{n-1} \right) \left(\frac{(YY^T)^{-1}}{n-1} \right) \left(\frac{\sum_{i=1}^n z_i [Y^T]_i}{n-1} \right)^T, \quad (5.21)$$

where \bar{x} in the first term is the mean of all the past data points, and $N = mn$, m : number of days, n : number of intraday data points. In the above estimated conditional variance (5.21), $z_i, i = 1, \dots, n$, are the only unknowns, i.e., the “Future Data Vector” $\mathbf{z} = (z_1, \dots, z_n)$ which can be found for the cases of minimum and maximum of the conditional variance, by a suitable optimization problem formulation. Since $z_i = X_i^F - e^F \sim N(0, S)$, it is reasonable to assume that $-4\sqrt{S} \leq z_i \leq 4\sqrt{S}$, $i = 1, \dots, n$.

We want to find the minimum and maximum values of (5.21), i.e., lower and upper bounds of the variance of $\mathbf{X}^F | \mathbf{X}^P = \mathbf{x}^P$ (volatility on the next day). For meaningful bounding values of (5.21), the future data points z_i , $i = 1, \dots, n$, and their relations to all given past data points at time i , $i = 1, \dots, n$ (i.e. $[Y^T]_i = [(x^P - e^P)^T]_i$, $i = 1, \dots, n$) are essential. Thus, all possible cases must be examined for a reasonable selection of “Future Data Vector” $\mathbf{z} = (z_1, \dots, z_n)$.

Let f be p.d.f. of $X_i^F - e^F$, normally distributed with mean 0 and variance S for all $i = 1, \dots, n$. Let $z^{(k)}$ denotes k th largest element in (z_1, \dots, z_n) , i.e., $z^{(n)} \leq z^{(n-1)} \leq \dots \leq z^{(1)}$. Similarly, let $y_j^{(k)}$ denotes the k th largest element (i.e., $y_j^{(n)} \leq y_j^{(n-1)} \leq \dots \leq y_j^{(1)}$), i.e., the ordering of the components of vector $y_j = (x_{j1} - e_j, \dots, x_{jn} - e_j)$: n intraday data points on day j , $j = 1, \dots, m$. Let e_j^P and σ_j^P denote the expectation and standard deviation of the normal random variable X_j^P , on day $j = 1, \dots, m$. Let g_j denote p.d.f. of $X_j^P - e_j^P \sim N(0, \sigma_j^2)$, $j = 1, \dots, m$. Then we define the following probabilities $p_j^{(i)}$ on day j , $j = 1, \dots, m$ by:

$$\begin{aligned} p_j^{(0)} &= \int_{y_j^{(1)}}^{\infty} g_j(t) dt, \\ p_j^{(i)} &= \int_{y_j^{(i+1)}}^{y_j^{(i)}} g_j(t) dt, \quad i = 1, \dots, n-1, \\ p_j^{(n)} &= \int_{-\infty}^{y_j^{(n)}} g_j(t) dt, \end{aligned} \quad (5.22)$$

where $y_j^{(n)} \leq y_j^{(n-1)} \leq \dots \leq y_j^{(1)}$ such that $\sum_{i=0}^n p_j^{(i)} = 1$. Figure 5.3 can be referred for the description of $p_j^{(i)}$, $i = 0, \dots, n$.

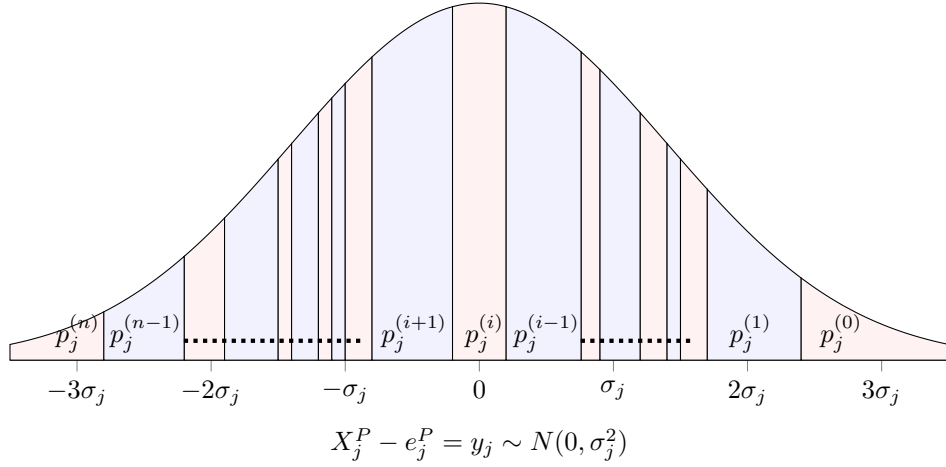


Figure 5.3: Description of data points and their corresponding probabilities

$p_j^{(0)} = \int_{y_j^{(1)}}^{\infty} f(t)dt$, $p_j^{(i)} = \int_{y_j^{(i+1)}}^{y_j^{(i)}} g_j(t)dt$, $i = 1, \dots, n-1$ and $p_j^{(n)} = \int_{-\infty}^{y_j^{(n)}} g_j(t)dt$ where g_j is normal p.d.f. of $\xi_j^P - e_j^P$ on day j in the past, which is assumed to have mean zero and variance σ_j^2 .

Based on the stock price data for the past m days, we want to give some reasonable intervals to probabilities regarding “each” of the next day’s data points z_i , $i = 1, \dots, n$. For this reason, by (5.22), we set lower and upper bounds for probabilities regarding the ordered future data points $z^{(n)} \leq \dots \leq z^{(1)}$, out of z_i , $i = 1, \dots, n$ of in the following way:

$$\begin{aligned}
 l_0 &= \min\{p_1^{(0)}, p_2^{(0)}, \dots, p_m^{(0)}\} \leq \int_{-\infty}^{\infty} f(t)dt \leq \max\{p_1^{(0)}, p_2^{(0)}, \dots, p_m^{(0)}\} = u_0 \\
 l_1 &= \min\{p_1^{(1)}, p_2^{(1)}, \dots, p_m^{(1)}\} \leq \int_{z^{(2)}}^{z^{(1)}} f(t)dt \leq \max\{p_1^{(1)}, p_2^{(1)}, \dots, p_m^{(1)}\} = u_1 \\
 &\vdots \\
 l_n &= \min\{p_1^{(n)}, p_2^{(n)}, \dots, p_m^{(n)}\} \leq \int_{-\infty}^{z^{(n)}} f(t)dt \leq \max\{p_1^{(n)}, p_2^{(n)}, \dots, p_m^{(n)}\} = u_n,
 \end{aligned} \tag{5.23}$$

where $z^{(k)}$ denotes k th largest element in (z_1, \dots, z_n) , and f the p.d.f. of $z_i = X_i^F - e^F$, normally distributed with mean 0 and variance S for all $i = 1, \dots, n$.

With the condition (5.23), we can write min-max problem formulation as follows.

$$\begin{aligned}
& \min(\max) \frac{\sum_{i=1}^N (x_i^P - \mu)^2}{N-1} - \left(\frac{\sum_{i=1}^n z_i [Y^T]_i}{n-1} \right) \left(\frac{(YY^T)^{-1}}{n-1} \right) \left(\frac{\sum_{i=1}^n z_i [Y^T]_i}{n-1} \right)^T \\
& \text{subject to} \\
& l_0 \leq \int_{z^{(1)}}^{\infty} f(t) dt \leq u_0 \\
& l_1 \leq \int_{z^{(2)}}^{z^{(1)}} f(t) dt \leq u_1 \\
& \quad \vdots \\
& l_n \leq \int_{-\infty}^{z^{(n)}} f(t) dt \leq u_n \\
& -4\sqrt{S} \leq z_i \leq 4\sqrt{S}, \quad i = 1, \dots, n \\
& z^{(k)} = k\text{th largest element in } (z_1, \dots, z_n),
\end{aligned} \tag{5.24}$$

where f is the p.d.f. of z_i , $i = 1, \dots, n$ and $S = \frac{\sum_{i=1}^N (x_i^P - \mu)^2}{N-1}$ (the first term of the objective function), the estimated variance for z_1, \dots, z_n .

The above formulation (5.24) can be said to be a stochastic combinatorial optimization problem due to the last constraint, which requires the necessity of counting of all outcomes and a proper selection among them. We know decision making in a finite sample space is very often considered as a counting problem. There are $n!$ permutation of the set $\{z_1, \dots, z_n\}$, i.e., $n!$ choices of a total ordering of the set of intraday data points $\{z_1, \dots, z_n\}$. In order to find minimum and maximum of the problem (5.24), all $n!$ -permutations should be counted and examined (all must simultaneously satisfy all the constraints for feasibility.). In order to count all of the $n!$ cases, among various methods we propose a set representation of the n intraday data points. In this way, all possible cases can be counted in a more systematic way, allowing a faster subsequent computation. In general, for any sequence of finite sets, the principle of inclusion and exclusion can be used in counting problems

This is the fundamental motivation for implementing the set theoretical approach, and the formulation (5.24) can be written up in a more (mathematically) compact form by the use of an ordered partition of a finite set, and binomial coefficients. By a set representation of the “Future Data Points” $\mathbf{z} = (z_1, \dots, z_n)$ as “ordered partitions

with n nonempty blocks” of the set $\{z_1, \dots, z_n\}$, we can reformulate the problem as a modified binomial moment problem, which effectively counts all possible cases without actual counting *per se*.

5.4 Binomial moment problem formulation

Utilizing a binomial moment scheme, the problem (5.24) can be reformulated in a systematic form. For details about the binomial moment scheme, we refer the reader to the literature, e.g., Prékopa (1988; 1995; 2003), etc. For completeness we present some basic definitions here.

Let ν designate the number of events from A_1, \dots, A_n that occur. Let $v_i = P(\nu = i)$, $i = 1, \dots, n$. Then

$$\sum_{i=0}^n \binom{i}{k} v_i = S_k, \quad k = 0, 1, \dots, n, \quad (5.25)$$

where, by definition, $S_k = E \left[\binom{\nu}{k} \right]$, $k = 0, \dots, n$. Essentially, the binomial moment problem formulation is to optimize an objective function with a counting method leveraging the inclusion-exclusion principle.

For a suitable representation of the future data points $\mathbf{z} = (z_1, \dots, z_n)$, let us define the sets as follows:

$$A_j = \{t \mid t \leq z^{(j)}\}, \quad j = 1, \dots, n, \quad (5.26)$$

where $z^{(j)}$ is the j th largest among the next day's n intraday data points z_1, z_2, \dots, z_n , i.e., $z^{(n)} \leq z^{(n-1)} \leq \dots \leq z^{(1)}$. Note that $A_n \subseteq A_{n-1} \subseteq \dots \subseteq A_1$.

Let us introduce the functions, for $k = 1, \dots, n$:

$$\begin{aligned} S_k(\mathbf{z}) &= \sum_{1 \leq i_1 < \dots < i_k \leq n} P(z^{(i_1)} \geq \eta, \dots, z^{(i_k)} \geq \eta) \\ &= \sum_{1 \leq i_1 < \dots < i_k \leq n} P(A_{i_1} \dots A_{i_k}), \end{aligned} \quad (5.27)$$

where $\eta = X^F - e^F$, and A_{i_j} , $j = 1, \dots, k$ are defined as in (5.26). Due to the special shape of sets of (5.26), i.e., $A_n \subseteq A_{n-1} \subseteq \dots \subseteq A_1$, the probabilities v_i , $i = 1, \dots, n$,

can be defined, for the binomial moment problem formulation, by

$$\begin{aligned} v_0 &= 1 - P(A_1), \\ v_i &= P(A_i \setminus A_{i+1}), i = 1, \dots, n-1, \\ v_n &= P(A_n), \end{aligned} \tag{5.28}$$

where $v_i = P(\nu = i)$, $i = 1, \dots, n$, and ν designates the number of events out of A_1, \dots, A_n that occur.

Equivalently, if the random variable $X^F - e^F$ has a normal p.d.f. f , then we can write:

$$\begin{aligned} v_0 &= \int_{z^{(1)}}^{\infty} f(t) dt, \\ v_i &= \int_{z^{(i+1)}}^{z^{(i)}} f(t) dt, i = 1, \dots, n-1, \\ v_n &= \int_{-\infty}^{z^{(n)}} f(t) dt, \end{aligned} \tag{5.29}$$

where $z^{(n)} \leq z^{(n-1)} \leq \dots \leq z^{(1)}$ such that $\sum_{i=0}^n v_i = 1$. Probabilities v_0, v_1, \dots, v_n are also described in Figure 5.4.

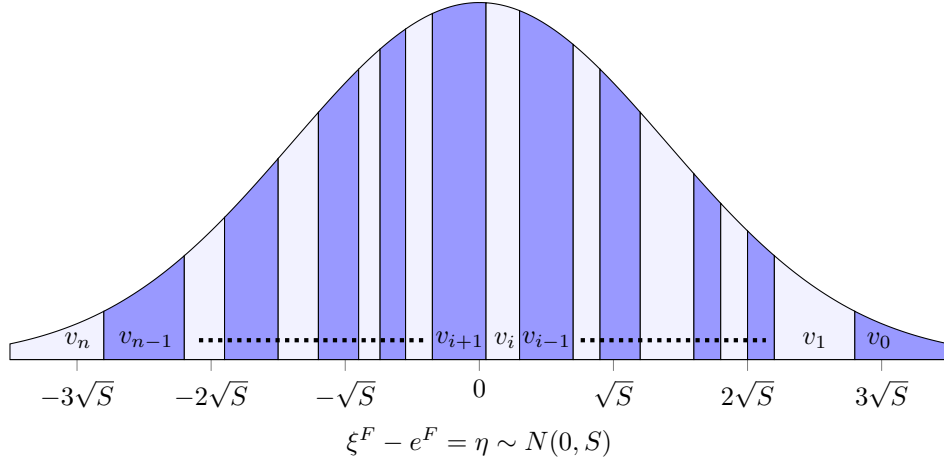


Figure 5.4: Description of data points and their corresponding probabilities

$v_0 = \int_{z^{(1)}}^{\infty} f(t) dt$, $v_i = \int_{z^{(i+1)}}^{z^{(i)}} f(t) dt, i = 1, \dots, n-1$ and $v_n = \int_{-\infty}^{z^{(n)}} f(t) dt$ where f is normal p.d.f. of $X^F - e^F$ which is assumed to have mean zero and variance S .

In the same way in (5.24), from the past stock prices moving tendency, we set

reasonable lower and upper bounds for v_i , $i = 0, \dots, n$ of (5.29) in the following way:

$$\begin{aligned}
l_0 &= \min\{p_1^{(0)}, p_2^{(0)}, \dots, p_m^{(0)}\} \leq v_0 \leq \max\{p_1^{(0)}, p_2^{(0)}, \dots, p_m^{(0)}\} = u_0 \\
l_1 &= \min\{p_1^{(1)}, p_2^{(1)}, \dots, p_m^{(1)}\} \leq v_1 \leq \max\{p_1^{(1)}, p_2^{(1)}, \dots, p_m^{(1)}\} = u_1 \\
&\vdots \\
l_n &= \min\{p_1^{(n)}, p_2^{(n)}, \dots, p_m^{(n)}\} \leq v_n \leq \max\{p_1^{(n)}, p_2^{(n)}, \dots, p_m^{(n)}\} = u_n,
\end{aligned} \tag{5.30}$$

where $p_j^{(k)}$, $j = 1, \dots, m$, $k = 1, \dots, n$ are defined (5.22), and described in Figure 5.3.

Using (5.30), the stochastic combinatorial problem (5.24) can be reformulated, together with (5.27) in the RHS of the binomial constraints, as a binomial moment problem formulation:

$$\begin{aligned}
&\min(\max) \frac{\sum_{i=1}^N (x_i^P - \mu)^2}{N-1} - \left(\frac{\sum_{i=1}^n z_i [Y^T]_i}{n-1} \right) \left(\frac{(YY^T)^{-1}}{n-1} \right) \left(\frac{\sum_{i=1}^n z_i [Y^T]_i}{n-1} \right)^T \\
&\text{subject to} \\
&\sum_{i=0}^n \binom{i}{k} v_i = S_k(\mathbf{z}), \quad k = 0, 1, \dots, m \leq n \\
&l_i \leq v_i \leq u_i, \quad i = 0, 1, \dots, n \\
&-4\sqrt{S} \leq z_i \leq 4\sqrt{S}, \quad i = 1, \dots, n.
\end{aligned} \tag{5.31}$$

Since the sample variance $S = \frac{\sum_{i=1}^N (x_i^P - \mu)^2}{N-1}$ is a constant and independent of decision variables z_i , $i = 1, \dots, n$ and v_i , $i = 0, \dots, n$, it can be removed from the formulation (5.31). Now the min-max problem (5.31) becomes max-min problem as:

$$\begin{aligned}
&\max(\min) \left(\frac{\sum_{i=1}^n z_i [Y^T]_i}{n-1} \right) \left(\frac{(YY^T)^{-1}}{n-1} \right) \left(\frac{\sum_{i=1}^n z_i [Y^T]_i}{n-1} \right)^T \\
&\text{subject to} \\
&\sum_{i=0}^n \binom{i}{k} v_i = S_k(\mathbf{z}), \quad k = 0, 1, \dots, m \leq n \\
&l_i \leq v_i \leq u_i, \quad i = 0, 1, \dots, n \\
&-4\sqrt{S} \leq z_i \leq 4\sqrt{S}, \quad i = 1, \dots, n.
\end{aligned} \tag{5.32}$$

Without the condition (5.30) for the probabilities related to the future data points,

the optimal objective function values (lower and upper bounds of the next day's volatility) would be less dependent on the past. If we assume a more independent stochastic structure between past and future, then we can remove the condition of (5.30) from the constraints. However, we still need a positivity restriction on probabilities v_i , $i = 0, \dots, n$, just to have a reasonable vector $\mathbf{z} = (z_1, \dots, z_n)$.

Simply,

$$\begin{aligned} & \max(\min) \left(\frac{\sum_{i=1}^n z_i [Y^T]_i}{n-1} \right) \left(\frac{(YY^T)^{-1}}{n-1} \right) \left(\frac{\sum_{i=1}^n z_i [Y^T]_i}{n-1} \right)^T \\ & \text{subject to} \\ & \sum_{i=0}^n \binom{i}{k} v_i = S_k(\mathbf{z}), \quad k = 0, 1, \dots, m \leq n \\ & v_i \geq p_i, \quad i = 0, 1, \dots, n \\ & -4\sqrt{S} \leq z_i \leq 4\sqrt{S}, \quad i = 1, \dots, n, \end{aligned} \tag{5.33}$$

where $0 \leq p_i \ll 1$, $i = 1, \dots, n$ are some fixed (very small) probabilities that can be chosen in various ways (e.g., $p_i = 0.001$, $i = 1, \dots, n$). Still, constraints on probabilities v_i , $i = 0, \dots, n$, ensure reasonable placement of future data points at optimality of the problem, as described in Figure 5.4.

As a result of the special structure of the sets regarding future data points as in (5.26) (i.e., $A_n \subseteq A_{n-1} \subseteq \dots \subseteq A_1$), the calculations of S_k , $k = 1, \dots, n$, are manageable (in general, it is very expensive computationally). This is due to the fact that, for the calculations of S_k , $k = 1, \dots, n$, in particular, we need to find only the minimum value among k data points over all cases of $\binom{n}{k}$, $k = 1, \dots, n$. We then calculate their CDF values, followed by summing up the values of all $\binom{n}{k}$, $k = 1, \dots, n$, cases to calculate the binomial functions S_k , $k = 0, \dots, n$ as follows:

$$\begin{aligned}
S_0(\mathbf{z}) &\equiv 1 \\
S_1(\mathbf{z}) &= P(\eta \leq z_1) + P(\eta \leq z_2) + \cdots + P(\eta \leq z_n) \\
S_2(\mathbf{z}) &= P(\eta \leq z_1, \eta \leq z_2) + P(\eta \leq z_1, \eta \leq z_3) + \cdots + P(\eta \leq z_{n-1}, \eta \leq z_n) \\
&= P(\eta \leq \min\{z_1, z_2\}) + P(\eta \leq \min\{z_1, z_3\}) + \cdots + P(\eta \leq \min\{z_{n-1}, z_n\}) \\
S_3(\mathbf{z}) &= P(\eta \leq z_1, \eta \leq z_2, \eta \leq z_3) + P(\eta \leq z_1, \eta \leq z_2, \eta \leq z_4) + \cdots + P(\eta \leq z_{n-2}, \eta \leq z_{n-1}, \eta \leq z_n) \\
&= P(\eta \leq \min\{z_1, z_2, z_3\}) + P(\eta \leq \min\{z_1, z_2, z_4\}) + \cdots + P(\eta \leq \min\{z_{n-2}, z_{n-1}, z_n\}) \\
S_4(\mathbf{z}) &= P(\eta \leq \min\{z_1, z_2, z_3, z_4\}) + P(\eta \leq \min\{z_1, z_2, z_3, z_5\}) + \cdots + P(\eta \leq \min\{z_{n-3}, z_{n-2}, z_{n-1}, z_n\}) \\
&\vdots \\
S_{n-1}(\mathbf{z}) &= P(\eta \leq \min\{z_1, \dots, z_{n-1}\}) + P(\eta \leq \min\{z_1, \dots, z_{n-2}, z_n\}) + \cdots + P(\eta \leq \min\{z_2, \dots, z_n\}) \\
S_n(\mathbf{z}) &= P(\eta \leq \min\{z_1, \dots, z_n\}),
\end{aligned} \tag{5.34}$$

where the binomial functions S_k have $\binom{n}{k}$ number of terms for all $k = 0, \dots, n$. We note that the calculations of S_k , $k = 0, \dots, n$ are computationally expensive when n is large, despite the fact that calculation of each term for the addition is relatively easy as a result of the special shape of sets in (5.26).

In (5.34), the random variable η is normally distributed (setting $\eta = \mathbf{X}^{\mathbf{F}} - \mathbf{e}^{\mathbf{F}}$ implies $\eta \sim N(0, S)$, where S is the next day's estimated variance, defined by (5.14).). These detailed binomial functions are equivalent to (5.27), and \mathbf{z} is an n -tuple vector (vector of n intraday-data points); z_i designates the data point at time i on the next trading day for $i = 1, \dots, n$ (i.e. the realization of random variable $\eta \sim N(0, S)$ at time $i = 1, \dots, n$).

With the binomial functions of (5.34) in the RHS of the constraints, we can write the following detailed formulation:

$$\begin{aligned}
& \max(\min) \left(\frac{\sum_{i=1}^n z_i [Y^T]_i}{n-1} \right) \left(\frac{(YY^T)^{-1}}{n-1} \right) \left(\frac{\sum_{i=1}^n z_i [Y^T]_i}{n-1} \right)^T \\
& \text{subject to} \\
& \begin{aligned}
\begin{pmatrix} 0 \\ 0 \end{pmatrix} v_0 + \begin{pmatrix} 1 \\ 0 \end{pmatrix} v_1 + \begin{pmatrix} 2 \\ 0 \end{pmatrix} v_2 + \cdots + \begin{pmatrix} n-1 \\ 0 \end{pmatrix} v_{n-1} + \begin{pmatrix} n \\ 0 \end{pmatrix} v_n &= S_0 \equiv 1 \\
\begin{pmatrix} 1 \\ 1 \end{pmatrix} v_1 + \begin{pmatrix} 2 \\ 1 \end{pmatrix} v_2 + \cdots + \begin{pmatrix} n-1 \\ 1 \end{pmatrix} v_{n-1} + \begin{pmatrix} n \\ 1 \end{pmatrix} v_n &= S_1(\mathbf{z}) \\
\begin{pmatrix} 2 \\ 2 \end{pmatrix} v_2 + \cdots + \begin{pmatrix} n-1 \\ 2 \end{pmatrix} v_{n-1} + \begin{pmatrix} n \\ 2 \end{pmatrix} v_n &= S_2(\mathbf{z}) \\
&\vdots \\
\begin{pmatrix} n-1 \\ n-1 \end{pmatrix} v_{n-1} + \begin{pmatrix} n \\ n \end{pmatrix} v_n &= S_{n-1}(\mathbf{z}) \\
\begin{pmatrix} n \\ n \end{pmatrix} v_n &= S_n(\mathbf{z}) \\
l_i \leq v_i \leq u_i, \quad i = 0, \dots, n \\
-4\sqrt{S} \leq z_i \leq 4\sqrt{S}, \quad i = 1, \dots, n,
\end{aligned}
\end{aligned} \tag{5.35}$$

where n is the number of intraday data points; $S_k(\mathbf{z}), k = 1, \dots, n$ are defined by (5.34); where l_i and $u_i, i = 0, \dots, n$ are the lower and upper bounds that can be found by (5.30); Y^T denotes the “Past Data Matrix” of (5.16), and $[Y^T]_i$ the i th row of that matrix; S is the sample variance for X^F defined in (5.14).

The decision variables of the above formulation (5.35) are $v_i, i = 0, \dots, n$, and $z_i, i = 1, \dots, n$. The optimal objective function value will be used for the calculation of σ^C in order to determine the width of the bands as in Definition 5.2.1. The optimal solution $z_i, i = 1, \dots, n$, are used for the calculation of the conditional mean, e^C , since with those optimal solutions we can find the estimated value of U by (5.20). Thus, the conditional mean $e^C = \mathbf{e}^F + UT^{-1}(\mathbf{x}^P - \mathbf{e}^P)$ can be calculated with the estimated mean \mathbf{e}^F (moving average, front-weighted, or exponential average, etc.) as in Bollinger Bands (see Table 5.1.).

Using the “Past Data Matrix” Y^T in (5.16), together with (5.19) and (5.20), we can estimate the second term of e^C as

$$UT^{-1}(\mathbf{x}^P - \mathbf{e}^P) = \frac{1}{n-1} \left\{ \frac{\sum_{i=1}^n z_i [Y^T]_i}{n-1} \frac{(YY^T)^{-1}}{n-1} Y^T \right\}. \tag{5.36}$$

Table 5.2: The Prékopa-Lee Bands

By the maximization problem of (5.35)	By the minimization problem of (5.35)
Price-bands with the minimum volatility	Price-bands with the maximum volatility
Upper bands = $e_{\min}^C + k\sigma_{\min}^C$	Upper bands = $e_{\max}^C + k\sigma_{\max}^C$
Middle bands = e_{\min}^C	Middle bands = e_{\max}^C
Lower bands = $e_{\min}^C - k\sigma_{\min}^C$	Lower bands = $e_{\max}^C - k\sigma_{\max}^C$
<i>the lowest volatility level at a given setting</i>	<i>the highest volatility level at a given setting</i>

σ_{\min}^C can be calculated by the maximization problem of (5.35), i.e., $\sigma^C = \sqrt{S - UT^{-1}U^T}$; e_{\min}^C is determined by its optimal solution. Similarly, σ_{\max}^C is from the minimization problem of (5.35), and by its solution e_{\max}^C can be found. The multiplier k can be chosen depending on the time periods n . We recommend width parameters with time periods are $k = 2$ if $n = 10$, $k = 2.1$ if $n = 20$, $k = 2.2$ if $n = 50$, etc.

By adding the value calculated from (5.36) to the estimation of \mathbf{e}^F (same as the mid-band of Bollinger), we get conditional expectation e^C to find the “Middle band” in Table 5.2. e^C is the central tendency of a stock price in a time period. The width of the bands is determined by the optimal objective function value of (5.35), and the central tendency of the bands can be determined by the optimal solution of (5.35).

As we solve both maximization and minimization problems, two different bands are formed as in Table 5.2. The two different bands can be interpreted as follows. The optimal objective function value (i.e. the conditional variance) from the maximization problem is the lowest volatility level that we can expect under a reasonable setting (i.e. the constraints of the problem (5.35)). The maximization problem provides us with the tightest bands. Thus, investors with a higher risk tolerance (i.e. more aggressive trader) may want to use the tightest bands constructed by solving the maximization problem (minimization of σ^C). The bands from the minimization problem can be considered analogous.

The constraints of (5.35) can be customized to incorporate investment preference among various risk tolerance levels. For example, we can modify the lower bounds of v_i $i = 0, \dots, n$, or the upper and lower bounds of z_i , $i = 1, \dots, n$. One simple band-customization example is that by controlling the lower bounds of v_0 and v_n , we can limit the next day’s highest and lowest stock price, respectively. This is because the

lower bounds of v_0 and v_n determine the positions of “the upper bound of the highest” and “the lower bound of the lowest” price of stock for the optimization problem in the following way:

$$\begin{aligned} v_0 &= \int_{z^{(1)}}^{\infty} f(t) dt \\ v_n &= \int_{-\infty}^{z^{(n)}} f(t) dt, \end{aligned} \tag{5.37}$$

where $z^{(1)}$ and $z^{(n)}$ are the largest and the smallest stock price of the next business day.

The proper usage of our bands is as follows. More aggressive investment strategy pairs with the bands constructed by solving the maximization problem (minimization of σ^C), and conversely for the bands from the minimization problem (maximization of σ^C). If the indicated actions agree, the result of price-bands analysis lends confidence to a potential investment decision. On the other hand, if price-bands analysis indicates a different action (e.g. “take no action”), this may suggest further verification of the potential investment decision-making process before execution.

5.5 Numerical examples and discussion

We construct the price-bands of Apple Inc. (NASDAQ: APPL), Verizon Communications Inc. (NYSE: VZ), Yahoo! Inc. (NASDAQ: YHOO) and The Walt Disney Company (NYSE: DIS over the same time period from May 1 to July 16 in 2013, by solving the binomial moment problem (5.35). We do not only consider the price data for the bands construction, but we also take into account that “trading volume” plays a meaningful role in the stock price movement. Trading volume can be thought as a measure of investors’ interest in the stock. There is intrinsic duality between buying and selling (shares cannot be bought unless they are sold), and so the calculation of daily average price can be formulated by

$$e_j^P = \frac{\sum_{i=1}^n P_i^{(j)} V_i^{(j)}}{\sum_{i=1}^n V_i^{(j)}}, \tag{5.38}$$

where $P_i^{(j)}$ and $V_i^{(j)}$ denote the stock price and its volume at time i on day j , $i = 1, \dots, n$ and $j = 1, \dots, m$. This is called “volume weighted average price” (VWAP).

For all stocks, we use 30-minute intraday data points, i.e., the number of data points

per day is $n = 12$. Using the historical stock prices data of the past 10 days, we calculate S and e^F . For S , let us use the same variance as in the Bollinger Bands for the next day:

$$S = \frac{\sum_{i=1}^N (x_i^P - \mu)^2}{N}, \quad (5.39)$$

where x_i^P 's are all the given data points, μ their mean, and N the number of data points, i.e., $N = 12 \times 10 = 120$.

It is reasonable and widely accepted that recent time periods influence price movements more than earlier periods, and several measures are widely used in practice, e.g. m -day moving average, exponential average and front-weighted average. Here and in what follows we will use the front-weighted average. We calculate e^F as a front-weighted average given by

$$e^F = \frac{\sum_{i=1}^m i e_i^P}{\sum_{i=1}^m i}, \quad (5.40)$$

where $m = 10$ since we looked back over past 10 business days, and $e_i^P, i = 1, \dots, 10$ are calculated by (5.38).

Then, with the next business day variance S from (5.39), we solve the following maximization and minimization problems:

$$\begin{aligned} & \max(\min) \left(\frac{\sum_{i=1}^n z_i [Y^T]_i}{n-1} \right) \left(\frac{(YY^T)^{-1}}{n-1} \right) \left(\frac{\sum_{i=1}^n z_i [Y^T]_i}{n-1} \right)^T \\ & \text{subject to} \\ & \sum_{i=0}^n \binom{i}{k} v_i = S_k(\mathbf{z}), \quad k = 0, 1, \dots, 12 \\ & l_i \leq v_i \leq u_i, \quad i = 0, 1, \dots, 12 \\ & -4\sqrt{S} \leq z_i \leq 4\sqrt{S}, \quad i = 1, \dots, 12, \end{aligned} \quad (5.41)$$

where $S_k(\mathbf{z}), k = 1, \dots, 12$ are defined by

$$S_k(\mathbf{z}) = \sum_{1 \leq i_1 < \dots < i_k \leq n} P(z^{(i_1)} \geq \eta, \dots, z^{(i_n)} \geq \eta), \quad (5.42)$$

where $\eta \sim N(0, S)$.

Using the optimal solutions to (5.41), e^F from (5.40) and by the use of (5.36), we find the conditional mean vector

$$e^C = e^F + UT^{-1}(x^P - e^P). \quad (5.43)$$

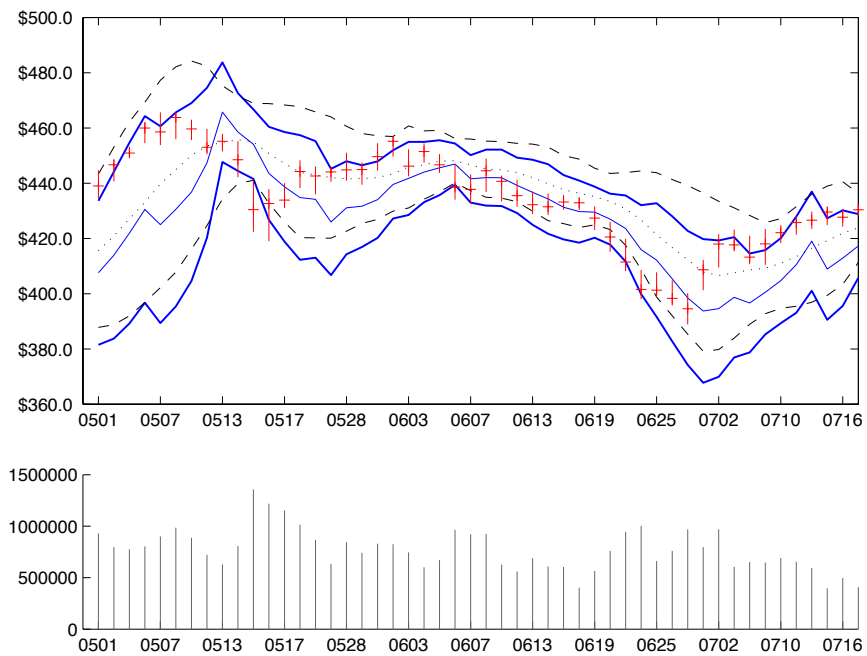
By the results of (5.41) and (5.43), Prékopa-Lee Bands are constructed by the formulae in Table 5.2, and depicted in Figures 5.5, 5.6, 5.7 and 5.8. As the bar chart is widely used and easy to follow, we utilize it in those Figures. The thin vertical line segments (red) are drawn to the high and low of the day, and the intersecting horizontal lines (red) represent closing prices. Let us refer to the results described in Figures and the summary of numerical results presented in Table 5.3 in terms of the number of chances to make profit by short-selling or buying. The blue piecewise linear bands are the Prékopa-Lee Bands, while the dashed and dotted black bands are the Bollinger Bands. The lower figure of each subfigure represents the total trading volume of each day. Volume represents the amount of trading activity, and is a main indicator of investors' interest.

We are especially interested in finding opportunities for short-term profit realization—simply put, “buy low sell high.” Given that stock prices are out of the bands, we can expect that the stock price will be back in the range of the bands in the same time period. If the red vertical line (daily price range of stocks) is fully in the bands, then this price-band strategy suggests taking no action. Otherwise, there exists the opportunity to make a profitable decision: Above the upper bands is indication to sell short and then buy as the price descends into the bands. Similarly, we can buy a stock if it's below the lower bands and sell it at a higher price upon entering the bands.

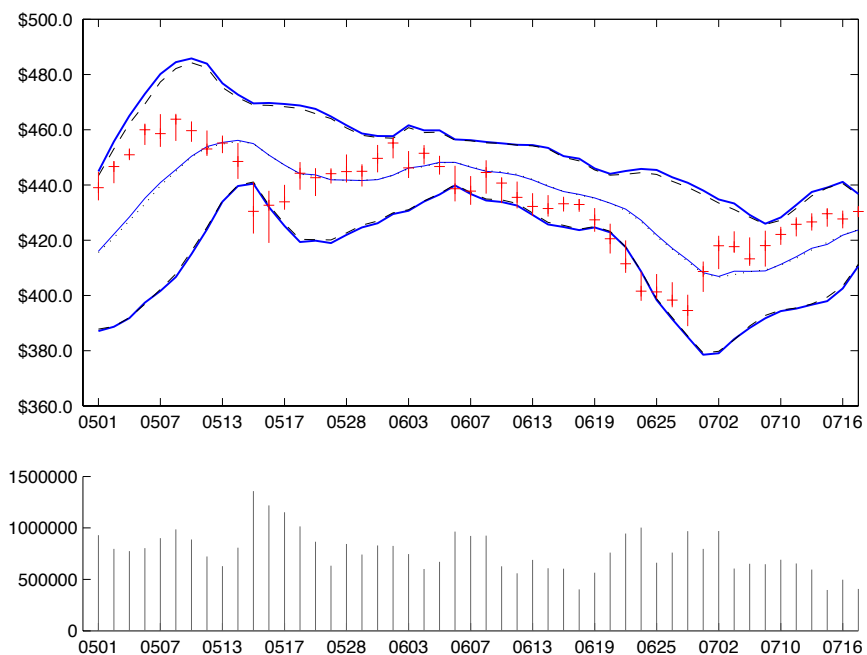
Table 5.3: Comparison of the performances of Bollinger and Prékopa-Lee bands

	The Bollinger bands			The Prekopa-Lee bands		
	Buying	Short-selling	Total number	Total number	Buying	Short-selling
NASDAQ: AAPL	10	0	10	26	7	19
NYSE: VZ	8	3	11	23	11	12
NASDAQ: YHOO	5	6	11	14	5	9
NYSE: DIS	7	10	17	20	7	13

For all the cases of the stocks over the same time span, May 1 to July 16 in 2013,

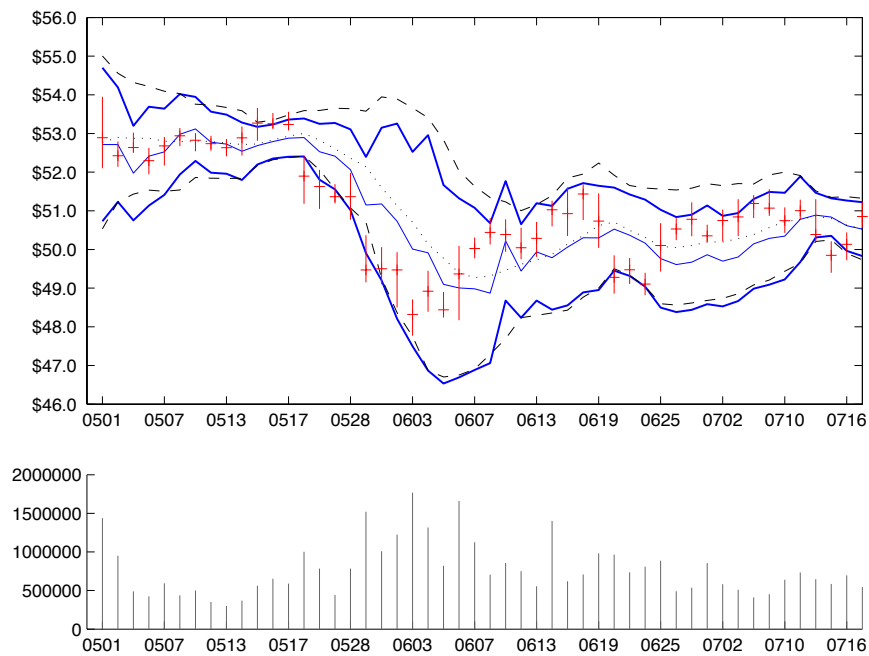


(a) From max problem (5.41); Prékopa-Lee bands in blue; black dotted are Bollinger-bands

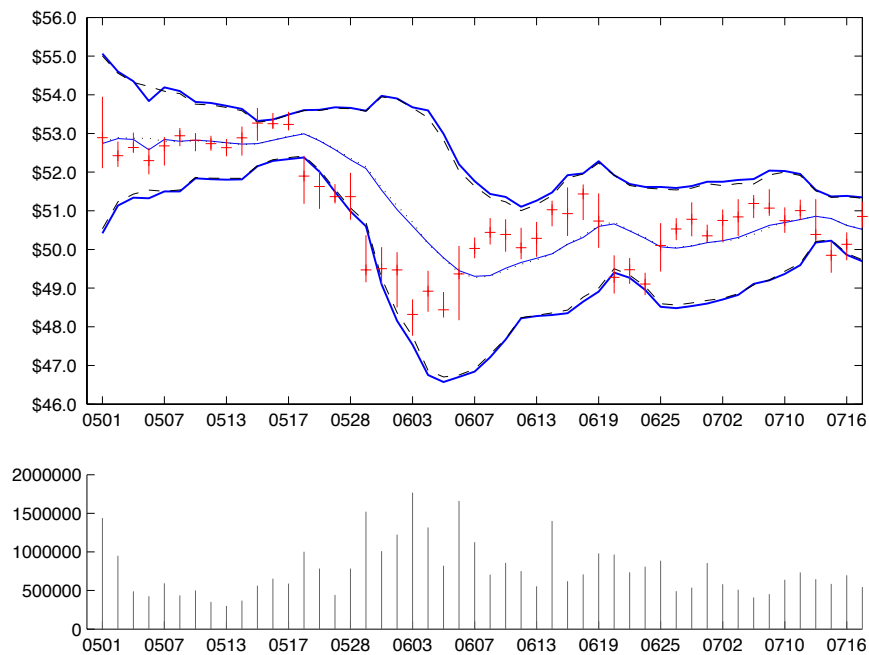


(b) From min problem (5.41); Prékopa-Lee bands in blue; black dotted are the Bollinger-bands

Figure 5.5: Prékopa-Lee bands on Apple Inc. (NASDAQ: AAPL) from May 1 to July 16, 2013

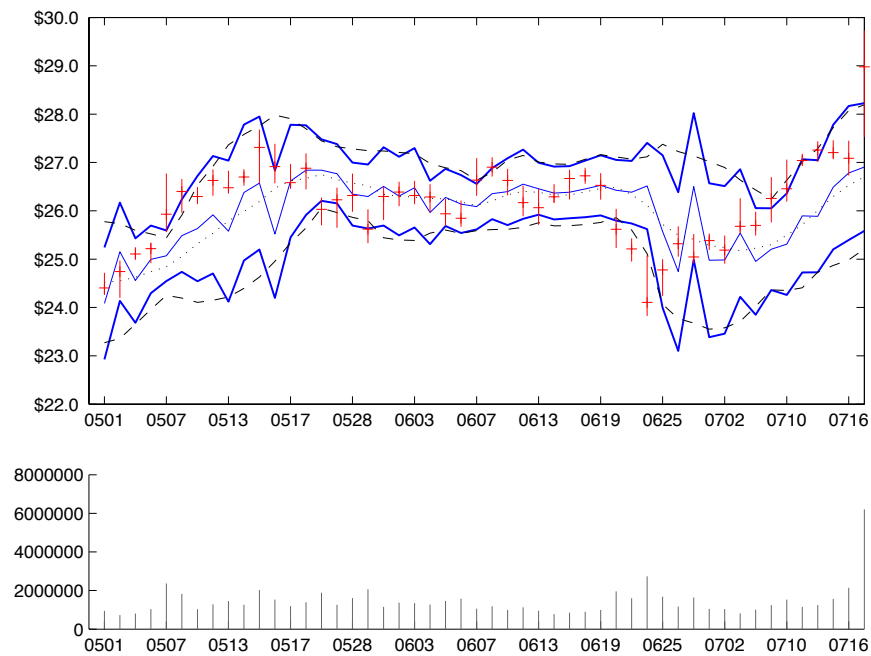


(a) From max problem (5.41); the Prékopa-Lee bands in blue; black dotted are the Bollinger-bands

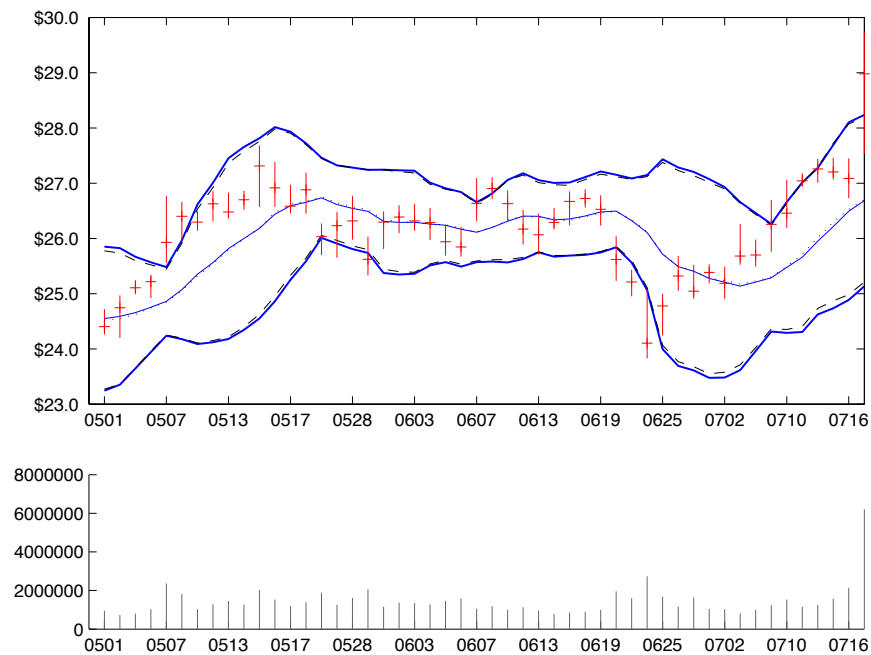


(b) From min problem (5.41); the Prékopa-Lee bands in blue; black dotted are the Bollinger-bands

Figure 5.6: Prékopa-Lee bands on Verizon Communications Inc. (NYSE: VZ) from May 1 to July 16, 2013

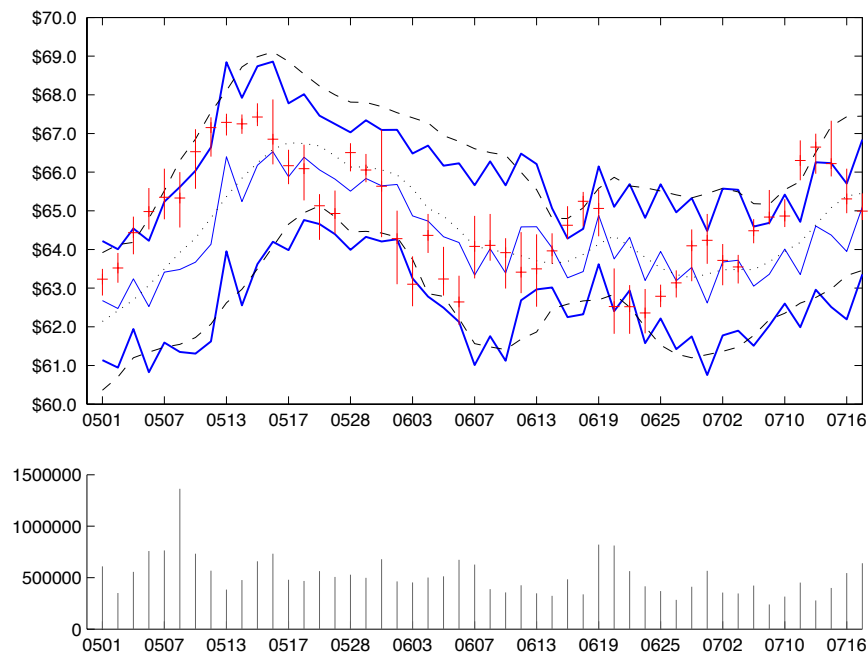


(a) From max problem (5.41); the Prékopa-Lee bands in blue; black dotted are the Bollinger-bands

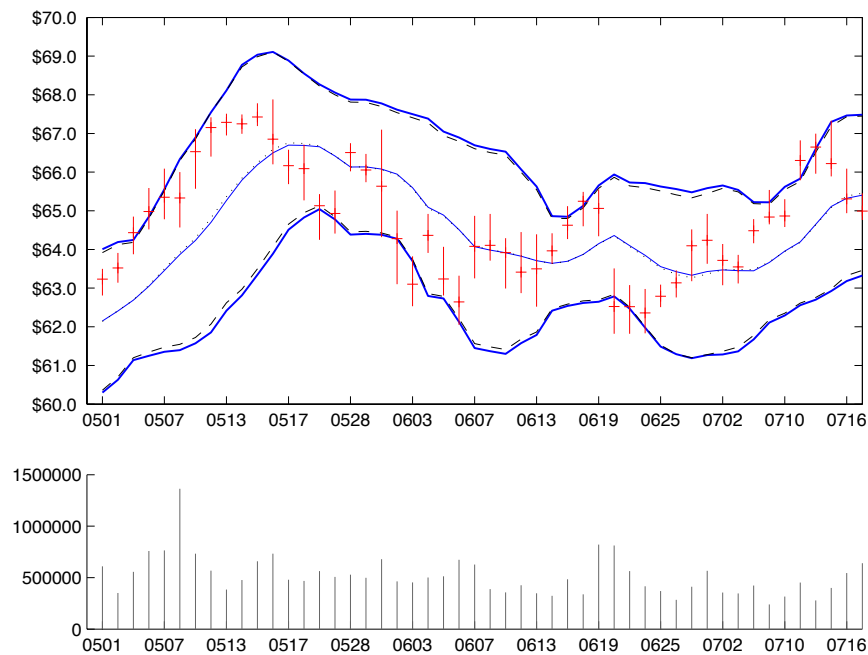


(b) From min problem (5.41); the Prékopa-Lee bands in blue; black dotted are the Bollinger-bands

Figure 5.7: Prékopa-Lee bands on Yahoo! Inc. (NASDAQ: YHOO) from May 1 to July 16, 2013



(a) From max problem (5.41); the Prékopa-Lee bands in blue; black dotted are the Bollinger-bands



(b) From min problem (5.41); the Prékopa-Lee bands in blue; black dotted are the Bollinger-bands

Figure 5.8: Prékopa-Lee bands on The Walt Disney Company (NYSE: DIS) from May 1 to July 16, 2013

Prékopa-Lee Bands (from the max problem (5.41)) provide more opportunities for making profit than Bollinger Bands.

In the case of Apple Inc. (NASDAQ: AAPL), as depicted in Figure 5.5, the Bands from the maximization problem (5.41) look appealing—26 red vertical lines intersect either upper or lower bands, indicating good opportunities on 26 days out of 50 days. Another Prékopa-Lee Bands construction (by solving the minimization problem (5.41)) is depicted below in Figure 5.5 and it is almost identical to the Bollinger Bands. For this example, the Prékopa-Lee Bands constructed from the max problem (5.41) perform much better than the other. In other words, a more aggressive trading strategy works better in case of NASDAQ: APPL in that time period. We note that there is a performance gap between two different Prékopa-Lee Bands. We observe that the price-bands with highest volatility behave almost identically as Bollinger bands.

For the other cases, Verizon Communications Inc. (NYSE: VZ), Yahoo! Inc. (NASDAQ: YHOO) and The Walt Disney Company (NYSE: DIS), over the same time frame, our price-bands all performed at a good level of effectiveness. Note that in case of “The Walt Disney Company (NYSE: DIS),” Prékopa-Lee Bands marginally outperformed the Bollinger Bands.

We do not insist that our model (i.e., the bands created from the maximization problem (5.41)) outperforms Bollinger Bands in more cases, because sometimes less aggressive trading strategy works better. Note that our model (from the minimization problem (5.41)) also provides us with almost identical bands as Bollinger bands. What we hope for in our research is to provide practitioners with another useful tool of technical analysis for stock trading. By spotting stock price trends with various functional tools, including Prékopa-Lee Bands, Bollinger Bands, and others, we believe that it would be possible to execute more winning trades than losing trades.

5.6 Concluding remarks

In the current financial climate, low interest rates make stock investment more attainable, since low-cost borrowing is possible for most individuals (i.e. money is less

expensive). However, successful investment remains elusive. Although it is simply determined by only four words: buy low, sell high, there is, ironically, no clear way for generating consistent profits from a stock trade, largely due to the mixture of the complexity and efficiency of the market and irrationality of its participants. Indeed, the emotional reactions of investors often (but not always) lead them to make poor real-time investment decisions. Price-bands are certainly helpful to deter investors from *entirely* following their feelings, and such tools have been widely used in practice, especially for short term investment, to help people validate their investment decisions. As one variant, we construct new price-bands via binomial moment problem formulation under the assumption that stock prices follow a Gaussian process. Usage of conditional probability distributions is the key attribute that differentiates our model. We hope that our model will pique the interest of many for both theoretical aspects and its applicability to stock trading businesses.

Chapter 6

Conclusion

As we move toward a data-rich society, the number of newly tractable decision-making problems are increasing and finding the key to determining the best outcome becomes increasingly complex and challenging. For such data-driven decision-making problems, we need a systematic way of finding solutions – decision support modeling. In reality, all information pertinent to all possible circumstances cannot, even in principle, be fully known, and hence risk is intrinsically part of any decision with future outcomes. Thus risk analysis is instrumental in decision making.

For quantitative decision support modeling, we need to express risk in meaningful numerical values. To this end, stochastic optimization is a natural and superior skill for finding a key to optimal decision-making, and multivariate risk measures are crucial ingredients for the decision-making processes. Each chapter of this dissertation is devoted to the development of tools and methods useful in various aspects of decision support modeling.

I hope my research to be useful to industry, academia, and to my colleagues as we work for a more efficient and “Smarter Planet.”

Bibliography

- Acerbi C, Tasche D (2001) On the coherence of expected shortfall. *Journal of Banking & Finance* 26:1487–1503
- Ané T, Kharoubi C (2003) Dependence Structure and Risk Measure. *Journal of business* 76(3):411–438
- Artzner P, Delbaen F, Eber J, Heath D (1999) Coherent Measures of Risk. *Mathematical Finance* 9(3):203–228
- Atkinson KE (1988) An introduction to numerical analysis, Second Edition. John Wiley and Sons, Inc.
- Ben-Tal A, Ben-Israel A (1991) A recourse certainty equivalent for decisions under uncertainty. *Annals of Operations Research* 30:3–44
- Ben-Tal A, Teboulle M (1986) Expected utility, penalty functions, and duality in stochastic nonlinear programming. *Management Science* 32(11):1445–1466
- Ben-Tal A, Teboulle M (2007) An old-new concept of convex risk measures: the optimized certainty equivalent. *Mathematical Finance* 17(3):449–476
- Ben-Tal A, Ben-Israel A, Teboulle M (1991) Certainty equivalents and generalized information measures: Duality and extremal principles. *J Math Anal Appl* 157:211–236
- Berger A, Saunders A, Scalise J, Udell GF (1998) The effects of bank mergers and acquisitions on small business lending. *Journal of Financial Economics* 50:187–229
- Bollinger J (2002) Bollinger on Bollinger Bands. McGraw-Hill
- Boros E, Elbassioni K, Gurvich V, Khachiyan L, Makino K (2003) An intersection

- inequality for discrete distributions and related generation problems, vol 2719. Berlin: Springer
- Bowers NLJ, Gerber HU, Hickman JC, Jones DA, Nesbitt CJ (1997) Actuarial Mathematics. The society of Actuaries, Schaumburg, Illinois
- Busemann H (2008) Convex Surfaces. Dover, Mineola, New York
- Cousin A, Di Bernadino E (2011) A multivariate extension of Value-at-Risk and Conditional-Tail-Expectation. Université Lyon 1, ISFA, Laboratoire SAF
- Damodaran A (2012) Investment Valuation: Tools and Techniques for Determining the Value of Any Asset. Wiley Finance
- De Pril N (1986a) Improved recursions for some compound Poisson distributions. Insurance: Mathematics and Economics 5:129–132
- Dentcheva D, Ruszczyński A (2009) Optimization with multivariate stochastic dominance constraints. Mathematical Programming 117:111–127
- Dentcheva D, Prékopa A, Ruszczyński A (2000) Concavity and efficient points of discrete distributions in probabilistic programming. Mathematical Programming Ser. A(89):55–77
- Embrechts P, McNeil A, Straumann D (1999) Correlation and dependency in risk management : properties and pitfalls. Departement of Mathematik, ETHZ, Zürich, Working Paper
- Embrechts P, Lindskog F, McNeil A (2001) Modeling dependence with copulas and applications to risk management. Departement of Mathematik, ETHZ, Zürich, Working Paper
- Erickson M (1998) The Effect of Taxes on the Structure of Corporate Acquisitions. Journal of Accounting Research 36(2):279–298
- Ernst&Young (2013) M&A Tracker Insights. EYGM Limited

- Fama E (1970) Efficient Capital Markets: A Review of Theory and Empirical Work. *Journal of Finance* 25(2):383–417
- Föllmer H, Penner I (2006) Convex risk measures and the dynamics of their penalty functions. *Statistics & Decisions* 24:61–96
- Föllmer H, Schied A (2002) Convex measures of risk and trading constraints. *Finance and Stochastics* 6:429–447
- Föllmer H, Schied A (2010) Convex and Coherent Risk Measures. In: R. Cont, Ed., *Encyclopedia of Quantitative Finance*, John Wiley & Sons, Hoboken
- Frittelli M, Gianin ER (2002) Putting order in risk measures. *Finance Research Letters* 26:1473–1486
- Gasca M, Sauer T (2000) On the history of multivariate polynomial interpolation. *Journal of Computational and Applied Mathematics* 122:23–35
- Gregoriou GN, Renneboog L (2007) International mergers and acquisitions activity since 1990: Recent research and quantitative analysis. Elsevier Finance
- Grimes A (2012) The Art and Science of Technical Analysis: Market Structure, Price Action & Trading Strategies. John Wiley & Sons
- Hagedoorn J, Duysters G (2002) External Sources of Innovative Capabilities: The Preferences for Strategic Alliances or Mergers and Acquisitions. *Journal of Management Studies* 39:167–188
- Hodges SD, M SS (1977) A model for bond portfolio improvement. *Journal of Financial and Quantitative Analysis* 12:243–260
- Jarrow R, Purnanandam AK (2005) A generalized coherent risk measure: The firms perspective. *Finance Research Letters* 2:23–29
- Jaruzelski B, Mueller M, Conway P (2009) The top 10 fallacies and self-deceptions M&A integration. Booz&co.

- Joe H (1997) *Multivariate Models and Multivariate Dependence Concepts*. Chapman and Hall/CRC
- Jorion P (2006) *Value at Risk: The New Benchmark for Managing Financial Risk*, Third Edition. McGraw-Hill
- Junker M, May A (2005) Measurement of aggregate risk with copulas. *Econometrics Journal* 8:428–454
- Kahneman D, Tversky A (1979) Prospect theory: an analysis of decision under risk. *Econometrica* 47(2):263–279
- Kou S, X Peng C, Heyde C (2013) External Risk Measures and Basel Accords. *Mathematics of Operations Research* 38(3):393–417
- Lee J, Prékopa A (2013) Properties and Calculations of Multivariate risk measures: MVaR and MCVaR. *Annals of Operations Research* 211(1):225–254
- Luciano U, Vecchiato E, Cherubini W (2004) *Copula Methods in Finance*. John Wiley and Sons Ltd.
- Müller A, Stoyan D (2002) *Comparison Methods for Stochastic Models and Risks*. Wiley, Chichester
- Nahavandi A, Malekzadeh A (1988) Acculturation in mergers and acquisitions. *Academy of management review* 13(1):79–90
- Noyan N, Rudolf G (2013) Optimization with multivariate conditional value-at-risk constraints. *Operations Research* 61(4):990–1013
- Pflug GC (2000) Some remarks on the value-at-risk and conditional Value at Risk. *Probabilistic constrained optimization* pp 272–281
- Prékopa A (1970) On probabilistic constrained programming. *Proceedings of the Princeton Symposium on Mathematical Programming* (Princeton University Press, NJ) pp 113–138

- Prékopa A (1971) Logarithmic concave measures with applications to stochastic programming. *Acta Scientiarum Mathematicarum* (Szeged) 32:301–316
- Prékopa A (1973a) Contributions to the theory of stochastic programming. *Mathematical Programming* 4:202–221
- Prékopa A (1973b) On logarithmic concave measures and functions. *Acta Scientiarum Mathematicarum* (Szeged) 34:335–343
- Prékopa A (1988) Boole-Bonferroni inequalities and linear programming. *Operations Research* 36(1):145–162
- Prékopa A (1990) Dual method for a one-stage stochastic programming problem with random rhs obeying a discrete probability distribution. *Operations Research* 34:441–461
- Prékopa A (1990a) Sharp bounds on probabilities using linear programming. *Operations Research* 38(2):227–239
- Prékopa A (1990b) The discrete moment problem and linear programming. *Discrete Applied Mathematics* 27:235–254
- Prékopa A (1995) *Stochastic Programming*. Kluwer Academic Publishers
- Prékopa A (2003) Probabilistic Programming. *Hand books in Operations Research and Management Science* (Ruszczynski, A and Shapiro, A, Eds) 10:267–351
- Prékopa A (2012) Multivariate value at risk and related topics. *Annals of Operations Research* 193(1):49–69
- Prékopa A, Vizvári B, Badics T (1998) *Programming Under Probabilistic Constraint with Discrete Random Variable*. Kluwer, Dordrecht, Boston
- Rose T, Frame T (2011) *The 10 steps to successful M&A integration*. Bain & Company, Inc.
- Saita F (2007) *Value at Risk and Bank Capital Management*. Elsevier Inc.

- Sauerbrei W, Meier-Hirmer C, Benner A, Royston P (2006) Multivariate regression model building by using fractional polynomials: Description of SAS, STATA and R programs. *Computational Statistics and Data Analysis* 50:3464–3485
- Shimizu K, Hitt M, Vaidyanath V D, Pisano (2004) Theoretical foundations of cross-border mergers and acquisitions: A review of current research and recommendations for the future. *Journal of International Management* 10:307–353
- Shleifer A (2000) *Inefficient Markets: An Introduction to Behavioral Finance*. Clarendon Lectures in Economics
- Strang G (2007) *Computational Science and Engineering*. Wellesley-Cambridge Press
- Szegö G (2002) Measures of risk. *Journal of Banking & Finance* 26:1253–1272
- Uryasev S, Rockafellar R (2000) Optimization of conditional Value at Risk. *The Journal of Risk* 2:21–41
- Vizvári B (1987) Beiträge zum Frobenius Problem. Dr. sc. nat. Dissertation. Technische Hochschule Carl Schorlemmer, Leuna-Merseburg, Germany
- Walsh J (1988) Top management turnover following mergers and acquisitions. *Strategic Management Journal* 9:173–183
- Wulf J, Singh H (2011) How do acquirers retain successful target CEOs? The role of governance. *Management Science* 57(12):2101–2114