BSc thesis

Department of Finance
Stockholm School of Economics

Portfolio optimization with catastrophe bonds

Tianfang Zhang
24015@student.hhs.se

May 15, 2019

Tutor: Jungsuk Han

Abstract

This thesis investigates properties of the returns of catastrophe bonds and their risk diversification potential in portfolios. Ideas from existing literature, such as modeling the evolution of the outstanding principal as a compound Poisson process, are extended to take into account the times of occurrence of the loss events, which are important for the returns of catastrophe bonds with certain types of payment structure. Monte Carlo sampling of joint returns is used for optimization of risk measures such as value-at-risk and conditional value-at-risk, where a special method for rendering standard nonlinear optimization techniques applicable is developed. Numerical results performed for illustrative purposes showcase key characteristics of portfolios augmented by a single catastrophe bond as well as one exclusively consisting of catastrophe bonds.

Keywords: Catastrophe bonds, Poisson processes, portfolio theory, risk diversification, value-at-risk, conditional value-at-risk.

Contents

Chapter 1

Introduction

With the currently ongoing climate changes, natural catastrophes are occurring with increasingly higher frequency. In 2017, hurricanes in the Atlantic ocean alone caused damage worth over 280 billion USD and insured losses of over 80 billion USD [3]. If left unmanaged, the risks associated with such catastrophic events would be overwhelming for insurance companies, and as a consequence contracts are often written with so called *reinsurance* companies in order to relieve parts of these risks in exchange for a premium. Reinsurance is a central part of the risk management of insurance companies and have traditionally been handled by large global reinsurance companies such as Munich Re, Swiss Re and Hannover Re.

In recent years, however, a type of financial product known as *insurance-linked securities* (ILSs) has gained popularity in serving a similar purpose. ILSs are broadly defined as financial instruments whose values are driven by insurance loss events, and a particular type of ILSs are *catastrophe bonds* which are triggered by catastrophic events such as earthquakes or hurricanes. From the perspective of an investor, the risk associated with a catastrophe bond is vastly different in nature compared to that of e g stocks, bonds or options due to a highly skewed return distribution in which tail events are the main drivers of the bond price. As a consequence, parameters in loss models are often hard to estimate with satisfactory certainty, which makes accurate pricing hard.

Due to the fundamental independence between trigger events of catastrophe bonds and other market events, it is often claimed that catastrophe bonds are efficient in diversifying portfolio risks [23] and, furthermore, can form successful portfolios entirely on their own. Given this independence, it is perhaps not surprising that there is a gain from including the possibility of taking positions in one or multiple catastrophe bonds, but quantifying this gain will require modeling of their losses and risks.

1.1 Catastrophe bonds

1.1.1 Preliminaries

In 1984, the Swedish state-owned corporation Svensk Exportkredit launched what can be considered as the earliest known instance of a catastrophe bond [5]. This took the form of a private placement of earthquake bonds which were immediately redeemable if a major earthquake hit Japan, and Japanese insurers bought these and agreed to accept coupons with relatively low rate as premium for the right to put the bonds back to the issuer at face value should an earthquake occur [5]. Ever since then, catastrophe bonds have been issued in many different formats, varying in perils, trigger mechanisms, payment structures, etc. Unlike e g corporate bonds, which behave similarly as catastrophe bonds with the event of default analogous to that of a catastrophe occurring, the returns of catastrophe bonds cannot be replicated by only ordinary bonds and stocks and thus cannot be hedged by primitive instruments.

Starting with formalities, the issuer of a catastrophe bond—that is, the party whose risks are to be relieved, typically an insurance or reinsurance company—is called the *sponsor*, and a counterparty buying the bond and assuming the specified risk is called an *investor*. When issuing a catastrophe bond, a *contract size* is decided by the sponsor, which is the desired amount of cash to be raised from the capital market. There is, however, no guarantee that this amount will be met—this will depend on the market's interest in the specific contract. An investor buying the bond pays an amount called the *original principal* at the time of the purchase, and the degree of participation of one investor is given by the ratio between the original principal and the contract size [10].

The money raised from the market is placed in a *special purpose vehicle* (SPV), which can be seen as a third party whose sole purpose is to manage the money and handle eventual payments to the sponsor and the investors. The SPV, in turn, puts the money in a *collateral account*, usually consisting of low-risk investments such as government bonds. For the investors, the SPV comprises a protection from risks such as that of the sponsor going bankrupt, whereas for the sponsor, it has the effect of ear-marking the money in the collateral account. During the term of the bond, coupon payments are paid regularly to the investors from the SPV. If one or multiple catastrophic events which meet the criteria of the contract occur, the sponsor withdraws corresponding amounts from the collateral account. The remaining amount in the collateral account at a given time during the term of the bond is called the *outstanding principal*. At maturity, the outstanding principal, if any remaining, is repaid to the investors, and the SPV is closed. Of special importance is the fact that in most cases, the coupon payments are proportional to the outstanding principal [14, 10].

For each catastrophe bond issued, there is a legal document defining the circumstances under which money may be withdrawn from the SPV by the sponsor and how the severity of these circumstances translate into the size of the instrument loss. Typically, each catastrophe bond has a specific *peril*—that is, a type of catastrophe triggering the bond, such as earthquakes, hurricanes or windstorms—and a *risk exposure area*—that is, the geographic area in which catastrophes may lead to instrument losses, such as “US east coast”, “Japan” or “northern Europe”. For the calculation of the actual instrument loss associated with a certain event, there are different types of *trigger mechanisms*. In the case of an *indemnity* trigger, actual losses for the sponsor are determined with respect to definitions and standards stated in the contract (which may be a tedious process); for a *modeled loss* trigger, the instrument losses are instead based on statistics such as earthquake magnitude regardless of the actual losses incurred for the sponsor; for a *parametric* trigger, the instrument losses are determined from a formula based on continuous measurements of metrics such as air pressure or wind speed; and for an *industry loss index* trigger, the instrument losses are instead based on an index of the losses of the insurance industry following an event. The probability that instrument losses are incurred at all is usually called the *attachment probability*, and the probability that the outstanding principal is zero at maturity is called the *exhaustion probability* [14, 10].

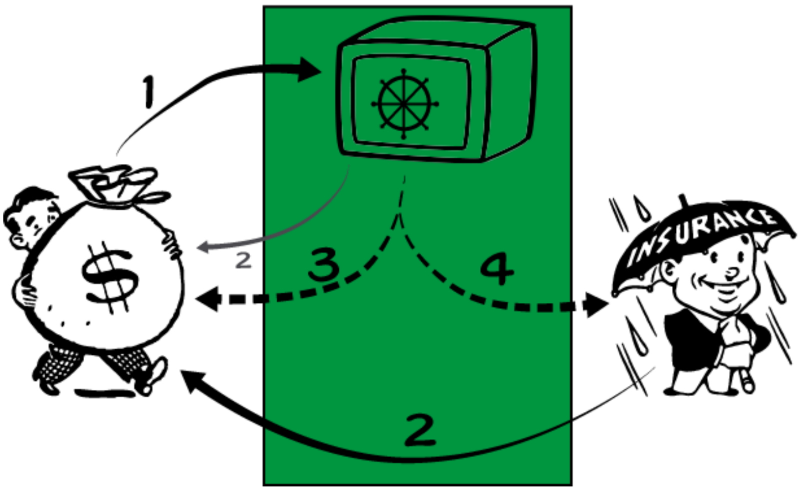

Figure 1.1: Illustration of cash flows in a typical catastrophe bond contract. (1) shows the payment of the initial principal by the investors, (2) shows the coupon (and interest) payments to the investor, (3) shows the eventual repayment of the outstanding principal at maturity and (4) shows the eventual payments from the SPV to the sponsor.

1.1.2 Restrictions

Due to the vast amount of different payment structures for catastrophe bonds in the market today, it is not possible to analyze losses and risks without first making some restrictions. In this thesis, only catastrophe bonds with coupons of constant rate and proportional to the outstanding principal at the time of payment will be considered. Eventual interest yields from the collateral account (which are normally paid forward to the investors) will be disregarded and joined with the coupons, and the canonical payment structure will be that demonstrated in Example 1 below. Each event qualified as a catastrophe according to the contract terms will be assumed to incur an immediate instrument loss—that is, an immediate reduction of outstanding principal—although other ways of loss calculation, such as annual aggregation of catastrophe event severities, are common [14]. Also, all types of transaction costs will be disregarded.

1.1.3 Examples of catastrophe bonds

Example 1 is taken from [14] and illustrates typical cash flow scenarios for an investor of a catastrophe bond, where it is showcased that both the degrees of severity of the catastrophes and their times of occurrence are important. Example 2 is also taken from [14] and is a real-world example of how the terms of a cat bond may be specified. Example 3 is a catastrophe bond with somewhat different and more complex properties, taken from [5].

Example 1 (Cash flow). *Consider a one-year catastrophe bond with original principal* 100 *and quarterly coupon payments, where the coupon rate is* $10\ \%$ *per annum, quarterly compounded. If no catastrophe occurs, the outstanding principal is constantly equal to* 100 *over the whole term, and the cash flow for the investor is as in Table 1.1 below. The nominal net cash flow is* 10.

Table 1.1: Cash flow for investor if no catastrophe occurs.

t (years)	0	0.25	0.5	0.75	1
Cash flow	-100	2.5	2.5	2.5	102.5

If, instead, a catastrophe occurs at $t = 0.6$ *such that the sponsor immediately withdraws the whole original principal from the collateral account, the cash flow would be as in Table 1.2 below. The nominal net cash flow is* -95.

Table 1.2: Cash flow for investor if a large catastrophe occurs.

t (years)	0	0.25	0.5	0.75	1
Cash flow	-100	2.5	2.5	0	0

Lastly, consider the case of a smaller catastrophe occurring at $t = 0.6$ *leading to a reduction of the outstanding principal by* 50. *Since the coupon payments are proportional to the outstanding principal, the cash flow would be as in Table 1.3. The nominal net cash flow is* 42.5.

Table 1.3: Cash flow for investor if a smaller catastrophe occurs.

t (years)	0	0.25	0.5	0.75	1
Cash flow	-100	2.5	2.5	1.25	51.25

Example 2 (Calypso Capital II Ltd). *The following catastrophe bond is a real-world example with the following specifications [14]:*

Table 1.4: Specifications of the Calypso Capital II Ltd catastrophe bond.

SPV:	Calypso Capital II Ltd, class A
Sponsor:	AXA Global P&C
Trigger mechanism:	Industry loss index
Peril:	European windstorm
Size:	185 million EUR
Risk period:	Jan 1, 2014–Dec 31, 2016
Annual attachment probability:	1.45 %
Annual exhaustion probability:	0.61 %

Example 3 (Winterthur windstorm bonds). *In 1998, the Swiss insurance company Winterthur, named after the city in which it is based, issued three-year annual coupon bonds with face value* $4\,700$ *Swiss francs. The coupon rate of* $2.25\ \%$ *was subject to the risk of windstorm damage of automobiles owned by the company's insurance customers during a specified period each year—if the number of automobile windstorm claims during the period exceeded* $6\,000$*, the coupon for the corresponding year was not paid. Furthermore, at maturity, each bond was convertible to five shares of Winterthur common stock. The deal was described in the trade press and was valued in an article in 1999 [5].*

1.2 Catastrophe bonds in portfolios

1.2.1 Comparison with equity and bond market

Given the risk diversification benefits of taking positions in catastrophe bonds for insurance and reinsurance companies, it is natural to investigate whether they are also able to benefit "ordinary" investors in terms of return and risk. The market has grown steadily over the years, with about 32 billion USD of public deals outstanding as of 2018 [22]. An overview of the returns from the catastrophe bond market after 2006 is seen in Figure 1.2, where catastrophe bond market indices are compared to equity and bond indices. The dark blue line shows the *Swiss Re Global Cat Bond Total Return Index*, which is a market-value weighted index of cat bonds, excluding life and health bonds. The yellow line is the *Eurekahedge ILS Advisers Index*, which is an equally weighted index of 31 constituent funds, designed to provide a broad measure of the performance of underlying hedge fund managers who explicitly allocate to insurance-linked investments. The light blue line is the *MSCI World Net Total Return Index*, representing the equity market, and the green line is the *Bloomberg Barclays Capital Global Aggregate Bond Index*, representing the bond market. Figures 1.3, 1.4 show, respectively, descriptive statistics and cross-correlations.

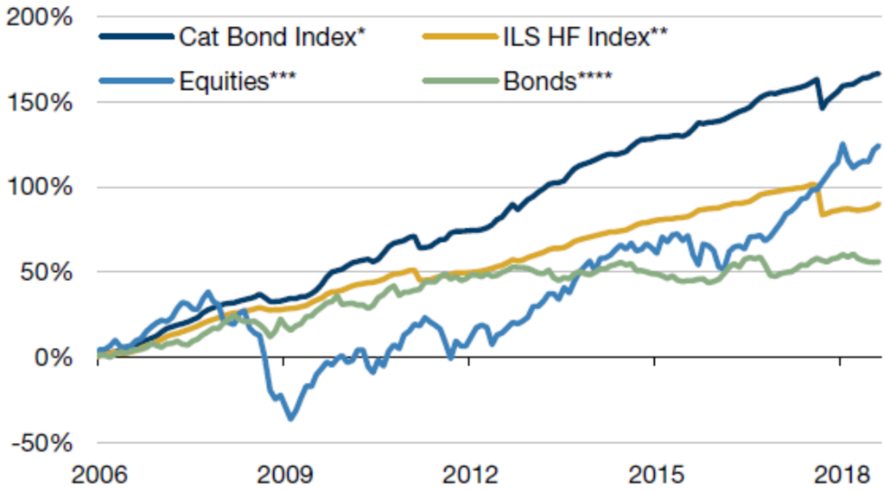

Figure 1.2: Comparison of returns for two market indices for catastrophe bonds (dark blue and yellow) against those for equity and bonds, 2006–2019 [22]

	CAT BOND INDEX*	ILS HF INDEX**	EQUITIES***	BONDS****
Total Return	166.4%	89.9%	124.3%	55.9%
Volatility	3%	3%	15%	5%
Annulised return	7.9%	5.1%	6.5%	3.5%
Sharpe Ratio	2.39	1.69	0.45	0.69

Figure 1.3: Comparison of descriptive statistics for two market indices for catastrophe bonds against those for equity and bonds [22].

	CAT BOND INDEX*	ILS HF INDEX**	EQUITIES***	BONDS****
Cat Bond Index*	**1.00**			
ILS HF Index**	0.87	**1.00**		
Equities***	0.18	0.10	**1.00**	
Bonds****	0.17	0.14	0.39	**1.00**

Figure 1.4: Cross-correlations between two market indices for catastrophe bonds and those for equity and bonds [22].

By inspection of the figures provided, it is easy to see that catastrophe bonds present appealing alternatives for investors. First, Figure 1.2 shows higher as well as less volatile returns for the catastrophe bond market, which is confirmed in Figure 1.3 where Sharpe ratios of 2.39 and 1.69 are contrasted to 0.45 and 0.69. Second, Figure 1.4 provides evidence of low correlations between catastrophe bond returns and the equity and bond markets, which motivates their often claimed diversification potential.

1.2.2 Literature review

The topic of the risk diversification benefits of catastrophe bonds in portfolios, either as complement to other assets or entirely on their own, has been addressed in numerous published studies as well as online articles. A general discussion of the topic is done in [23], as well as a review of the contemporary literature on catastrophe bond pricing and risk modeling. A more technical study of the quantitative effects of including catastrophe bonds in portfolios is done in [21]. [11], [4] and [22] provide short presentations of the topic accessible for the general reader and illustrate clearly why catastrophe bonds are interesting from an investor's perspective.

In particular, a variety of valuation models for catastrophe bonds have been proposed since the early 2000s. [12] studied in 1999 the general concept of catastrophe bonds from a quantitative finance perspective. Soon thereafter, [5] developed in 2000 a pricing model based on a term structure model and certain probability features for the catastrophe risk exposure, which became one of the most frequently cited papers on the subject. Many authors have focused on the problem of finding relevant factors and developing some regression model for the price. For example, a model based on factors such as interest rate, credit rating and expected loss is proposed in [15], and an analysis is done in [2] based on a series of least-squares regressions corrected for heteroskedasticity and autocorrelation, where significances of additional factors such as covered territory, sponsor and reinsurance cycle are determined. [7] investigated the effect of a major catastrophic event, such as the hurricane Katrina, on catastrophe bond prices, and [16] also examined the impact of financial market volatility. For a thorough evaluation of similar models in literature, see e g [13] and [18].

Although the majority of the existing literature has been focused on evaluating the significance of price determinants, other authors have addressed the modeling of the probability distributions of catastrophe occurrences and their influence on the price and risk of catastrophe bonds. A simplistic yet illustrative model is developed in [1] based on somewhat different assumptions than those used in this thesis, where the possibility of one catastrophic event occurring is modeled by a Poisson-distributed stopping time. A more sophisticated pricing model is presented in [20], based on a Cox–Ingersoll–Ross interest rate and an aggregated loss specified as a compound Poisson process with log-normal jumps; these authors, however, also use a somewhat different payoff structure than in this thesis, although the underlying loss process is almost identical to that used here. Poisson processes also occur in the loss models in [14].

Other miscellaneous topics on catastrophe bonds addressed in literature include moral hazard issues between the sponsor and the investors [20, 6] and role of credit rating classes [9].

1.3 Aim and research questions

The thesis at hand aims to investigate the risk diversification potential of catastrophe bonds in equity and bond portfolios, as well as the performance of portfolios only consisting of catastrophe bonds. More precisely, by developing a loss model for catastrophe bonds and optimizing on risk measures such as value-at-risk and conditional value-at-risk in a one-period portfolio optimization problem, the risk diversification effect is quantified for given model parameters. Focus will be on theoretical and computational aspects of the models rather than on making inferences from real-world data. The thesis is based on the following research questions:

- *How can one model the return of a catastrophe bond given specifications of the distributions of catastrophe frequency and severity?*
- *How can one minimize the risk associated with a portfolio of ordinary assets and/or catastrophe bonds subject to some mean return constraint?*
- *How is the tradeoff between mean return and risk affected by the inclusion of a catastrophe bond in a portfolio of ordinary assets?*
- *What are the characteristics of portfolios consisting of only catastrophe bonds?*

Chapter 2

Theoretical framework

In order to precisely describe the models used in this thesis, it is necessary to first acquaint the reader to some of the necessary mathematical theory. The following sections are brief summaries of chosen topics and are by no means exhaustive of all theory used in this thesis.

2.1 General mathematical notation

A *set* S with elements $s_1, ..., s_n$ is denoted by $S = \{s_1, ..., s_n\}$, for which we shall often use the shorthand notation $S = \{s_i\}_{i=1}^n$. Similarly, a collection $\{X(t) : t \in \mathbb{R}, t \geq 0\}$ will be denoted by $\{X(t)\}_{t\geq 0}$. A *vector* v with components $v_1, ..., v_n$ will be denoted by $v = (v_1, ..., v_n)^{\mathrm{T}}$, where the superscript T indicates *transpose* and where the convention is that vectors are column vectors. Similarly, a *matrix* A where A_{ij} is the component in row i and column j will be denoted by $A = (A_{ij})_{i,j}$. For a vector v, $\operatorname{diag} v$ is the matrix with the components of v on its diagonal and zero elsewhere. The *one-vector* $\mathrm{e} = (1, ..., 1)^{\mathrm{T}}$ and the *identity matrix* $I = \operatorname{diag} \mathrm{e}$ are often used when formulating problems on vector form.

For a multivariate, scalar-valued, twice differentiable function $f : \mathbb{R}^n \to \mathbb{R}$ written as $f(x)$, the *gradient*

$$\frac{\partial f(x)}{\partial x} = \left(\frac{\partial f(x)}{\partial x_1}, ..., \frac{\partial f(x)}{x_n} \right)^{\mathrm{T}}$$

is a vector collecting all partial derivatives, and the *Hessian*

$$\frac{\partial^2 f(x)}{\partial x^2} = \left(\frac{\partial^2 f(x)}{\partial x_i \, \partial x_j} \right)_{i,j=1}^{n}$$

is the matrix collecting all second partial derivatives. If $f(x) = (f_1(x), ..., f_m(x))^{\mathrm{T}}$ is vector-valued, the *Jacobian* $\partial f(x)/\partial x$ is the matrix

$$\frac{\partial f(x)}{\partial x} = \left(\frac{\partial f_i(x)}{\partial x_j} \right)_{i,j=1}^{m,n}.$$

In calculations, it is often convenient to introduce the *indicator function* $1(A)$, which takes the value 1 if the predicate A is true and 0 otherwise—in particular, we have the *Kronecker delta* $\delta_{ij} = 1(i = j)$.

2.2 Probability theory

The following is a short summary of the most important parts—for a thorough mathematical treatment of probability theory, see e g [19].

2.2.1 Probability spaces and random variables

A *probability space* is a triplet $(\Omega, \mathcal{F}, \mathrm{P})$, where Ω is the *sample space*, a set containing all possible events, where $\mathcal{F}$ is the associated *σ-algebra*, consisting of subsets of Ω, and where $\mathrm{P} : \mathcal{F} \to [0, 1]$ is a *probability measure*, assigning for each event $A \in \Omega$ a probability $\mathrm{P}\,A$. σ-algebras and the probability measures have to satisfy certain properties (see [19] for details). Of particular interest is the *Borel σ-algebra $\mathcal{B}$*, defined as the smallest σ-algebra on $\mathbb{R}$ containing all open intervals.

A real-valued *random variable* on a probability space $(\Omega, \mathcal{F}, \mathrm{P})$ is a function $X : \Omega \to \mathbb{R}$ such that $X^{-1}(B) = \{\omega : X(\omega) \in B\} \in \mathcal{F}$ for each $B \in \mathcal{B}$. Its distribution is described by the *cumulative distribution function* (cdf) $F_X(x) = \mathrm{P}(X \leq x)$, which is non-decreasing and ranges from 0 to 1. If $F_X(x)$ is differentiable, the distribution of X is also determined by the *probability density function* (pdf) $f_X(x) = \partial F_X(x)/\partial x$. Else, if X only takes values in a discrete set S, its distribution can also be described by the *probability mass function* (pmf) $p_X(x) = \mathrm{P}(X = x)$, $x \in S$. Two random variables X, Y are *independent* if and only if their joint cdf $F_{X,Y}(x, y) = \mathrm{P}(X \leq x, Y \leq y)$ can be written as $F_{X,Y}(x, y) = F_X(x)F_Y(y)$.

The *expectation* $\mathrm{E}\,X$ of a random variable X is given by the Lebesgue–Stieltjes integral (for the precise meaning of this, see e g [19])

$$\mathrm{E}\,X = \int_{-\infty}^{\infty} x\, dF_X(x) = \int_{-\infty}^{\infty} x f_X(x)\, dx,$$

where the latter is well-defined only if $f_X(x)$ exists. The *variance* $\mathrm{Var}\,X$ of X is given by

$$\mathrm{Var}\,X = \mathrm{E}\left[(X - \mathrm{E}\,X)^2\right] = \mathrm{E}[X^2] - \mathrm{E}[X]^2,$$

and the *covariance* $\mathrm{Cov}[X, Y]$ between two random variables X, Y is given by

$$\mathrm{Cov}[X, Y] = \mathrm{E}\left[(X - \mathrm{E}\,X)(Y - \mathrm{E}\,Y)\right] = \mathrm{E}[XY] - \mathrm{E}\,X\,\mathrm{E}\,Y.$$

If $X = (X_1, ..., X_n)^{\mathrm{T}}$ is a vector, $\mathrm{E}\,X$ is the componentwise expectation $\mathrm{E}\,X = (\mathrm{E}\,X_1, ..., \mathrm{E}\,X_n)^{\mathrm{T}}$. The *covariance matrix* $\mathrm{Var}\,X = (\mathrm{Var}\,X)_{i,j=1}^n$ is defined componentwise by $(\mathrm{Var}\,X)_{ij} = \mathrm{Cov}[X_i, X_j]$ or in vector form by $\mathrm{Var}\,X = \mathrm{E}[XX^{\mathrm{T}}] - \mathrm{E}[X]\,\mathrm{E}[X]^{\mathrm{T}}$.

2.2.2 Common distributions

Below are some common distributions of random variables, described by their cdf's and pdf's/pmf's.

Definition 1 (Uniform). *A random variable X is uniformly distributed on the interval (a, b), denoted $X \sim \mathrm{U}(a, b)$, if*

$$F_X(x) = \begin{cases} 0, & x < a, \\ \dfrac{x-a}{b-a}, & a \le x \le b, \\ 1, & x > b, \end{cases}$$

and

$$f_X(x) = \begin{cases} 0, & x < a, x > b, \\ \dfrac{1}{b-a}, & a \le x \le b. \end{cases}$$

Definition 2 (Normal). *A scalar-valued random variable X is normally distributed with mean μ and variance σ^2, denoted $X \sim \mathrm{N}(\mu, \sigma^2)$, if*

$$F_X(x) = \frac{1}{2}\left(1 + \operatorname{erf}\left(\frac{x-\mu}{\sigma\sqrt{2}}\right)\right)$$

and

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right).$$

A vector-valued random variable X is multivariate normal with mean μ and covariance matrix Σ, denoted by $X \sim \mathrm{N}(\mu, \Sigma)$, if $a^{\mathrm{T}}X \sim \mathrm{N}(a^{\mathrm{T}}\mu, a^{\mathrm{T}}\Sigma a)$ for all a.

Definition 3 (Log-normal). *A random variable X is log-normally distributed with mean μ and variance σ^2, denoted $X \sim \mathrm{LN}(\mu, \sigma^2)$, if*

$$F_X(x) = \frac{1}{2}\left(1 + \operatorname{erf}\left(\frac{\ln x-\mu}{\sigma\sqrt{2}}\right)\right)$$

and

$$f_X(x) = \frac{1}{\sigma x\sqrt{2\pi}} \exp\left(-\frac{(\ln x-\mu)^2}{2\sigma^2}\right).$$

In particular, a scalar-valued random variable X is $\mathrm{LN}(\mu, \sigma^2)$ if and only if $\ln X \sim \mathrm{N}(\mu, \sigma^2)$, and a vector-valued random variable X is $\mathrm{LN}(\mu, \Sigma)$ if and only if $\ln X \sim \mathrm{N}(\mu, \Sigma)$.

Definition 4 (Binomial). *A random variable X is binomially distributed with parameters $n \in \mathbb{N}$ and $0 < p < 1$, denoted $X \sim \mathrm{Bin}(n, p)$, if*

$$F_X(x) = \begin{cases} 0, & x < 0, \\ \displaystyle\sum_{k=0}^{\lfloor x \rfloor} \binom{n}{k} p^k (1-p)^{n-k}, & x \ge 0, \end{cases}$$

and

$$p_X(k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k \in \mathbb{N} \cup \{0\}.$$

Definition 5 (Poisson). *A random variable X is Poisson distributed with rate λ, denoted $X \sim \mathrm{Po}(\lambda)$, if*

$$F_X(x) = \begin{cases} 0, & x < 0, \\ e^{-\lambda} \sum_{k=0}^{\lfloor x \rfloor} \frac{\lambda^k}{k!}, & x \geq 0, \end{cases}$$

and

$$p_X(k) = e^{-\lambda} \frac{\lambda^k}{k!}, \quad k \in \mathbb{N} \cup \{0\}.$$

Definition 6 (Exponential). *A random variable X is exponentially distributed with rate λ, denoted $X \sim \mathrm{Exp}(\lambda)$, if*

$$F_X(x) = \begin{cases} 0, & x < 0, \\ 1 - e^{-\lambda x}, & x \geq 0, \end{cases}$$

and

$$f_X(x) = \begin{cases} 0, & x < 0, \\ \lambda e^{-\lambda x}, & x \geq 0. \end{cases}$$

2.2.3 Poisson processes

A *Poisson process* can be thought of as a counting process of events which occur with constant rate and independently of each other. Due to its convenient properties and relatively weak assumptions, it is widely used for modeling stochastic phenomena in numerous areas in science and engineering. Over time, it is nondecreasing and piecewise constant, with discontinuities at certain jump points.

Example 4 (Bicycle counter). *Consider a bridge on which a bicycle counter has been installed, displaying the number of bicycles that have crossed the bridge that day. Disregarding the fact that the bicycle traffic is likely to be heavier at certain times of the day, e g rush hour, and assuming that the event of one cyclist crossing the bridge is independent of those of all other cyclists crossing, we can model the count at each time of the day as a Poisson process.*

Formally, Poisson processes are defined as follows:

Definition 7 (Poisson process). *A Poisson process with rate λ is a time-indexed collection $\{M(t)\}_{t \geq 0}$ of random variables such that*

(i) $M(0) = 0$,

(ii) for each finite collection $0 = t_0 < t_1 < t_2 < \ldots$*, the increments* $\{M(t_{i+1}) - M(t_i)\}_{i \geq 0}$ *are independent, and*

(iii) $M(t) \sim \mathrm{Po}(\lambda t)$ *for all* t*.*

Poisson processes have many interesting properties, one of which is *memorylessness* in the sense that at any given time, the distribution of the time of the next jump does not depend on the time since the last jump. Letting $0 = t_0 < t_1 < t_2 < \ldots$ be the times of the jumps, i e $\{t_i\}_{i \geq 1} = \{t \geq 0 : M(t^-) \neq M(t^+)\}$, and letting $\Delta t_i = t_i - t_{i-1}$ for all i, we have that

$$\mathrm{P}(\Delta t_i > t + t' \mid \Delta t_i > t') = \mathrm{P}(\Delta t_i > t).$$

One can prove that the *interarrival times* $\{\Delta t_i\}_{i \geq 0}$ are in fact independent and identically $\mathrm{Exp}(\lambda)$-distributed, where λ is the rate of the Poisson process. Furthermore, we have the following property about the distribution of $\{t_i\}_{i \geq 1}$, which will be important for sampling Poisson processes:

Theorem 1. *Let* $T > 0$*, let* $N \sim \mathrm{Po}(\lambda T)$ *for some constant* $\lambda > 0$*, and let* $\{\tau_i\}_{i=1}^N$ *be conditionally independent given* N *and* $\mathrm{U}(0, T)$*-distributed. Then the collection* $\{M(t)\}_{t \in [0,T]}$*, where*

$$M(t) = \sum_{i=1}^{N} 1(\tau_i \leq t)$$

for each t*, is a Poisson process with rate* λ *on* $[0, T]$*.*

Proof. Properties (i) and (ii) in Definition 7 are trivially satisfied due to the indicator functions. For (iii), note that $M(t) \mid N = n \sim \mathrm{Bin}(t/T, n)$ since the event $M(t) = k$ amounts to finding k of the τ_i below and $n - k$ above t. Hence, by the tower property, direct calculation of the pmf gives

$$\begin{aligned}
p_{M(t)}(k) &= \mathrm{P}(M(t) = k) \\
&= \mathrm{E}\left[\mathrm{P}(M(t) = k \mid N)\right] \\
&= \mathrm{E}\left[\binom{N}{k}\left(\frac{t}{T}\right)^k \left(\frac{T-t}{T}\right)^{N-k}\right] \\
&= \sum_{n=0}^{\infty} \binom{n}{k}\left(\frac{t}{T}\right)^k \left(\frac{T-t}{T}\right)^{n-k} \frac{(\lambda T)^n}{n!} e^{-\lambda T} \\
&= e^{-\lambda T} \frac{(\lambda t)^k}{k!} \sum_{n=k}^{\infty} \frac{(\lambda(T-t))^{n-k}}{(n-k)!} \\
&= e^{-\lambda t} \frac{(\lambda t)^k}{k!},
\end{aligned}$$

which is identified as the pmf of a $\mathrm{Po}(\lambda t)$ distribution.

As a generalization of the Poisson process, which only counts the number of jumps until a certain time, the *compound Poisson process* has variable jumps which are iid random variables. The compound Poisson process can be defined as follows:

Definition 8 (Compound Poisson process). *Let $\{M(t)\}_{t\geq 0}$ be a Poisson process with rate λ. A compound Poisson process with rate λ is a collection $\{P(t)\}_{t\geq 0}$ of random variables defined by*

$$P(t) = \sum_{i=1}^{M(t)} D_i$$

for each t, where $\{D_i\}_{i\geq 1}$ are iid random variables independent of $\{M(t)\}_{t\geq 0}$.

Note that by virtue of Theorem 1, we can write

$$M(t) = \sum_{i=1}^{M(T)} 1(\tau_i \leq t)$$

for each t, $\{\tau_i\}_{i=1}^{M(T)}$ being conditionally independent given $M(T)$ and $\mathrm{U}(0, T)$-distributed. The process $\{P(t)\}_{t\in[0,T]}$ can then be represented as

$$P(t) = \sum_{i=1}^{M(T)} D_i 1(\tau_i \leq t)$$

for each t, where $\{D_i\}_{i=1}^{M(T)}$ are conditionally independent given $M(T)$, independent of $\{\tau_i\}_{i=1}^{M(T)}$ and identically distributed. This will be important for simulating compound Poisson processes.

2.3 Portfolio theory and risk valuation

This section briefly summarizes some of the most important parts from standard portfolio theory and risk valuation—for an introduction to the topic, see e g [17].

2.3.1 Preliminaries

Let $[0, T]$ be a given time period, and consider n risky assets with prices $\{S_j(t)\}_{j=1}^{n}$ and a risk-free asset with price $S_0(t)$ for $0 \leq t \leq T$. Writing $S(t) = (S_1(t), ..., S_n(t))^{\mathrm{T}}$, a position in these assets—a *portfolio*—is represented by a vector

$$h(t) = (h_1(t), ..., h_n(t))^{\mathrm{T}}$$

and a scalar $h_0(t)$, where the investor is $h_i(t)$ units long in asset j at time t for each j, and where the total portfolio value $V(t)$ at time t is given by

$$V(t) = \sum_{j=0}^{n} h_j(t) S_j(t) = h_0(t) S_0(t) + h(t)^{\mathrm{T}} S(t).$$

Here, a negative value of $h_j(t)$ represents a short position.

For the purposes of this thesis, we will assume that the positions are fixed during the whole time period and write $h(t) = h$ and $h_0(t) = h_0$—in other words, we will restrict ourselves to solve the one-period optimal asset allocation problem. Without loss of generality, suppose that the initial value $V(0)$ of the portfolio is 1. It is often convenient to let $w_j = h_j S_j(0)$ for each j and write $V(t)$ on the form

$$V(t) = \sum_{j=0}^{n} w_j \frac{S_j(t)}{S_j(0)} = w_0 R_0(t) + w^{\mathrm{T}} R(t),$$

where $R_0(t) = S_0(t)/S_0(t)$ and

$$R(t) = \left(\frac{S_1(t)}{S_1(0)}, ..., \frac{S_n(t)}{S_n(0)} \right)^{\mathrm{T}}$$

is the vector of (geometric) returns over the period $[0, t]$ and where $\sum_{i=0}^{n} w_i = 1$.

2.3.2 Mean-variance analysis

In *mean-variance analysis*, the aim is to find the optimal allocation (w_0, w) of resources such that the expected portfolio value $\mathrm{E}\, V(T)$ at the end of the time period is maximized and such that the variance $\mathrm{Var}\, V(T)$ is minimized. The latter can be interpreted as a quantification of the risk carried by the investor. This fundamental tradeoff between the two objectives can be illustrated by drawing an *efficient frontier* in a plot of $\sqrt{\mathrm{Var}\, V(T)}$ against $\mathrm{E}\, V(t)$—that is, a curve on which each point is optimal in the sense that one cannot improve one objective without worsening the other. Such an efficient frontier can be produced by solving multiple optimization problems minimizing the variance subject to a constraint on the mean return, varying the latter between suitable values.

Consider first the case when all assets are risky, which corresponds to the restriction $w_0 = 0$. Letting $\mu = \mathrm{E}\, R(T)$ be the mean return vector of the risky assets and $\Sigma = \mathrm{Var}\, R(T) = \mathrm{E}[R(T)R(T)^{\mathrm{T}}] - \mu\mu^{\mathrm{T}}$ be the covariance matrix of returns, we have $\mathrm{E}\, V(T) = w^{\mathrm{T}}\mu$ and $\mathrm{Var}\, V(T) = w^{\mathrm{T}}\Sigma w$, so our optimization problem may be formulated as

$$\begin{aligned} \underset{w}{\text{minimize}} \quad & w^{\mathrm{T}}\Sigma w \\ \text{subject to} \quad & \mu^{\mathrm{T}} w \geq \mu_{\mathrm{req}}, \\ & \mathrm{e}^{\mathrm{T}} w \leq 1, \end{aligned} \tag{2.1}$$

where μ_{req} is the required minimum mean return and e is the one-vector. The constraint $\mathrm{e}^{\mathrm{T}} w \leq 1$ rather than $\mathrm{e}^{\mathrm{T}} w = 1$ is due to the possibility that one might not want to invest all initial capital if no risk-free asset is available. The solutions to (2.1) are given by the following result [17]:

Theorem 2. *All solutions to (2.1) can be written on the form*

$$w = \frac{1}{\lambda}\Sigma^{-1}\left(\mu - \frac{\max\{\mathrm{e}^{\mathrm{T}}\Sigma^{-1}\mu - \lambda, 0\}}{\mathrm{e}^{\mathrm{T}}\Sigma^{-1}\mathrm{e}}\mathrm{e}\right),$$

where $\lambda > 0$ is a parameter.

In a graph with $\sqrt{\operatorname{Var} V(T)} = \sqrt{w^{\mathrm{T}}\Sigma w}$ on the horizontal axis and $\operatorname{E} V(T) = w^{\mathrm{T}}\mu$ on the vertical axis, one can see from this that for $\lambda \geq \mathrm{e}^{\mathrm{T}}\Sigma^{-1}\mu$, the efficient frontier is a straight line, whereas for $\lambda < \mathrm{e}^{\mathrm{T}}\Sigma^{-1}\mu$, the frontier curves down as a hyperbola (see Figure 2.1). The point for which $\lambda = \mathrm{e}^{\mathrm{T}}\Sigma^{-1}\mu$ (marked with a hollow circle in Figure 2.1) corresponds to the *minimum-variance portfolio* where the mean return constraint is ignored completely, and the straight line beyond it can be interpreted as discarding a fraction of the initial capital.

Now, introducing a risk-free asset with rate r_0 and return $e^{r_0 T}$ over $[0, T]$, we have instead $\operatorname{E} V(T) = w_0 e^{r_0 T} + w^{\mathrm{T}}\mu$ (the variance is unchanged) and are to solve

$$\begin{aligned} \underset{w_0, w}{\text{minimize}} \quad & w^{\mathrm{T}}\Sigma w \\ \text{subject to} \quad & w_0 e^{r_0 T} + \mu^{\mathrm{T}} x \geq \mu_{\text{req}}, \\ & w_0 + \mathrm{e}^{\mathrm{T}} w = 1. \end{aligned} \tag{2.2}$$

Similarly, one can show the following [17]:

Theorem 3. *All solutions to (2.2) are on the form*

$$w = \frac{1}{\lambda}\Sigma^{-1}\left(\mu - e^{r_0 T}\right), \quad w_0 = 1 - \mathrm{e}^{\mathrm{T}} w.$$

Thus, it is easy to see that the entire efficient frontier is now a straight line. Comparing this to the previous efficient frontier (dashed line in Figure 2.1), it is not surprising to see that the opportunity to invest in a risk-free asset improves the efficient frontier. It can be shown that the dashed line is always tangent to the solid curve, and the tangency point corresponds to what is called the *tangency portfolio*. The dashed line is sometimes called the *capital market line*, and different points on this line correspond to different risk-aversion preferences. In particular, only the ratio of capital invested in risk-free assets and capital invested in risky assets is changed, while the relative capital allocation of the risky assets remains the same.

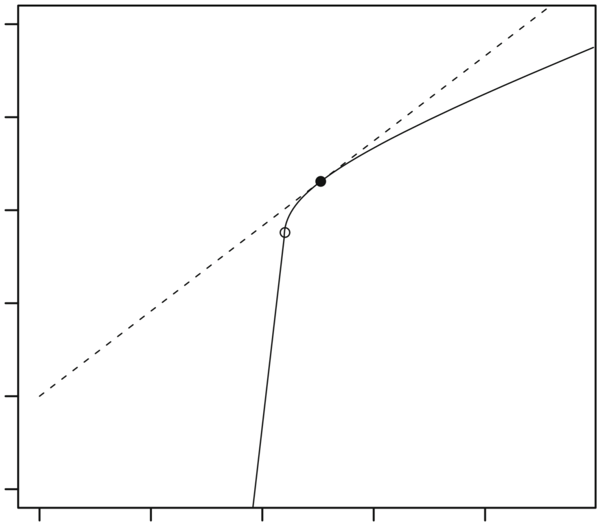

Figure 2.1: Schematic plot of an efficient frontier with (dashed) and without (solid) a risk-free asset, with standard deviation on the horizontal axis and mean return on the vertical axis. The solid dot is the tangency portfolio and the hollow dot is the minimum-variance portfolio.

2.3.3 Stock processes and geometric Brownian motions

Stock processes are commonly modeled by *geometric Brownian motions*, implying log-normal marginal distributions of returns. The model is crude but standard due to its relatively weak assumptions and the existence of closed formulas, and is used in e g Black-Scholes option pricing. Below is a brief outline of the underlying theory—for a more rigorous treatment, see [8].

Consider n stocks with prices $S(t) = (S_1(t), ..., S_n(t))^{\mathrm{T}}$ for $t \in [0, T]$. Built on the idea that the relative increment of $S(t)$ at each time should be split into a drift term and a volatility term, we write

$$dS_i(t) = S_i(t)\,(\mu_i\,dt + \sigma_i\,dW_i(t))$$

for $1 \leq i \leq n$, where $\mu = (\mu_1, ..., \mu_n)^{\mathrm{T}}$ and $\sigma = (\sigma_1, ..., \sigma_n)^{\mathrm{T}}$ are constant parameters and where $\{W(t)\}_{t\in[0,T]} = \{(W_1(t), ..., W_n(t))^{\mathrm{T}}\}_{t\in[0,T]}$ is a multidimensional Brownian motion on $\mathbb{R}^n$ with correlation $\mathrm{E}[dW_i(t)\,dW_j(t)] = \varrho_{ij}\,dt$. We then have the following:

Theorem 4. *The return* $R(T) = (S_1(T)/S_1(0), ..., S_n(T)/S_n(0))^{\mathrm{T}}$ *is distributed as*

$$R(T) \sim \mathrm{LN}\left(\left(\mu - \frac{1}{2}\,\mathrm{diag}(\sigma)\sigma\right)T, \mathrm{diag}(\sigma)\varrho\,\mathrm{diag}(\sigma)T\right)$$

Proof. By Itô's lemma, we have

$$\begin{aligned} d\ln S_i(t) &= \left(\frac{1}{S_i(t)}S_i(t)\mu_i - \frac{1}{2}\frac{1}{S_i(t)^2}S_i(t)^2\sigma_i^2\right)\,dt + \frac{1}{S_i(t)}S_i(t)\sigma_i\,dW_i(t) \\ &= \left(\mu_i - \frac{1}{2}\sigma_i^2\right)\,dt + \sigma_i\,dW_i(t) \end{aligned}$$

or, integrating over $[0, T]$,

$$\ln R_i(T) = \ln \frac{S_i(T)}{S_i(0)} = \left(\mu_i - \frac{1}{2}\sigma_i^2\right) T + \sigma_i W_i(T).$$

On vector form, this amounts to

$$\ln R(T) = \left(\mu - \frac{1}{2}\operatorname{diag}(\sigma)\sigma\right) T + \operatorname{diag}(\sigma) W(T),$$

and the result follows by noting that $W(T) \sim \mathrm{N}(0, \varrho T)$.

2.3.4 Risk measures

In classical mean-variance analysis, variance or standard deviation is used to quantify the risk associated with the return of a portfolio. Being simply the expected squared deviation from the mean, however, the variance does not differentiate between positive and negative deviations and only works in principle when the distribution of the return is completely parametrized by its mean and variance, e g in the case of a normal distribution. This motivates the use of more sophisticated risk measures, which can take into account tail risks and highly skewed return distributions. Formally, we have the following:

Definition 9 (Risk measure). *Let $X \in \mathcal{X}$ be the return at $t = T$ of a portfolio for some linear vector space $\mathcal{X}$ of random variables. A risk measure is a function $\rho : \mathcal{X} \to \mathbb{R} \cup \{\infty\}$, assigning a (possibly infinite) real number to each X representing the risk associated with X.*

The number $\rho(X)$ can be interpreted as the minimum amount of capital that needs to be invested in the risk-free asset in order to make the position X acceptable. Equivalently, for our purposes, it might be convenient to instead consider the discounted loss $L = 1 - V(T)e^{-r_0 T}$ where $V(0) = 1$ and $V(T) = x_0 R_0(T) + x^{\mathrm{T}} R(T)$ are the portfolio values at $t = 0$ and $t = T$, respectively. The following is a list of desirable properties of a risk measure:

1. *Translation invariance.* $\rho(X + ce^{r_0 T}) = \rho(X) - c$ for all $c \in \mathbb{R}$. The interpretation of this is that adding an amount of c in risk-free investments will reduce the risk by the same amount.

2. *Monotonicity.* If $X_1 \geq X_2$, then $\rho(X_1) \leq \rho(X_2)$. That is, if some position surely has higher return than another, then the first position is considered less risky.

3. *Convexity.* $\rho(\lambda_1 X_1 + \lambda_2 X_2) \leq \lambda_1 \rho(X_1) + \lambda_2 \rho(X_2)$ for all $\lambda_1, \lambda_2 \geq 0$ such that $\lambda_1 + \lambda_2 = 1$.

4. *Normalization.* $\rho(0) = 0$. This tells us that not investing anything will also imply zero risk.

5. *Positive homogeneity.* $\rho(\lambda X) = \lambda\rho(X)$ for all $\lambda \geq 0$. That is, if we scale the size of the position by a factor $\lambda \geq 0$, the risk will be scaled by the same amount.

6. *Subadditivity.* $\rho(X_1 + X_2) \leq \rho(X_1) + \rho(X_2)$.

There are numerous commonly used risk measures, some of which are more sophisticated than others and each having its own advantages and disadvantages. For a thorough presentation of different examples of risk measures and their properties, see e g [17]. In this thesis, the study of risk measures will be restricted to *value-at-risk* (VaR) and *conditional value-at-risk* (CVaR).

Definition 10 (Value-at-risk). *The value-at-risk at level $\alpha \in (0,1)$ of a portfolio X with loss L is defined as*

$$\mathrm{VaR}_\alpha(X) = \inf\left\{x \in \mathbb{R} : F_L(x) \geq 1-\alpha\right\},$$

where $F_L(x)$ is the cdf of L. In particular, if F_L^{-1} exists, then one can write $\mathrm{VaR}_\alpha(X)$ more conveniently as

$$\mathrm{VaR}_\alpha(X) = F_L^{-1}(1-\alpha).$$

The VaR can be interpreted as the smallest amount of money required to be invested in the risk-free asset at $t = 0$ such that the probability of a positive loss at $t = T$ is at most α. In practical applications, one commonly sets α to 0.01, 0.05 or 0.1. The VaR captures to some extent information about the right tail of the distribution of L—that is, loosely speaking, optimizing on VaR corresponds to controlling the presence of abnormally large losses. It is easy to show that VaR is translation-invariant, monotonous, and positive homogeneous; however, it is neither convex nor subadditive, which is one of its drawbacks. Another serious drawback is that VaR ignores the right tail of L beyond the $1-\alpha$ level quantile, which makes it possible to miss risks which are highly unlikely but very large in magnitude. This is especially relevant in the context of catastrophe bonds, where tail risks make up the main characteristics of the instrument.

As an improvement of VaR, one can instead use CVaR:

Definition 11 (Conditional value-at-risk). *The conditional value-at-risk at level $\alpha \in (0,1)$ of a portfolio X with loss L is defined as*

$$\mathrm{CVaR}_\alpha(X) = \frac{1}{\alpha}\int_0^\alpha \mathrm{VaR}_{\alpha'}(X)\,d\alpha'.$$

The motivation behind CVaR that it, instead of measuring where the "boundary" to the α-level upper tail lies, measures the mean in the whole tail. Indeed, when $F_L(x)$ is continuous, it is easy to show that

$$\mathrm{CVaR}_\alpha(X) = \mathrm{E}\left[L \mid L \geq \mathrm{VaR}_\alpha(X)\right].$$

It is intuitive that measuring the mean tail loss better incorporates extreme loss events. Furthermore, one can show that CVaR, as opposed to VaR, is also subadditive and convex apart from being translation-invariant, monotonous and positive homogeneous.

2.4 Numerical methods

2.4.1 Monte Carlo sampling

Apart from in very special cases, for a general portfolio with value X at $t = T$, the expression for $\rho(X)$ is analytically intractable and not possible to evaluate directly. Instead, the standard way to proceed is by *Monte Carlo sampling* of the returns, from which $\rho(X)$ can be estimated. In particular, suppose that we have independent and identically distributed (iid) realizations $\{X_i\}_{i=1}^N$ of X, which we may obtain if the distribution of X is known. For a portfolio with weight vector w and known joint distribution of the return vector $R(T)$ corresponding to the asset values $\{S_j(T)\}_{j=0}^n$ at $t = T$, we can obtain iid samples $\{R_i\}_{i=1}^N$ of $R(T)$ where $X_i = w_0 e^{r_0 T} + w^{\mathrm{T}} R_i$ for each i. In turn, letting $\{L_i\}_{i=1}^N$ be such that $L_i = -X_i e^{-r_0 T}$ for each i, the cdf $F_L(x)$ of the discounted loss L may be estimated by the *empirical cdf*

$$F_L^{\mathrm{emp}}(x) = \frac{1}{N} \sum_{i=1}^{N} 1(L_i \leq x).$$

Note that with this, we may estimate the VaR of X as

$$\mathrm{VaR}_\alpha(X) \approx \inf\{x \in \mathbb{R} : F_L^{\mathrm{emp}}(x) \geq 1 - \alpha\}$$

and the CVaR of X as

$$\mathrm{CVaR}_\alpha(X) \approx \frac{1}{\alpha} \int_0^\alpha \inf\{x \in \mathbb{R} : F_L^{\mathrm{emp}}(x) \geq 1 - \alpha'\} \, d\alpha'.$$

We shall, however, refrain from presenting details of explicitly computing these, as an arguably better method will be presented in Section 3.2.

Next, let $R = (R_{ij})_{i,j}$ be an $N \times (n+1)$ matrix of sampled returns, where R_{ij} is the return of asset j for the ith realization for all $0 \leq j \leq n$ and $1 \leq i \leq N$—that is, each row R_i is the outcome of a realization of the return vector $(R_0(T), R(T))$ including the risk-free asset. Thus, the components of $\ell = \mathrm{e} - e^{-r_0 T} R w$, where we have included w_0 in w, can be seen as outcomes of iid realizations of L with the resource allocation w, and the corresponding risk may be estimated by a function $\rho(\ell)$. For convenience, we shall adopt a slight abuse of notation and use $\rho(\ell)$ and $\rho(X)$ interchangeably, where the former is the sample estimated risk measure and the latter a random variable defining the risk measure.

2.4.2 Optimization

In principle, we are now ready to state our main optimization problem formulation. The inherent tradeoff between a low estimated risk $\rho(\ell)$ and a high estimated expected return $\mathrm{e}^{\mathrm{T}} R w$ can be visualized by computing an efficient frontier similar to that in the case of mean-variance analysis. Again, we choose to optimize on the risk measure subject to a minimum mean return constraint $\mathrm{e}^{\mathrm{T}} R w \geq \mu_{\text{req}}$ and repeating for different values of μ_{req}. The main optimization problem is thus stated as follows:

$$\begin{aligned} \underset{w}{\text{minimize}} \quad & \rho\left(\mathrm{e} - e^{-r_0 T} R w\right) \\ \text{subject to} \quad & \frac{1}{N} \mathrm{e}^{\mathrm{T}} R w \geq \mu_{\text{req}}, \\ & \mathrm{e}^{\mathrm{T}} w = 1. \end{aligned} \tag{2.3}$$

The problem (2.3) is a *nonlinear optimization* problem and typically solved numerically using some optimization solver package. For the purposes of this thesis, we shall use the `fmincon` function in MATLAB, which by default implements an *interior-point algorithm*. In short, the interior-point algorithm casts a generic nonlinear optimization problem

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & f(x) \\ \text{subject to} \quad & h(x) = 0, \\ & g(x) \leq 0 \end{aligned}$$

to an approximate problem

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & f(x) - \mu^{\mathrm{T}} \ln s \\ \text{subject to} \quad & h(x) = 0, \\ & g(x) + s = 0, \end{aligned}$$

for all $\mu > 0$, where s is a vector of *slack variables* and the logarithmic term is called the *barrier function*. At each iteration, the algorithm attempts to minimize a *merit function* on the form

$$f(x) - \mu^{\mathrm{T}} \ln s + \nu \left(h(x)^{\mathrm{T}} h(x) + (g(x) + s)^{\mathrm{T}} (g(x) + s) \right),$$

where violation of the constraints is taken into account by the ν-term. The algorithm relies on specifying the objective function and its gradient—in particular, one must be able to easily compute the gradient $\partial\rho/\partial w$, which is not the case for the above definitions of VaR and CVaR. This issue is addressed in detail and resolved in Section 3.2. For more on optimization theory in general, see e g [24].

Chapter 3

Main work

The main contribution of this thesis is twofold: the loss modeling of catastrophe bonds and the method of direct gradient-based optimization of value-at-risk and conditional value-at-risk. As for loss modeling, we present first a simplistic model—referred to as the *discontinuous uniform model*—intended to showcase the main characteristics of catastrophe bond returns, and then a more advanced model—referred to as the *compound Poisson model*. The former is mainly a special case of the two-point Kumaraswamy model outlined in [14], whereas the latter takes into account the possibility of multiple catastrophic events (which is highly probable for e g bonds with parametric or industry index loss trigger) and the influence of their times of occurrence on the total return of the bond.

As for the direct gradient-based optimization on VaR and CVaR, we are solving the problem of both being computationally intractable in the sense that they are discontinuous and nondifferentiable in their empirical forms, preventing the application of standard optimization algorithms for which a gradient to the function must be supplied. Numerous methods of circumventing this problem are used in literature, such as the one described in the original article [25] in which CVaR was first introduced where the problem is cast into a linear programming form; [14], for example, uses the Lagrangian dual of this formulation. In this thesis, the problems are overcome by the introduction of an auxiliary (fictional) measurement noise making VaR as well as CVaR smooth with easily computable gradients. The method is equivalent to the technique of kernel density estimation and resembles the work in [26].

3.1 Loss modeling of cat bonds

3.1.1 Discontinuous uniform model

Consider Example 2 and, in particular, the fact that the attachment and exhaustion probabilities are given. As a crude initial attempt at modeling the return distribution of a catastrophe bond, we shall exploit this information and divide the outcome at maturity into three distinct cases: (i) the outstanding principal equals the original principal, (ii) the outstanding principal is non-zero but smaller than the original principal, and (iii) the outstanding principal is zero. Denoting by $p_{\rm a}$ and $p_{\rm e}$, respectively, the attachment and exhaustion probabilities, (i) occurs with probability $1 - p_{\rm a}$, (ii) with probability $p_{\rm a} - p_{\rm e}$ and (iii) with probability $p_{\rm e}$.

Let $t \in [0, T]$ be the term of a catastrophe bond and r_0 the risk-free rate, where we are interested in the (nominal) value $V(T)$ at maturity of a portfolio containing only the catastrophe bond and the risk-free asset. Suppose that the original principal is 1 and that the position in the risk-free asset is 0 at $t = 0$, and assume continuous compounding and continuous coupon yield with rate c. Furthermore, in order to keep the model as simple as possible, suppose that at most one catastrophe may occur during the term and that, if it occurs, it occurs immediately after the contract has been entered (that is, at $t = 0^+$). Finally, suppose that given (ii), the principal reduction due to the catastrophe is $\mathrm{U}(0, 1)$. The model built on these assumptions will be referred to as the *discontinuous uniform model*.

For (iii), it is clear that $V(T) = 0$ since the bond loses all of its value as soon as the collateral account is exhausted. For (i), note that the continuously paid coupons are constant in nominal value over the whole term. During each infinitestimal interval $[t, t + dt)$, a coupon of $c\,dt$ is received and immediately invested in the risk-free asset, the value of which will be $ce^{r_0(T-t)}\,dt$ at $t = T$. Integrating, we get that $V(T)$ in this case will amount to

$$V(T) = 1 + \int_0^T ce^{r_0(T-t)}\,dt = 1 + \frac{c}{r_0}\left(e^{r_0 T} - 1\right),$$

where the added 1 is the repayment of the whole original principal. Similarly, for (ii), if the outstanding principal after the catastrophe and at maturity is $P(T)$, we have that

$$V(T) = P(T)\left(1 + \frac{c}{r_0}\left(e^{r_0 T} - 1\right)\right).$$

In other words, in terms of the random variable $P(T)$, where $\mathrm{P}(P(T) = 1) = 1 - p_{\rm a}$, $\mathrm{P}(P(T) = 0) = p_{\rm e}$ and $P(T) \mid 0 < P(T) < 1 \sim \mathrm{U}(0, 1)$, the cdf $F_{V(T)}(x)$ of $V(T)$ may be written as

$$F_{V(T)}(x) = \begin{cases} 0, & x < 0, \\ \dfrac{p_{\rm a} - p_{\rm e}}{1 + c\chi(0,T)}x + p_{\rm e}, & 0 \le x < 1 + c\chi(0,T), \\ 1, & x \ge 1 + c\chi(0,T), \end{cases}$$

where we have introduced the function $\chi(t_1, t_2)$, defined for $0 \le t_1 < t_2 \le T$ as

$$\chi(t_1, t_2) = \int_{t_1}^{t_2} e^{r_0(T-t)}\,dt = \frac{e^{r_0(T-t_1)} - e^{r_0(T-t_2)}}{r_0}.$$

In particular, we have the following:

Proposition 1. *The expected value* $\operatorname{E} V(T)$ *at maturity of a catastrophe bond as defined in the discontinuous uniform model with unit original principal is given by*

$$\operatorname{E} V(T) = (1 + c\chi(0,T)) \left(1 - \frac{p_\mathrm{a} + p_\mathrm{e}}{2}\right).$$

Proof. Direct calculation gives

$$\begin{aligned}\operatorname{E} V(T) &= (1 + c\chi(0,T)) \operatorname{E} P(T) \\ &= (1 + c\chi(0,T)) \left(0p_\mathrm{e} + 1(1 - p_\mathrm{a}) + \operatorname{E}[P(T) \mid 0 < P(T) < 1](p_\mathrm{a} - p_\mathrm{e})\right) \\ &= (1 + c\chi(0,T)) \left(1 - \frac{p_\mathrm{a} + p_\mathrm{e}}{2}\right),\end{aligned}$$

using the fact that $\operatorname{E}[P(T) \mid 0 < P(T) < 1] = \operatorname{E} \mathrm{U}(0,1) = 1/2$.

For illustration, a plot of the cdf of $V(T)$ for parameters $T = 1$, $r_0 = 0.05$, $c = 0.25$, $p_\mathrm{a} = 0.4$ and $p_\mathrm{e} = 0.2$ is shown in Figure 3.1, where $\operatorname{E} V(T) \approx 0.88$.

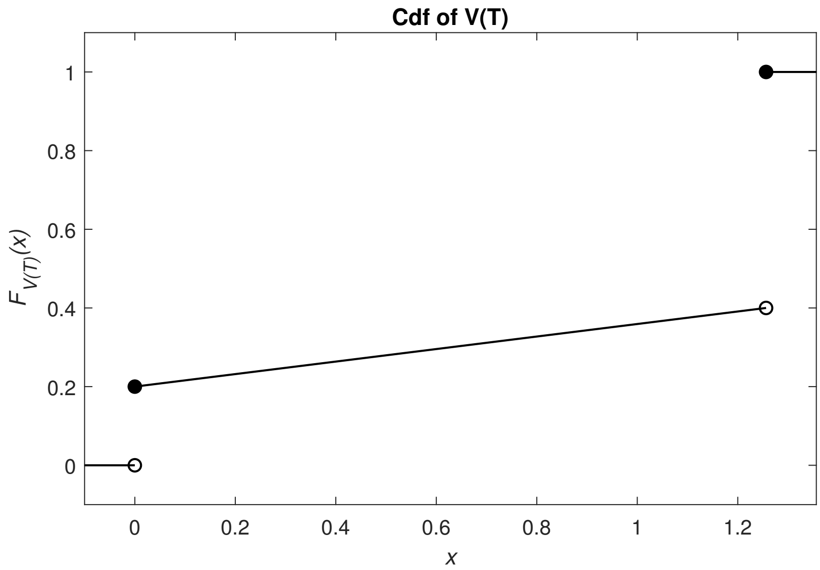

Figure 3.1: Plot of the cdf of $V(T)$ for the discontinuous uniform model with parameters $T = 1$, $r_0 = 0.05$, $c = 0.25$, $p_\mathrm{a} = 0.4$ and $p_\mathrm{e} = 0.2$.

3.1.2 Compound Poisson process

In this model, we shall try to take into account the times of occurrence of the underlying catastrophe events more carefully than for the discontinuous uniform model. Since the coupons are central for the profitability of owning the bond and are affected by the arrival times of the events, as illustrated in Example 1, it is important not only to incorporate the possibility that several catastrophic events may occur but also model their arrival time distributions. Furthermore, it is reasonable to assume that the catastrophic events vary in degree of severity.

Let $\{P(t)\}_{t\in[0,T]}$ be the outstanding principal process over the term $[0,T]$ of a catastrophe bond with unit original principal $P(0)=1$. Motivated by the above, suppose that $\{P(t)\}_{t\in[0,T]}$ follows a compound Poisson process with rate λ_{f}, where $\{M(t)\}_{t\in[0,T]}$ is the associated Poisson process with the same rate counting the number $M(t)$ of catastrophes that have occurred in the time interval $[0,t]$ for each $t\in[0,T]$. Suppose, furthermore, that the catastrophes ($M(T)$ in total) can be represented by pairs $\{(\tau_i,D_i)\}_{i=1}^{M(T)}$, where the τ_i are the arrival times of the catastrophes and D_i are the corresponding reductions in outstanding principal (severities) of the catastrophes. Finally, suppose that the severities $\{D_i\}_{i=1}^{M(T)}$ are conditionally independent given $M(T)$ and identically $\mathrm{Exp}(\lambda_{\mathrm{s}})$-distributed.

Just as for the discontinuous uniform model, we seek the nominal value $V(T)$ of a portfolio only containing the catastrophe bond and the risk-free asset, with initial value $V(0)=1$ all invested in the former. Note that the $\{\tau_i\}_{i=1}^{M(T)}$ are the (not necessarily sorted) jump times in the processes $\{M(t)\}_{t\in[0,T]}$ and $\{P(t)\}_{t\in[0,T]}$, which by the properties of the Poisson process are conditionally independent given $M(T)$ and $\mathrm{U}(0,T)$-distributed. We can thus for each $t\in[0,T]$ write $M(t)$ and $P(t)$ as

$$M(t)=\sum_{i=1}^{M(T)} 1(\tau_i\le t)$$

and

$$P(t)=\max\left\{1-\sum_{i=1}^{M(T)} D_i 1(\tau_i\le t),0\right\},$$

which leads to the following:

Proposition 2. *Letting* $0=\tau_{(0)}\le\tau_{(1)}\le\ldots\le\tau_{(M(T))}\le\tau_{(M(T)+1)}=T$ *be the sorted arrival times, we can write*

$$V(T)=\max\left\{1-\sum_{j=1}^{M(T)} D_j,0\right\}+c\sum_{i=0}^{M(T)}\max\left\{1-\sum_{j=1}^{i} D_{(j)},0\right\}\chi\left(\tau_{(i)},\tau_{(i+1)}\right).$$

Proof. In each interval $\left[\tau_{(i)},\tau_{(i+1)}\right)$, $0\le i\le M(T)$, the outstanding principal is constantly $P(\tau_{(i)})$, and in each infinitesimal interval $[t,t+dt)\subset\left[\tau_{(i)},\tau_{(i+1)}\right)$ we have that the contribution to $V(T)$ due to coupon yield is $cP\left(\tau_{(i)}\right)e^{r_0(T-t)}\,dt$, where c is the continuously compounded coupon rate. Integrating and summing over all subintervals, we get that

$$\begin{aligned}
V(T) &= P(T) + \sum_{i=0}^{M(T)} \int_{\tau_{(i)}}^{\tau_{(i+1)}} cP\left(\tau_{(i)}\right) e^{r_0(T-t)}\, dt \\
&= P(T) + c \sum_{i=0}^{M(T)} P\left(\tau_{(i)}\right) \chi\left(\tau_{(i)}, \tau_{(i+1)}\right) \\
&= \max\left\{1 - \sum_{j=1}^{M(T)} D_j, 0\right\} + c \sum_{i=0}^{M(T)} \max\left\{1 - \sum_{j=1}^{i} D_{(j)}, 0\right\} \chi\left(\tau_{(i)}, \tau_{(i+1)}\right),
\end{aligned}$$

where $P(T)$ is added as the repayment of the outstanding principal at maturity.

Mainly due to the positive part operators $x \mapsto \max\{x, 0\}$, even the expectation of $V(T)$ is intractable, but it is on the other hand easy to sample from $V(T)$ since all distributions are well-defined. In practice, a realization of $V(T)$ is obtained by the following procedure:

1. Draw $M(T) \sim \mathrm{Po}(\lambda_{\mathrm{f}} T)$.
2. Draw $\{\tau_i\}_{i=1}^{M(T)}$ as iid realizations of a $\mathrm{U}(0, T)$-distributed variable and sort in increasing order to obtain $\{\tau_{(i)}\}_{i=1}^{M(T)}$.
3. Draw $\{D_i\}_{i=1}^{M(T)}$ as iid realizations of a $\mathrm{Exp}(\lambda_{\mathrm{s}})$-distributed variable.
4. Compute $V(T)$ according to the above.

3.2 Gradient-based optimization on VaR and CVaR

3.2.1 Variance as risk measure

In order to demonstrate the problem with optimizing directly on VaR and CVaR with their current definitions, consider the case when variance is used as "risk measure"—that is,

$$\rho(\ell) = \frac{1}{N-1} \sum_{i=1}^{N} \left(\ell_i - \frac{1}{N} \sum_{i'=1}^{N} \ell_{i'}\right)^2 = \frac{1}{N-1} \ell^{\mathrm{T}} \left(I - \frac{1}{N} \mathrm{e}\mathrm{e}^{\mathrm{T}}\right) \ell.$$

The gradient $\partial\rho(\ell)/\partial\ell$ with respect to ℓ is given by

$$\frac{\partial\rho(\ell)}{\partial\ell} = \left(\frac{\partial\rho(\ell)}{\partial\ell_1}, \ldots, \frac{\partial\rho(\ell)}{\partial\ell_N}\right)^{\mathrm{T}} = \frac{2}{N-1} \left(I - \frac{1}{N} \mathrm{e}\mathrm{e}^{\mathrm{T}}\right) \ell,$$

and by the chain rule, the gradient $\partial\rho(\ell)/\partial w$ with respect to the portfolio weights w is given by

$$\frac{\partial\rho(\ell)}{\partial w} = \frac{\partial\ell}{\partial w} \frac{\partial\rho(\ell)}{\partial\ell} = -\frac{2e^{r_0 T}}{N-1} R^{\mathrm{T}} \left(I - \frac{1}{N} \mathrm{e}\mathrm{e}^{\mathrm{T}}\right) \ell,$$

using the Jacobian $\partial\ell/\partial w = -e^{r_0 T} R^{\mathrm{T}}$. Hence, for each ℓ, both the function value $\rho(\ell)$ and its gradient $\partial\rho(\ell)/\partial w$ are easily computed, which is necessary for using the interior-point algorithm. Moreover, it is apparent that $\rho(\ell)$ is convex and infinitely differentiable in w, guaranteeing that the algorithm will be well-behaved.

For VaR and CVaR, however, we do not have this convenience. In fact, since the empirical cdf $F_L^{\mathrm{emp}}(x)$ is discontinuous and piecewise constant, both VaR and CVaR are discontinuous functions of w and nowhere differentiable. This is a serious problem since it makes it impossible to proceed with well-established gradient-based nonlinear optimization algorithms. Instead, one has to resort to more or less *ad hoc* methods such as that first presented by [25], where one introduces auxiliary variables to transform the optimization problem into linear programming form (which can be more computationally expensive).

3.2.2 Smoothing by auxiliary noise

In this thesis, we shall employ a somewhat different approach to overcome this problem. Noting that the discontinuities and general intractabilities of VaR and CVaR arise as a consequence of $F_L^{\mathrm{emp}}(x)$ being piecewise constant, we seek a way of "smoothing out" the function and making its properties more mathematically convenient—in particular, ideally, we would like our estimate of the cdf to be infinitely differentiable and strictly increasing everywhere, guaranteeing the existence of a corresponding pdf. Note that estimating $F_L(x)$ by $F_L^{\mathrm{emp}}(x)$ is equivalent to exchanging L with a *discrete* random variable L_{disc} with possible outcomes $\{\ell_i\}_{i=1}^N$ with $\mathrm{P}(L = \ell_i) = 1/N$ for each i.

Now, instead of replacing L by L_{disc}, we shall replace L by $L' = L_{\mathrm{disc}} + \varepsilon$ where $\varepsilon \sim \mathrm{N}(0, \epsilon^2)$ is some normally distributed noise independent of L_{disc}. The role of $F_L^{\mathrm{emp}}(x)$ is thus replaced by the cdf $F_{L'}(x)$ of L', which is given by

$$
\begin{aligned}
F_{L'}(x) &= \mathrm{P}(L' \le x) \\
&= \int_{-\infty}^{\infty} \mathrm{P}(L' \le x \mid \varepsilon = x') f_\varepsilon(x')\, dx' \\
&= \int_{-\infty}^{\infty} \frac{1}{N} \sum_{i=1}^{N} 1(\ell_i + x' \le x) k(x')\, dx' \\
&= \frac{1}{N} \sum_{i=1}^{N} \int_{-\infty}^{x-\ell_i} k(x')\, dx' \\
&= \frac{1}{N} \sum_{i=1}^{N} K(x - \ell_i).
\end{aligned}
$$

Here, $k(x)$ and $K(x)$ are the pdf and the cdf, respectively, of a $\mathrm{N}(0, \epsilon^2)$-distributed random variable. Note that with this, $F_{L'}(x)$ becomes infinitely differentiable and everywhere strictly increasing, and, in particular, one also has a corresponding pdf

$$
f_{L'}(x) = \frac{\partial F_{L'}(x)}{\partial x} = \frac{1}{N} \sum_{i=1}^{N} k(x - \ell_i)
$$

which satisfies $f_{L'}(x) > 0$ for all $x \in \mathbb{R}$. All of these properties will turn out to be important in the following. The above smoothing step amounts to making a *kernel density estimation* using a Gaussian kernel with bandwidth ϵ, and illustrations of the effect of this are shown in Figures 3.2, 3.3.

For convenience, we shall in the following use L and L' interchangeably as the discounted loss associated with X, eliminating the use of eventual approximation signs. For sufficiently small noise variances ϵ^2, one can argue that the addition of the noise ε has little effect on the loss distribution.

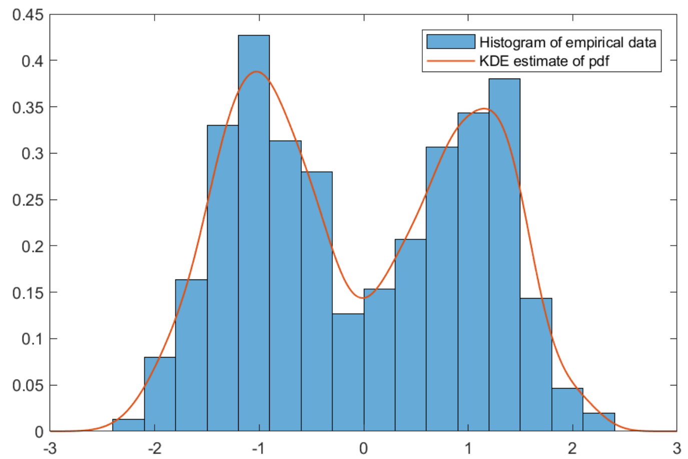

Figure 3.2: Illustration of the effect of kernel density estimation compared to the histogram method, with data drawn from a $\mathrm{N}(\pm 1, 1/2)$ mixture distribution. Note the smooth nature of the KDE estimate.

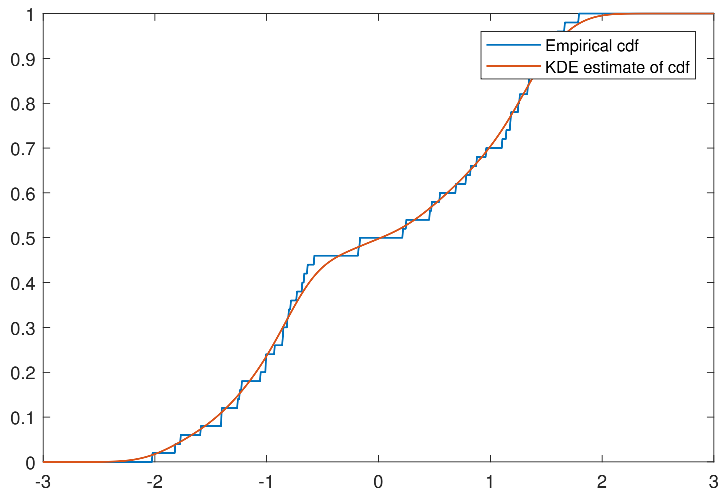

Figure 3.3: Comparison between the empirical cdf and the kernel density estimated cdf with the same data as in 3.2. Again, note the smooth nature of the latter.

For VaR, as a direct consequence of $F_{L'}(x)$ being everywhere strictly increasing, its inverse exists and we have $\mathrm{VaR}_\alpha(\ell) = F_{L'}^{-1}(1-\alpha)$. One can thus in a straightforward fashion evaluate $\mathrm{VaR}_\alpha(\ell)$ numerically, e g by the *Newton-Raphson algorithm*, and it turns out that once this has been done the gradient is also easily evaluated. This is shown in the following:

Proposition 3. *Using L' as substitute for L, $\mathrm{VaR}_\alpha(\ell)$ is the unique solution to the equation*

$$\frac{1}{N}\sum_{i=1}^{N} K(\mathrm{VaR}_\alpha(\ell) - \ell_i) = 1 - \alpha.$$

Furthermore, for each $1 \le i \le N$, we have

$$\frac{\partial \mathrm{VaR}_\alpha(\ell)}{\partial \ell_i} = \frac{k(\mathrm{VaR}_\alpha(\ell) - \ell_i)}{\displaystyle\sum_{i'=1}^{N} k(\mathrm{VaR}_\alpha(\ell) - \ell_{i'})}.$$

Proof. Differentiation with respect to ℓ_i gives

$$\begin{aligned}
0 &= \frac{\partial}{\partial \ell_i}\frac{1}{N}\sum_{i'=1}^{N} K(\mathrm{VaR}_\alpha(\ell) - \ell_{i'}) \\
&= \frac{1}{N}\sum_{i'=1}^{N} k(\mathrm{VaR}_\alpha(\ell) - \ell_{i'})\left(\frac{\partial \mathrm{VaR}_\alpha(\ell)}{\partial \ell_i} - \delta_{ii'}\right) \\
&= \frac{\partial \mathrm{VaR}_\alpha(\ell)}{\partial \ell_i}\frac{1}{N}\sum_{i'=1}^{N} k(\mathrm{VaR}_\alpha(\ell) - \ell_{i'}) - \frac{1}{N}k(\mathrm{VaR}_\alpha(\ell) - \ell_i)
\end{aligned}$$

or

$$\frac{\partial \mathrm{VaR}_\alpha(\ell)}{\partial \ell_i} = \frac{k(\mathrm{VaR}_\alpha(\ell) - \ell_i)}{\displaystyle\sum_{i'=1}^{N} k(\mathrm{VaR}_\alpha(\ell) - \ell_{i'})},$$

where the denominator is clearly always positive as $k(x) > 0$ for all x.

As for CVaR, we may proceed similarly. The main result is summarized in the following:

Proposition 4. *Using L' as substitute for L, we have*

$$\mathrm{CVaR}_\alpha(\ell) = \frac{1}{\alpha N}\sum_{i=1}^{N}\Big(\ell_i K(\ell_i - \mathrm{VaR}_\alpha(\ell)) + \epsilon^2 k(\ell_i - \mathrm{VaR}_\alpha(\ell))\Big)$$

and

$$\frac{\partial \mathrm{CVaR}_\alpha(\ell)}{\partial \ell_i} = \frac{1}{\alpha N}K(\ell_i - \mathrm{VaR}_\alpha(\ell))$$

for each $1 \le i \le N$.

Proof. By using the representation

$$\mathrm{CVaR}_\alpha(\ell) = \frac{1}{\alpha}\int_0^\alpha F_{L'}^{-1}(1-\alpha')\,d\alpha' = \frac{1}{\alpha}\int_{\mathrm{VaR}_\alpha(\ell)}^\infty x f_{L'}(x)\,dx,$$

where the last step is by substituting $1-\alpha' = F_{L'}(x)$, we can integrate by parts to obtain

$$\begin{aligned}
\mathrm{CVaR}_\alpha(\ell) &= \frac{1}{\alpha}\left(x(F_{L'}(x)-1)\Big|_{\mathrm{VaR}_\alpha(\ell)}^\infty - \int_{\mathrm{VaR}_\alpha(\ell)}^\infty (F_{L'}(x)-1)\,dx \right) \\
&= \frac{1}{\alpha}\left(0 - \mathrm{VaR}_\alpha(\ell)(1-\alpha-1) + \frac{1}{N}\sum_{i=1}^N \int_{\mathrm{VaR}_\alpha(\ell)}^\infty K(\ell_i - x)\,dx\right) \\
&= \mathrm{VaR}_\alpha(\ell) - \frac{1}{\alpha N}\sum_{i=1}^N \Big((\ell_i - x)K(\ell_i - x) + \epsilon^2 k(\ell_i - x)\Big)\Big|_{\mathrm{VaR}_\alpha(\ell)}^\infty \\
&= \mathrm{VaR}_\alpha(\ell) + \frac{1}{\alpha N}\sum_{i=1}^N (\ell_i - \mathrm{VaR}_\alpha(\ell))K(\ell_i - \mathrm{VaR}_\alpha(\ell)) \\
&\qquad + \frac{\epsilon^2}{\alpha N}\sum_{i=1}^N k(\ell_i - \mathrm{VaR}_\alpha(\ell)) \\
&= \frac{1}{\alpha N}\sum_{i=1}^N \Big(\ell_i K(\ell_i - \mathrm{VaR}_\alpha(\ell)) + \epsilon^2 k(\ell_i - \mathrm{VaR}_\alpha(\ell))\Big).
\end{aligned}$$

Here, we have used the fact that $\partial(xK(x) + \epsilon^2 k(x))/\partial x = K(x)$ as a property of k and K and that $F_{L'}(\mathrm{VaR}_\alpha(\ell)) = 1-\alpha$. Moreover,

$$\begin{aligned}
\frac{\partial \mathrm{CVaR}_\alpha(\ell)}{\partial \ell_i} &= \frac{1}{\alpha N}\sum_{i'=1}^N \delta_{ii'} K(\ell_{i'} - \mathrm{VaR}_\alpha(\ell)) \\
&\qquad + \frac{1}{\alpha N}\sum_{i'=1}^N \ell_{i'} k(\ell_{i'} - \mathrm{VaR}_\alpha(\ell))\left(\delta_{ii'} - \frac{\partial \mathrm{VaR}_\alpha(\ell)}{\partial \ell_i}\right) \\
&\qquad - \frac{\epsilon^2}{\alpha N}\sum_{i'=1}^N \frac{\ell_{i'} - \mathrm{VaR}_\alpha(\ell)}{\epsilon^2} k(\ell_{i'} - \mathrm{VaR}_\alpha(\ell))\left(\delta_{ii'} - \frac{\partial \mathrm{VaR}_\alpha(\ell)}{\partial \ell_i}\right) \\
&= \frac{1}{\alpha N} K(\ell_i - \mathrm{VaR}_\alpha(\ell)) + \frac{1}{\alpha N}\mathrm{VaR}_\alpha(\ell) k(\ell_i - \mathrm{VaR}_\alpha(\ell)) \\
&\qquad - \frac{1}{\alpha N}\frac{\partial \mathrm{VaR}_\alpha(\ell)}{\partial \ell_i}\mathrm{VaR}_\alpha(\ell)\sum_{i'} k(\ell_{i'} - \mathrm{VaR}_\alpha(\ell)) \\
&= \frac{1}{\alpha N} K(\ell_i - \mathrm{VaR}_\alpha(\ell)),
\end{aligned}$$

where we have used $\partial k(x)/\partial x = -(x/\epsilon^2)k(x)$ and substituted in the expression for $\partial \mathrm{VaR}_\alpha(\ell)/\partial \ell_i$.

Thus, once $\mathrm{VaR}_\alpha(\ell)$ is obtained, both the function value and the gradient of $\mathrm{CVaR}_\alpha(\ell)$ are easily computed. It is worth noting that the partial derivative $\partial \mathrm{CVaR}_\alpha(\ell)/\partial \ell_i$ comes out surprisingly clean despite the somewhat cumbersome expression for $\partial \mathrm{VaR}_\alpha(\ell)/\partial \ell_i$.

As a side note, one can also show that the convexity of $\mathrm{CVaR}_\alpha(\ell)$ is preserved by exchanging L with L'. This is important for practical reasons as it guarantees uniqueness of the global minimum of the function, which is not the case for VaR and which is one of the main benefits with using CVaR.

Proposition 5. $\mathrm{CVaR}_\alpha(\ell)$ *is a convex function in* ℓ*.*

Proof. It is easy to show that $\mathrm{CVaR}_\alpha(\ell)$ is twice differentiable in ℓ, and it is well-known that such a function is convex if and only if its Hessian is positive semi-definite. Componentwise, we have

$$\begin{aligned}\frac{\partial^2 \mathrm{CVaR}_\alpha(\ell)}{\partial \ell_i \, \partial \ell_{i'}} &= \frac{1}{\alpha N} k(\ell_i - \mathrm{VaR}_\alpha(\ell)) \left(\delta_{ii'} - \frac{\partial \mathrm{VaR}_\alpha(\ell)}{\partial \ell_{i'}} \right) \\ &= \frac{1}{\alpha N \mathrm{e}^{\mathrm{T}} \kappa} \left(\mathrm{e}^{\mathrm{T}} \kappa \kappa_i \delta_{ii'} - \kappa_i \kappa_{i'} \right),\end{aligned}$$

where $\kappa = (\kappa_1, ..., \kappa_N)^{\mathrm{T}} = (k(\ell_1 - \mathrm{VaR}_\alpha(\ell)), ..., k(\ell_N - \mathrm{VaR}_\alpha(\ell)))^{\mathrm{T}}$, so for each $\xi \in \mathbb{R}^N$,

$$\begin{aligned}\xi^{\mathrm{T}} \frac{\partial^2 \mathrm{CVaR}_\alpha(\ell)}{\partial \ell^2} \xi &= \frac{1}{\alpha N \mathrm{e}^{\mathrm{T}} \kappa} \xi^{\mathrm{T}} \left(\mathrm{e}^{\mathrm{T}} \kappa \, \mathrm{diag}(\kappa) - \kappa \kappa^{\mathrm{T}} \right) \xi \\ &= \frac{1}{\alpha N \mathrm{e}^{\mathrm{T}} \kappa} \left(\left(\mathrm{e}^{\mathrm{T}} \kappa \right) \left(\xi^{\mathrm{T}} \, \mathrm{diag}(\kappa) \xi \right) - \xi^{\mathrm{T}} \kappa \kappa^{\mathrm{T}} \xi \right) \\ &= \frac{1}{\alpha N \mathrm{e}^{\mathrm{T}} \kappa} \left(\|\mathrm{e}\|^2 \|\xi\|^2 - |\langle \mathrm{e}, \xi \rangle|^2 \right) \\ &\geq 0,\end{aligned}$$

showing positive semi-definiteness. The last inequality is due to Cauchy–Schwarz using the inner product $\langle \xi_1, \xi_2 \rangle = \xi_1^{\mathrm{T}} \, \mathrm{diag}(\kappa) \xi_2$ and corresponding norm $\| \cdot \|$, which is well-defined as $\kappa_i > 0$ for all i.

Chapter 4

Numerical results

We present here a selection of numerical results obtained for different settings of portfolio assets. In order to solve multiple problems on the form (2.3) for varying values of the minimum expected return μ_{req} to produce efficient frontiers and corresponding optimal solutions, a short program was built in MATLAB R2017b. The built-in nonlinear constrained optimization solver `fmincon`, implementing an interior-point algorithm, was used to solve the optimization problems numerically. For all of the setup cases below, a Monte Carlo sample size of $N = 50000$ was used to form the empirical return matrix R, and the auxiliary noise $\varepsilon \sim \mathrm{N}(0, \epsilon^2)$ was assumed to have variance $\epsilon^2 = 0.05^2$. In the assumed absence of transaction costs, short-sellings were treated as negative positions.

4.1 Mean-variance analysis of stock portfolio

In order to test the validity of the algorithm before proceeding with more advanced setups, we first tried to reproduce the efficient frontiers in Figure 2.1 where variance was used as portfolio risk measure. The stocks $\{S_i(t)\}_{i=1}^n$ were assumed to follow a multidimensional geometric Brownian motion on the form

$$\frac{dS(t)}{S(t)} = b\,dt + s\,dW(t),$$

where $\{W(t)\}_{t\in[0,T]}$ is a Brownian motion, yielding log-normal marginal distributions of returns (the quotient on the left-hand side is taken componentwise). For our purposes, we used $n = 10$ stocks with parameter values

$$b = (0.0100, 0.0133, 0.0167, 0.0200, 0.0233, 0.0267, 0.0300, 0.0333, 0.0367, 0.0400)^{\mathrm{T}},$$

$$s = (0.0800, 0.0844, 0.0889, 0.0933, 0.0978, 0.1022, 0.1067, 0.1111, 0.1156, 0.1200)^{\mathrm{T}}$$

and

$$\varrho = \begin{bmatrix} 1 & 0.5 & \dots & 0.5 \\ 0.5 & 1 & \dots & 0.5 \\ \vdots & \vdots & \ddots & \vdots \\ 0.5 & 0.5 & \dots & 1 \end{bmatrix}$$

where ϱ is such that $\operatorname{Var} dW(t) = \varrho\, dt$. By Theorem 4, the return vector $R(T)$ satisfies

$$R(T) \sim \mathrm{LN}(\mu, \Sigma)$$

where $\mu = (b - (1/2)\operatorname{diag}(s)s)T$ and $\Sigma = \operatorname{diag}(s)\varrho\operatorname{diag}(s)T$, from which we can easily sample. Marginal distributions of returns for this setup are shown in Figure 4.1.

Figure 4.2 shows the efficient frontiers obtained from this setup, where the maturity was set to $T = 3$ and where the risk-free rate was assumed to be $r_0 = 0.001$. One can see that the behaviour described in Section 2.3.2 is indeed reproduced. In particular, the tangency portfolio shows at a minimum mean return of around 1.3—in fact, this rather low value of r_0 was used only to make the tangency portfolio show in the plot window.

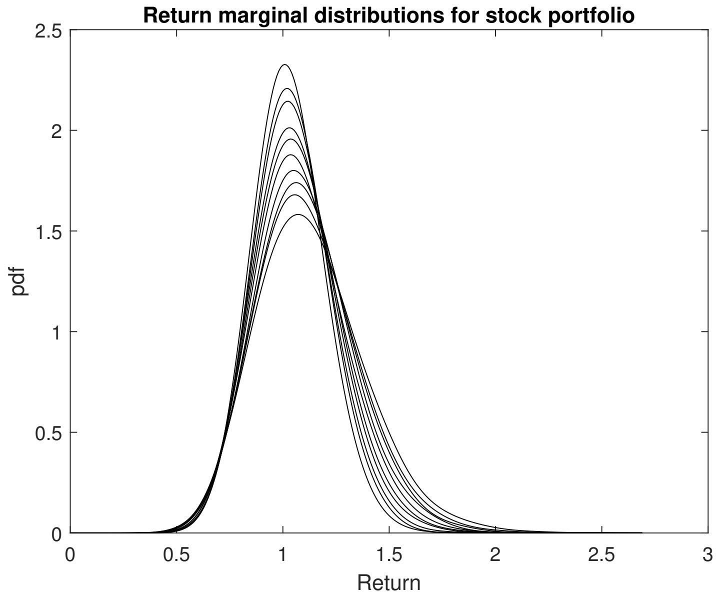

Figure 4.1: Plot of the pdf's of the marginal distributions of returns for the given stock portfolio, all log-normal.

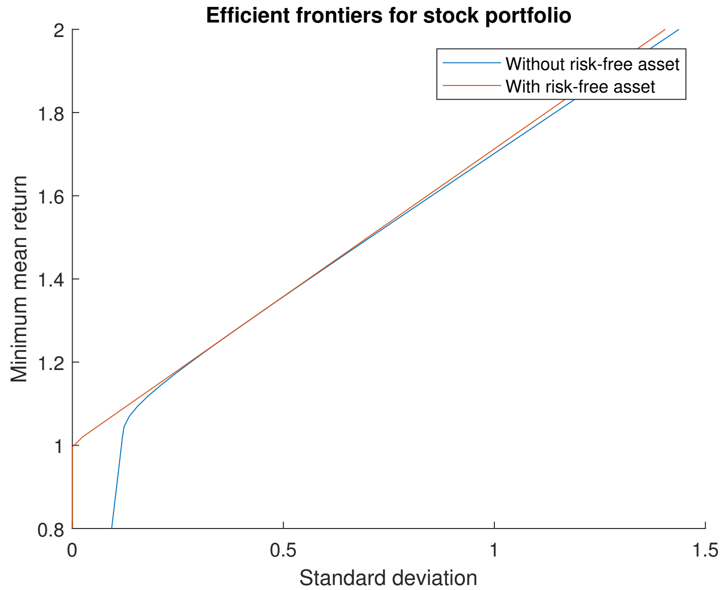

Figure 4.2: Plot of obtained efficent frontiers for the above described setup, with and without the possibility of investing in a risk-free asset. Note especially the tangency portfolio occurring at a minimum mean return of around 1.3.

4.2 Single catastrophe bond in stock portfolio

4.2.1 Discontinuous uniform model

Consider the same stock portfolio as in Section 4.1 and maturity $T = 3$, but with a more realistic risk-free rate of $r_0 = 0.02$ and parameters

$$b = (0.0100, 0.0167, 0.0233, 0.0300, 0.0367, 0.0433, 0.0500, 0.0567, 0.0633, 0.0700)^{\mathrm{T}},$$
$$s = (0.1000, 0.1056, 0.1111, 0.1167, 0.1222, 0.1278, 0.1333, 0.1389, 0.1444, 0.1500)^{\mathrm{T}}.$$

We investigate the effect of also including a cat bond specified by the discontinuous uniform model, with a coupon rate of $c = 0.08$ and exhaustion and attachment probabilities $p_{\mathrm{e}} = 0.03$, $p_{\mathrm{a}} = 0.07$. By Proposition 1, the mean return of this catastrophe bond over $[0, 3]$ is 1.1850, which lies in the middle of the mean marginal returns of the stock given by

$$(1.0316, 1.0521, 1.0727, 1.0951, 1.1172, 1.1386, 1.1627, 1.1854, 1.2107, 1.2361)^{\mathrm{T}},$$

although higher than the return $e^{0.02 \cdot 3} = 1.0618$ of the risk-free asset. Figure 4.3 shows the return profile of this catastrophe bond in comparison to those of the stocks.

We optimize on this portfolio using VaR at level $\alpha = 0.05$ as risk measure. Figure 4.4 shows the evolution of portfolio weights for the risk-free asset, the stocks and the catastrophe bond as the minimum mean return increases. One can see that after a minimum mean return of around 1.2, less is invested in the risk-free asset and more in the catastrophe bond. Furthermore, Figure 4.5 shows efficient frontiers with respect to $\mathrm{VaR}_{0.05}$ and $\mathrm{CVaR}_{0.05}$ and compares risks before and after introducing the possibility of investing in the catastrophe bond. Perhaps not very surprisingly, the efficient frontiers are shifted outwards by the catastrophe bond, showing that it adds value to the portfolio.

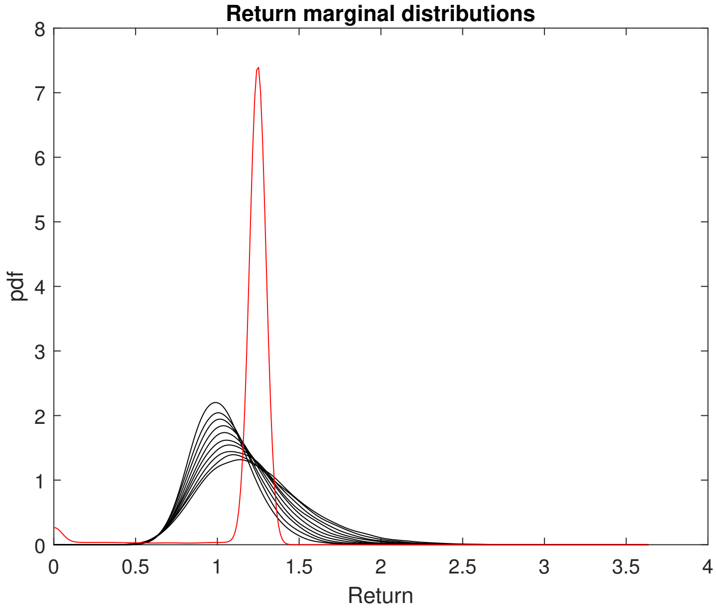

Figure 4.3: Comparison of return profiles between catastrophe bond (red) and stocks (black).

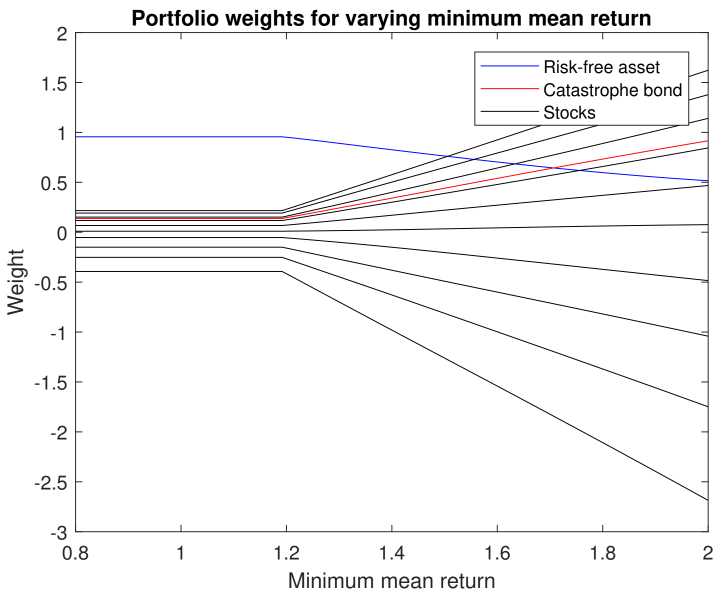

Figure 4.4: Plot of portfolio weights for risk-free asset, stocks, and catastrophe bond when varying minimum mean returns.

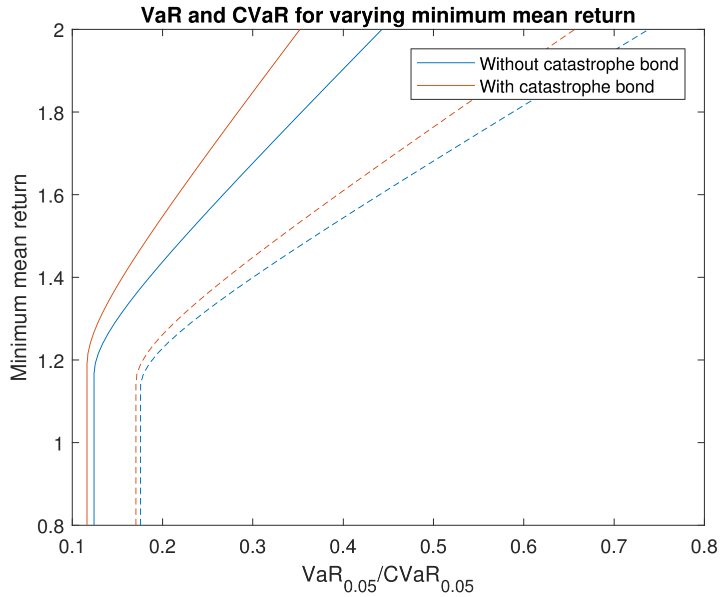

Figure 4.5: Plot of efficient frontiers for $\mathrm{VaR}_{0.05}$ (solid) and $\mathrm{CVaR}_{0.05}$ (dashed), before and after including a catastrophe bond.

4.2.2 Compound Poisson model

We can repeat the above using the same stock portfolio but the compound Poisson model. We set $r_0 = 0.02$ and $T = 3$. In order to reflect the contrast between the models, suppose that the catastrophe bond triggers relatively often but with relatively small instrument losses each time—more precisely, we use $\lambda_\mathrm{f} = 0.5$ and $\lambda_\mathrm{s} = 5$, which amounts to the assertion that a catastrophe occurs once every two years on average and that each catastrophe incurs an instrument of 0.2 on average. These are arguably reasonable assumptions for e g a catastrophe bond with parametric trigger. In compensation, we also set a relatively high coupon rate of $c = 0.2$. This gives a mean return over $[0, 3]$ of 1.2458, comparable to that of the stock with the highest mean return. The return profile is compared to those of the stocks in Figure 4.6.

As above, we optimize on this portfolio using VaR at level $\alpha = 0.05$ as risk measure. Figure 4.7 shows the evolution of portfolio weights for the risk-free asset, the stocks and the catastrophe bond as the minimum mean return increases. One can see that after a minimum mean return of around 1.1, less is invested in the risk-free asset and more in the catastrophe bond. Furthermore, Figure 4.8 shows efficient frontiers with respect to $\mathrm{VaR}_{0.05}$ and $\mathrm{CVaR}_{0.05}$ and compares risks before and after introducing the possibility of investing in the catastrophe bond. Again, the efficient frontiers are shifted outwards by the catastrophe bond.

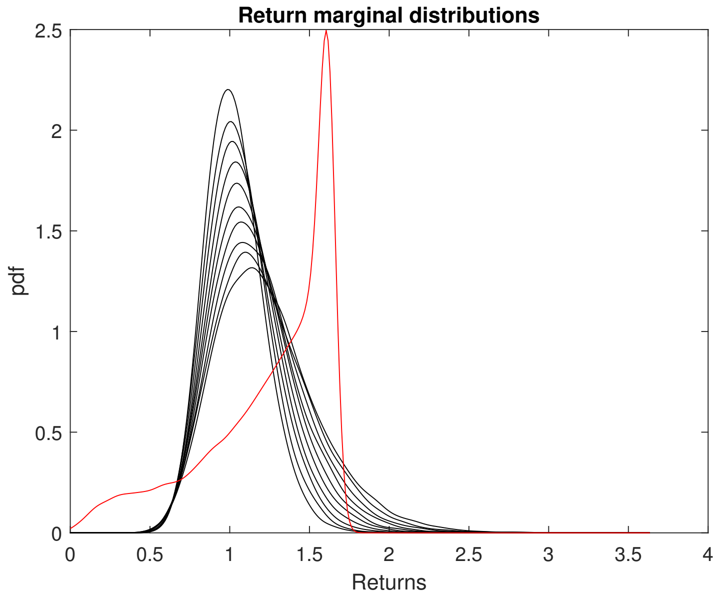

Figure 4.6: Comparison of return profiles between catastrophe bond (red) and stocks (black).

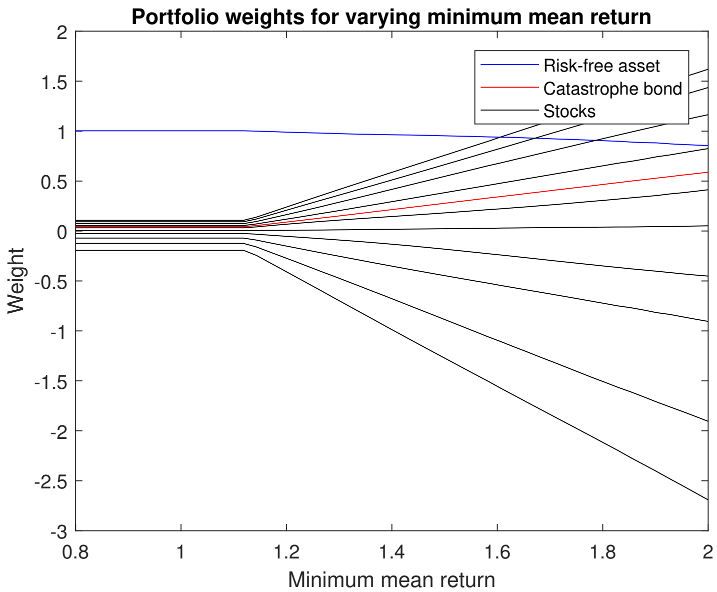

Figure 4.7: Plot of portfolio weights for risk-free asset, stocks, and catastrophe bond when varying minimum mean returns.

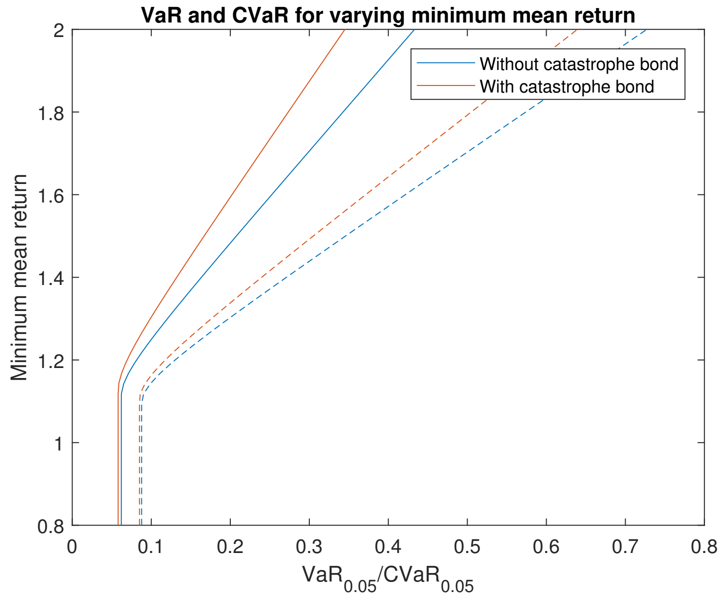

Figure 4.8: Plot of efficient frontiers for $\mathrm{VaR}_{0.05}$ (solid) and $\mathrm{CVaR}_{0.05}$ (dashed), before and after including a catastrophe bond.

4.3 Catastrophe bond portfolio

As a concluding experiment, we try the case of a portfolio only consisting of catastrophe bonds and a risk-free asset. More precisely, we use the compound Poisson model on n independent catastrophe bonds with coupon rates all equal to 0.2 and frequency/severity rates $(\lambda_{\mathrm{f}}, \lambda_{\mathrm{s}})$ given according to

$$(\lambda_{\mathrm{f}}, \lambda_{\mathrm{s}}) \in \left\{ \left(0.1 + 0.4\frac{i-1}{n-1}, 1 + 4\frac{i-1}{n-1} \right) \right\}_{i=1}^{n}.$$

For example, if $n = 5$, catastrophe bond i would be such that the mean frequency of catastrophes is $0.1i$ per year and that the mean severity is $1/i$ for each $1 \leq i \leq 5$. Again, we use $r_0 = 0.02$ and $T = 3$.

Figure 4.9 shows efficient frontiers with respect to $\mathrm{CVaR}_{0.05}$ for portfolios consisting of $n = 2, 4, 6, 8, 10$ catastrophe bonds, testing for minimum mean returns between 0.8 and 2.0. One can observe outward shifting of the efficient frontier when n increases, which is expected since the catastrophe bond returns are all independent. Figure 4.10 shows, perhaps more interestingly, return profiles of the optimal portfolios corresponding to a minimum mean return of 1.4 for varying n. From a highly skewed distribution for $n = 2$, the estimated pdf becomes more symmetric and Gaussian-like as n increases. Indeed, this can be made rigorous by the central limit theorem, although some care is required as the returns are not identically distributed (more precisely, one can use the Lyapunov version of the central limit theorem).

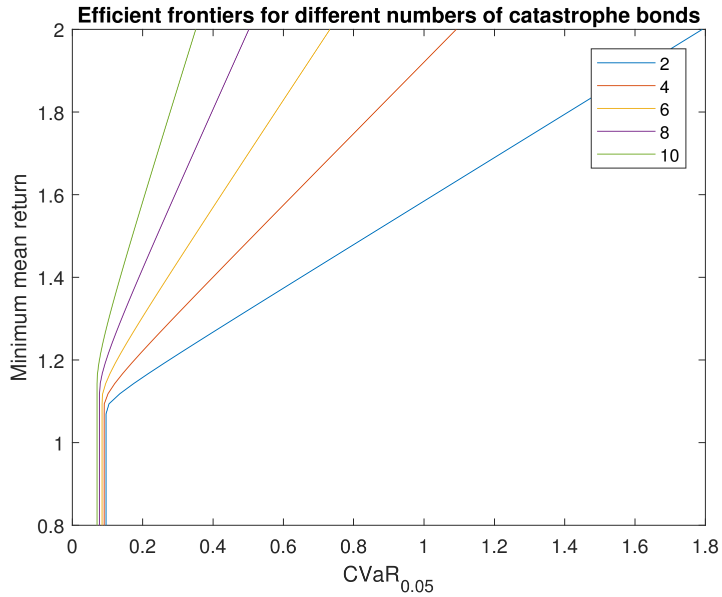

Figure 4.9: Plot of efficient frontiers for $\mathrm{CVaR}_{0.05}$ for different numbers of independent catastrophe bonds.

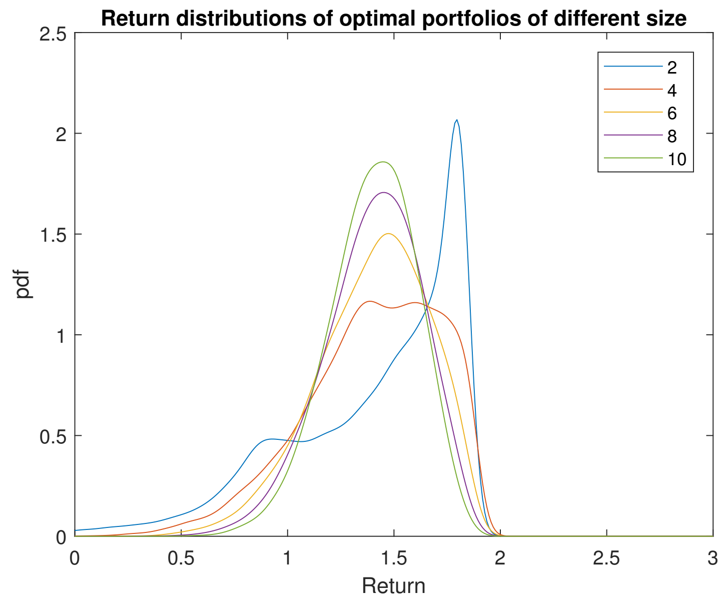

Figure 4.10: Plot of return profiles of optimal portfolios corresponding to a minimum mean return of 1.4 for varying n.

Chapter 5

Discussion

In this thesis, the risk diversification potential of catastrophe bonds in portfolios as well as the performance of portfolios exclusively consisting of catastrophe bonds have been investigated. We account here for the main points treated, discuss the novelty of the findings and give suggestions for further work.

5.1 Summary and conclusions

For the loss modeling, while the discontinuous uniform model was used to showcase essential characteristics of the return of catastrophe bonds, the compound Poisson model was developed to also take into account the times of occurrence of the underlying losses. In particular, given specifications of the distributions of catastrophe frequency and severity, a compound Poisson process with exponentially distributed increments was employed to model the evolution of the outstanding principal during the term. With the assumed coupon payment structure of constant coupon rate, proportionality to the outstanding principal and continuous coupon yield, the total return over the term for the catastrophe bond could be computed given each realization of the compound Poisson process.

With fully specified joint return distributions of all assets, we used Monte Carlo sampling to obtain the return matrix R on which all optimization was based. Standard optimization techniques on risk measures such as value-at-risk and conditional value-at-risk were made applicable by the introduction of an auxiliary noise variable, mathematically equivalent to kernel smoothing. This rendered value-at-risk as well as conditional value-at-risk continuously differentiable, enabling derivations of explicit and computationally tractable formulas for objective function values and gradients. As such, the implementation in code was straightforward and the computation times were relatively short.

Numerical results demonstrated the risk diversification value of a catastrophe bond in a hypothetical portfolio of correlated stocks. This was done with both the discontinuous uniform and compound Poisson models, where the former case was intended to reflect a catastrophe bond triggering very seldomly and the latter a catastrophe bond triggering more often but with relatively small instrument losses. For both cases, the expected outward shift of the efficient frontiers with respect to value-at-risk and conditional value-at-risk was confirmed. Moreover, it was shown that the return profile of a portfolio exclusively consisting of catastrophe bonds is highly skewed for few assets but converges to a normal distribution when the number of assets increases, which is also expected.

5.2 Novelty and impact

The main contributions of this thesis are related to the compound Poisson model for the return distribution of a catastrophe bond and the formulas for function values and gradients of risk measures derived from introducing an auxiliary noise. As for the former, although models accounting for the possibility of multiple independent catastrophe events by a compound Poisson outstanding principal process have been previously mentioned in literature [20, 14], they address catastrophe bonds with somewhat different payment structure where the arrival of the catastrophe events do not matter. The model presented here thus combines the time aspect of the assumed payment structure with these jump process models of the outstanding principal, which to the best of the author's knowledge is novel in literature. For the formulas for risk measures, despite the fact that it is well-known that kernel smoothing renders VaR differentiable by virtue of the implicit function theorem, derivations of explicit formulas have not been accounted for in existing literature. Instead, the work here is mainly based on that done in [26], which uses the same idea but in an essentially different setting.

On the other hand, the impact of the results are not as obvious. This is mainly due to the lack of real-world data used in the test cases, which was intentional due to the limited scope of the project. Instead, the numerical tests were solely performed in the purpose of demonstrating the validity of the portfolio optimization procedure and illuminating key features of the loss models. Although the tests successfully fulfill this purpose, further inferences cannot be made as to, for instance, how the models compare to existing ones and how real-world catastrophe bond portfolios perform in comparison to other portfolios. However, it is believed that the developed method of applying gradient-based optimization of VaR and CVaR will be an interesting alternative to current standard approaches.

5.3 Further work

A natural extension of the work in this thesis would be to apply it to realistic examples, e g by using historical data on stock returns and the events underlying catastrophe bonds, which would strengthen the impact. Furthermore, the compound Poisson model could be adapted to other types of payment structures, which is straightforward given the corresponding specifications. The impact of this thesis work would be further strengthened by some comparison to other existing catastrophe bonds loss models. Lastly, the usefulness of the method of gradient-based optimization could be evaluated by comparing e g behavior and computational times to other standard methods.

Chapter 6

References

[1] Baryshnikov Y, Mayo A, Taylor D R. Pricing of cat bonds. http://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=BA9EF3FBFF0696271638C18BC10ED284?doi=10.1.1.42.5811&rep=rep1&type=pdf, accessed 2019-05-15.

[2] Braun A. Pricing in the primary market for cat bonds: New empirical evidence. *Journal of Risk and Insurance*. 2016;83(4):811–847.

[3] Business Insurance. US insurers well prepared for expected average hurricane season: Fitch. https://www.businessinsurance.com/article/00010101/NEWS06/912321670/US-insurers-well-prepared-for-expected-average-hurricane-season-Fitch, accessed 2019-05-15.

[4] Carnegie. Diversifiera med Irma och Maria. https://www.carnegie.se/private-banking/Aktuellt/cat-bonds-i-portfoljen/, accessed 2019-05-15.

[5] Cox S H, Pedersen H W. Catastrophe risk bonds. *North Americal Actuarial Journal*. 2000;4(4):56–82.

[6] Cummins J D, Weiss M A. Convergence of insurance and financial markets: hybrid and securitized risk-transfer solutions. *Journal of Risk and Insurance*. 2009;76(3):493–545.

[7] Dieckmann S. By force of nature: explaining the yield spread on catastrophe bonds. Working paper, University of Pennsylvania, 2009.

[8] Djehiche B. Stochastic calculus: An introduction with applications. Lecture notes, KTH Royal Institute of technology, 2000.

[9] Elton E J, Gruber M J, Agrawal D, Mann C. Explaining the rate spread on corporate bonds. *Journal of Finance*. 2001;56(1):247–277.

[10] Entropics Asset Management. How cat bonds work. http://en.entropics.se/cat-bonds/how-cat-bonds-work/, accessed 2019-05-15.

[11] Entropics Asset Management. How is a portfolio affected by the inclusion of cat bonds? http://en.entropics.se/blog/portfolio-affected-inclusion-cat-bonds/, accessed 2019-05-15.

[12] Froot K A. The market for catastrophe risk: A clinical examination. *Journal of Financial Economics*. 2001;60(2–3):529–571.

[13] Galeotti M, Gürtler M, Winkelvos C. Accuracy of premium calculation models for cat bonds: An empirical analysis. *Journal of Risk and Insurance*. 2013;80(2):401–421.

[14] Giertz F. Analysis and optimization of a portfolio of catastrophe bonds. MSc thesis, KTH Royal Institute of Technology, 2014.

[15] Gomez L, Carcamo U. A multifactor pricing model for cat bonds in the secondary market. *Journal of Business, Economics and Finance*. 2014;3(2):247–258.

[16] Gürtler M, Hibbeln M T, Winkelvos C. The impact of the financial crisis and natural catastrophes on cat bonds. *Journal of Risk and Insurance*. 2016;83(3):579–612.

[17] Hult H, Lindskog F, Hammarlid O, Rehn C J. *Risk and portfolio analysis*. New York: Springer; 2012.

[18] Jaeger L, Müller S, Scherling S. Insurance-linked securities: What drives their returns? *Journal of Alternative Investments*. 2010;13(2):9–34.

[19] Koski T. Probability and random processes at KTH. Lecture notes, KTH Royal Institute of Technology, 2017.

[20] Lee J, Yu M. Pricing default-risky cat bonds with moral hazard and basis risk. *Journal of Risk and Insurance*. 2002;69(1):25–44.

[21] Loubergé H, Kellezi E, Gilli M. Using catastrophe-linked securities to diversify insurance risk: A financial analysis of cat bonds. *Journal of Insurance Issues*. 1999;22(2):125–146.

[22] Man Group. Catastrophe bonds: Investing with impact. https://www.man.com/catastrophe-bonds-investing-with-impact, accessed 2019-05-15.

[23] Mariani M, Amoruso P. The effectiveness of catastrophe bonds in portfolio diversification. *International Journal of Economics and Financial Issues*. 2016;6(4):1760–1767.

[24] Nash S G, Sofer A. *Linear and nonlinear programming*. New York: McGraw-Hill; 1996.

[25] Rockafellar R T, Uryasev S. Optimization of conditional value-at risk. *Journal of Risk*. 1997;2:21–-41.

[26] Zhang T. Direct optimization of dose-volume-histogram metrics in intensity-modulated radiation therapy treatment planning. BSc thesis, KTH Royal Institute of Technology, 2018.