MINIMIZATION OF CONDITIONAL VALUE AT RISK FOR SPREAD OPTIONS UNDER CAPITAL CONSTRAINTS

by

Chingis Maksimov

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

in

Mathematical Finance

Department of Mathematical and Statistical Sciences
University of Alberta

Abstract

The thesis deals with the problem of minimization of Conditional Value at Risk within the context of Margrabe market under constraints on the initial capital available. We propose to approximate the distribution of the difference between two lognormal random variables using normal distribution and derive a closed-form pricing formula for spread options. We use this idea along with the existing spread option pricing formulas to develop a new methodology for determining Conditional Value at Risk-efficient portfolios. We conclude that the approaches considered provide comparable results given that the parameters of the market are of particular form. Theoretical results are supported by numerical examples based on real financial data.

To Constant Search for Knowledge

Acknowledgements

I would like to express my sincere gratitude to professor Alexander Melnikov for his trust and supervision. I am very grateful to professor Christoph Frei and professor Felipe Aguerrevere for the invaluable advice I received from them. I thank my friends and loved ones for the constant support throughout my studies at the University of Alberta.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

In their seminal 1973 paper Black and Scholes laid groundwork for what is now estimated to be a half quadrillion dollar derivatives market (see Bank for International Settlements, 2017). In the Black-Scholes theory the price behaviour of any contingent claim in a complete arbitrage-free market can be viewed as a weighted portfolio of risky securities underlying the contract and a non-defaultable zero-coupon bond. At least in theory an investor can perfectly replicate the payoff of any contingent claim by constantly adjusting the weights in this portfolio and the Black-Scholes formula provides us with the means of calculating these weights at any point in time. Ever since the initial publication, Black and Scholes' insightful idea was successfully applied to contingent claims with various payoff structures and the problem of hedging of contingent claims in complete arbitrage-free markets no longer attracts much of scientific interest.

When a martingale measure is not unique, the market is considered incomplete. Many experts (see, for instance, El Karoui and Quenez, 1995) studied the problem of hedging of contingent claims in such markets and came to a conclusion that for any given contingent claim there is a whole range of fair or arbitrage-free prices. The minimum price of setting up a replicating portfolio that guarantees no underheging at maturity is equal to the maximum price of the arbitrage-free range. Hence, hedging in such markets might require a larger amount of initial capital outlay than what would be deemed optimal for practical purposes. Similar situation can arise when a financial institution or an investor operates in a complete market but is constrained by the amount of initial capital available or is reluctant to put up the amount required for complete hedging. One of the possible motivations could be the possibility of using the extra funds saved on hedging to earn additional return in excess of the expected shortfall, the expected amount by which a replicating portfolio underhedges an option payoff at maturity. The problem attracted a lot of scientific attention: see, for example, Kulldorff (1993), Spivak and Cvitanic (1999), Browne (1999). Foellmer and Leukert (1999) proposed to use quantile hedging strategies whereby an investor is aiming at maximizing the probability of a successful hedge. By following this strategy an investor is using a dynamic version of static Value at Risk (VaR) concept. The major drawback of this approach is that the size of potential shortfall is not taken into account. To address this issue Foellmer and Leukert (2000) proposed a new methodology to minimize the amount of expected shortfall, where an investor's attitude to the size of shortfall is measured by some loss function l. The central idea is to use the results from Neyman-Pearson lemma to modify the original claim in a special way so that this modified claim can be completely hedged. The authors proceed to prove that the strategy of complete hedging for the modified claim will also be the optimal strategy for the original contingent claim.

Under previous Basel II accord VaR was the preferred method of estimating and reporting market risk exposure by financial institutions. By definition, for a chosen confidence level a, VaR_a is the smallest amount β such that the probability of incurring losses in excess of β is less than or equal to $1-a$. In other words, VaR_a is the a-quantile of the distribution function of portfolio losses. There are three common approaches to measuring VaR: historical simulation approach, Monte Carlo (MC) simulation approach and parametric approach, where the distributions of the returns of risky securities held in a portfolio are assumed to be jointly normal (see JP Morgan, 1996). While all three approaches provide a quick and intuitive way of assessing market risk exposure, VaR, as a risk measure, lacks some of the desired mathematical properties and was severely criticized as it failed to predict the scope of the losses during the global financial crisis of 2008. Artzner et al. (1999) showed that VaR does not satisfy the subadditivity property of coherent risk measures and it does not encourage diversification as VaR "does not take into account the economic consequences of the events the probabilities of which it controls". Further, the successful implementation of VaR methodology is constrained by the assumption of normality of stock returns, which is not always in line with empirical evidence, where the distributions of stock returns are leptokurtic. (Hebner, 2014). The most recent Basel III framework encourages the shift from VaR to $CVaR$, also known as Average Value at Risk ($AVaR$), or Tail Value at Risk ($TVaR$), which is a spectral (Acerbi, 2002) and coherent measure of risk as it satisfied the following four desirable properties: monotonicity, subadditivity, positive homogeneity and translation invariance. (Artzner et al., 1999). For continuous distributions, $CVaR_a$ is defined as the conditional expectation of losses, given that losses exceed VaR_a. For discrete distributions and more general distributions, $CVaR$ is determined as the weighted average between VaR_a and losses strictly greater than VaR_a.

Melnikov and Smirnov (2012) considered the dual problem: minimization of $CVaR$ of an investment portfolio within the Black-Scholes market model with constraints on the initial capital available and the minimization of hedging costs subject to constraints on $CVaR$. The authors used the results of Rockafellar and Uryasev (2002) for the alternative representation of $CVaR$ that allows to minimize both VaR and $CVaR$ at the same time. Closed-form solutions related to construction of $CVaR$-efficient portfolios have been derived. The aim of this thesis is to consider the problem of a similar kind - minimization of $CVaR$ within the context of Margrabe market. The contingent claim of interest is a plain vanilla option to exchange one asset for another. We will review the formula to price such a contingent claim as well as approximating formulas for spread options, options to exchange one asset for another with payoffs containing non-stochastic components. The latter class of options does not have an exact solution as one is dealing with the distribution of the difference between two lognormal random variables which is not lognormal. However, we propose to use normal distribution as an approximate distribution of the difference between two lognormal random variables, which allows to price spread options in closed-form. We compare the proposed approximation with the ones already existing. We will proceed by solving the initial problem and deriving closed-form formulas for the construction of $CVaR$-efficient replicating portfolios by using the proposed normal approximation as well as the approximation by Bjerksund and Stensland (2006). The problem of this kind has not been considered so far and we believe the results of this thesis will carry a lot of benefits to the practitioners in view of the most recent developments in risk management regulations.

The thesis is structured in the following way: Chapter 2 contains the necessary theoretical background related to spread option pricing, Conditional Value at Risk, Neyman-Pearson lemma and expected shortfall minimization. We show how these theoretical developments are unified to provide means for minimizing $CVaR$ of a portfolio subject to constrains on the initial capital available. Chapter 3 is devoted to the application of the methodology within the Margrabe market. The methodology is then tested on both hypothetical and real data portfolios in Chapter 4. Chapter 5 provides the limitations as well as recommendations for further research in this direction, and finally Chapter 6 concludes the thesis.

Chapter 2

Theoretical Background

This chapter is devoted to the statements and proofs of some of the theorems that lay groundwork for the solution to the problem of minimization of $CVaR$.

2.1 Two-Asset Lemma

Lemma 1. *Let* $X \sim N(\mu_x, \sigma_x^2), Y \sim N(\mu_y, \sigma_y^2)$ *and* $Z \sim N(\mu_z, \sigma_z^2)$ *be three normally distributed random variables with correlations* $\rho_{XY}, \rho_{XZ}, \rho_{YZ}$*. Then:*

$$E(exp\left\{-Z\right\} I_{\{X<x\}} I_{\{Y<y\}}) = exp\left\{-\mu_z + \frac{\sigma_z^2}{2}\right\} \Phi^2(\hat{x}, \hat{y}, \rho_{XY}) \tag{2.1}$$

where:

$$\hat{x} = \frac{x - \mu_x}{\sigma_x} + \sigma_z \rho_{XZ}$$
$$\hat{y} = \frac{y - \mu_y}{\sigma_y} + \sigma_z \rho_{YZ}$$

and Φ^2 denotes the two-dimensional normal cumulative distribution function (CDF). (Melnikov and Romanyuk, 2008).

2.2 Spread Option Pricing

The pricing formula for options to exchange one asset for another was derived independently by Margrabe (1978) and Fischer (1978). Suppose that $S_1 = (S_1(t) : t \in [0,T])$ and $S_2 = (S_2(t) : t \in [0,T])$ are the price processes of two correlated non-dividend paying stocks and are solutions to the following stochastic differential equations (SDEs) under the unique risk-neutral probability measure Q:

$$dS_1(t) = S_1(t)\sigma_1 dW_1^Q(t)$$
$$dS_2(t) = S_2(t)\sigma_2 dW_2^Q(t)$$

where $W_1^Q = (W_1^Q(t) : t \in [0,T])$ and $W_2^Q = (W_2^Q(t) : t \in [0,T])$ are standard Brownian motion processes with the correlation coefficient ρ. The corresponding standard deviations are σ_1 and σ_2, which we assume to be constant. The original Margrabe's model of the market assumed only the existence of two risky assets and no bank account. For the purposes of this thesis we assume that the interest rate $r = 0$. Consider a European type spread option which gives the holder the right to exchange the second asset for the first one with the following payoff:

$$(S_1(T) - S_2(T))^+$$

where T is the maturity of the contract. The price of such a contingent claim is given by:

$$p = S_1(0)\Phi(d_1) - S_2(0)\Phi(d_2) \tag{2.2}$$

where:

$$\begin{aligned}
d_1 &= \frac{ln(\frac{S_1(0)}{S_2(0)}) + \frac{\sigma^2 T}{2}}{\sigma\sqrt{T}} \\
d_2 &= d_1 - \sigma\sqrt{T} \\
\sigma &= \sqrt{\sigma_1^2 + \sigma_2^2 - 2\sigma_1\sigma_2\rho}
\end{aligned}$$

and Φ denotes the CDF of a normal distribution. However, if the option's payoff structure is modified to have a deterministic strike price in addition (we will refer to options with such payoffs as spread options to distinguish with options to exchange one asset for another):

$$(S_1(T) - S_2(T) - K)^+ \tag{2.3}$$

no exact closed-form solution has been found. The problem arises because pricing such an option requires the knowledge of the distribution of the difference between two lognormal random variables, which is not lognormal. More generally, any linear combination of correlated lognormal random variables is not lognormal. (Poulsen, 2010). We note that closed-form approximations for the distribution of sums of lognormal random variables exist in the literature, see for example Mehta et al. (2007), Cobb and Rumi (2012), Hcine and Bouallegue (2015), Rook and Kerman (2015).

Less is known about the distribution of the difference between correlated lognormal random variables. Lo (2012) proposed to use Lie-Trotter operator splitting method and showed that both the sum and the difference of two correlated lognormal random variables follow a shifted lognormal process. For the evolution of the joint distribution function of the difference between two lognormal random variables in time the author considered the following Kolmogorov backward equation (KBE):

$$\left\{\frac{\partial}{\partial t_0} + \hat{L}_+ + \hat{L}_0 + \hat{L}_-\right\} p(S_-, t; S_+(0), S_-(0), t_0) = 0 \tag{2.4}$$

with the boundary condition:

$$p(S_-, t; S_+(0), S_-(0), t_0 \to t) = \delta(S_-(0) - S_-)$$

where:

$$\hat{L}_+ = \frac{1}{8}\left(\sigma_+^2\Big(S_+(0)\Big)^2 + 2\Big(\sigma_1^2 - \sigma_2^2\Big)S_+(0)S_-(0) + \sigma_-^2\Big(S_-(0)\Big)^2\right)\frac{\partial^2}{\partial S_+^2(0)}$$
$$\hat{L}_0 = \frac{1}{4}\left((\sigma_1^2 - \sigma_2^2)\Big((S_+(0))^2 + (S_-(0))^2\Big) + (\sigma_1^2 + \sigma_2^2)S_+(0)S_-(0)\right)\frac{\partial^2}{\partial S_+(0)\partial S_-(0)}$$
$$\hat{L}_- = \frac{1}{8}\left(\sigma_+^2\Big(S_-(0)\Big)^2 + 2(\sigma_1^2 - \sigma_2^2)S_+(0)S_-(0) + \sigma_-^2\Big(S_+(0)\Big)^2\right)\frac{\partial^2}{\partial S_-^2(0)}$$
$$\sigma_\pm = \sqrt{\sigma_1^2 + \sigma_2^2 \pm 2\rho\sigma_1\sigma_2}$$
$$S_\pm(0) = S_1(0) \pm S_2(0)$$
$$S_- = S_1(t) - S_2(t)$$

and δ is the Dirac delta function such that:

$$\delta(x) = \begin{cases} +\infty & \text{, if } x = 0 \\ 0 & \text{, otherwise} \end{cases}$$

The following theorem helps to better interpret the form in which the above problem is stated:

Theorem 2. *Denote by $p(y,t;x,s)$ the transition probability density function (PDF) from state x at time s to state y at time t. Let $u(x,s) = p(y,t;x,s)$. Then:*

$$\begin{cases} u_s(x,s) + \mathcal{L}\left\{u(x,s)\right\} = 0, \;\; for\ (x,s) \in \mathbb{R} \times (0,t) \\ \lim_{s\to t} u(x,s) = \delta(x-y) \end{cases}$$

More informally, in problem (2.4) one is trying to reconstruct the transition probability density function for the difference between two lognormal random variables by moving backward in time and making sure that the boundary condition is satisfied. Solution to this problem exists and is of the following form:

$$p(S_-,t;S_+(0),S_-(0),t_0) = exp\Big((t-t_0)(\hat{L}_+ + \hat{L}_0 + \hat{L}_-)\Big)\delta(S_-(0)-S_-)$$

The difficulties arise with exponentiating $(t-t_0)(\hat{L}_+ + \hat{L}_0 + \hat{L}_-)$ term. The solution is to use the Lie-Trotter splitting method (Trotter, 1959), which is a generalization of the Lie product formula for arbitrary real and complex matrices. Trotter showed that:

$$exp\left\{\epsilon C\right\} = exp\left\{\epsilon A\right\} exp\left\{\epsilon B\right\} + O(\epsilon^2) \tag{2.5}$$

where ϵ is some parameter and C is an operator that can be decomposed into subcomponents:

$$C = A + B$$

Splitting methods are widely used for numerical solutions to partial differential equations (PDEs), where we aim to split a given differential operator into operators that are simpler to deal with. Applying this to (2.5) we get an approximate solution:

$$p^{approx}(S_-, t; S_+(0), S_-(0), t_0) = exp\Big((t - t_0)\hat{L}_-\Big)\delta(S_-(0) - S_-)$$

Now, having derived the approximate transition PDF of the difference between two lognormal random variables, Lo proceeds to show that the difference is distributed as a shifted lognormal process.

Coming back to (2.3), the first actual attempt to price an option with such a payoff was done by Phelim Boyle (1988), who extended the lattice binomial approach proposed by Cox, Ross and Rubinstein (1979) to the current setting. The major benefit of such an approach is that pricing American-style options becomes possible. Later Kirk (1995) proposed the following closed-form approximation:

$$p = e^{-rT}\Big(S_1(0)\Phi(d_1) - (S_2(0) + K)\Phi(d_2)\Big) \tag{2.6}$$

where:

$$\begin{aligned}
d_1 &= \frac{ln\Big(\frac{S_1(0)}{S_2(0)+K}\Big) + \frac{\sigma^2 T}{2}}{\sigma\sqrt{T}} \\
d_2 &= d_1 - \sigma\sqrt{T} \\
\sigma &= \sqrt{\sigma_1^2 - \frac{2S_2(0)}{S_2(0) + K}\sigma_1\sigma_2\rho + \Big(\frac{S_2(0)}{S_2(0) + K}\Big)^2\sigma_2^2}
\end{aligned}$$

The idea of Kirk was to approximate the true price through the following expectation:

$$p = E_Q\left(\left(S_1(T) - \frac{aS_2^b(T)}{E_Q(S_2^b(T))}\right)^+\right) \tag{2.7}$$

where:

$$a = S_2(0) + K$$
$$b = \frac{S_2(0)}{S_2(0) + K}$$

Refer to Appendix 1 for derivations.

Bjerksund and Stensland (2006) elaborated upon the idea proposed by Kirk and derived the following approximation to the price of a spread option:

$$p = e^{-rT}\Big(S_1(0)\Phi(d_1) - S_2(0)\Phi(d_2) - K\Phi(d_3)\Big) \tag{2.8}$$

where:

$$d_1 = \frac{ln(\frac{S_1(0)}{a}) + (\frac{\sigma_1^2}{2} - \sigma_1\sigma_2 b\rho + \frac{\sigma_2^2 b^2}{2})T}{\sigma\sqrt{T}}$$
$$d_2 = \frac{ln(\frac{S_1(0)}{a}) + (\frac{\sigma_1^2}{2} + \sigma_1\sigma_2\rho + \frac{\sigma_2^2 b^2}{2} - \sigma_2^2 b)T}{\sigma\sqrt{T}}$$
$$d_3 = \frac{ln(\frac{S_1(0)}{a}) + (-\frac{\sigma_1^2}{2} + \frac{\sigma_2^2 b^2}{2})T}{\sigma\sqrt{T}}$$
$$\sigma = \sqrt{\sigma_1^2 - 2\sigma_1\sigma_2 b\rho + \sigma_2^2 b^2}$$
$$a = S_2(0) + K$$
$$b = \frac{S_2(0)}{S_2(0) + K}$$

The authors consider the exact same strategy to exercise the option if and only if $S_1(T)$ exceeds a power function of $S_2(T)$ multiplied by a scalar $a/E_Q(S_2^b(T))$ but do not modify the option payoff as in (2.7). Then the price is obtained by evaluating the following expectation:

$$p = E_Q\Big((S_1(T) - S_2(T) - K)I_{\left\{S_1(T) \geq \frac{aS_2^b(T)}{E_Q(S_2^b(T))}\right\}}\Big)$$

The choice of parameters a and b yields quite satisfactory results for the approximation of the true price as we will see later, however the authors claim that the precision could still be further improved if one optimizes the option price with respect to these parameters. We derive (2.8) in Appendix 2.

A somewhat different approach was undertaken by Carmona and Durrleman (2003) where the authors proposed to express the correlation coefficient between the two Brownian motion processes through trigonometric functions. The values of the price processes of the two stocks at any time t are given by:

$$S_1(t) = S_1(0)exp\left\{-\frac{\sigma_1^2 t}{2} + \Big(\epsilon_1 \sin(\phi) + \epsilon_2 \cos(\phi)\Big)\sigma_1\sqrt{t}\right\}$$
$$S_2(t) = S_2(0)exp\left\{-\frac{\sigma_2^2 t}{2} + \epsilon_2\sigma_2\sqrt{t}\right\}$$

where ϵ_1 and ϵ_2 are independent standard normal random variables, $\cos(\phi) = \rho, \phi \in [0, \pi]$. The authors propose to exercise the option whenever the following condition is satisfied:

$$\epsilon_1 \sin(\theta^*) - \epsilon_2 \cos(\theta^*) \leq d^*$$

where $\theta^* \in [\pi, 2\pi]$ and d^* are found numerically by maximizing the option value. The value obtained by following this strategy represents the lower bound of the true price of this option and is equal to:

$$p = e^{-rT}\Big(S_1(0)\Phi(d^*+\sigma_1\sqrt{T}\cos(\theta^*+\phi))-S_2(0)\Phi(d^*+\sigma_2\sqrt{T}\cos(\theta^*))-K\Phi(d^*)\Big) \tag{2.9}$$

Along with the approximation for the lower price bound on the price of a spread option Carmona and Durrleman point out that one may approximate the distribution between two lognormal random variables by means of a normal distribution, in which case one can derive a closed-form solution for the price of such a contract. Indeed, if one was to consider the difference between two lognormal prices at maturity T under risk neural probability measure:

$$S_1(T) - S_2(T) = S_1(0)exp\left\{-\frac{\sigma_1^2 T}{2} + \sigma_1 W_1^Q(T)\right\} - S_2(0)exp\left\{-\frac{\sigma_2^2 T}{2} + \sigma_2 W_2^Q(T)\right\}$$

Using Taylor series expansion we can rewrite the above in the following way:

$$\begin{aligned} S_1(T)-S_2(T) &= S_1(0)\sum_{n=0}^{\infty}\frac{z_1^n}{n!} - S_2(0)\sum_{n=0}^{\infty}\frac{z_2^n}{n!} \\ &= S_1(0) - S_2(0) + S_1(0)z_1 - S_2(0)z_2 + S_1(0)\sum_{n=2}^{\infty}\frac{z_1^n}{n!} - S_2(0)\sum_{n=2}^{\infty}\frac{z_2^n}{n!} \end{aligned}$$

where:

$$\begin{aligned} z_1 &= -\frac{\sigma_1^2 T}{2} + \sigma_1 W_1^Q(T) \\ z_2 &= -\frac{\sigma_2^2 T}{2} + \sigma_2 W_2^Q(T) \end{aligned}$$

Noting that in practice the initial prices of the two stocks are standardized to be equal, the above simplifies to the following:

$$S_1(T) - S_2(T) = S(0)(z_1 - z_2) - S(0)\Big(\sum_{n=2}^{\infty} \frac{z_1^n}{n!} - \sum_{n=2}^{\infty} \frac{z_2^n}{n!}\Big) \tag{2.10}$$

where: $S(0) = S_1(0) = S_2(0)$. The first term is normally distributed. So, the difference between the prices of two risky securities at maturity follows a normal distribution minus some additional error term in the amount of $S(0)\Big(\sum_{n=2}^{\infty} \frac{z_1^n}{n!} - \sum_{n=2}^{\infty} \frac{z_2^n}{n!}\Big)$. Ideally, one would like to know the distribution of the error term, however Berg (1988) showed that a normal random variable raised to a power p, for $p \geq 3$, has an indeterminate distribution (in the Hamburger sense). Indeed, knowing just the moments of a distribution does not help in managing the exposure to risk and, on the contrary, can lead to problems by providing a false sense of security. Suppose that we have estimated 3 moments of a given random variable, which, for example, represents the loss function on some exposure: mean of 0, variance of $5/3$ and skewness of 0. The distribution of this random variable could be that of normal. However, using this incomplete knowledge about the distribution of the loss function for risk management purposes is incorrect as there are many other distribution functions that have the same 3 moments but fail to match the moments of higher order. To illustrate our point consider figure 2.1.

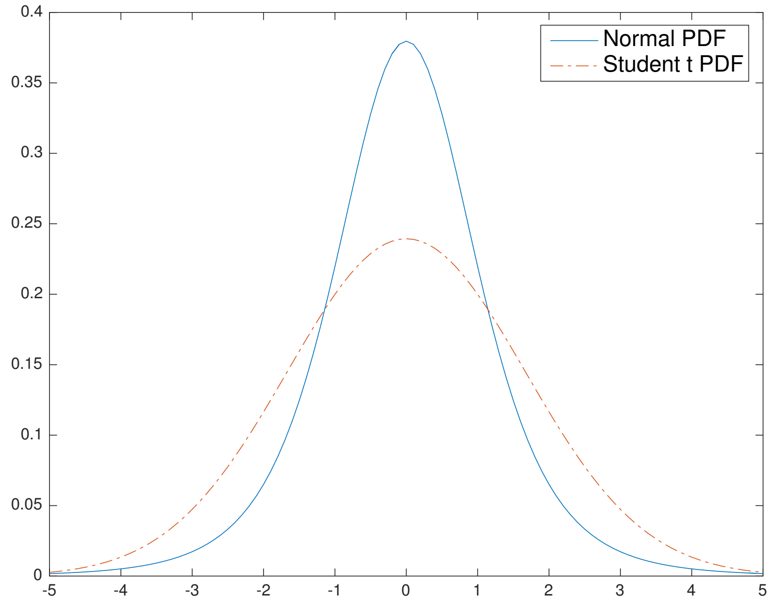

Figure 2.1: Comparison of Normal and Student t Distributions

The figure plots a PDF of a normal distribution with the first three moments as described above against a PDF of a student's t distribution with $v = 5$ degrees of freedom. We note that for a given random variable $X \sim t(v)$ we can estimate the first 4 moments in the following way:

$$
\begin{aligned}
E(X) &= 0, && v > 1 \\
E(X^2) &= \frac{v}{v-2}, && v > 2 \\
E(X^3) &= 0, && v > 3 \\
E(X^4) &= \frac{6}{v-4} + 3, && v > 4
\end{aligned}
$$

So, the two distributions have the first three moments matching. However, the kurtosis of $X \sim t(5)$ is equal to 9, which is higher than the corresponding kurtosis of 3 for a normal distribution, making it a leptokurtic distribution, also known as a fat-tailed distribution. Thus, the probability of observing more extreme losses is going to be higher if the actual distribution is student's t rather than the normal distribution. Coming back to (2.10), we observe that the significance of each additional random variable in the error term quickly diminishes due to the presence of a factorial function in the denominator, which gives us some hope that for a certain choice of initial model parameters the approximation might yield satisfactory results. Since we are interested in the error term to be minimized, we want $z_1^n = z_2^n \quad \forall\ n \geq 2$. This can only be achieved if $\sigma_1 \approx \sigma_2$ and if the variances of the Brownian motion processes $W_1^Q(T)$ and $W_2^Q(T)$ are minimized as this reduces the probability that the two Brownian motions end up further apart at maturity T. Hence, these differences are dependent on the initial parameters of our market model. To illustrate this dependence we have simulated 100,000 stock price values for two stocks at maturity T, found and standardized the differences and compared the obtained PDFs and CDFs of the differences with the one of a standard normal random variable. We have also used the one-sample Kolmogorov-Smirnov test at the 5% significance level with the following hypotheses:

H_0 : the difference follows a standard normal distribution

H_1 : the difference does not come from a standard normal distribution

Refer to figures 2.1 - 2.5 for the results of this simulation.

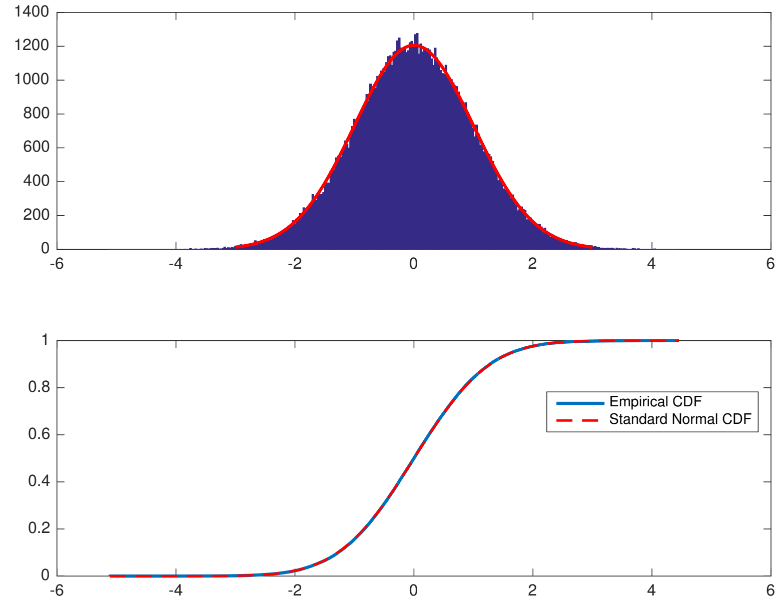

Figure 2.2: Parameters: $S(0) = 100, \sigma_1 = 0.1, \sigma_2 = 0.1, T = 1, \rho = 0.5$
Kolmogorov Smirnov Test: 0, p = 0.58051

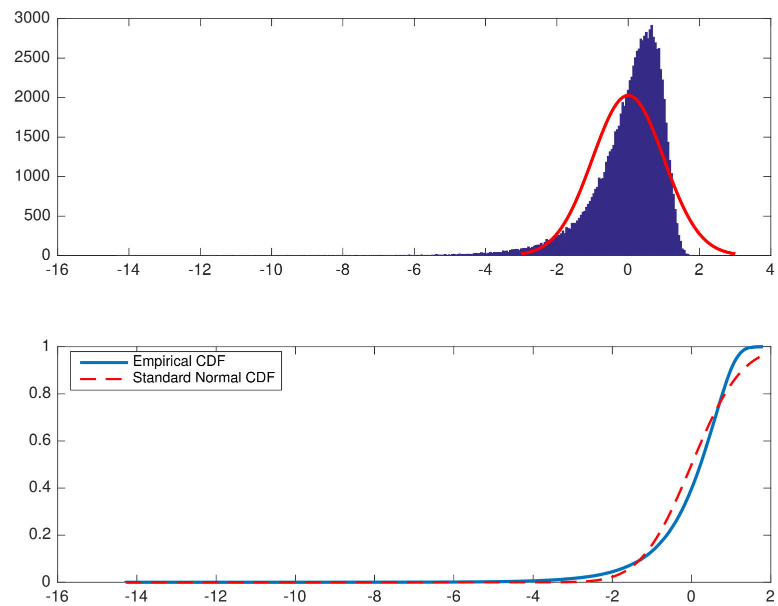

Figure 2.3: Parameters: $S(0) = 100, \sigma_1 = 0.1, \sigma_2 = 0.5, T = 1, \rho = 0.5$
Kolmogorov Smirnov Test: 1, p = 0

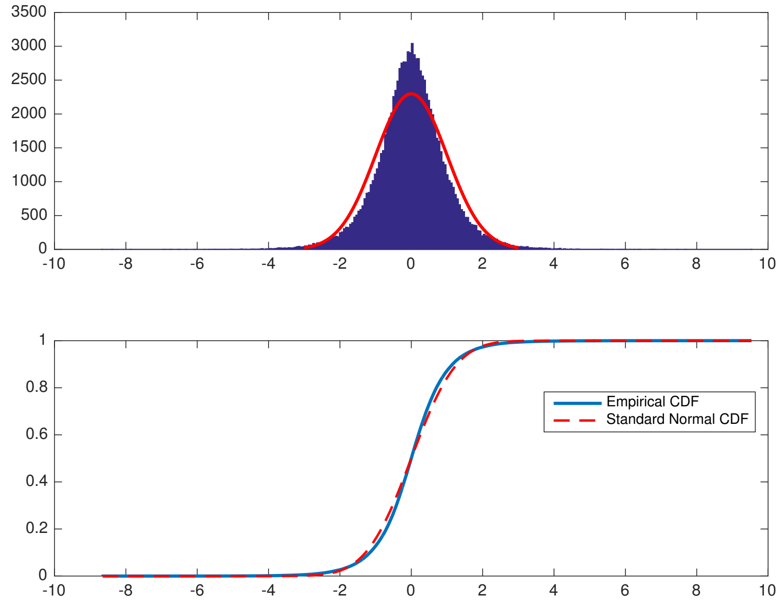

Figure 2.4: Parameters: $S(0) = 100, \sigma_1 = 0.1, \sigma_2 = 0.1, T = 20, \rho = 0.5$
Kolmogorov Smirnov Test: 1, p = 2.2038e-156

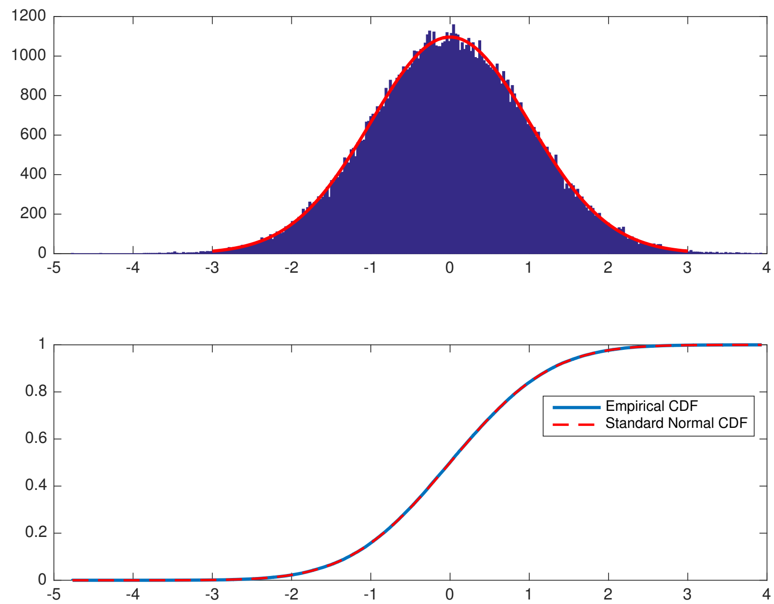

Figure 2.5: Parameters: $S(0) = 100, \sigma_1 = 0.01, \sigma_2 = 0.01, T = 20, \rho = 0.5$
Kolmogorov Smirnov Test: 0, p = 0.73933

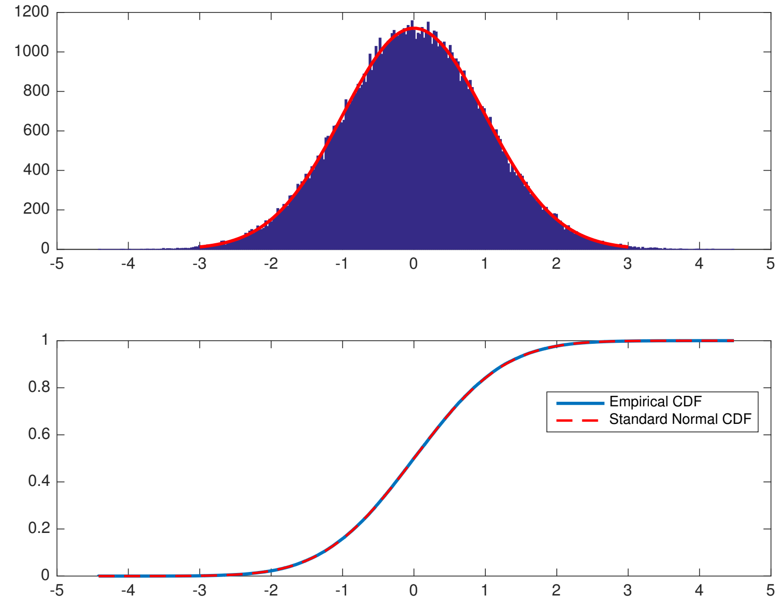

Figure 2.6: Parameters: $S(0) = 100, \sigma_1 = 0.1, \sigma_2 = 0.1, T = 1, \rho = -0.9$
Kolmogorov Smirnov Test: 0, p = 0.88763

The results of the simulation confirm our previous claims. We can observe that for $\sigma_1 = \sigma_2$ and low values of T we fail to reject the null hypothesis at 5% significance level as per figure (2.2). If there is a significant difference in volatilities, the distribution of the differences standardized is no longer standard normal as in figure (2.3) but is negatively skewed. From figure (2.4) we can infer that even for $\sigma_1 = \sigma_2$ but with maturity T extending far into the future, we can no longer accept the null hypothesis as the variances of each of the two Brownian motion processes are increased. We also note that correlation coefficient between W_1^Q and W_2^Q does not affect the distribution of the differences standardized as can be seen from figure (2.6).

Suppose that an investor or a financial institution is willing to sacrifice precision in favour of quickness of implementation and ease of understanding, in which case it is possible to price a contingent claim with payoff as in (2.3) by working with a normally distributed random variable. Denote $S_1(T) - S_2(T)$ by γ and suppose that γ is normally distributed with mean m and variance σ^2. Then the price p of a contingent claim with the payoff as in (2.3) can be approximated by considering the following expectation:

$$
\begin{aligned}
E_Q((\gamma - K)^+) &= E_Q\Big((\gamma - K)I_{\{\gamma \geq K\}}\Big) \\
&= E_Q(\gamma I_{\{\gamma \geq K\}}) - E_Q(K I_{\{\gamma \geq K\}}) \\
&= \int_K^{\infty} \frac{u}{\sigma\sqrt{2\pi}} exp\left\{-\frac{(u-m)^2}{2\sigma^2}\right\} du - K P_Q\Big(\epsilon \leq \frac{m-K}{\sigma}\Big)
\end{aligned}
$$

where: $\epsilon \sim N(0,1)$. By evaluating the above integral:

$$
p = \frac{m}{2}\Big(1 - erf\Big(\frac{K-m}{\sigma\sqrt{2}}\Big)\Big) + \frac{\sigma}{\sqrt{2\pi}} exp\left\{-\frac{(K-m)^2}{2\sigma^2}\right\} - K\Phi\Big(\frac{m-K}{\sigma}\Big) \quad (2.11)
$$

where:

$$
erf(x) = \frac{2}{\sqrt{\pi}} \int_0^x exp\left\{-u^2\right\} du
$$

Alternatively, since we use normal distribution only as an approximation, we can improve the precision of (2.11) considering an expectation expression of the following form:

$$
E_Q\Big((S_1(T) - S_2(T) - K)^+\Big) = E_Q(S_1(T)I_{\{\gamma > K\}}) - E_Q(S_2(T)I_{\{\gamma > K\}}) - K E_Q(I_{\{\gamma > K\}})
$$

The price of a spread option in this case is estimated as:

$$
p = S_1(0)\Phi(\hat{x}_1) - S_2(0)\Phi(\hat{x}_2) - K\Phi(\hat{x}_3) \quad (2.12)
$$

where:

$$
\begin{aligned}
\hat{x}_1 &= \frac{-K-m}{\sigma} + \sigma_1\rho_1\sqrt{T} \\
\hat{x}_2 &= \frac{-K-m}{\sigma} + \sigma_2\rho_2\sqrt{T} \\
\hat{x}_3 &= \frac{-K-m}{\sigma} \\
\rho_1 &= \frac{\Big(S_1(0)\sigma_1 - S_2(0)\sigma_2\rho\Big)\sqrt{T}}{\sigma} \\
\rho_2 &= \frac{\Big(S_1(0)\sigma_1\rho - S_2(0)\sigma_2\Big)\sqrt{T}}{\sigma} \\
S_1(T) &- S_2(T) \sim N(m, \sigma^2)
\end{aligned}
$$

Refer to Appendix 3 where we derive a similar but more advanced version of the above formula that we will encounter later in the thesis.

Carmona and Durrleman (2003) also note that the parameters m and σ^2 can be calculated directly by requiring that at least the first two moments of the true distribution and the approximating normal distribution coincide. In this case the mean rate of return m can be estimated as follows:

$$
\begin{aligned}
&E_Q\Big(S_1(T) - S_2(T)\Big) = \\
&= E_Q\left(S_1(0)exp\left\{-\frac{\sigma_1^2 T}{2} + \sigma_1 W_1^Q(T)\right\} - S_2(0)exp\left\{-\frac{\sigma_2^2 T}{2} + \sigma_2 W_2^Q(T)\right\}\right) \\
&= E_Q\left(S_1(0)exp\left\{-\frac{\sigma_1^2 T}{2} + \sigma_1 W_1^Q(T)\right\}\right) - E_Q\left(S_2(0)exp\left\{-\frac{\sigma_2^2 T}{2} + \sigma_2 W_2^Q(T)\right\}\right) \\
&= S_1(0) - S_2(0)
\end{aligned}
$$

To calculate variance we note that: $Var(x) = E(x^2) - (E(x))^2$. Let us first calculate the expectation of the difference squared:

$$E_Q\Big(\Big(S_1(T)-S_2(T)\Big)^2\Big) = E_Q\Big(S_1^2(0)exp\Big\{-\sigma_1^2T+2\sigma_1W_1^Q(T)\Big\}\Big)-$$

$$E_Q\Big(2S_1(0)S_2(0)exp\Big\{-\Big(\frac{\sigma_1^2+\sigma_2^2}{2}\Big)T+\sigma_1W_1^Q(T)+\sigma_2W_2^Q(T)\Big\}\Big)$$

$$+E_Q\Big(S_2^2(0)exp\Big\{-\sigma_2^2T+2\sigma_2W_2^Q(T)\Big\}\Big)$$

$$= S_1^2(0)exp\left\{\sigma_1^2T\right\} + S_2^2(0)exp\left\{\sigma_2^2T\right\} - 2S_1(0)S_2(0)exp\Big\{\Big(-\frac{\sigma_1^2+\sigma_2^2}{2}\Big)T\Big\}$$

$$+E_Q\Big(exp\Big\{\sigma_1W_1^Q(T)+\sigma_2W_2^Q(T)\Big\}\Big)$$

Since $\sigma_1W_1^Q(T)+\sigma_2W_2^Q(T) \sim N(0, \sigma_1^2T+\sigma_2^2T+2\sigma_1^2\sigma_2^2\rho T)$, the above sums up to:

$$E_Q\Big(S_1(T)-S_2(T)\Big)^2 =$$

$$S_1^2(0)exp\left\{\sigma_1^2T\right\} + S_2^2(0)exp\left\{\sigma_2^2T\right\} -$$

$$2S_1(0)S_2(0)exp\Bigg\{-\Bigg(\Big(\frac{\sigma_1^2+\sigma_2^2}{2}\Big)T - \frac{\sigma_1^2T+\sigma_2^2T+2\sigma_1^2\sigma_2^2\rho T}{2}\Bigg)\Bigg\}$$

Now we subtract the mean squared to finally get:

$$Var\Big(S_1(T)-S_2(T)\Big) = S_1^2(0)exp\left\{\sigma_1^2T\right\} + S_2^2(0)exp\left\{\sigma_2^2T\right\} -$$

$$2S_1(0)S_2(0)exp\Big\{-\Big(\frac{\sigma_1^2+\sigma_2^2}{2}\Big)T + \frac{\sigma_1^2T+\sigma_2^2T+2\sigma_1^2\sigma_2^2\rho T}{2}\Big\} - \Big(S_1(0)-S_2(0)\Big)^2$$

We would like to note that while these approximations for the mean and variance of the approximating normal distribution are good for certain choices of initial market model parameters, for the rest of the thesis we are using the parameters inferred from Monte Carlo simulations whenever we use (2.12) and its derivatives.

To compare the different pricing formulas for spread options we have estimated the prices by first varying the volatility of the first stock σ_1 and the time to maturity T parameters. The results are presented in table 2.1. Tables 2.2 and 2.3 show the values of absolute and percentage errors when compared with the results of Monte Carlo simulations. To check if there is any dependence of the accuracy of our estimates on the strike price K we have estimated the prices by varying the non-stochastic component and the results are summarized in table 2.4.

T / σ_1	0.5	1	3	5
0.1	4.8886	6.9092	11.937	15.372
	4.8853	6.9049	11.933	15.37
	4.885	6.9041	11.928	15.361
	4.885	6.9041	11.928	15.361
	4.8399	6.8061	11.404	14.228
0.15	5.1504	7.2788	12.573	16.188
	5.145	7.2715	12.563	16.178
	5.1447	7.2708	12.559	16.17
	5.1447	7.2708	12.559	16.17
	5.1124	7.1577	12.103	15.121
0.2	5.7911	8.1831	14.127	18.179
	5.7835	8.1726	14.111	18.159
	5.7833	8.172	14.108	18.153
	5.7833	8.172	14.108	18.153
	5.7503	8.0437	13.445	16.897
0.25	6.7008	9.4663	16.327	20.99
	6.6926	9.4547	16.307	20.963
	6.6925	9.4544	16.305	20.959
	6.6925	9.4544	16.305	20.959
	6.6062	9.2315	15.358	18.908
0.3	7.7852	10.994	18.937	24.313
	7.7771	10.983	18.914	24.279
	7.7771	10.983	18.913	24.277
	7.7771	10.983	18.913	24.277
	7.7153	10.651	17.074	20.852

Table 2.1: Spread Option: Value Approximation
Parameters: $S_1(0) = 105, S_2(0) = 100, K = 5, \sigma_2 = 0.2, \rho = 0.5$
The different formulas are from top to bottom: MC simulations, Kirk's approximation, Bjerksund and Stensland approximation, Carmona and Durrleman approximation, Normal approximation

T / σ_1	0.5	1	3	5
0.1	0 -0.0033 -0.0036 -0.0036 -0.0487	0 -0.0043 -0.0051 -0.0051 -0.1031	0 -0.004 -0.009 -0.009 -0.533	0 -0.002 -0.011 -0.011 -1.144
0.15	0 -0.0054 -0.0057 -0.0057 -0.038	0 -0.0073 -0.008 -0.008 -0.1211	0 -0.01 -0.014 -0.014 -0.47	0 -0.01 -0.018 -0.018 -1.067
0.2	0 -0.0076 -0.0078 -0.0078 -0.0408	0 -0.0105 -0.0111 -0.0111 -0.1394	0 -0.016 -0.019 -0.019 -0.682	0 -0.02 -0.026 -0.026 -1.282
0.25	0 -0.0082 -0.0083 -0.0083 -0.0946	0 -0.0116 -0.0119 -0.0119 -0.2348	0 -0.02 -0.022 -0.022 -0.969	0 -0.027 -0.031 -0.031 -2.082
0.3	0 -0.0081 -0.0081 -0.0081 -0.0699	0 -0.011 -0.011 -0.011 -0.343	0 -0.023 -0.024 -0.024 -1.863	0 -0.034 -0.036 -0.036 -3.461

Table 2.2: Spread Option: Absolute Error

T σ_1	0.5	1	3	5
0.1	0.00%	0.00%	0.00%	0.00%
	-0.07%	-0.06%	-0.03%	-0.01%
	-0.07%	-0.07%	-0.08%	-0.07%
	-0.07%	-0.07%	-0.08%	-0.07%
	-1.00%	-1.49%	-4.47%	-7.44%
0.15	0.00%	0.00%	0.00%	0.00%
	-0.10%	-0.10%	-0.08%	-0.06%
	-0.11%	-0.11%	-0.11%	-0.11%
	-0.11%	-0.11%	-0.11%	-0.11%
	-0.74%	-1.66%	-3.74%	-6.59%
0.2	0.00%	0.00%	0.00%	0.00%
	-0.13%	-0.13%	-0.11%	-0.11%
	-0.13%	-0.14%	-0.13%	-0.14%
	-0.13%	-0.14%	-0.13%	-0.14%
	-0.70%	-1.70%	-4.83%	-7.05%
0.25	0.00%	0.00%	0.00%	0.00%
	-0.12%	-0.12%	-0.12%	-0.13%
	-0.12%	-0.13%	-0.13%	-0.15%
	-0.12%	-0.13%	-0.13%	-0.15%
	-1.41%	-2.48%	-5.93%	-9.92%
0.3	0.00%	0.00%	0.00%	0.00%
	-0.10%	-0.10%	-0.12%	-0.14%
	-0.10%	-0.10%	-0.13%	-0.15%
	-0.10%	-0.10%	-0.13%	-0.15%
	-0.90%	-3.12%	-9.84%	-14.24%

Table 2.3: Spread Option: Percentage Error

K	Option Price	Absolute Error	Percentage Error
5	8.1831	0	0.00%
	8.1726	-0.0105	-0.13%
	8.172	-0.0111	-0.14%
	8.172	-0.0111	-0.14%
	8.0114	-0.1717	-2.10%
10	8.0068	0	0.00%
	7.9965	-0.0103	-0.13%
	7.9956	-0.0112	-0.14%
	7.9956	-0.0112	-0.14%
	7.8902	-0.1166	-1.46%
20	7.7043	0	0.00%
	7.694	-0.0103	-0.13%
	7.6929	-0.0114	-0.15%
	7.6929	-0.0114	-0.15%
	7.5557	-0.1486	-1.93%
40	7.3263	0	0.00%
	7.3143	-0.012	-0.16%
	7.314	-0.0123	-0.17%
	7.314	-0.0123	-0.17%
	7.2108	-0.1155	-1.58%

Table 2.4: Spread Option: Value Approximation for Varying K
Parameters: $S_1(0) = 105, \sigma_1 = 0.2, \sigma_2 = 0.2, \rho = 0.5$

We can see that the percentage errors vary significantly depending on the choice of initial parameters with the lowest percentage error being -0.70% for $\sigma_1 = 0.2$ and $T = 0.5$ and the highest being -14.24% for $\sigma_1 = 0.3$ and $T = 5$. There seems to be some dependence on the choice of the strike price K, however this relationship is not linear. All other formulas are invariant to changes in the deterministic strike price component. We also observe that all the formulas tend to undervalue the price of a spread option for all the choices of market parameters. For the purposes of this thesis we will use the Bjerksund and Stensland approximation (2.8) as well as the proposed normal approximation (2.12) to tackle the problem of minimizing $CVaR$.

2.3 Conditional Value at Risk

In this section we mathematically define VaR and $CVaR$ as well as provide the necessary theorems related to $CVaR$ minimization as was proposed by Rockafellar and Uryasev (2002). Let $(\Omega, \mathscr{F}, P)$ be a probability space and $L = L(x, w)$ be the function of losses, which depends on the chosen strategy $x \in X$, and a random outcome $w \in \Omega$. Note that throughout the thesis we consider losses as positive quantities. Assume that $L(x, w)$ is $\mathscr{F}$- measurable and $\sup_x E(L(x, w)) < \infty$. Let $\Phi(x, \bullet)$ be the CDF of L for a fixed strategy x:

$$\Phi(x, z) = P(w : L(x, w) \leq z)$$

Now we fix some confidence level $a \in (0, 1)$, usually $a = 0.9, 0.95$ or 0.99 for practical purposes, and definde the corresponding VaR_a and $CVaR_a$.

Definition 3. *For a given strategy x VaR_a is equal to the a-quantile of the loss function:*

$$VaR_a(x) = inf \{z \in \mathbb{R} : \Phi(x, z) \geq a\} \tag{2.13}$$

Since $z \in \mathbb{R}$, the infimum in (2.13) is always attained. We note that VaR does not consider the "information" above a certain cutoff point.

Definition 4. *For a given strategy x $CVaR_a$ of the loss function is equal to:*

$$CVaR_a(x) = \int_{-\infty}^{\infty} z \, d\Phi^a(x, z), \ \ \forall \ \ a \in (0, 1) \tag{2.14}$$

where:

$$\Phi^a(x, z) = \begin{cases} 0 & , \textit{if } \ z < \ \ VaR_a(x) \\ \frac{\Phi(x,z)-a}{1-a} & , \textit{otherwise} \end{cases}$$

Alternatively, we can express $CVaR_a$ associated with VaR_a in the following form:

$$CVaR_a(x) = E\big(L(x) : L(x) \geq VaR_a\big)$$

In other words, $CVaR_a$ is the conditional expectation of the losses given that these losses exceed VaR_a. Refer to figure 2.7 for graphical representation of the two concepts, where $CVaR_a$ is the expectation of the tail distribution of $\Phi(x, z)$, coloured orange.

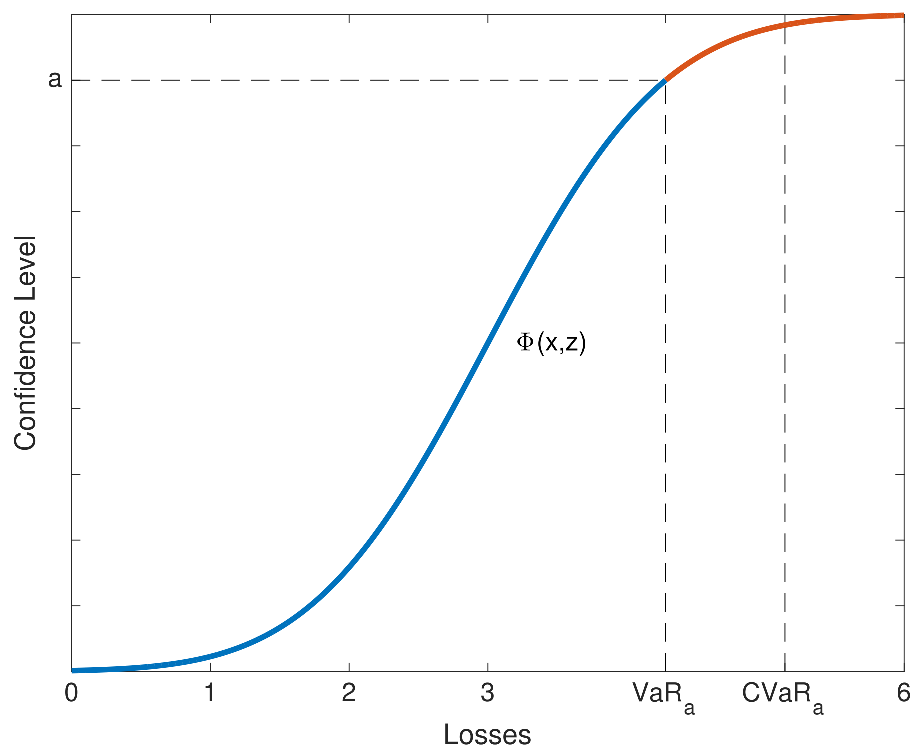

Figure 2.7: Value at Risk and Conditional Value at Risk

Artzner et al. (1999) established the axiomatics of coherent risk measures for a general type of market, both complete and incomplete. Let $\mathscr{G}$ be the set of all risks, i.e. the set of all real-valued functions on Ω. Denote by $\mathscr{A}$ the acceptance set, the set of risks with the acceptable level of future net worth (see the original article for details). In practice the set is specified by regulatory, accounting and risk management concerns. Define a risk measure:

Definition 5. *Given the total rate of return on a reference instrument r_T the risk measure associated with the acceptance set $\mathscr{A}$ is the mapping $\kappa_{\mathscr{A}}$ from $\mathscr{G}$ to $\mathbb{R}$ defined by:*

$$\kappa_{\mathscr{A}}(L) = inf \{m : mr_T + L \in \mathscr{A}\}$$

A risk-free zero-coupon bond is usually taken as the reference instrument. A positive value of $\kappa_{\mathscr{A}}(L)$ is interpreted as the amount of extra capital that an institution needs to invest at the risk-free rate of interest to ensure that the risk is acceptable and can be taken on by the institution. A negative value of $\kappa_{\mathscr{A}}(L)$ signifies the amount m that can be safely withdrawn. Let us now introduce the axioms.

Axiom 6. *Translation invariance: for all $L \in \mathscr{G}$ and real numbers c:*

$$\kappa(L + cr_T) = \kappa(L) - c$$

Adding an amount of money c and investing it at a risk-free rate of interest reduces the overall risk by this amount c. In other words, risk-free bond does not contribute to the overall risk of a portfolio.

Axiom 7. *Subadditivity: for all L_1 and $L_2 \in \mathscr{G}$:*

$$\kappa(L_1 + L_2) \leq \kappa(L_1) + \kappa(L_2)$$

This is a variant of the triangle inequality applied within the context of financial markets. Risk measure that satisfies axiom 7 benefits from diversification. It is possible to reduce an overall risk of a portfolio by adding an additional risky asset as long as it is not perfectly positively correlated to the ones already present.

Axiom 8. *Positive Homogeneity: for all* $\lambda \geq 0$ *and all* $L \in \mathscr{G}$*:*

$$\kappa(\lambda L) = \lambda \kappa(L)$$

Axiom 8 connects the overall risk of a portfolio to its size.

Axiom 9. *Monotonicity: for all* L_1 *and* $L_2 \in \mathscr{G}$ *with* $L_1 \leq L_2$*:*

$$\kappa(L_1) \leq \kappa(L_2)$$

A coherent risk measure is defined as follows:

Definition 10. *A risk measure satisfying the four axioms of translation invariance, subadditivity, positive homogeneity and monotonicity is called coherent.*

VaR as a risk measure satisfies the following three properties: translation invariance, positive homogeneity and monotonicity. Unless the distribution of portfolio risks is from an elliptical family of distributions, VaR does not satisfy the subadditivity property and thus may, in theory, discourage diversification - "the only free lunch in finance" (Harry Markowitz).

There has been extensive research done in the past years that challenges the theory of random walks for stock prices first put forward by Louis Bachelier, where stock returns are modelled as normally distributed random variables. Benoit Mandelbrot (1963), the father of fractal geometry, studied the stock price data from the year 1900 and concluded that the empirical distributions tend to be more "peaked" as compared to what would be expected under the normality assumption and proposed to model stock returns using the Pareto distribution. Fama (1965) confirmed the findings of Mandelbrot and found out that the empirical stock returns are riskier than what the standard deviations of normal distributions tend to predict. For a more recent overview refer to LeBaron (2008). Generally, empirical stock returns tend to be drawn from fat-tailed distributions. Figure 2.8 plots a PDF of a sinh-arcsinh distribution with parameters $\epsilon = 0$ and $\delta = 1.5$, which has a kurtosis of 4.7577, against a standard normal distribution with the kurtosis of 3. We have implemented the sinh-asrcsinh transformation proposed by Jones and Pewsey (2009), where, given a generating distribution (standard normal in our case), one can generate a family of 4-parameter distributions and control the kurtosis and skewness parameters by changing the values of ϵ and δ.

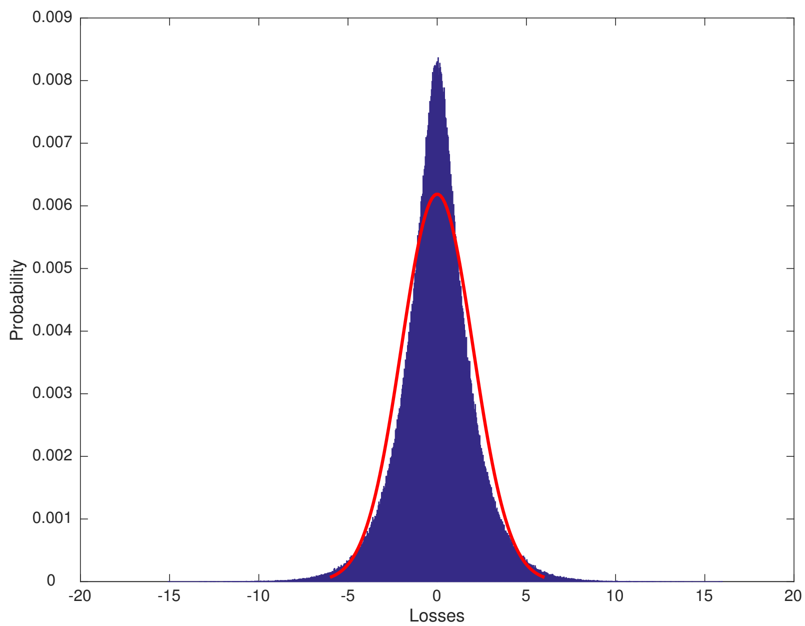

Figure 2.8: Sinh-Arcsinh Distribution PDF. Parameters: $\epsilon = 0$, $\delta = 1.5$

To provide an example of non-subadditivity of VaR, consider a portfolio consisting of two zero-coupon bonds B_1 and B_2 selling at par. For simplicity assume that there are only 3 scenarios that can occur, where the portfolio and each of the bonds will have the following losses L:

$$\text{Losses of } B_1 = \begin{cases} 20, & \text{if scenario 1 occurs} \\ 10, & \text{if scenario 2 occurs} \\ 0, & \text{otherwise} \end{cases}$$

$$\text{Losses of } B_2 = \begin{cases} 10, & \text{if scenario 1 occurs} \\ 20, & \text{if scenario 2 occurs} \\ 0, & \text{otherwise} \end{cases}$$

$$\text{Portfolio losses} = \begin{cases} 30, & \text{if scenario 1 occurs} \\ 30, & \text{if scenario 2 occurs} \\ 0, & \text{otherwise} \end{cases}$$

The probabilities assigned to each scenario are $5\%, 5\%, 90\%$ respectively. Let us first consider $VaR_{0.95}$ of B_1. According to definition 3, the set of $z \in \mathbb{R}$ for which $P(L \leq z) \geq 0.95$ is an open set $(10, \infty)$. The greatest lower bound of this set is 10 and hence its infimum is 10, which is the corresponding $VaR_{0.95}$ of B_1. Likewise, $VaR_{0.95}$ of B_2 is 10. However, the portfolio $VaR_{0.95}$ is equal to 30. The combined risk of this portfolio is greater than the individual risks of its constituents.

Artzner et al. (1999) also point out that VaR does not recognize concentration of risks and fails to encourage an optimal distribution of wealth between agents in the market. On the other hand, Embrechts and Wang (2015) provided seven proofs that $CVaR$ is a sub-additive measure of risk, which in addition to the other three axioms makes $CVaR$ a coherent risk measure in the sense just discussed.

Definition 4 points to a direct link between VaR and $CVaR$, which we hope would make the computations of the two risk measures connected to each other. Indeed, Rockafellar and Uryasev (2002) showed that it is possible to estimate both VaR and $CVaR$ associated with a given loss function $L(x, w)$ at the same time by solving a one-dimensional convex optimization problem. To do that we need to introduce the following auxiliary function:

$$F_a(x, z) = z + \frac{1}{1-a} E\Big((L(x, w) - z)^+\Big)$$

Theorem 11. *$F_a(x, z)$ is finite with respect to all $z \in \mathbb{R}$ and is convex with respect to z. Also, the following relationships are true:*

$$CVaR_a(x) = \min_{z \in \mathbb{R}} F_a(x, z) \tag{2.15}$$

$$VaR_a(x) = \min \Big\{ y : y \in \arg\min_{z \in \mathbb{R}} F_a(x, z) \Big\} \tag{2.16}$$

In particular one always gets:

$$CVaR_a(x) = F_a\Big(x, VaR_a(x)\Big) \tag{2.17}$$

Proof. The fact that $F_a(x, z)$ is finite at any $z \in \mathbb{R}$ is a direct consequence of our assumption that $\sup_x E(L, w) < \infty$. The fact that $F_a(x, z)$ is convex follows from the fact that $(L(x, w))^+$ is a convex function with respect to z and expectation is a linear operator that preserves the convexity property.

The proof of assertion (2.17) is taken from the original article. We first note that $F_a(x, z)$ has finite right and left derivatives at any point z since $F_a(x, z)$ is convex and finite with respect to z (for proof see Rockafellar, 1970). Let us first derive these one-sided derivatives. Consider the following:

$$\frac{F_a(x,z') - F_a(x,z)}{z'-z} = 1 + \frac{1}{1-a} E\Big(\frac{(L(x,w)-z')^+ - (L(x,w)-z)^+}{z'-z}\Big)$$

For the case when $z' > z$:

$$\frac{(L(x,w)-z')^+ - (L(x,w)-z)^+}{z'-z} = \begin{cases} 1, & \text{if } \ L(x,w) \geq z' \\ 0, & \text{if } \ L(x,w) \leq z \\ \in (-1,0), & \text{if } \ z < L(x,w) < z' \end{cases}$$

Observe that:

$$P(w|L(x,w) > z') = 1 - \Phi(x,z')$$

and:

$$P(w|z < L(x,w) \leq z') = \Phi(x,z') - \Phi(x,z)$$

Then, there should exist some value $p(z,z') \in [0,1]$ for which:

$$E\Big(\frac{(L(x,w)-z')^+ - (L(x,w)-z)^+}{z'-z}\Big) = -(1 - \Phi(x,z') - p(z,z')((x,z') - \Phi(x,z)$$

Since $\Phi(x,z') \searrow \Phi(x,z)$ as $z' \searrow z$:

$$\lim_{z' \searrow z} E\Big(\frac{(L(x,w)-z')^+ - (L(x,w)-z)^+}{z'-z}\Big) = -(1 - \Phi(x,z))$$

Applying this to (2.3), one obtains:

$$\lim_{z' \searrow z} \frac{F_a(x, z') - F_a(x, z)}{z' - z} = 1 + \frac{1}{1-a}(\Phi(x, z) - 1) = \frac{\Phi(x, z) - a}{1-a}$$

Following the same line of reasoning for the case $z' < z$, one can show that the left derivative is equal to:

$$\lim_{z' \nearrow z} \frac{F_a(x, z') - F_a(x, z)}{z' - z} = 1 + \frac{1}{1-a}(\Phi(x, z^-) - 1) = \frac{\Phi(x, z^-) - a}{1-a}$$

Since function $F_a(x, z)$ is a convex function, its derivatives are increasing with respect to z. By taking limits:

$$\lim_{z \to \infty} \frac{\partial^+ F_a(x, z)}{\partial z} = \lim_{z \to \infty} \frac{\partial^- F_a(x, z)}{\partial z} = 1$$

and

$$\lim_{z \to -\infty} \frac{\partial^+ F_a(x, z)}{\partial z} = \lim_{z \to -\infty} \frac{\partial^- F_a(x, z)}{\partial z} = -\frac{a}{1-a}$$

These limits indicate that the sets $\{z | F_a(x, z) \leq c\}$ are bounded for any $c \in \mathbb{R}$ and hence the minimum in (2.15) is attainable. Also, the set of arguments of the minima in (2.16) is a closed bounded interval, where the elements are such that:

$$\frac{\partial^- F_a(x, z)}{\partial z} \leq 0 \leq \frac{\partial^+ F_a(x, z)}{\partial z}$$

Relationship (2.17) immediately follows.

□

Theorem 11 reveals why $CVaR$ is a more robust measure of risk than VaR:

the optimal value in an optimization problem has better properties as a function of parameter than the interval of arguments of the minima.

Theorem 12. *Minimizing $CVaR_a(x)$ over all $x \in X$ is equivalent to minimizing $F_a(x,z)$ over all $(x,z) \in X \times \mathbb{R}$:*

$$\min_{x \in X} CVaR_a(x) = \min_{(x,z) \in X \times \mathbb{R}} F_a(x,z) \tag{2.18}$$

Proof. The proof follows from the fact that for each $x \in X$ the minimum with respect to z of $F_a(x,z)$ is attainable. □

2.4 Neyman-Pearson Lemma

This section is devoted to the statements of the classical results of Neyman-Pearson theory, which first appeared in the article by Jerzy Neyman and Egon Pearson (1933). Before stating the lemma itself, let us introduce some relevant definitions:

Definition 13. *Given a random sample drawn from a population distribution with parameter θ, simple hypothesis is the hypothesis that uniquely determines the population distribution. Any hypothesis that is not simple is a composite hypothesis.*

For example, given a random sample from a normal distribution with a known variance parameter σ^2, a simple hypothesis will be $H : \mu = 10$ as it would completely determine the population distribution. On the other hand, $H : \mu > 10$ is a composite hypothesis.

Definition 14. *Most powerful test is the test which, given some fixed confidence level $a \in (0,1)$, has the highest probability of rejecting the null hypothesis H_0 when it is indeed false.*

Neyman-Pearson lemma gives us the answer to the question of what is the most powerful test for a hypothesis test at a given confidence level.

Lemma 15. *Given some random sample x from a population distribution with parameter θ, suppose one is using the likelihood-ratio test to compare a simple hypothesis $H_0 : \theta_0$ against another simple hypothesis $H_1 : \theta_1$. The significance level a at which the null hypothesis is rejected is:*

$$a = P\left(\frac{\mathcal{L}(\theta_0|x)}{\mathcal{L}(\theta_1|x)} \leq \eta \mid H_0\right)$$

where $\mathcal{L}(\theta|x)$ is the likelihood function and η is the threshold of the likelihood-ratio test. In this case, $\frac{\mathcal{L}(\theta_0|x)}{\mathcal{L}(\theta_1|x)}$ is the most powerful test.

Due to its power and uniformity many extensions of the Neyman-Pearson lemma have been developed following the initial publication. As Lehmann (1992) put it: "Nevertheless, despite their shortcomings, the new paradigm formulated in the 1933 paper, and the many developments carried out within its framework continue to play a central role in both the theory and practice of statistics and can be expected to do so in the foreseeable future". In what follows we stick to the methodology proposed by Cvitanic and Karatzas (2001). Let Q and P be two probability measures on some measurable space $(\Omega, \mathscr{F})$. Suppose that we want to construct a hypothesis test that would distinguish between the two measures. Hence, we consider the following two simple hypotheses: $H_0 : Q$ and $H_1 : P$. Consider a randomized test, a continuous random variable $X : \Omega \to [0, 1]$. If we observe $w \in \Omega$, then we reject H_0 with probability $X(w)$. The probability of accepting H_0 is correspondingly $1 - X(w)$. Type I error, the probability of rejecting H_0 when it is correct, is:

$$E_Q(X) = \int X(w)Q(dw)$$

The probability of accepting H_0 when it is incorrect, also know as type II error is:

$$1 - E_P(X) = 1 - \int X(w)P(dw)$$

The power of this test is:

$$E_P(X) = \int X(w)P(dw)$$

In practice, it is impossible to minimize both types of errors at the same time. Indeed, suppose a trader is setting up an automatic trading strategy where he is receiving buy signals whenever the market is believed to go up. By trying to minimize type I error, i.e. the error of missing out on a rise in the market level, he is more likely make purchases before the market level drops, type II error. The compromise is to minimize the probability of making type II error while keeping the probability of type I error less than or equal to some fixed significance level a. Then, one needs to solve the following problem:

$$\begin{cases} 1 - E_P(X) \to min \\ E_Q(X) \leq a \end{cases} \tag{2.19}$$

which is equivalent to maximizing the power of the test while keeping type I error less than or equal to the significance level a:

$$\begin{cases} E_P(X) \to max \\ E_Q(X) \leq a \end{cases} \tag{2.20}$$

Let us introduce an auxiliary measure v such that: $P << v$ and $Q << v$, i.e. both P and Q are absolutely continuous with respect to the third measure v. Define the following Radon-Nikodym derivatives:

$$G := \frac{dP}{dv}$$
$$J := \frac{dQ}{dv}$$

Theorem 16. *Problem (2.20) has a solution for any $a \in (0,1)$:*

$$\widetilde{X} = I_{\{\widetilde{z}J<G\}} + \gamma I_{\{\widetilde{z}J=G\}} \tag{2.21}$$

where:

$$\widetilde{z} = inf\,\{z \geq 0 : Q(zJ < G) \leq a\} \tag{2.22}$$

$$\gamma = \frac{a - Q(\widetilde{z}J < G)}{Q(\widetilde{z}J = G)} = \frac{a - Q(\widetilde{z}J < G)}{Q(\widetilde{z}J \leq G) - Q(\widetilde{z}J < G)} \in [0,1] \tag{2.23}$$

Theorem 16 will play key role for the problem of minimizing expected shortfall.

2.5 Minimization of Expected Shortfall

In this section we state the results for the problem of expected shortfall minimization as in Foellmer and Leukert (2000). Let $S = (S(t) : t \in [0,T])$ be a discounted stock price process on a standard stochastic basis $(\Omega, \mathscr{F}, \mathscr{F}(t), P)$, where $\mathscr{F}(t)$ is a sigma-algebra of events or filtration, which we assume satisfies the usual conditions, and $\mathscr{F}(0) = \{\Omega, \emptyset\}$. We assume that the market is complete and thus there exists a unique martingale measure Q such that $E_Q(S(t) \mid \mathscr{F}(s)) = S(s), \forall\, s < t$. In other words, the discounted stock price process is a martingale under Q. Let us define a process of capital $V = (V(t) : t \in [0,T])$ for a class of self-financing strategies:

$$V(t) = V(0) + \int_0^t \xi(s) dS(s), \forall\, t \in [0,T]$$

where $V(0) > 0$ is the initial capital and $\xi = (\xi(t) : t \in [0,T])$ is a predictable process which represents the number units of the underlying asset held at any time $t \in [0,T]$. The strategy is self-financing meaning that there is no external inflow or outflow of capital in the system. All the transactions are financed by the sale of the existing assets in the portfolio. Together $V(0)$ and ξ form a portfolio $\pi = (V(0), \xi)$.

Definition 17. *A given self-financing strategy $\pi = (V(0), \xi)$ is admissible if the following condition is satisfied:*

$$V(t) \geq 0, \forall\, t \in [0,T], P - a.s.$$

Denote by $\mathscr{A}$ be the set of all admissible trading strategies. Assume also that there is an $\mathscr{F}(T)$-measurable contingent claim $H \in L^1(Q)$. Since the market is complete, we can construct the perfect hedge. In other words, there exists a predictable process ξ^H:

$$E_Q\Big(H|\mathscr{F}(t)\Big) = H(0) + \int_0^t \xi^H(s) dS(s), \forall\, t \in [0,T], P - a.s.$$

where $H(0)$ is the price of setting up a replicating portfolio and is determined as:

$$H(0) = E_Q(H)$$

So, H can be perfectly replicated with a self-financing strategy $\pi^H = (H(0), \xi^H)$.

If an investor only has $\hat{V}(0) < H(0)$ or is not willing to invest the whole amount into a hedging portfolio, it is not possible to construct the perfect hedge and the investor will be faced with the possibility of shortfall at maturity. The criterion of optimality that we are going to use is $E\Big(H - V(T)\Big)^+$, which we are going to minimize over all strategies $x \in X$ with the initial capital $V(0) \leq \hat{V}(0)$. Foellmer and Leukert (2000) also weighted the shortfall by a power loss function $l(z) = z^p$ to show the investor's attitude to risk, however for the purposes of this thesis we assume that $p = 1$. So, we get the following optimization problem to solve:

$$\begin{cases} E\Big(H - V(T)\Big)^+ = E\Big(H - V(0) - \int_0^T \xi(s) dS(s)\Big)^+ \to \min\limits_{\pi} \\ \pi \in \mathscr{A}, V(0) \leq \hat{V}(0) \end{cases} \tag{2.24}$$

Let us introduce the success ratio φ associated with a given admissible strategy π:

$$\varphi(\pi) = I_{\{V(T) \geq H\}} + \frac{V(T)}{H} I_{\{V(T) < H\}}$$

So, $\varphi(\pi)$ is $\mathscr{F}(T)$-measurable random variable and one can easily deduce that $\varphi(\pi^H) = 1$. We also note that:

$$\begin{aligned} \varphi(\pi)H &= H I_{\{V(T) \geq H\}} + \frac{V(T)}{H} I_{\{V(T) < H\}} H \\ &= H I_{\{V(T) \geq H\}} + V(T) I_{\{V(T) < H\}} \\ &= min \{H, V(T)\} \\ &= V(T) \wedge H \end{aligned}$$

The size of a shortfall for a given strategy π can be expressed via its success ratio $\varphi(\pi)$ in the following way:

$$(H - V(T))^+ = H - V(T) \wedge H = \Big(1 - \varphi(\pi)\Big) H \tag{2.25}$$

Denote by $\mathscr{M}$ the class of all $\mathscr{F}$-measurable random variables that take values in $(0, 1)$. This class will include all the random variables which correspond to the success ratios of all the admissible self-financing strategies. Now we reformulate problem (2.24) using the derived relationship in (2.25):

$$\begin{cases} E\Big((1-\varphi)H\Big) \to \min\limits_{\varphi \in \mathscr{M}} \\ E_Q(\varphi H) \le \hat{V}(0) \end{cases}$$

Which is equivalent to:

$$\begin{cases} E(\varphi H) \to \max\limits_{\varphi \in \mathscr{M}} \\ E_Q(\varphi H) \le \hat{V}(0) \end{cases} \tag{2.26}$$

Introduce the auxiliary probability measures P^* and Q^* as Radon-Nikodym derivatives in the following way:

$$\frac{dP^*}{dP} = \frac{H}{E(H)}$$
$$\frac{dQ^*}{dQ} = \frac{H}{E_Q(H)}$$

The expectations in (2.26) can be expressed through the newly defined measures P^*

and Q^* as follows:

$$\begin{aligned}
E(\varphi H) &= E_{P^*}\Big(\varphi H \frac{dP}{dP^*}\Big) \\
&= E_{P^*}\Big(\varphi H \frac{E(H)}{H}\Big) \\
&= E_{P^*}\Big(\varphi E(H)\Big) \\
&= E_{P^*}(\varphi) E(H)
\end{aligned}$$

and:

$$\begin{aligned}
E_Q(\varphi H) &= E_{Q^*}\Big(\varphi H \frac{dQ}{dQ^*}\Big) \\
&= E_{Q^*}\Big(\varphi H \frac{E_Q(H)}{H}\Big) \\
&= E_{Q^*}\Big(\varphi E_Q(H)\Big) \\
&= E_{Q^*}(\varphi) E_Q(H) \\
&= E_{Q^*}(\varphi) H(0)
\end{aligned}$$

Then, we can rewrite (2.26) in the following way:

$$\begin{cases} E_{P^*}(\varphi) \to \max\limits_{\varphi \in \mathscr{M}} \\ E_{Q^*}(\varphi) \leq \frac{\hat{V}(0)}{H(0)} = \beta \end{cases} \tag{2.27}$$

where $\beta \in (0,1)$ since $0 < \hat{V}(0) < H(0)$. By introducing a notion of a success ratio φ associated with a given strategy π, we have transformed the original optimization problem (2.24) to the class of problems considered in the previous section (see (2.20)), for which the existence of a solution is guaranteed by theorem 16. According to the theorem we should take a third measure. Let the third measure be the martingale measure Q. It satisfies the required conditions of absolute continuity:

$Q^* << Q$ and, since $Q \approx P$, we have that $P^* << Q$. So, for this problem the relation (2.22) takes the following form:

$$\widetilde{a} = \inf\left\{a \geq 0 : Q^*\Big(a\frac{dQ^*}{dQ} < \frac{dP^*}{dQ}\Big) \leq \beta\right\}$$

Let us simplify the above expectation expression:

$$\begin{aligned} Q^*\Big(a\frac{dQ^*}{dQ} < \frac{dP^*}{dQ}\Big) &= Q^*\Big(a\frac{dQ^*}{dQ} < \frac{dP^*}{dP}\frac{dP}{dQ}\Big) = Q^*\Big(a\frac{H}{E_Q(H)} < \frac{H}{E(H)}\frac{dP}{dQ}\Big) \\ &= Q^*\Big(a\frac{E(H)}{E_Q(H)} < \frac{dP}{dQ}\Big) = Q^*\Big(\frac{dP}{dQ} > b\Big) = E_{Q^*}\Big(I_{\{\frac{dP}{dQ}>b\}}\Big) \\ &= E_Q\Big(I_{\{\frac{dP}{dQ}>b\}}\frac{dQ^*}{dQ}\Big) = E_Q\Big(I_{\{\frac{dP}{dQ}>b\}}\frac{H}{E_Q(H)}\Big) \\ &= \frac{1}{E_Q(H)}E_Q\Big(HI_{\{\frac{dP}{dQ}>b\}}\Big) = \frac{1}{H(0)}E_Q\Big(HI_{\{\frac{dP}{dQ}>b\}}\Big) \end{aligned}$$

where:

$$b = a\frac{E(H)}{H(0)} \geq 0$$

Then:

$$\begin{aligned} \widetilde{a} &= \inf\left\{a \geq 0 : Q^*\Big(a\frac{dQ^*}{dQ} < \frac{dP^*}{dQ}\Big) \leq \beta\right\} \\ &= \inf\left\{a \geq 0 : E_Q\Big(HI_{\{\frac{dP}{dQ}>a\}}\Big) \leq \hat{V}(0)\right\} \end{aligned}$$

To transform (2.23) we follow the same procedure as for (2.22) to get:

$$\gamma = \frac{\hat{V}(0) - E_Q\Big(HI_{\{\frac{dP}{dQ}>\widetilde{a}\}}\Big)}{E_Q\Big(HI_{\{\frac{dP}{dQ}=\widetilde{a}\}}\Big)}$$

Finally, the solution to problem (2.27) takes the following form:

$$\begin{cases} \widetilde{\varphi} = I_{\left\{\frac{dP}{dQ} > \widetilde{a}\right\}} + \gamma I_{\left\{\frac{dP}{dQ} = \widetilde{a}\right\}}, \text{ if } Q\Big(\left\{\frac{dP}{dQ} = \widetilde{a}\right\} \cap \{H > 0\}\Big) > 0 \\ \widetilde{\varphi} = I_{\left\{\frac{dP}{dQ} > \widetilde{a}\right\}} \qquad\qquad\quad , \text{ if } Q\Big(\left\{\frac{dP}{dQ} = \widetilde{a}\right\} \cap \{H > 0\}\Big) = 0 \end{cases}$$

So, we have found $\widetilde{\varphi} \in \mathscr{M}$ - the solution to problem (2.26). Now consider a random variable of the following form:

$$\widetilde{H} = \widetilde{\varphi} H$$

It is obvious that $\widetilde{H}$ is $\mathscr{F}(T)$-measurable, nonnegative and integrable random variable. Hence, we can consider it as a contingent claim. The price of this claim or, equivalently, the price of setting up a hedging portfolio is given by:

$$E_Q(\widetilde{H}) = \widetilde{V}(0)$$

In a complete market the above contingent claim is replicable with a strategy $\widetilde{\pi} = (\widetilde{V}(0), \widetilde{\xi})$:

$$E_Q\Big(\widetilde{H}|\mathscr{F}(t)\Big) = \widetilde{V}(0) + \int_0^t \widetilde{\xi}(s) dS(s), \forall\, t \in [0, T], P - a.s.$$

Theorem 18. *The perfect hedge $\widetilde{\pi} = (\widetilde{V}(0), \widetilde{\xi})$ for a contingent claim $\widetilde{H} = \widetilde{\varphi} H$, where $\widetilde{\varphi}$ is the solution to (2.26), is the optimal solution to problem (2.24). What is more, if we assume that $\widetilde{\varphi} = 1$ on $\{H = 0\}$, then: $\widetilde{\varphi} = \varphi(\widetilde{V}, \widetilde{\xi})$, P-a.s.*

Proof. Let $\pi = (V(0), \xi)$ be any admissible trading strategy such that $V(0) \leq \hat{V}(0)$ with the corresponding success ratio $\varphi = \varphi(\pi)$. Its value process satisfies:

$$V(t) = V(0) + \int_0^t \xi(s) dS(s), \forall\, t \in [0, T], P - a.s.$$

The corresponding shortfall is equal to:

$$(H - V(T))^+ = H - V(T) \wedge H = (1 - \varphi)H$$

Under the risk-neutral probability measure Q the corresponding value process is a supermartingale:

$$E_Q(\varphi H) = E_Q\Big(V(T) \wedge H\Big) \leq E_Q(V(T)) \wedge E_Q(H) \leq E_Q(V(T)) \leq V(0) \leq \hat{V}(0)$$

Thus, the success ratio satisfies the constraints in (2.26). The following inequality is further satisfied:

$$E(H - V(T))^+ = E((1 - \varphi)H) \geq E((1 - \widetilde{\varphi})H)$$

since $\widetilde{\varphi}$ is the optimal solution to problem (2.26). Now we are left to show that $\widetilde{\varphi} = \varphi(\widetilde{\pi})(P - a.s.)$, if we assume that $\widetilde{\varphi} = 1$ on $\{H = 0\}$. This assumption is of technical kind and does change the random variable $\widetilde{H} = \widetilde{\varphi} H$ and the solution to problem (2.26). We have:

$$\varphi(\widetilde{\pi}) = \widetilde{V}(T) \wedge H \geq \widetilde{\varphi} H \;\text{ on}\, \{H > 0\}\, (P - a.s.) \tag{2.28}$$

Since $\widetilde{\varphi}$ is the optimal solution to problem (2.26) and $\varphi(\widetilde{\pi})$ is an admissible solution:

$$E(\widetilde{\varphi} H) \geq E(\varphi(\widetilde{\pi})) \tag{2.29}$$

From (2.28) and (2.29) we deduce that:

$$\varphi(\widetilde{\pi})H = \widetilde{\varphi}H \text{ on } \{H > 0\}\,(P - a.s.)$$

Moreover:

$$\varphi(\widetilde{\pi}) = \widetilde{\varphi} = 1 \text{ on } \{H = 0\}\,(P - a.s.)$$

It follows that $\varphi(\widetilde{\pi}) = \widetilde{\varphi}\ \ (P - a.s.)$. $\square$

2.6 Minimization of Conditional Value at Risk

In this section we combine the results of the previous sections to derive the theorems related to the problem of minimizing $CVaR$. The methodology was first explained by Melnikov and Smirnov (2012). Again let $S = (S(t) : t \in [0, T])$ be a discounted stock price processes on a standard stochastic basis $(\Omega, \mathscr{F}, \mathscr{F}(t), P)$ with filtration $\mathscr{F}(t)$ that satisfies the usual conditions and $\mathscr{F}(0) = \{\Omega, \emptyset\}$. Let Q be a unique equivalent martingale measure which is absolutely continuous with respect to measure P. Denote by $\mathscr{A}$ the class of admissible self-financing strategies such that $\pi = (V(0), \xi)$. Let there be a contingent claim, which is $\mathscr{F}(T)$-measurable, nonnegative random variable $H \in L^1(Q)$. Further suppose that the initial capital available $\hat{V}(0)$ is strictly less than the required amount $H(0)$. We define the loss in the following way:

$$\begin{aligned} L(\pi) &= H - V(T) \\ &= H - V(0) - \int_0^T \xi(s)dS(s) \end{aligned}$$

Then, $CVaR_a(\pi)$ is determined by (2.14), where a is a fixed confidence level. We will be minimizing $CVaR_a$ over all strategies $\pi \in X$ with the restriction on the amount of initial capital available $V(0) \leq \hat{V}(0)$:

$$\begin{cases} CVaR_a(\pi) \to \min\limits_{\pi} \\ \pi \in \mathscr{A}, V(0) \leq \hat{V}(0) \end{cases} \tag{2.30}$$

The solution to problem (2.30) will be a hedging strategy $\widetilde{\pi} = (\widetilde{V}(0), \widetilde{\xi})$, optimal in the sense of $CVaR_a$, over all admissible self-financing strategies with initial capital $V(0) \leq \hat{V}(0)$. Problem (2.30) can be rewritten:

$$CVaR_a(\pi) \to \min_{\pi \in \mathscr{A}(\hat{V}(0))}$$

Applying theorem 12, where instead of set X we take set $\mathscr{A}(\hat{V}(0))$:

$$\begin{aligned} \min_{\pi \in \mathscr{A}(\hat{V}(0))} CVaR_a(\pi) &= \min_{\pi \in \mathscr{A}(\hat{V}(0))} \left(\min_{z \in \mathbb{R}} \left\{ z + \frac{1}{1-a} E(H - V(T) - z)^+ \right\} \right) \\ &= \min_{z \in \mathbb{R}} \left(\min_{\pi \in \mathscr{A}(\hat{V}(0))} \left\{ z + \frac{1}{1-a} E(H - V(T) - z)^+ \right\} \right) \end{aligned} \tag{2.31}$$

Define the auxiliary function $c(z)$ in the following way:

$$c(z) = \min_{\pi \in \mathscr{A}(\hat{V}(0))} \left\{ z + \frac{1}{1-a} E(H - V(T) - z)^+ \right\}$$

Problem (2.31) can be reformulated in the following form:

$$\min_{\pi \in \mathscr{A}(\hat{V}(0))} CVaR_a(\pi) = \min_{z \in \mathbb{R}} c(z) \tag{2.32}$$

Let the minimum of function $c(z)$ in (2.32) with respect to z be achieved using strategy $\widetilde{\pi}(z) = (\widetilde{V}(0,z), \widetilde{\xi}(z))$:

$$\min_{\pi \in \mathscr{A}(\hat{V}(0))} E(H - V(T) - z)^+ = E(H - \widetilde{V}(T,z) - z)^+$$

where:

$$\widetilde{V}(T,z) = \widetilde{V}(0,z) + \int_0^T \widetilde{\xi}(s,z)\, dS(s)$$

Let the global minimum of function $c(z)$ in (2.32) be achieved at point $\widetilde{z}$:

$$\min_{z \in \mathbb{R}} c(z) = c(\widetilde{z})$$

Then the optimal solution to the problem of $CVaR_a$ minimization over all $\pi \in \mathscr{A}(\hat{V}(0))$ is the strategy:

$$\widetilde{\pi}(\widetilde{z}) = (\widetilde{V}(0,\hat{z}), \widetilde{\xi}(0,\hat{z}))$$

Then according to theorem 11:

$$\begin{cases} CVaR_a(\widetilde{\pi}) = c(\widetilde{z}) \\ VaR_a(\widetilde{\pi}) = \widetilde{z} \end{cases}$$

It follows that if we can find the strategy $\widetilde{\pi}$ in an explicit form, then the problem of $CVaR_a$ minimization will be reduced to the problem of minimization of function $c(z)$. For each z strategy $\widetilde{\pi}$ is a solution to the following problem:

$$E(H - V(T) - z)^+ \to \min_{\pi \in \mathscr{A}(\hat{V}(0))} \tag{2.33}$$

Note that:

$$(H - V(T) - z)^+ = \Big((H - z)^+ - V(T)\Big)^+$$

Denote $(H - z)^+$ by $H(z)$. It is obvious that $H(z)$ is $\mathscr{F}$-measurable random variable, $H(z) \in L^1(Q)$ and $H(z) \geq 0$, which means that (2.33) can be restated in the following form:

$$E\Big(H(z) - V(T)\Big)^+ \to \min_{\pi \in \mathscr{A}(\hat{V}(0))} \tag{2.34}$$

Problem (2.34) can be interpreted as the problem of expected shortfall minimization over the strategy set $\mathscr{A}(\hat{V}(0))$ of contingent claim $H(z)$, to which we can apply theorem 18.

Proposition 19. *The optimal solution $\widetilde{\pi} = (\widetilde{V}(0), \widetilde{\xi})$ of problem (2.33) is the perfect hedge of contingent claim $\widetilde{H}(z) = \widetilde{\varphi}(z)(H - z)^+$ or, equivalently, $\widetilde{H}(z) = \widetilde{\varphi}(z)H(z)$:*

$$E_Q(\widetilde{H}(z)|\mathscr{F}(t)) = \widetilde{V}(0, z) + \int_0^t \widetilde{\xi}(s, z)\, dS(s),\ \ \forall\, t \in [0, T], (P - a.s.)$$

where:

$$\widetilde{\varphi}(z) = I_{\left\{\frac{dP}{dQ} > \widetilde{a}(z)\right\}} + \gamma(z) I_{\left\{\frac{dP}{dQ} = \widetilde{a}(z)\right\}} \tag{2.35}$$

$$\widetilde{a} = \inf\Big\{a \geq 0 : E_Q\Big((H - z)^+ I_{\left\{\frac{dP}{dQ} > a\right\}}\Big) \leq \hat{V}(0)\Big\} \tag{2.36}$$

$$\gamma = \frac{\hat{V}(0) - E_Q\Big((H - z)^+ I_{\left\{\frac{dP}{dQ} > \widetilde{a}(z)\right\}}\Big)}{E_Q\Big((H - z)^+ I_{\left\{\frac{dP}{dQ} = \widetilde{a}(z)\right\}}\Big)} \tag{2.37}$$

Proposition 20. *Function $c(z)$ is determined in the following way:*

$$c(z) = \begin{cases} z + \frac{1}{1-a} E\Big((1 - \widetilde{\varphi}(z))(H - z)^+\Big), & z < \hat{z} \\ z & , \ z \geq \hat{z} \end{cases} \tag{2.38}$$

Equivalently:

$$c(z) = \begin{cases} z + \frac{1}{1-a} E\Big(H(z) - \widetilde{H}(z)\Big), & z < \hat{z} \\ z & , \ z \geq \hat{z} \end{cases}$$

where $\hat{z}$ is the solution to the following equation:

$$\hat{V}(0) = E_Q\Big((H - \hat{z})^+\Big) \tag{2.39}$$

Proposition 21. *Let $\widetilde{z}$ be the global minimum of function $c(z)$, then strategy $\widetilde{\pi} = \widetilde{\pi}(\widetilde{z})$ is the optimal solution to the original problem (2.30) of $CVaR_a$ minimization and:*

$$\begin{cases} CVaR_a(\widetilde{\pi}) = c(\widetilde{z}) \\ VaR_a(\widetilde{\pi}) = \widetilde{z} \end{cases}$$

Note that in problem (2.34) we assumed that the initial capital available $\hat{V}(0)$ is strictly less than $E_Q(H(z))$. If this was not the case, we could use the strategy of complete hedging for $H(z)$. Then, $E(H(z) - V(T))^+ = 0$. From this (2.38) and (2.39) follow. It is also worth noting that when $z \geq \hat{z}$ from (2.35), (2.36) and (2.37) we have:

$$\widetilde{a}(z) = -\infty$$

and:

$$\widetilde{\varphi}(z) = 1$$

Hence for $z \geq \hat{z}$ the optimal strategy $\widetilde{\pi}(z) = (\widetilde{V}(0,z), \widetilde{\xi}(z))$ for problem (2.33) will be the perfect hedge of contingent claim $\widetilde{H}(z)$.

Chapter 3

Applications to Margrabe Market

Under the Margrabe market model we assume the existence of only two risky correlated assets $S_1 = (S_1(t) : t \in [0, T])$ and $S_2 = (S_2(t) : t \in [0, T])$ that satisfy the following stochastic differential equations:

$$\begin{cases} dS_1(t) = S_1(t)(\sigma_1 dW_1(t) + \mu_1 dt) \\ dS_2(t) = S_2(t)(\sigma_2 dW_2(t) + \mu_2 dt) \end{cases}$$

where $W_1 = (W_1(t) : t \in [0, T])$ and $W_2 = (W_2(t) : t \in [0, T])$ are standard Brownian motion processes with the correlation coefficient ρ. The mean rates of return μ_1 and μ_2 and volatilities σ_1 and σ_2 are assumed to be constant. The original model does not assume the existence of a bank account: $r = 0$, however the extension to the case where $r > 0$ is straightforward. The market is arbitrage-free and complete, hence we can introduce the equivalent martingale measure $Q \approx P$ via the Radon-Nikodym derivative (see, for instance, Melnikov, 2004):

$$Z(T) = \frac{dQ}{dP}$$

In an explicit form:

$$Z(t) = exp\left\{\phi_1 W_1(t) + \phi_2 W_2(t) - \frac{\sigma_\phi^2}{2} t\right\}$$

where:

$$\phi_1 = \frac{\sigma_1 \mu_2 \rho - \sigma_2 \mu_1}{\sigma_1 \sigma_2 (1 - \rho^2)}$$
$$\phi_2 = \frac{\sigma_2 \mu_1 \rho - \sigma_1 \mu_2}{\sigma_1 \sigma_2 (1 - \rho^2)}$$
$$\sigma_\phi^2 = \phi_1^2 + \phi_2^2 + 2\rho \phi_1 \phi_2$$

Under the risk-neutral measure Q the dynamics of the two stock price processes' satisfy the following SDEs:

$$dS_1(t) = S_1(t) \sigma_1 dW_1^Q(t)$$

$$dS_2(t) = S_2(t) \sigma_2 dW_2^Q(t)$$

where $W_1^Q = (W_1^Q(t) : t \in [0, T])$ and $W_2^Q = (W_2^Q(t) : t \in [0, T])$ are standard Brownian motion processes under measure Q with correlation coefficient ρ. These processes are defined in the following way:

$$W_1^Q(t) = W_1(t) + \theta_1 t$$

$$W_2^Q(t) = W_2(t) + \theta_2 t$$

where:

$$\theta_1 = \frac{\mu_1}{\sigma_1}$$
$$\theta_2 = \frac{\mu_2}{\sigma_2}$$

We can rewrite the process Z under measure Q as follows:

$$Z(t) = exp\left\{\phi_1 W_1^Q(t) + \phi_2 W_2^Q(t) - \Big(\frac{\sigma_\phi^2}{2} + \phi_1\theta_1 + \phi_2\theta_2\Big)t\right\}$$

Let us consider a European style contingent claim H with the following payoff:

$$H = (S_1(T) - S_2(T))^+$$

The initial capital $H(0)$, which is required for the perfect hedging of this contingent claim, is determined via the Margrabe formula (2.2):

$$H(0) = S_1(0)\Phi(d_1) - S_2(0)\Phi(d_2)$$

where:

$$d_1 = \frac{ln(\frac{S_1(0)}{S_2(0)}) + \frac{\sigma^2 T}{2}}{\sigma\sqrt{T}}$$
$$d_2 = d_1 - \sigma\sqrt{T}$$
$$\sigma = \sqrt{\sigma_1^2 + \sigma_2^2 - 2\sigma_1\sigma_2\rho}$$

Assuming that the initial capital available $\hat{V}(0)$ is less than $H(0)$, we will be minimizing $CVaR_a$ over all strategies with the initial capital less than or equal to $\hat{V}(0)$, i.e. $\pi = (V(0), \xi, \eta)$ such that $V(0) \le \hat{V}(0)$. Let us construct function $c(z)$

as in (2.31):

$$c(z) = z + \frac{1}{1-a}\left(\min_{\pi\in\mathscr{A}(\hat{V}(0))} E\Big(((S_1(T) - S_2(T))^+ - V(T) - z)^+\Big)\right)$$

Now, using proposition 20 we can rewrite function $c(z)$ in the following way:

$$c(z) = \begin{cases} z + \frac{1}{1-a}E((1-\widetilde{\varphi}(z))(H-z)^+), \; z < \hat{z} \\ z \qquad\qquad\qquad\qquad\qquad\qquad\quad , \; z \geq \hat{z} \end{cases} \tag{3.1}$$

where $\hat{z}$ is the solution to (2.39) and we will be solving for it using the proposed normal approximation as in (2.12) and the approximation put forward by Bjerksund and Stensland (2.8). Having determined the unique value $\hat{z}$, we can minimize $c(z)$ numerically using Monte Carlo simulation approach. Suppose that $\widetilde{z}$ is the point of global minimum of function $c(z)$, then proposition 19 tells us how to construct the $CVaR$-efficient portfolio as well as provides information related to the Greeks of the option and the optimal amount of initial capital required, since it is possible to find a strategy that is optimal but requires less capital than what is available. Noting that the distribution of Brownian motion processes is atomless, the problem is reduced to evaluating the following expectation:

$$E_Q(\widetilde{H}(\widetilde{z})) = E_Q(S_1(T) - S_2(T) - \widetilde{z})^+ I_{\{\frac{dP}{dQ} > \widetilde{a}\}} \tag{3.2}$$

Approximating the distribution of the difference between two lognormal random variables as using the normal distribution, the initial price p of replicating portfolio for a spread option can be estimated as follows:

$$p = S_1(0)\Phi^2(\hat{x}_1, \hat{y}_1, \rho_3) - S_2(0)\Phi^2(\hat{x}_2, \hat{y}_2, \rho_3) - \widetilde{z}\Phi^2(\hat{x}_3, \hat{y}_3, \rho_3) \tag{3.3}$$

where:

$$\hat{x}_1 = \frac{m - \widetilde{z}}{\sigma} + \sigma_1 \rho_1 \sqrt{T}$$
$$\hat{y}_1 = \widetilde{K} + \sigma_1 \rho_4 \sqrt{T}$$
$$\hat{x}_2 = \frac{m - \widetilde{z}}{\sigma} + \sigma_2 \rho_2 \sqrt{T}$$
$$\hat{y}_2 = \widetilde{K} + \sigma_2 \rho_5 \sqrt{T}$$
$$\hat{x}_3 = \frac{m - \widetilde{z}}{\sigma}$$
$$\hat{y}_3 = \widetilde{K}$$
$$\rho_1 = \frac{\sqrt{T}}{\sigma}\Big(S_1(0)\sigma_1 - S_2(0)\sigma_2\rho\Big)$$
$$\rho_2 = \frac{\sqrt{T}}{\sigma}\Big(S_1(0)\sigma_1\rho - S_2(0)\sigma_2\Big)$$
$$\rho_3 = \frac{a - b}{\sigma}$$
$$\rho_4 = -\frac{\rho\phi_2 + \phi_1}{\sigma_\phi}$$
$$\rho_5 = -\frac{\phi_2 + \phi_1\rho}{\sigma_\phi}$$
$$\sigma_\phi = \sqrt{\phi_1^2 + \phi_2^2 + 2\rho\phi_1\phi_2}$$
$$a = S_2(0)\frac{\sigma_2\sqrt{T}}{\sigma_\phi}\Big(\phi_1\rho + \phi_2\Big)$$
$$b = S_1(0)\frac{\sigma_1\sqrt{T}}{\sigma_\phi}(\phi_1 + \phi_2\rho)$$
$$\widetilde{K} = \frac{(\frac{\sigma_\phi^2}{2} + \phi_1\theta_1 + \phi_2\theta_2)T + ln(\frac{1}{\overline{a}})}{\sigma_\phi\sqrt{T}}$$

$$S_1(T) - S_2(T) \sim N(m, \sigma^2)$$

The derivation of the above formula is in Appendix 3. Alternatively, following the methodology proposed by Bjerksund and Stensland:

$$p = S_1(0)\Phi^2(\hat{x}_1, \hat{y}_1, \rho_3) - S_2(0)\Phi^2(\hat{x}_2, \hat{y}_2, \rho_3) - \widetilde{z}\Phi^2(\hat{x}_3, \hat{y}_3, \rho_3) \tag{3.4}$$

where:

$$\begin{aligned}
\hat{x}_1 &= \hat{K} + \sigma_1\rho_1\sqrt{T} \\
\hat{y}_1 &= \widetilde{K} + \sigma_1\rho_4\sqrt{T} \\
\hat{x}_2 &= \hat{K} + \sigma_2\rho_2\sqrt{T} \\
\hat{y}_2 &= \widetilde{K} + \sigma_2\rho_5\sqrt{T} \\
\hat{x}_3 &= \hat{K} \\
\hat{y}_3 &= \widetilde{K} \\
\rho_1 &= \frac{(\sigma_1 - \sigma_2 b\rho)\sqrt{T}}{\sqrt{\sigma_1^2 T - 2\sigma_1\sigma_2 b\rho T + \sigma_2^2 b^2 T}} \\
\rho_2 &= \frac{(\sigma_1\rho - \sigma_2 b)\sqrt{T}}{\sqrt{\sigma_1^2 T - 2\sigma_1\sigma_2 b\rho T + \sigma_2^2 b^2 T}} \\
\rho_3 &= \frac{(\sigma_2 b\phi_1\rho + \sigma_2 b\phi_2 - \sigma_1\phi_1 - \sigma_1\phi_2\rho)\sqrt{T}}{\sigma_\phi\sqrt{\sigma_1^2 T + \sigma_2^2 b^2 T - 2\sigma_1\sigma_2 b\rho T}} \\
\rho_4 &= -\frac{\phi_2\rho + \phi_1}{\sigma_\phi} \\
\rho_5 &= -\frac{\phi_2 + \phi_1\rho}{\sigma_\phi} \\
\sigma_\phi &= \sqrt{\phi_1^2 + \phi_2^2 + 2\rho\phi_1\phi_2} \\
b &= \frac{S_2(0)}{a}
\end{aligned}$$

$$a = S_2(0) + \widetilde{z}$$

$$\widetilde{K} = \frac{(\frac{\sigma_\phi^2}{2} + \phi_1\theta_1 + \phi_2\theta_2)T + ln(\frac{1}{\widetilde{a}})}{\sigma_\phi\sqrt{T}}$$

$$\hat{K} = \frac{ln(\frac{S_1(0)}{a}) - \frac{\sigma_1^2 T}{2} + \frac{\sigma_2^2 b^2 T}{2}}{\sqrt{\sigma_1^2 T + \sigma_2^2 b^2 T - 2\rho\sigma_1\sigma_2 bT}}$$

Refer to Appendix 4, where we show how to derive the above formula.

Chapter 4

Empirical Tests

This chapter is devoted to implementing the methodology developed in the previous chapter using both hypothetical and real data. For the first hypothetical portfolio example consider two stocks S_1 and S_2 with the volatility of the first stock being $\sigma_1 = 0.3$ and the volatility of the second asset being $\sigma_2 = 0.1$. The first stock is perceived to be riskier than the second one and hence market participants are rewarded by higher rate of return $\mu_1 = 0.1$, the rate of return on the latter stock is $\mu_2 = 0.05$. The two stocks are assumed to be positively correlated with $\rho = 0.5$ and the initial prices of the two stocks are standardized to be equal to 100. A financial institution has sold an option to exchange the second asset for the first one with maturity of $T = 5$. Using the Margrabe formula we estimate the fair price of this option to be $p = 23.26$. However, suppose that the issuing institution decides not to invest all the proceeds from the sale of the option into a replicating portfolio. What are the associated $CVaRs$ for various levels of initial capital available as percentage of p at $90\%, 95\%$ and 99% confidence levels? We have estimated the $CVaRs$ using the normal approximation and the approximation proposed by Bjerksund and Stensland. Refer to figures 4.1 - 4.3 where we plot $CVaRs$ against the initial capital available for hedging purposes.

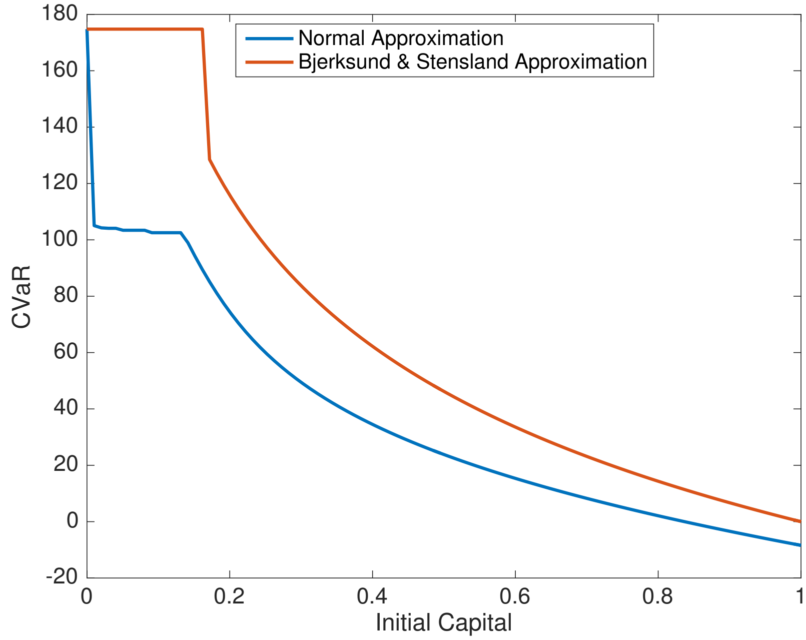

Figure 4.1: Hypothetical Portfolio 1: CVaR at 90%

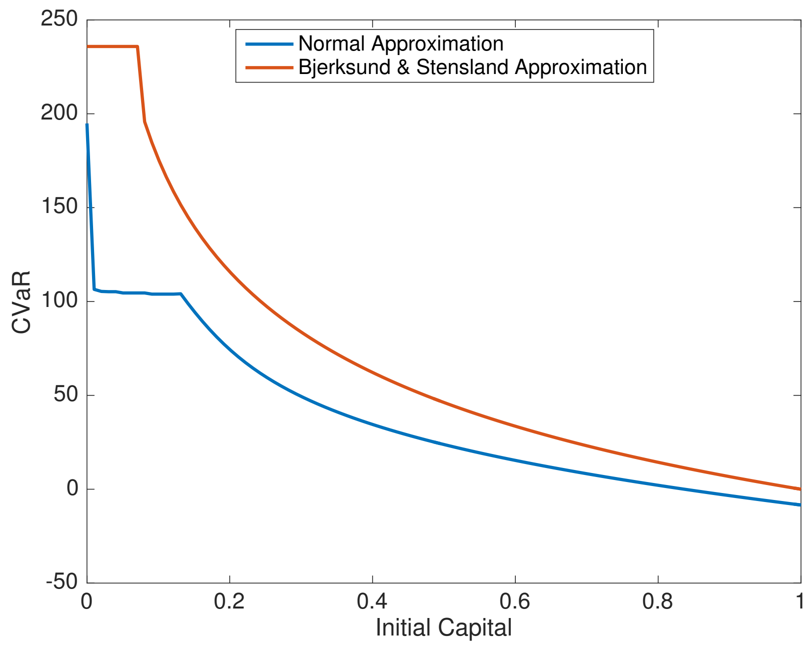

Figure 4.2: Hypothetical Portfolio 1: CVaR at 95%

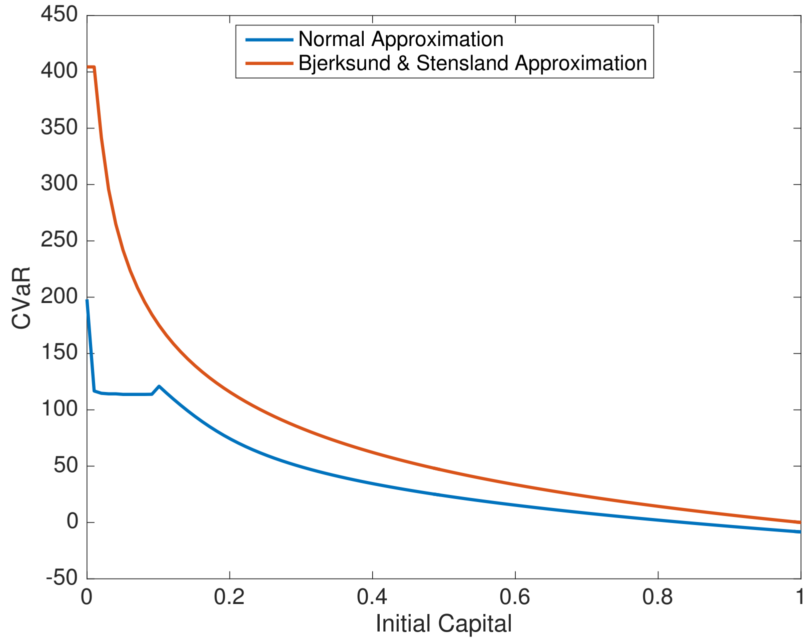

Figure 4.3: Hypothetical Portfolio 1: CVaR at 99%

We can infer from the plots that there is a big divergence between $CVaRs$ estimated using the proposed normal approximation as compared to the approximation of Bjerksund and Stensland, which we take as an accurate estiamte of $CVaR$. We attribute this drastic difference to the market parameters that we chose for this hypothetical portfolio: the first stock is perceived to be riskier by a factor of 3 and the maturity of the contract spans too far into the future. We also observe that the normal approximation that we have proposed underestimates $CVaRs$ at all the confidence levels and all the levels of initial capital, which is in line with the results represented in table 2.1 for the choice of market parameters. We can also see that there are some ranges over which the graph of $CVaR$ is horizontal. The natural interpretation would be that increasing (reducing) the amount of capital used for hedging over this range does not reduce (increase) the overall exposure. However, this could also be simply attributed to the estimation error: "the trade-off between estimation accuracy and computational efficiency is well known for Monte-Carlo simulation. When the underlying assets have high volatilities, stochastic volatilities, or time-dependent correlations or when the option has a long maturity or high dimensionality, precision deterioration becomes a major concern unless computation time is increased exponentially" (Li and Wu, 2006).

For the second hypothetical portfolio we consider a similar market but this time we compare stocks which are more 'equivalent' in terms of risk-reward characteristics. Thus, assume the following parameters for this market: $\sigma_1 = 0.12, \sigma_2 = 0.1, \mu_1 = 0.055$ and $\mu_2 = 0.05$. A financial institution has sold an option o exchange one asset for another to a client for the amount $p = 4.44$ with maturity of $T = 1$. Similarly, we estimate $CVaRs$ for varying levels of initial capital used for hedging and for the three different confidence levels. Figures 4.7 - 4.9 plot the results.

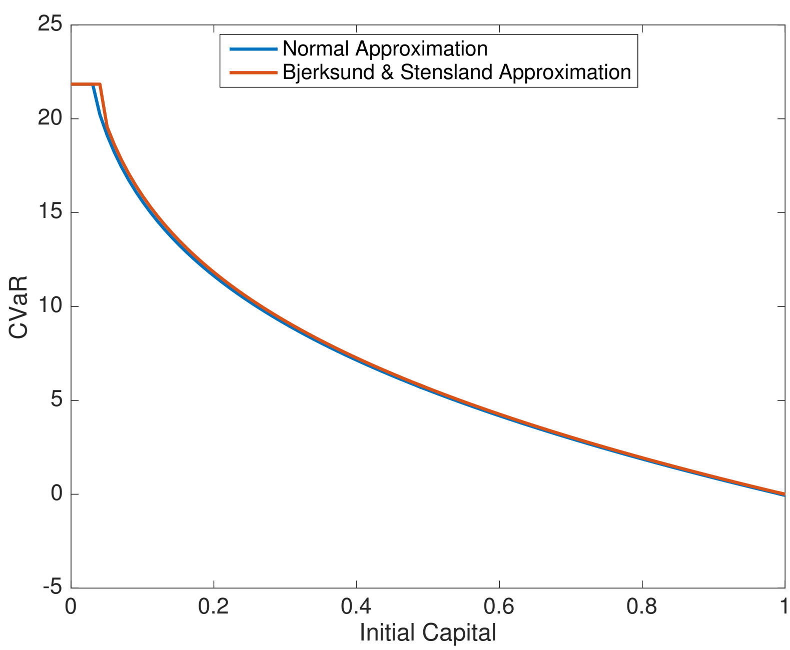

Figure 4.4: Hypothetical Portfolio 2: CVaR at 90%

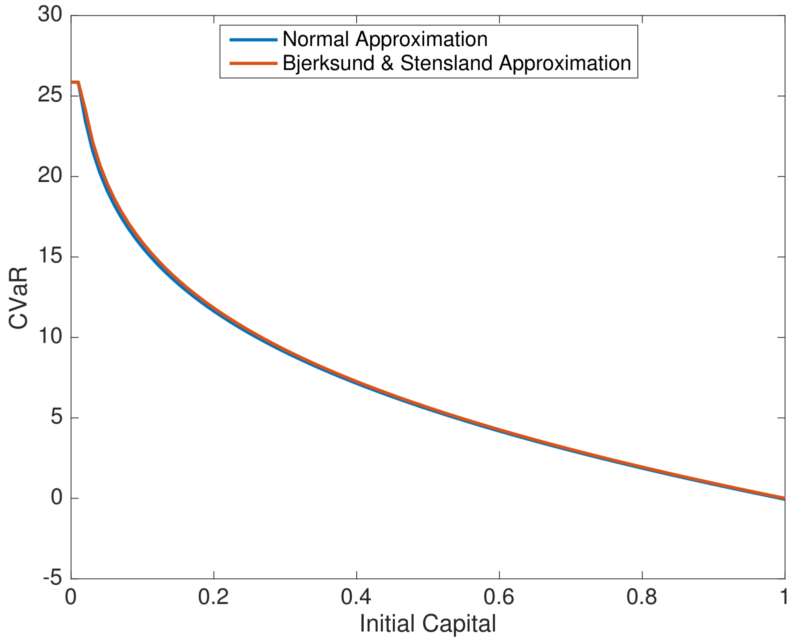

Figure 4.5: Hypothetical Portfolio 2: CVaR at 95%

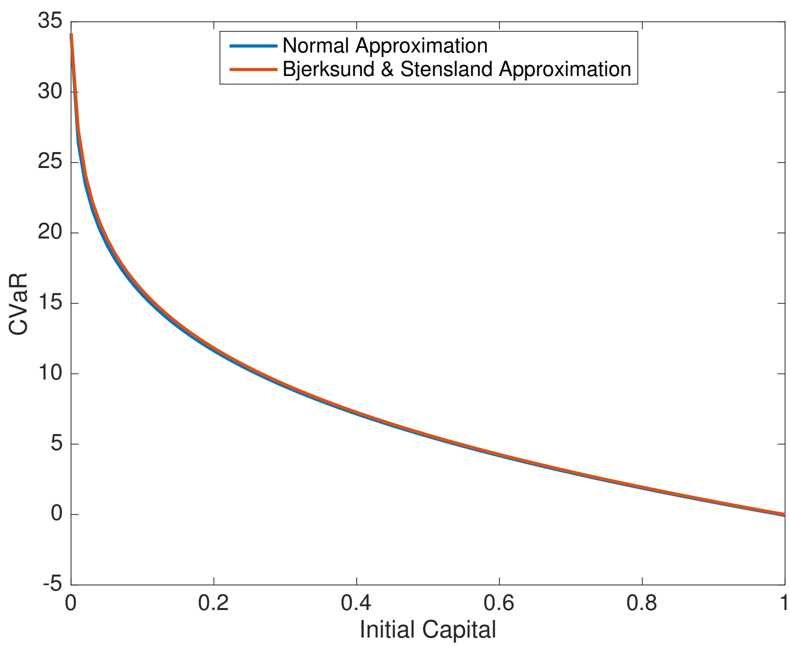

Figure 4.6: Hypothetical Portfolio 2: CVaR at 99%

As can be inferred from the plots the $CVaR$ levels under the two methods are almost identical given the market parameters we have chosen. Again, the normal approximation method does underestimate risk for all the cases, however this underestimation is of almost negligible magnitude. It also becomes evident that the estimation error of MC simulation is almost eliminated completely, which is attributed to lower standard deviations and lower maturity for the option under consideration. Finally, $CVaRs$ are inversely related to the amount of initial capital available for hedging, which is what one would expect.

Finally, to see how the methodology would apply to the real market data, we have downloaded the closing price data for Apple Inc. and SP500 index from 1st January 2013 to 28th March 2018, overall 1319 observations. Having transformed the prices to logarithmic returns and annualized the returns, we obtained the following market parameters: $\sigma_1 = 0.24, \sigma_2 = 0.12$, where subscript 1 refers to Apple Inc. and subscript 2 to S&P 500 index. The annualized returns are: $\mu_1 = 0.14, \mu_2 = 0.11$. The estimated correlation coefficient over the period was $\rho = 0.5068$. The initial prices as of 1st January 2013 are:

$$S_1(0) = 78.4329 \qquad\qquad S_2(0) = 1462.42 * \frac{S_1(0)}{S_2(0)}$$

We have standardized the initial prices to be equal so as to make the spread option feasible. The institution has sold an option to exchange the SP500 for the stock of Apple Inc. with the expiration date 1 year from now. The price that is required for complete hedging is determined via the Margrabe formula to be equal to $p = 6.49$. As usual we estimate $CVaRs$ that the issuing house faces following the sale of this option using both methodologies and for the three confidence levels.

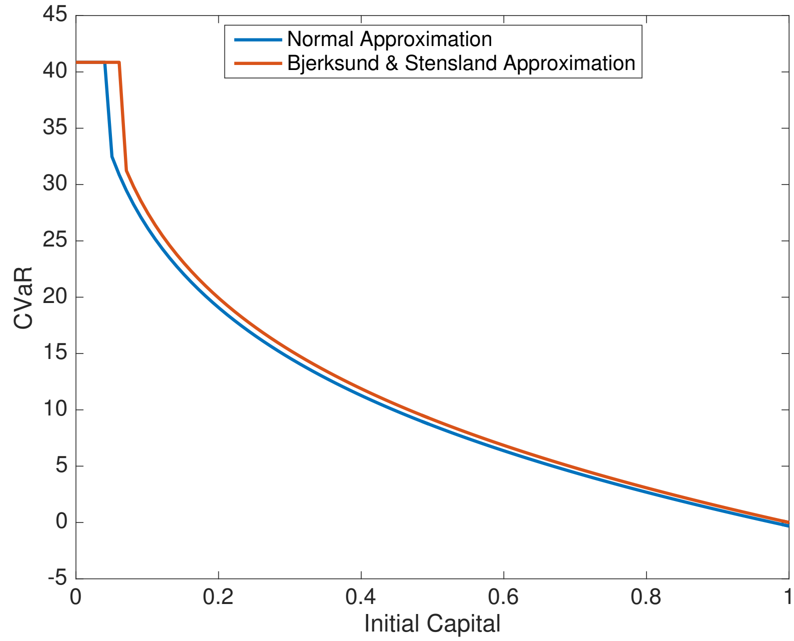

Figure 4.7: Apple Inc. and SP500 Portfolio: CVaR at 90%

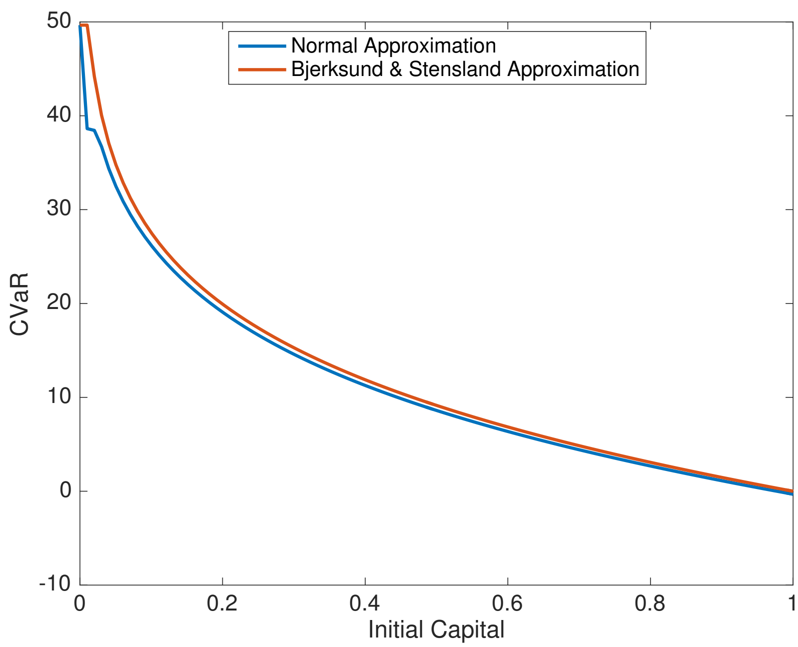

Figure 4.8: Apple Inc. and SP500 Portfolio: CVaR at 95%

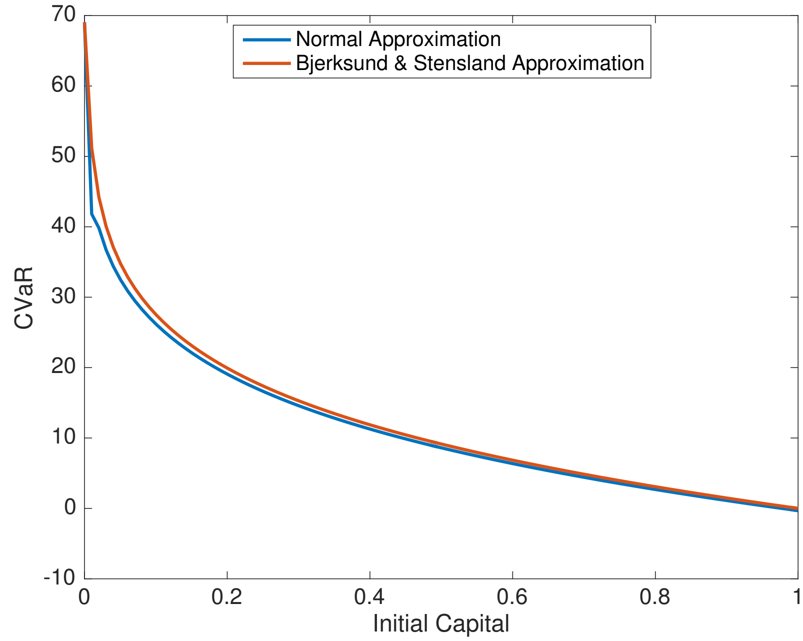

Figure 4.9: Apple Inc. and SP500 Portfolio: CVaR at 99%

The results plotted are in line with the conclusions we have drawn from the examples of the two hypothetical portfolios. To further investigate the risk management aspect of this portfolio, refer to tables 4.1 - 4.3.

	Normal Approximation		Bjerksund & Stensland	
Capital Available	VaR	CVaR	VaR	CVaR
0%	28.0807	40.9760	27.9631	40.8337
10%	26.0966	26.0966	27.4378	27.4378
20%	18.9853	18.9853	19.8209	19.8209
30%	14.4674	14.4674	15.1719	15.1719
40%	11.1456	11.1456	11.7549	11.7549
51%	8.4653	8.4653	9.0196	9.0196
61%	6.2721	6.2721	6.7189	6.7189
71%	4.3021	4.3021	4.7194	4.7194
81%	2.5291	2.5291	2.9413	2.9413
91%	0.9631	0.9631	1.3326	1.3326
100%	-0.3239	-0.3239	0.0000	0.0000

Table 4.1: Apple Inc. and S&P500 Portfolio: VaR and CVaR at 90%

	Normal Approximation		Bjerksund & Stensland	
Capital Available	VaR	CVaR	VaR	CVaR
0%	37.2960	49.6749	37.2442	49.5760
10%	26.1079	26.1079	27.4378	27.4378
20%	18.9376	18.9376	19.8209	19.8209
30%	14.4738	14.4738	15.1719	15.1719
40%	11.1210	11.1210	11.7549	11.7549
51%	8.4614	8.4614	9.0196	9.0196
61%	6.2273	6.2273	6.7189	6.7189
71%	4.2790	4.2790	4.7194	4.7194
81%	2.5496	2.5496	2.9413	2.9413
91%	0.9693	0.9693	1.3326	1.3326
100%	-0.3206	-0.3206	0.0000	0.0000

Table 4.2: Apple Inc. and S&P500 Portfolio: VaR and CVaR at 95%

	Normal Approximation		Bjerksund & Stensland	
Capital Available	VaR	CVaR	VaR	CVaR
0%	57.2583	68.9700	57.3636	69.0788
10%	26.0501	26.0501	27.4378	27.4378
20%	18.9578	18.9578	19.8209	19.8209
30%	14.4591	14.4591	15.1719	15.1719
40%	11.1611	11.1611	11.7549	11.7549
51%	8.4916	8.4916	9.0196	9.0196
61%	6.2363	6.2363	6.7189	6.7189
71%	4.2763	4.2763	4.7194	4.7194
81%	2.5743	2.5743	2.9413	2.9413
91%	0.9873	0.9873	1.3326	1.3326
100%	-0.3293	-0.3293	0.0000	0.0000

Table 4.3: Apple Inc. and S&P500 Portfolio: VaR and CVaR at 99%

The tables provide the estimated VaR and $CVaR$ values for the portfolio at the three different confidence levels. First of all, we observe that for most of the levels of initial capital available $VaRs$ and $CVaRs$ are equal. This is possible according to proposition 20 when $z \geq \hat{z}$. We naturally interpret this results as follows: since $CVaR$ is the average of losses in excess of VaR, it must be that the probability of those losses is low enough to make the mean of their expected values equal to the cutoff point beyond which they are estimated. As previously, we believe that the precision of these estimates could be improved by increasing the number of simulations in MC procedure. Secondly, since the normal approximation underestimates the price of a spread option, it naturally leads to lower estimates for the risks associated with this option. In our particular case this resulted in $VaRs$ and $CVaRs$ taking negative values, which is, of course, not possible as using 100% of capital required for perfect hedging should lead to no gains and no losses at maturity. Other than that, the two methods provide comparable results.

We can further supplement our analysis by looking at $CVaR$-efficient portfolios from a regulatory point of view. Suppose that a regulator in the market requires the member institutions to keep a certain amount of capital in reserves depending on the estimated level of $CVaR$. Let β be the amount of capital required per unit of $CVaR$ exposure. Denote by $\lambda_a(\hat{V}(0)) = \beta CVaR_a\Big(\hat{V}(0)\Big) + \hat{V}(0)$ the total amount of capital to be kept in reserves provided that an amount of $\hat{V}(0)$ has been used for hedging purposes at a significance level a. Then, the $CVaR$ of an unhedged position is $\lambda_a(0)$. Introduce the following ratio:

$$\Theta_a = \frac{\lambda_a(\hat{V}(0))}{\lambda_a(0)}$$

The ratio tells us the relative attractiveness of a $CVaR$-efficient portfolio. In the case where $\Theta_a < 1$ shows that engaging in $CVaR$-efficient hedging allows the institution to use less capital to meet the regulatory requirement as compared to an unhedged position and vice versa. We apply this line of analysis to our Apple Inc. and SP500 portfolio at 95% significance level and the results are shown in figure 4.10.

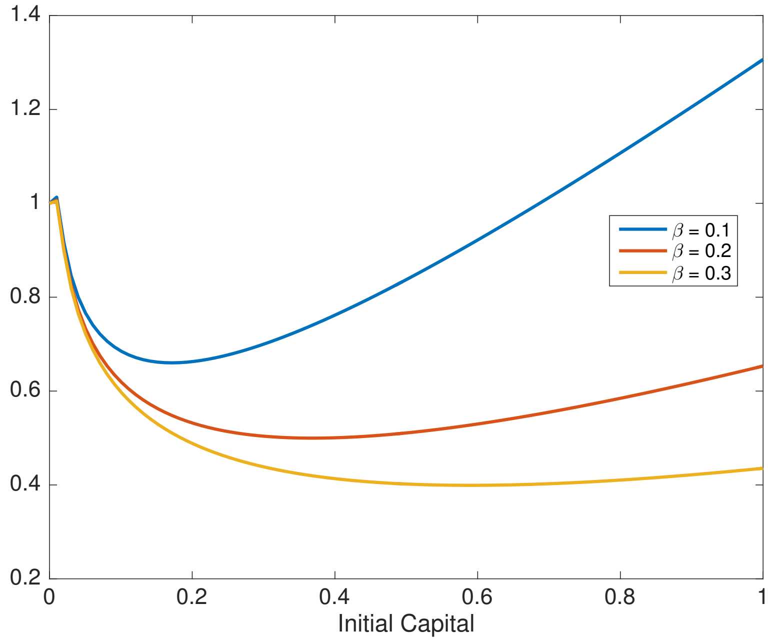

Figure 4.10: Relative Attractiveness of $CVaR$-efficient Portfolio at 95% confidence level

The above figure clearly indicates that the higher the regulatory requirements, the more attractive a $CVaR$-efficient portfolio is compared to a portfolio with no hedging.

Chapter 5

Recommendations

In drawing the conclusions about the comparative effectiveness of the two methods for minimizing $CVaR$ under capital constraints, we have assumed that the approximation of Bjerksund and Stensland is precise. In reality, this is not the case. The pricing formula developed by Bjerksund and Stensland provides a vey close lower bound to the true price of a spread option and thus, in theory, also underestimates the true risks associated with a given contingent claim. To be able to draw the ultimate conclusions about the effectiveness of the two methods, there should be an analytical pricing formula for a spread option or another approach to measuring $CVaR$ associated with a given portfolio. Both developments are currently perceived as incomplete. To be able to give a definite price for a spread option one needs to know the exact distribution of the difference between two lognormal random variables, a long standing problem with no definite answer. One possible way to improve the results in this direction is to use the results of Lo discussed in section 2.2, where we adopt a Lie-Trotter splitting method and work with a shifted lognormal process. With regards to the second development, a possible solution is to use the results of comparison theorem for stochastic processes investigated by Krasin et al. (2017).

The central idea is to find a process with a known distribution that will dominate over the process of interest, the process of difference between two lognormally distributed stock prices in our case. It would then be possible to find the $CVaR$ of a portfolio whose dynamics is governed by this dominating process and in this way determine the $CVaR$ of the portfolio of interest. Generally, "despite the advantages of ES, this measure is less frequently utilized than VaR because forecasting ES is challenging due to its complex mathematical definition" (Brutti Righi and Ceretta, 2016), where ES is used as an alternative name for $CVaR$.

Most of the developments in the financial theory are done within the Samuelson's model of the market, where stock price processes are modelled as geometric Brownian motions. A possible extension is to consider spread option pricing as well as minimization of risk functions within the context of Bachelier model of the market, where arithmetic Brownian motions describe the dynamic of the stock price processes. The reason why Bachelier model is rejected by many is that it allows for the stock prices to take on negative values, which in reality is impossible and is associated with bankruptcy of a given enterprise. The solution would be to introduce stopping times, random variables which are used to represent the default times of a given company. In this way we account for the problem of stock prices taking negative values and the transfer of the results of this thesis becomes a viable option.

Chapter 6

Summary and Conclusions

This thesis dealt with the problem of minimizing $CVaR$ under capital constraints within the Margrabe market. We have developed a methodology based on two spread option pricing methods: the normal approximation to the difference between two lognormal random variables and the pricing formula for spread options proposed by Bjerksund and Stensland. Both methods provide almost equivalent results provided the parameters of the market model under question satisfy certain constraints. Otherwise, there is a divergence in results and the normal approximation tends to significantly underestimate the risk exposure of a portfolio. The conclusions of the thesis might carry a lot of benefit to the practitioners of risk management in light or recent developments in regulations that call for more stringent requirements for managing exposure to risk. While we feel that the problem of the thesis has been addressed, there are still many improvements that can be done in this direction.

Bibliography

Acerbi, C. (2002). Spectral measures of risk: A coherent representation of subjective risk aversion. Journal of Banking Finance, 26(7), pp.1505-1518.

Artzner, P., Delbaen, F., Eber, J. and Heath, D. (1999). Coherent Measures of Risk. Mathematical Finance, 9(3), pp.203-228.

Bank for International Settlements (2017). Statistical release: OTC derivatives statistics at endJune 2017. BIS semiannual OTC derivatives statistics. [online]

Basel Committee on Banking Supervision (2016). Minimum capital requirements for market risk. [online]

Berg, C. (1988). The Cube of a Normal Distribution is Indeterminate. The Annals of Probability, 16(2), pp.910-913.

Bjerksund, P. and Stensland, G. (2006). Closed Form Spread Option Valuation. SSRN Electronic Journal.

Black, F. and Scholes, M. (1973). The Pricing of Options and Corporate Liabilities. Journal of Political Economy, [online] 81(3), pp.637-654.

Boyle, P. (1988). A Lattice Framework for Option Pricing with Two State Variables. The Journal of Financial and Quantitative Analysis, 23(1), p.1.

Browne, S. (1999). Reaching goals by a deadline: digital options and continuous-time active portfolio management. Advances in Applied Probability, 31(02), pp.551-577.

Brutti Righi, M. and Ceretta, P. (2016). Shortfall deviation risk: an alternative for risk measurement. The Journal of Risk, 19(2), pp.81-116.

Carmona, R. and Durrleman, V. (2003). Pricing and Hedging Spread Options. SIAM Review, 45(4), pp.627-685.

Cobb, B. and Rumi, R. (2012). Approximating the Distribution of a Sum of Log-normal Random Variables. In: Sixth European Workshop on Probabilistic Graphical Models. [online]

Cox, J., Ross, S. and Rubinstein, M. (1979). Option pricing: A simplified approach. Journal of Financial Economics, 7(3), pp.229-263.

Cvitanic, J. and Karatzas, I. (2001). Generalized Neyman-Pearson Lemma via Convex Duality. Bernoulli, 7(1), p.79.

El Karoui, N. and Quenez, M. (1995). Dynamic Programming and Pricing of

Contingent Claims in an Incomplete Market. SIAM Journal on Control and Optimization, 33(1), pp.29-66.

Embrechts, P. and Wang, R. (2015). Seven Proofs for the Subadditivity of Expected Shortfall. Dependence Modeling, 3(1).

Fama, E. (1965). The Behavior of Stock-Market Prices. The Journal of Business, 38(1), p.34.

Fischer, S. (1978). Call Option Pricing When the Exercise Price is Uncertain, and the Valuation of Index Bonds. The Journal of Finance, 33(1), p.169.

Foellmer, H. and Leukert, P. (1999). Quantile hedging. Finance and Stochastics, 3(3), pp.251-273.

Foellmer, H. and Leukert, P. (2000). Efficient hedging: Cost versus shortfall risk. Finance and Stochastics, 4(2), pp.117-146.

Fringuellotti, F. and Necula, C. (2015). A Generalized Bachelier Formula for Pricing Basket and Spread Options. SSRN Electronic Journal.

Hcine, M. and Bouallegue, R. (2015). On The Approximation Of The Sum Of Lognormals By A Log Skew Normal Distribution. International journal of Computer Networks Communications, 7(1), pp.135-151.

Hebner, M. (2014). QA with IFA: Are Monthly Stock Market Returns Normally

Distributed?. [online] Ifa.com.

J.P.Morgan/Reuters (1996). RiskMetrics. [online] New York.

Jones, M. and Pewsey, A. (2009). Sinh-arcsinh distributions. Biometrika, 96(4), pp.761-780.

Jorion, P. (2007). Value at risk. [New York]: McGraw Hill.

Kirk, E. (1995): Correlation in the Energy Markets, in Managing Energy Price Risk. London: Risk Publications and Enron.

Krasin, V., Smirnov, I. and Melnikov, A. (2017). Approximate option pricing and hedging in the CEV model via path-wise comparison of stochastic processes. Annals of Finance, 14(2), pp.195-209.

Kulldorff, M. (1993). Optimal Control of Favorable Games with a Time Limit. SIAM Journal on Control and Optimization, 31(1), pp.52-69.

LeBaron, B. (2008). Robust Properties of Stock Return Tails. [online]

Lehmann E.L. (1992) "Introduction to Neyman and Pearson (1933) On the Problem of the Most Efficient Tests of Statistical Hypotheses". In: Breakthroughs in Statistics, Volume 1, (Eds Kotz, S., Johnson, N.L.), Springer-Verlag. ISBN 0-387-94037-5

Li, X. and Wu, Z. (2005). A semi-analytic method for valuing high-dimensional options on the maximum and minimum of multiple assets. Annals of Finance, 2(2), pp.179-205.

Lo, C. (2012). The Sum and Difference of Two Lognormal Random Variables. SSRN Electronic Journal.

Mandelbrot, B. (1963). The Variation of Certain Speculative Prices. The Journal of Business, 36(4), p.394.

Macroption.com. (2018). Value At Risk (VAR) Limitations and Disadvantages - Macroption. [online]

Margrabe, W. (1978). The Value of an Option to Exchange One Asset for Another. The Journal of Finance, 33(1), p.177.

Mehta, N., Wu, J., Molisch, A. and Zhang, J. (2007). Approximating a Sum of Random Variables with a Lognormal. IEEE Transactions on Wireless Communications, 6(7), pp.2690-2699.

Melnikov, A. (2004). Risk analysis in finance and insurance. Boca Raton, Fla.: Chapman Hall/CRC.

Melnikov, A. and Romanyuk, Y. (2008). Efficient Hedging and Pricing of Equity-Linked Life Insurance Contracts on Several Risky Assets.. International Journal of Theoretical and Applied Finance, 11(03), pp.295-323.

Melnikov, A. and Smirnov, I. (2012). Dynamic hedging of conditional value-at-risk. Insurance: Mathematics and Economics, 51(1), pp.182-190.

Nakano, Y. (2004). Efficient hedging with coherent risk measure. Journal of Mathematical Analysis and Applications, 293(1), pp.345-354.

Neyman, J. and Pearson, E. (1933). On the Problem of the Most Efficient Tests of Statistical Hypotheses. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 231(694-706), pp.289-337.

Poulsen, R. (2010). The Margrabe Formula. [online] Centre for Finance, University of Gothenburg.

Rockafellar, R. and Uryasev, S. (2002). Conditional value-at-risk for general loss distributions. Journal of Banking Finance, 26(7), pp.1443-1471.

Rockafellar, T. (1970). Convex Analysis.

Rook, C. and Kerman, M. (2015). Approximating the Sum of Correlated Lognormals: An Implementation. SSRN Electronic Journal.

Spivak, G. and Cvitanic, J. (1999). Maximizing the probability of a perfect hedge. The Annals of Applied Probability, 9(4), pp.1303-1328.

Trotter, H. (1959). On the Product of Semi-Groups of Operators. Proceedings of the American Mathematical Society, 10(4), p.545.

Appendix 1

We are interested in the following expectation:

$$E_Q\left(\left(S_1(T)-\frac{aS_2^b(T)}{E_Q(S_2^b(T))}\right)^+\right) = E_Q\left(\left(S_1(T)-\frac{aS_2^b(T)}{E_Q(S_2^b(T))}\right)I_{\left\{S_1(T)\geq\frac{aS_2^b(T)}{E_Q(S_2^b(T))}\right\}}\right)$$

The term in the denominator is:

$$\begin{aligned}
E_Q(S_2^b(T)) &= S_2^b(0)exp\left\{-\frac{\sigma_2^2 bT}{2}\right\}E_Q\Big(exp\left\{\sigma_2 bW_2^Q(T)\right\}\Big)\\
&= S_2^b(0)exp\left\{\frac{\sigma_2^2 b(b-1)T}{2}\right\}
\end{aligned}$$

Let us now simplify the term in the indicator function:

$$\begin{aligned}
S_1(T) &\geq \frac{aS_2^b(T)}{S_2^b(0)exp\left\{\frac{\sigma_2^2 b(b-1)T}{2}\right\}}\\
S_1(0)exp\left\{-\frac{\sigma_1^2 T}{2}+\sigma_1 W_1^Q(T)\right\} &\geq a\ exp\left\{-\frac{\sigma_2^2 b^2 T}{2}+\sigma_2 bW_2^Q(T)\right\}\\
ln\Big(\frac{S_1(0)}{a}\Big)-\frac{\sigma_1^2 T}{2}+\frac{\sigma_2^2 b^2 T}{2} &\geq \sigma_2 bW_2^Q(T)-\sigma_1 W_1^Q(T)
\end{aligned}$$

Since:

$$\sigma_2 bW_2^Q(T)-\sigma_1 W_1^Q(T)\sim N(0,\sigma_1^2 T-2\sigma_1\sigma_2 b\rho T+\sigma_2^2 b^2 T)$$

the inequality above is equivalent to:

$$\epsilon \leq \widetilde{K}$$

where:

$$\epsilon = \frac{\sigma_2 b W_2^Q(T) - \sigma_1 W_1^Q(T)}{\sqrt{\sigma_1^2 T - 2\sigma_1\sigma_2 b\rho T + \sigma_2^2 b^2 T}} \sim N(0,1)$$

$$\widetilde{K} = \frac{ln\left(\frac{S_1(0)}{a}\right) - \frac{\sigma_1^2 T}{2} + \frac{\sigma_2^2 b^2 T}{2}}{\sqrt{\sigma_1^2 T - 2\sigma_1\sigma_2 b\rho T + \sigma_2^2 b^2 T}}$$

Consider the first term in the original expectation:

$$E_Q\left(S_1(T) I_{\left\{S_1(T) \geq \frac{aS_2^b(T)}{E_Q(S_2^b(T))}\right\}}\right)$$

$$= S_1(0) exp\left\{-\frac{\sigma_1^2 T}{2}\right\} E_Q\left(exp\left\{\sigma_1 W_1^Q(T)\right\} I_{\left\{\epsilon \leq \widetilde{K}\right\}}\right)$$

Applying the two-asset lemma (2.1) to the expectation term:

$$E_Q\left(exp\left\{\sigma_1 W_1^Q(T)\right\} I_{\left\{\epsilon \leq \widetilde{K}\right\}}\right) = exp\left\{\frac{\sigma_1^2 T}{2}\right\} \Phi\left(\widetilde{K} + \frac{(\sigma_1 - \sigma_2 b\rho)\sqrt{T}}{\sqrt{\sigma_1^2 T - 2\sigma_1\sigma_2 b\rho T + \sigma_2^2 b^2 T}}\right)$$

Which leads to:

$$E_Q\left(S_1(T) I_{\left\{S_1(T) \geq \frac{aS_2^b(T)}{E_Q(S_2^b(T))}\right\}}\right) = S_1(0)\Phi(d1)$$

where:

$$d_1 = \widetilde{K} + \sigma_1 \rho_1 \sqrt{T}$$

$$\rho_1 = \frac{(\sigma_1 - \sigma_2 b\rho)\sqrt{T}}{\sqrt{\sigma_1^2 T - 2\sigma_1\sigma_2 b\rho T + \sigma_2^2 b^2 T}}$$

Let us now consider the second term in the original expectation:

$$\begin{aligned} &E_Q\Big(\frac{aS_2^b(T)}{E_Q(S_2^b(T))} I_{\left\{S_1(T) \geq \frac{aS_2^b(T)}{E_Q(S_2^b(T))}\right\}}\Big) \\ &\quad = a\ exp\left\{-\frac{\sigma_2^2 b^2 T}{2}\right\} E_Q\Big(exp\left\{\sigma_2 b W_2^Q(T)\right\} I_{\left\{\epsilon \leq \widetilde{K}\right\}}\Big) \end{aligned}$$

Similarly, applying the two-asset lemma to the term under expectation operator:

$$E_Q\Big(\frac{aS_2^b(T)}{E_Q(S_2^b(T))} I_{\left\{S_1(T) \geq \frac{aS_2^b(T)}{E_Q(S_2^b(T))}\right\}}\Big) = a\Phi(d_2)$$

where:

$$\begin{aligned} d_2 &= \widetilde{K} + \sigma_2 b \rho_2 \sqrt{T} \\ \rho_2 &= \frac{(\sigma_1 \rho - \sigma_2 b)\sqrt{T}}{\sqrt{\sigma_1^2 T - 2\rho\sigma_1\sigma_2 T b + \sigma_2^2 b^2 T}} \end{aligned}$$

Finally:

$$E_Q\left(\Big(S_1(T) - \frac{aS_2^b(T)}{E_Q(S_2^b(T))}\Big)^+\right) = S_1(0)\Phi(d_1) - a\Phi(d_2)$$

which is exactly the formula of Kirk (2.6).

Appendix 2

Consider the following expression:

$$E_Q\Big((S_1(T) - S_2(T) - K)I_{\left\{S_1(T)\geq \frac{aS_2^b(T)}{E_Q(S_2^b(T))}\right\}}\Big)$$
$$= S_1(T)I_{\left\{S_1(T)\geq \frac{aS_2^b(T)}{E_Q(S_2^b(T))}\right\}} - S_2(T)I_{\left\{S_1(T)\geq \frac{aS_2^b(T)}{E_Q(S_2^b(T))}\right\}} - KI_{\left\{S_1(T)\geq \frac{aS_2^b(T)}{E_Q(S_2^b(T))}\right\}}$$

The first term is exactly the same as in Appendix 1. Let us consider the second term:

$$E_Q\Big(S_2(T)I_{\left\{S_1(T)\geq \frac{aS_2^b(T)}{E_Q(S_2^b(T))}\right\}}\Big)$$
$$= S_2(0)exp\left\{-\frac{\sigma_2^2 T}{2}\right\} E_Q\Big(exp\left\{\sigma_2 W_2^Q(T)\right\} I_{\left\{\epsilon\leq\widetilde{K}\right\}}\Big)$$

Applying the two-asset lemma:

$$E_Q\Big(S_2(T)I_{\left\{S_1(T)\geq \frac{aS_2^b(T)}{E_Q(S_2^b(T))}\right\}}\Big) = S_2(0)\Phi(d_2)$$

where:

$$d_2 = \widetilde{K} + \sigma_2\rho_2\sqrt{T}$$
$$\rho_2 = \frac{(\sigma_1\rho - \sigma_2 b)\sqrt{T}}{\sqrt{\sigma_1^2 T - 2\sigma_1\sigma_2 b\rho T + \sigma_2^2 b^2 T}}$$

Finally, the third term is simply:

$$KI_{\left\{S_1(T) \geq \frac{aS_2^b(T)}{E_Q(S_2^b(T))}\right\}} = K\Phi(\widetilde{K})$$

Combining the three terms together we get the approximation of Bjerksund and Stensland (2.8).

Appendix 3

$$E_Q(S_1(T) - S_2(T) - K)^+ I_{\left\{\frac{dP}{dQ} > \widetilde{a}\right\}}$$

Let us first consider the indicator function $I_{\left\{\frac{dP}{dQ} > \widetilde{a}\right\}}$:

$$\begin{aligned}
\left\{\frac{dP}{dQ} > \widetilde{a}\right\} &= \left\{\frac{1}{\widetilde{a}} > exp\left\{\phi_1 W_1^Q(T) + \phi_2 W_2^Q(T) - \Big(\frac{\sigma_\phi^2}{2} + \phi_1\theta_1 + \phi_2\theta_2\Big)T\right\}\right\} \\
&= \left\{\frac{\phi_1 W_1^Q(T) + \phi_2 W_2^Q(T)}{\sigma_\phi\sqrt{T}} < \frac{(\frac{\sigma_\phi^2}{2} + \phi_1\theta_1 + \phi_2\theta_2)T + ln(\frac{1}{\widetilde{a}})}{\sigma_\phi\sqrt{T}}\right\} \\
&= \left\{\epsilon < \widetilde{K}\right\}
\end{aligned}$$

where:

$$\begin{aligned}
\epsilon &= \frac{\widetilde{W}^Q(T)}{\sqrt{T}} \sim N(0,1) \\
\widetilde{W}^Q(T) &= \frac{\phi_1 W_1^Q(T) + \phi_2 W_2^Q(T)}{\sigma_\phi} \\
\widetilde{K} &= \frac{(\frac{\sigma_\phi^2}{2} + \phi_1\theta_1 + \phi_2\theta_2)T + ln(\frac{1}{\widetilde{a}})}{\sigma_\phi\sqrt{T}} \\
\phi_1 &= \frac{\sigma_1\mu_2\rho - \sigma_2\mu_1}{\sigma_1\sigma_2(1-\rho^2)} \\
\phi_2 &= \frac{\sigma_2\mu_1\rho - \sigma_1\mu_2}{\sigma_1\sigma_2(1-\rho^2)}
\end{aligned}$$

$$\sigma_\phi = \sqrt{\phi_1^2 + \phi_2^2 + 2\phi_1\phi_2\rho}$$
$$\theta_1 = \frac{\mu_1}{\sigma_1}$$
$$\theta_2 = \frac{\mu_2}{\sigma_2}$$

Now denote $S_1(T) - S_2(T)$ by $\gamma \sim N(m, \sigma^2)$. Then:

$$\begin{aligned} E_Q(S_1(T) - S_2(T) - K)^+ I_{\left\{\frac{dP}{dQ} > \tilde{a}\right\}} &= E_Q(S_1(T) I_{\{\gamma > K\}} I_{\left\{\epsilon < \widetilde{K}\right\}}) - \\ & E_Q(S_2(T) I_{\{\gamma > K\}} I_{\left\{\epsilon < \widetilde{K}\right\}}) - K E_Q(I_{\{\gamma > K\}} I_{\left\{\epsilon < \widetilde{K}\right\}}) \\ = S_1(0) exp\left\{-\frac{\sigma_1^2 T}{2}\right\} & E_Q\Big(exp\left\{\sigma_1 W_1^Q(T)\right\} I_{\{-\gamma < -K\}} I_{\left\{\epsilon < \widetilde{K}\right\}}\Big) - \\ S_2(0) exp\left\{-\frac{\sigma_2^2 T}{2}\right\} & E_Q\Big(exp\left\{\sigma_2 W_2^Q(T)\right\} I_{\{-\gamma < -K\}} I_{\left\{\epsilon < \widetilde{K}\right\}}\Big) - \\ & K E_Q(I_{\{-\gamma < -K\}} I_{\left\{\epsilon < \widetilde{K}\right\}}) \end{aligned}$$

Let us consider each of the terms above separately.

I.

$$\begin{aligned} S_1(0) exp\left\{-\frac{\sigma_1^2 T}{2}\right\} E_Q\Big(exp\left\{\sigma_1 W_1^Q(T)\right\} I_{\{-\gamma < -K\}} I_{\left\{\epsilon < \widetilde{K}\right\}}\Big) = \\ S_1(0) exp\left\{-\frac{\sigma_1^2 T}{2}\right\} E_Q\Big(exp\left\{-Z_1\right\} I_{\{X < -K\}} I_{\left\{Y < \widetilde{K}\right\}}\Big) \end{aligned}$$

where:

$$Z_1 = -\sigma_1 W_1^Q(T) \sim N(0, \sigma_1^2 T)$$
$$X = -\gamma \sim N(-m, \sigma^2)$$
$$Y = \frac{\phi_1 W_1^Q(T) + \phi_2 W_2^Q(T)}{\sigma_\phi \sqrt{T}} \sim N(0, 1)$$

Now we need to estimate the correlation coefficients ρ_{Z_1X}, ρ_{Z_1Y} and ρ_{XY}. Let us first consider ρ_{XY}:

$$\rho_{XY} = \frac{\sigma^2_{xy}}{\sigma_x \sigma_y}$$

Since $Y \sim N(0,1)$:

$$\begin{aligned}
\sigma^2_{xy} &= E_Q(XY) \\
&= E_Q\Big(\frac{\phi_1 W^Q_1(T) + \phi_2 W^Q_2(T)}{\sigma_\phi \sqrt{T}} S_2(T)\Big) - E_Q\Big(\frac{\phi_1 W^Q_1(T) + \phi_2 W^Q_2(T)}{\sigma_\phi \sqrt{T}} S_1(T)\Big) \\
&= S_2(0) exp\left\{-\frac{\sigma_2^2 T}{2}\right\} \frac{1}{\sigma_\phi\sqrt{T}} E_Q\Big(\phi_1 W^Q_1(T) exp\left\{\sigma_2 W^Q_2(T)\right\} + \phi_2 W^Q_2(T) exp\left\{\sigma_2 W^Q_2(T)\right\}\Big) \\
&- S_1(0) exp\left\{-\frac{\sigma_1^2 T}{2}\right\} \frac{1}{\sigma_\phi\sqrt{T}} E_Q\Big(\phi_1 W^Q_1(T) exp\left\{\sigma_1 W^Q_1(T)\right\} + \phi_2 W^Q_2(T) exp\left\{\sigma_1 W^Q_1(T)\right\}\Big)
\end{aligned}$$

Opening the brackets and calculating the expectations, the above yields:

$$S_2(0)\frac{\sigma_2\sqrt{T}}{\sigma_\phi}\Big(\phi_1\rho + \phi_2\Big) - S_1(0)\frac{\sigma_1\sqrt{T}}{\sigma_\phi}\Big(\phi_1 + \phi_2\rho\Big)$$

To get correlation we need to divide by $\sigma_x\sigma_y$ to finally get:

$$\rho_{XY} = \frac{a-b}{\sigma}$$

where:

$$\begin{aligned}
a &= S_2(0)\frac{\sigma_2\sqrt{T}}{\sigma_\phi}\Big(\phi_1\rho + \phi_2\Big) \\
b &= S_1(0)\frac{\sigma_1\sqrt{T}}{\sigma_\phi}\Big(\phi_1 + \phi_2\rho\Big)
\end{aligned}$$

Now consider ρ_{Z_1X}:

$$\rho_{Z_1X} = \frac{\sigma^2_{z_1x}}{\sigma_{z_1}\sigma_y}$$

Since $Z_1 \sim N(0, \sigma_1^2 T)$:

$$\begin{aligned}
\sigma^2_{z_1x} &= E_Q(Z_1X) \\
&= E_Q\Bigg(\sigma_1 W_1^Q(T)\Big(S_1(0)exp\left\{-\frac{\sigma_1^2 T}{2} + \sigma_1 W_1^Q(T)\right\} - S_2(0)exp\left\{-\frac{\sigma_2^2 T}{2} + \sigma_2 W_2^Q(T)\right\}\Big)\Bigg) \\
&= S_1(0)exp\left\{-\frac{\sigma_1^2 T}{2}\right\} E_Q\Big(\sigma_1 W_1^Q(T) exp\left\{\sigma_1 W_1^Q(T)\right\}\Big) \\
&- S_2(0)exp\left\{-\frac{\sigma_2^2 T}{2}\right\} E_Q\Big(\sigma_1 W_1^Q(T) exp\left\{\sigma_2 W_2^Q(T)\right\}\Big)
\end{aligned}$$

Opening the brackets and calculating the expectations, the above yields:

$$S_1(0)\sigma_1^2 T - S_2(0)\sigma_1\sigma_2\rho T$$

To get correlation we need to divide by $\sigma_x\sigma_y$ to finally get:

$$\rho_{Z_1X} = \frac{\sqrt{T}}{\sigma}\Big(S_1(0)\sigma_1 - S_2(0)\sigma_2\rho\Big)$$

Let us finally consider ρ_{Z_1Y}:

$$\rho_{Z_1Y} = \frac{\sigma^2_{z_1y}}{\sigma_{z_1}\sigma_y}$$

Since both random variables Z_1 and Y have zero expectation:

$$\begin{aligned}
\sigma^2_{z_1y} &= E_Q(Z_1Y) \\
&= E_Q\Bigg(-\sigma_1 W_1^Q(T)\Big(\frac{\phi_1 W_1^Q(T) + \phi_2 W_2^Q(T)}{\sigma_\phi\sqrt{T}}\Big)\Bigg)
\end{aligned}$$

Opening the brackets and calculating the expectations, the above yields:

$$\rho_{Z_1Y} = -\frac{\phi_1 + \rho\phi_2}{\sigma_\phi}$$

Finally, we can apply the Two-Asset lemma to I to get:

$$S_1(0)exp\left\{-\frac{\sigma_1^2 T}{2}\right\}E_Q\Big(exp\{-Z_1\}I_{\{X<-K\}}I_{\{Y<\widetilde{K}\}}\Big)$$
$$= S_1(0)\Phi^2(\hat{x}_1, \hat{y}_1, \rho_{XY})$$

where:

$$\hat{x}_1 = \frac{m-K}{\sigma} + \sigma_1\rho_{Z_1X}\sqrt{T}$$
$$\hat{y}_1 = \widetilde{K} + \sigma_1\sqrt{T}\rho_{Z_1Y}$$
$$\rho_{Z_1X} = \frac{\sqrt{T}}{\sigma}\Big(S_1(0)\sigma_1 - S_2(0)\sigma_2\rho\Big)$$
$$\rho_{Z_1Y} = -\frac{\phi_1 + \phi_2\rho}{\sigma_\phi}$$
$$\rho_{XY} = \frac{a-b}{\sigma}$$
$$a = S_2(0)\frac{\sigma_2\sqrt{T}}{\sigma_\phi}\Big(\phi_1\rho + \phi_2\Big)$$
$$b = S_1(0)\frac{\sigma_1\sqrt{T}}{\sigma_\phi}\Big(\phi_1 + \phi_2\rho\Big)$$
$$\widetilde{K} = \frac{(\frac{\sigma_\phi^2}{2} + \phi_1\theta_1 + \phi_2\theta_2)T + ln(\frac{1}{\tilde{a}})}{\sigma_\phi\sqrt{T}}$$

$II.$

$$S_2(0)exp\left\{-\frac{\sigma_2^2 T}{2}\right\}E_Q\Big(exp\left\{\sigma_2 W_2^Q(T)\right\}I_{\{-\gamma<-K\}}I_{\{\epsilon<\widetilde{K}\}}\Big) =$$
$$S_2(0)exp\left\{-\frac{\sigma_2^2 T}{2}\right\}E_Q\Big(exp\{-Z_2\}I_{\{X<-K\}}I_{\{Y<\widetilde{K}\}}\Big)$$

where:

$$Z_2 = -\sigma_2 W_2^Q(T) \sim N(0, \sigma_1^2 T)$$
$$X = -\gamma \sim N(-m, \sigma^2)$$
$$Y = \frac{\phi_1 W_1^Q(T) + \phi_2 W_2^Q(T)}{\sigma_\phi \sqrt{T}} \sim N(0,1)$$

Now we need to estimate the correlation coefficients $\rho_{Z_2 X}$ and $\rho_{Z_2 Y}$ since ρ_{XY} has already been estimated. Proceeding in the same manner as for I, we evaluate the correlation coefficients to be as follows:

$$\rho_{Z_2 X} = \frac{\sqrt{T}}{\sigma}\Big(S_1(0)\sigma_1\rho - S_2(0)\sigma_2\Big)$$
$$\rho_{Z_2 Y} = -\frac{\phi_1\rho + \phi_2}{\sigma_\phi}$$

Applying the Two-Asset lemma:

$$S_2(0)exp\left\{-\frac{\sigma_2^2 T}{2}\right\} E_Q\Big(exp\left\{-Z_2\right\} I_{\{X<-K\}} I_{\{Y<\widetilde{K}\}}\Big)$$
$$= S_2(0)\Phi^2(\hat{x}_2, \hat{y}_2, \rho_{XY})$$

where:

$$\hat{x}_2 = \frac{m-K}{\sigma} + \sigma_2\rho_{Z_2 X}\sqrt{T}$$
$$\hat{y}_2 = \widetilde{K} + \sigma_2\sqrt{T}\rho_{Z_2 Y}$$
$$\rho_{Z_2 X} = \frac{\sqrt{T}}{\sigma}\Big(S_1(0)\sigma_1\rho - S_2(0)\sigma_2\Big)$$
$$\rho_{Z_2 Y} = -\frac{\phi_1\rho + \phi_2}{\sigma_\phi}$$

The final term III is simply:

$$KE_Q(I_{\{-\gamma<-K\}}I_{\left\{\epsilon<\widetilde{K}\right\}}) = K\Phi^2(\hat{x}_3, \hat{y}_3, \rho_{XY})$$

where:

$$\hat{x}_3 = \frac{m-K}{\sigma}$$
$$\hat{y}_3 = \widetilde{K}$$

Combining all three terms together we get the stated formula (3.3).

Appendix 4

Consider the following expectation:

$$E_Q\Big((S_1(T) - S_2(T) - K)I_{\left\{\frac{dP}{dQ} > \tilde{a}\right\}} I_{\left\{S_1(T) \geq \frac{aS_2^b(T)}{E_Q(S_2^b(T))}\right\}}\Big)$$

The terms in the indicator functions have already been considered in Appendices 1 and 3. So, we can rewrite the above expectation in the following way:

$$E_Q\Big(S_1(T)I_{\left\{\epsilon_1 \leq \tilde{K}\right\}} I_{\left\{\epsilon_2 \leq \bar{K}\right\}}\Big) - E_Q\Big(S_2(T)I_{\left\{\epsilon_1 \leq \tilde{K}\right\}} I_{\left\{\epsilon_2 \leq \bar{K}\right\}}\Big) - KE_Q\Big(I_{\left\{\epsilon_1 \leq \tilde{K}\right\}} I_{\left\{\epsilon_2 \leq \bar{K}\right\}}\Big)$$

We can apply the Two-Asset lemma to each of the three terms in the above expression. Before that, however, we need to estimate the correlation coefficient between ϵ_1 and ϵ_2:

$$\begin{aligned}
\rho_{\epsilon_1\epsilon_2} &= E_Q\Bigg(\Big(\frac{\phi_1 W_1^Q(T) + \phi_2 W_2^Q(T)}{\sigma_\phi \sqrt{T}}\Big)\Big(\frac{\sigma_2 b W_2^Q(T) - \sigma_1 W_1^Q(T)}{\sqrt{\sigma_1^2 T - 2\sigma_1\sigma_2 b\rho T + \sigma_2^2 b^2 T}}\Big)\Bigg) \\
&= \frac{(\sigma_2 b\phi_1\rho + \sigma_2 b\phi_2 - \sigma_1\phi_1 - \sigma_1\phi_2\rho)\sqrt{T}}{\sigma_\phi\sqrt{\sigma_1^2 T + \sigma_2^2 b^2 T - 2\sigma_1\sigma_2 b\rho T}}
\end{aligned}$$

where we used the fact that $E(W^2(T)) = Var(W(T)) = T$ and $E(W_1(T)W_2(T)) = Cov(W_1(T)W_2(T)) = \rho T$. Combining this result with the results from Appendices

1 and 3 and applying the Two-Asset lemma to the first term:

$$E_Q\Big(S_1(T)I_{\{\epsilon_1\le\widetilde{K}\}}I_{\{\epsilon_2\le\bar{K}\}}\Big)=S_1(0)\Phi^2(\hat{x}_1,\hat{y}_1,\rho_3)$$

where:

$$\begin{aligned}
\hat{x}_1 &= \widetilde{K}+\sigma_1\rho_1\sqrt{T}\\
\hat{y}_1 &= \bar{K}+\sigma_1\rho_4\sqrt{T}\\
\rho_1 &= \frac{(\sigma_1-\sigma_2 b\rho)\sqrt{T}}{\sqrt{\sigma_1^2T-2\sigma_1\sigma_2 b\rho T+\sigma_2^2b^2T}}\\
\rho_3 &= \frac{(\sigma_2 b\phi_1\rho+\sigma_2 b\phi_2-\sigma_1\phi_1-\sigma_1\phi_2\rho)\sqrt{T}}{\sigma_\phi\sqrt{\sigma_1^2T+\sigma_2^2b^2T-2\sigma_1\sigma_2 b\rho T}}\\
\rho_4 &= -\frac{\phi_2\rho+\phi_1}{\sigma_\phi}
\end{aligned}$$

The second term evaluates to:

$$E_Q\Big(S_2(T)I_{\{\epsilon_1\le\widetilde{K}\}}I_{\{\epsilon_2\le\bar{K}\}}\Big)=S_2(0)\Phi^2(\hat{x}_2,\hat{y}_2,\rho_3)$$

where:

$$\begin{aligned}
\hat{x}_2 &= \widetilde{K}+\sigma_2\rho_2\sqrt{T}\\
\hat{y}_2 &= \bar{K}+\sigma_2\rho_5\sqrt{T}\\
\rho_2 &= \frac{(\sigma_1\rho-\sigma_2 b)\sqrt{T}}{\sqrt{\sigma_1^2T-2\sigma_1\sigma_2 b\rho T+\sigma_2^2b^2T}}\\
\rho_5 &= -\frac{\phi_2+\phi_1\rho}{\sigma_\phi}
\end{aligned}$$

Finally, the third term is:

$$KE_Q\Big(I_{\{\epsilon_1 \leq \widetilde{K}\}} I_{\{\epsilon_2 \leq \bar{K}\}}\Big) = K\Phi(\hat{x}_3, \hat{y}_3, \rho_3)$$

where:

$$\hat{x}_3 = \widetilde{K}$$

$$\hat{y}_3 = \bar{K}$$

Combining the three terms together we get the formula (3.4).