RISK MINIMIZING PORTFOLIO OPTIMIZATION AND HEDGING WITH CONDITIONAL VALUE-AT-RISK

by

Jing Li

A dissertation submitted to the faculty of
the University of North Carolina at Charlotte
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in
Applied Mathematics

Charlotte

2009

Approved by:

____

Dr. Mingxin Xu

____

Dr. Volker Wihstutz

____

Dr. You-lan Zhu

____

Dr. Dmitry Shapiro

ABSTRACT

JING LI. Risk Minimizing Portfolio Optimization and Hedging with Conditional Value-at-Risk. (Under the direction of DR. MINGXIN XU)

This thesis looks at the problem of finding the optimal investment strategy of a self-financing portfolio in a dynamic complete market setting so that the risk measured by Conditional Value-at-Risk (CVaR) is minimized under the condition that the expected return is bounded from below.

We start out with a CVaR minimization problem without expected return requirement. We find the exact optimal conditions and apply them to two classic complete market models: the Binomial model and the Black-Scholes model. In these cases, the procedures of finding the optimal strategies are given with exact formulas, and the resulting minimal CVaR values can be calculated.

We then add a minimal expected return constraint, and look for an optimal solution in a continuous-time setting. The optimal solution most likely does not exist if there is no upper bound on returns over time, but the infimum of CVaR can still be computed. However, when such a uniform upper bound is prescribed, we find the optimal conditions together with the optimal investment strategy and the resulting minimal CVaR.

ACKNOWLEDGMENTS

Thanks are first due to Dr. Mingxin Xu for persevering with me as my advisor through out the time of my PhD research and dissertation writing. Her ideas and advice are crucial for the success of my research, and I am grateful for her guidance and encouragement, without which, this dissertation would not have been written. The members of my dissertation committee, Dr. Volker Wihstutz, Dr. You-lan Zhu, and Dr. Dmitry Shapiro, have generously given their time during busy semesters. I thank them for their patience and contribution. The faculty members in the mathematics program and mathematical finance program have helped greatly in nurturing my academic development and critical thinking during my entire PhD program. I am grateful too to the staff of the program and especially, the graduate coordinator, Dr. Joel Avrin, who is always supportive of my work. Special thanks are given to Dr. Lloyd Blenman from the Belk college of business and Dr. Roger Lee from the mathematics department at University of Chicago for their helpful comments to better the research.

I shall express my gratitude to my colleagues at Evergreen Investments of Wachovia Corporation, now a Wells Fargo company, for sharing their comments and suggestions. I thank Abe Riazati and Seyi Olurotimi, my supervisors, for their patience in showing me the risk management practice, and for their tolerance of my flexible working hours during the seek of my PhD degree. The resources and information that are made available to me are extremely valuable in expanding my knowledge base and deepening my understanding of the research problem.

Lastly, I want to thank my husband, Wei Fu, for his comfort during those frustrating moments, my parents and parents-in-law for their care, and my baby boy for his cooperation. I am also grateful to many of my friends who assisted, supported my research and writing efforts over the years.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1: INTRODUCTION

1.1 Risk Measures and Risk Management

More than half a century ago, Markowitz proposed a method of ranking and selecting investments in his Nobel Prize winning work [14]. He used standard deviation, also called volatility, to gauge the risk level associated with each investment, and worked out the investment combination that yields minimal risk-taking at each return level. The positively sloped portion of this famous trade-off curve of minimal risk given return or maximum return given risk is named efficient frontier in that all the investment combination on the curve is risk-return efficient. This portfolio selection methodology is written into the standard finance textbooks and regarded as the foundation of modern portfolio theory.

While it is true that risk can be interpreted as uncertainty, and volatility is naturally called upon to measure this level, it is argued why the uncertainty in upward profit swing should be of worry. Investors will certainly welcome superior returns, and be concerned with the possibility of dwindling returns dipping below their expectation. Many risk measures, such as semi-standard deviation and (maximum) drawdown, are developed afterwards to include only the downward uncertainty.

Measuring and reporting risk exposures is merely the first step of the risk management process. Risk managers usually raise the questions: first, whether the risk profile is in compliance with the portfolio strategy as prescribed in the prospectus; second, whether the returns justify the levels of risk-taking; and third, what corrective actions can be suggested to a risk-return mismatched portfolio. It then interacts with portfolio manager's decision making process of whether these suggested corrective actions are indeed necessary, and if so, then how to strategically design and systematically execute the actions. From the perspective of regulatory bodies and compliance officers, they need to find out what level of cash cushion would be considered sufficient to back up an enterprize's risk position, and whether the enterprize has this level of cash in reserve. Neither standard deviation nor its variants provides a confident answer to the question upfront: what is the risk level?!

In 1994, J.P.Morgan in RiskMetrics system proposed a quantile based risk measure: Value-at-Risk (VaR). It answers the following question: how much a position is expected to lose during a measurement period with a given probability? For a given measurement period and a probability level λ, VaR is simply the loss that is exceeded with probability $1-\lambda$ during this period. It soon gains wide acceptance in the financial industry for its clarity in concept and it is later adopted by Basel Banking Supervisory Committee (BASEL) in the calculation of capital mandate that is required to back up risk position.

Despite its still popular use in the industry and regulatory bodies, its inadequacy as a risk measure surfaces. Being a quantile measure, VaR gives the threshold that the loss of a portfolio will exceed in the worst λ situations, but it fails to give the magnitude of loss should such situations realize. Optimizing a portfolio by minimizing VaR as a risk measure is a formidable task from an implementation perspective, because VaR is generally not convex. And this optimization leads to non-smooth results because of the discontinuous nature of the quantiles. Thus as a frequent reporting measure of market risk exposure, potential large change of value in VaR during a very short period of time is not a desirable property for financial stability. Moreover, reducing VaR may also thin out the tail, i.e., even though less loss is expected to be exceeded, but once it's exceeded, the magnitude is disastrous! The most notorious shortcoming of VaR that draws lots of criticism is that it discourages diversification, a well known way of reducing risk.

Recent research in the area of mathematical finance, as for the economic theory of utility functions, developed an axiomatic approach for risk measure. With the axioms of coherence parallel to those of rational investors, Artzner et al. [4] and [5] first proposed coherent risk measures and derived their representation theorems. Conditional Value-at-Risk (CVaR), sometimes called Shortfall Risk, is a distribution-based coherent risk measure first studied by Rockafellar and Uryasev [16] and Acerbi and Tasche [1]. It is known as the expected loss during a certain period of time, conditional that the loss is greater than a loss threshold corresponding to a certain confidence level λ. CVaR is a vast improvement over VaR in producing smooth portfolio optimization results. The wide use of VaR and the advantage of CVaR have lead many financial institutions to consider supplementing VaR with CVaR for internal risk control.

We have so far discussed various risk measures along the historical line, and recognized the superior properties CVaR possesses. Nonetheless, an adequate risk management system should look at several risk measures all at once while bearing their limitations in mind, and supplement these measures with stress testing and scenarios analysis.

1.2 Problems and Assumptions

Parallel to the problem of finding a static minimal variance portfolio solved by Markovitz, with the positively sloped portion being defined as efficient frontier, the problem of finding a static minimal CVaR portfolio is numerically solved by Rockafellar and Uryasev [16]. Bielecki et. al.[2] put one step further, and solve the dynamic version of mean-variance portfolio selection with bankruptcy prohibition in a complete market. In this thesis, we attempt to replace variance with CVaR, and find a mean-CVaR efficient dynamic portfolio with uniform bounds on returns over time.

A good reference for the risk measure CVaR is in the book written by Föllmer and Schied [10] where CVaR is given a third name Average VaR. We define VaR of a random variable Z with finite expectation at level λ to be

$$VaR_\lambda(Z) = \inf\{\, m \mid P(Z+m<0) \le \lambda \,\}, \quad \lambda \in (0,1),$$

and CVaR to be

$$CVaR_\lambda(Z) = \frac{1}{\lambda}\int_0^\lambda VaR_\gamma(Z)d\gamma, \quad \lambda \in (0,1). \tag{1.1}$$

With this definition, it is easy to see why CVaR is different from VaR: it is smooth with respect to the change of the confidence level λ. We have to be a little careful when we write down the equivalent form of (1.1) for the case when the probability space has atoms

$$CVaR_\lambda(Z) = -\frac{1}{\lambda}\Big(E[Z1_{\{Z<q_\lambda\}}] + q_\lambda(\lambda - P(Z<q_\lambda))\Big), \quad q_\lambda = -VaR_\lambda(Z).$$

Numerical implementation of an optimization problem with quantile-based constrains instead of variance does not have to be easy. The major contribution of Rochafellar and

Uryasev [17] is that they found an equivalent formula for CVaR as a convex function, thus opening the door for convex programming methods. Using Monte-Carlo simulation for a one-time step model with multiple assets, they formulated the portfolio optimization problem into a linear programming problem which can be efficiently implemented with standard programming software. For $\lambda \in (0,1)$, definition (1.1) is equivalent to

$$CVaR_\lambda(Z) = \frac{1}{\lambda} \inf_{x \in \mathcal{R}} (E[(x-Z)^+] - \lambda x). \tag{1.2}$$

The self-financing portfolio X_t under consideration comprises two investments: a money market account and a risky asset S_t. Suppose the interest rate is a constant r and the S_t is a real-valued semimartingale process on the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F})_{0\leq t\leq T}, P)$ that satisfies the usual conditions where $\mathcal{F}_0$ is trivial and $\mathcal{F}_T = \mathcal{F}$. With ξ_t shares invested in the risky asset, the value of the portfolio evolves according to the dynamics

$$dX_t = \xi_t dS_t + r(X_t - \xi_t S_t)dt,\, X_0 = x_0,$$

where x_0 is the initial portfolio value.

The question is how we should trade the shares throughout a finite holding period $[0, T]$ so that we can achieve minimal risk, measured by CVaR, at time T, while keeping the return within acceptable range? In the classic setup of portfolio optimization, expected returns are maximized given limits on the risk of the portfolio. It can be made formal with some technical conditions that the above problem is equivalent to minimizing the risk of the portfolio given requirements for its return. Our setup will focus on the later approach, which focuses on risk minimization as the objective.

An attempt to employ the dynamic programming method for multi-period models was made by Ruszczyński and Shapiro [21], whose approach is to modify the risk measure CVaR into a dynamic version "*conditional risk mappings for CVaR*". In this thesis, we will keep the original measure of CVaR at a fixed time horizon, and let the portfolio composition adjusts dynamically. This is similar to maximizing expected utility on the outcomes of a dynamic portfolio. Ruszczyński and Shapiro's choice of optimizing "conditional risk mappings for

CVaR" at each time period yields very different results than ours, where we optimize CVaR of the final wealth of a dynamic portfolio.

Mathematically, we are looking for a strategy $(\xi_t)_{0\leq t\leq T}$ to minimize the conditional Value-at-Risk at level $0 < \lambda < 1$ of the final portfolio value: $\inf_{\xi_t} CVaR_\lambda(X_T)$, while requiring the expected return to remain above constant z: $E[X_T] \geq z$. In addition, we allow uniform bounds on the value of the portfolio over time: $x_d \leq X_t \leq x_u$, $\forall t \in [0,T]$, where the constants satisfy $-\infty < x_d < x_0 < x_u \leq \infty$. Therefore, our **Main Problem** is

$$\begin{aligned} \xi_t^* &= \arg\inf_{\xi_t} CVaR_\lambda(X_T), \\ \text{subject to} \quad & E[X_T] \geq z, \\ & x_d \leq X_t \leq x_u \, a.s., \ \text{for all } t \in [0,T]. \end{aligned} \tag{1.3}$$

When $x_u = \infty$, there is practically no upper bound for X_t throughout time t, but we will not have the situation $P(X_T = \infty) > 0$ because we will exclude arbitrage in Assumption 1.1. If we further set $x_d = 0$, then we have the no bankruptcy condition.

The Main Problem (1.3) is equivalent to the problem of minimizing CVaR on the return R_T

$$\begin{aligned} \xi_t^* &= \arg\inf_{\xi_t} CVaR_\lambda(R_T), \\ \text{subject to} \quad & E[R_T] \geq z_r, \\ & r_d \leq R_t \leq r_u \, a.s., \ \text{for all } t \in [0,T], \end{aligned}$$

whether it be percentage return $R_T = \frac{X_T - X_0}{X_0}$ or log return $R_T = \ln \frac{X_T}{X_0}$ because we only need to identify the one-to-one correspondence between the quantiles of X_T and R_T. We are requiring the realized return to be above r_d in all cases if we take $r_u = \infty$.

When we have an existing portfolio H_t consisting of investments in securities, we can ask a second question as how to hedge our risk with a self-financing admissible portfolio. Let H be the random variable representing the final value of the existing portfolio. It is

$\mathcal{F}_T$-measurable since $H = H_T$. The optimal hedging problem is to solve

$$\xi_t^* = \arg\inf_{\xi_t} CVaR_\lambda(H + X_T), \tag{1.4}$$
$$\text{subject to} \quad E[X_T] \geq z$$
$$x_d \leq X_t \leq x_u \, a.s., \text{ for all } t \in [0, T].$$

Thus when we combine the original portfolio and the hedging portfolio, the risk is minimized. It is straight-forward to see that if X_T^* is the final wealth of the optimal portfolio for problem (1.3), by which we mean that

$$\min_{\xi_t} CVaR_\lambda(X_T) = CVaR_\lambda(X_T^*),$$

then $X_T^* - H$ is the optimal solution to problem (1.4) if it is the final wealth of some self-financing strategy. In a complete market model, this is not an issue because all derivatives can be replicated. Therefore, we will first focus on solving problem (1.3) in a complete market model.

In the search of solution to the Main Problem (1.3), we first solved the following Main Problem without the condition on the expectation $E[X_T] \geq z$:

$$\xi_t^* = \arg\inf_{\xi_t} CVaR_\lambda(X_T), \tag{1.5}$$
$$\text{subject to} \quad x_d \leq X_t \leq x_u \, a.s., \text{ for all } t \in [0, T].$$

Assumption 1.1. *Assume there is no arbitrage and the market is complete with a unique equivalent local martingale measure $\tilde{P}$.*

Assumption 1.2. *The Radon-Nikodým derivative $\frac{d\tilde{P}}{dP}$ has a continuous distribution.*

Under Assumption 1.1, any $\mathcal{F}$-measurable random variable can be replicated by a dynamic portfolio. The dynamic optimization problem (1.3) can be reduced to a static one

$$
\begin{aligned}
& \inf_{X \in \mathcal{F}} CVaR_\lambda(X) && (1.6)\\
\text{subject to} \quad & E[X] \geq z, \quad (\textit{return constraint}) \\
& \tilde{E}[X] = x_r, \quad (\textit{capital constraint}) \\
& x_d \leq X \leq x_u \, a.s.;
\end{aligned}
$$

and problem (1.5) becomes

$$
\begin{aligned}
& \inf_{X \in \mathcal{F}} CVaR_\lambda(X) && (1.7)\\
\text{subject to} \quad & \tilde{E}[X] = x_r, \quad (\textit{capital constraint}) \\
& x_d \leq X \leq x_u \, a.s..
\end{aligned}
$$

Here the expectation E is taken under the physical probability measure P, and the expectation $\tilde{E}$ is taken under the risk neutral probability measure $\tilde{P}$, while $x_r = x_0 e^{rT}$. To solve the main problem in an incomplete market setting, the exact hedging argument that translate the dynamic problem (1.3) into the static problem (1.6) has to be replaced by a super-hedging argument. This is done for expected shortfall minimization in Föllmer and Leukert [9], and for convex risk minimization in Rudloff [18]. Similarly, the hedging result can be easily adapted for S_t to be $\mathbb{R}^d$-valued, where the dimension d is a natural number. Assumption 1.2, namely the Radon-Nikodým derivative $\frac{d\tilde{P}}{dP}$ has a continuous distribution, is also made not because of technical impossibility, but because of the simplification it brings to the presentation for its lengthy discussion does not bring additional new insight to the main topic of this thesis. In fact, we'll consider the case of discontinuous Radon-Nikodým derivative in the solution to the Main Problem without expected return requirement.

1.3 Overview

Chapter 2 details the approach to the risk minimization problem without expected return constraint, finds the closed-form solution and applies it two popular complete market models: the Binomial model and the Black-Scholes model. Chapter 3 adds expected return constraint to the optimization problem solved in Chapter 2, addresses the case when there is no uniform upper bound on returns, and finds the solution to the case when such an upper bound is prescribed. Chapter 4 concludes.

CHAPTER 2: AN OPTIMIZATION PROBLEM WITHOUT EXPECTED RETURN REQUIREMENT

2.1 One-Constraint Optimization Problem

This chapter focuses on the Main Problem without expected return constraint (1.5), formulated as below:

$$\xi_t^* = \arg\inf_{\xi_t} CVaR_\lambda(X_T), \tag{2.1}$$

$$\text{subject to} \quad x_d \le X_t \le x_u \, a.s., \text{ for all } t \in [0,T].$$

And the related optimal hedging problem looks like:

$$\xi_t^* = \arg\inf_{\xi_t} CVaR_\lambda(H + X_T), \tag{2.2}$$

$$\text{subject to} \quad x_d \le X_t \le x_u \, a.s., \text{ for all } t \in [0,T].$$

2.2 Static Formulation of the Dynamic Problem

Our solution will anchor on duality methods based on risk neutral measures, similar to those employed in option pricing and utility maximization problems. This martingale approach is well-studied in recent mathematical finance research partly because it allows finding solutions to a wider ranges of problems which does not possess Markovian property and thus do not meet dynamic programming principles.

As mentioned in Chapter 1, CVaR minimization problem is complicated because the objective function involves quantile function and the corresponding numerical methods will have to involve ordering the position values. Rockafellar and Uryasev ([16] and [17]) found CVaR to be the Fenchel-Legendre dual of expected shortfall

$$CVaR_\lambda(Z) = \frac{1}{\lambda} \inf_{x \in \mathcal{R}} (E[(x-Z)^+] - \lambda x),$$

thus standard convex analysis applies.

Recall that under Assumption 1.1, the space of final outcomes of self-financing strategies are those $\mathcal{F}_T$-measurable random variables X such that $\tilde{E}[X] = x_r$. And the dynamic problem (1.5) has the following static form, as mentioned in Chapter 1:

$$\begin{aligned} & \inf_{X \in \mathcal{F}} CVaR_\lambda(X) \\ \text{subject to} \quad & \tilde{E}[X] = x_r, \\ & x_d \leq X \leq x_u \, a.s.. \end{aligned} \tag{2.3}$$

Now we can further reformulate the above static version into a more tractable static convex optimization problem

$$\begin{aligned} CVaR(X^*) = \inf_X \frac{1}{\lambda} \inf_{x \in \mathcal{R}} (E[(x - X)^+] - \lambda x), \\ \text{subject to} \quad \tilde{E}[X] = x_r, \quad x_d \leq X \leq x_u \, a.s.. \end{aligned} \tag{2.4}$$

Let $X_T^* = X^*$, then X_T^* is the final value of the optimal portfolio for problem (1.5):

$$CVaR_\lambda(X_T^*) = \inf_{\xi_t} CVaR_\lambda(X_T) = \inf_X \frac{1}{\lambda} \inf_{x \in \mathcal{R}} (E[(x-X)^+] - \lambda x) = \frac{1}{\lambda} \inf_{x \in \mathcal{R}} (E[(x-X^*)^+] - \lambda x).$$

Martingale representation theorem applied to $X_t^* = \tilde{E}[X_T^* | \mathcal{F}_t]$ will produce the optimal hedging strategy ξ_t^* for problem (1.5).

Problem (2.4) is intrinsically much simpler than problem (1.5) because it looks for an optimal random variable X^* with convex objective function. The above simplification steps we have taken is based on classic duality theory (martingale approach) in mathematical finance. By duality, we mean there are two important spaces: the primal space consisting of dynamics of self-financing portfolios and the dual space consisting of risk neutral measures. The optimization problem in the primal space is translated into an optimization problem in the dual space, where a solution is always easier to obtain in a complete market since the dual space consists of a singleton.

2.3 Solution to the Static Formulation

After rewriting the above static problem (2.4) by interchanging the order of infimum:

$$\inf_{\xi_t} CVaR_\lambda(X_T) = \frac{1}{\lambda}\inf_{x\in\mathcal{R}}\left(\inf_X E[(x-X)^+] - \lambda x\right) \tag{2.5}$$
$$\text{subject to} \quad \tilde{E}[X] = x_r, \quad x_d \le X \le x_u \, a.s.,$$

where the constants satisfy $-\infty < x_d < x_r < x_u \le \infty$, we arrive at the final form of the optimization problem (1.5) where we provide a direct solution in two steps:

One-Constraint Problem

Step 1 Minimization of Expected Shortfall

$$v(x) = \inf_X E[(x-X)^+] \tag{2.6}$$
$$\text{subject to} \quad \tilde{E}[X] = x_r,\, x_d \le X \le x_u \, a.s.,$$

Step 2 Minimization of CVaR

$$\inf_{\xi_t} CVaR_\lambda(X_T) = \frac{1}{\lambda}\inf_{x\in\mathcal{R}}(v(x) - \lambda x). \tag{2.7}$$

Schied [22] solved a general law invariant risk minimization problem of the type (2.6). We solve the CVaR minimization with the above two-step approach where we do not require the probability space to be atomless so the tree models are included. We also allow the upper bound to be infinity so there is no cap for how large the wealth can possibly be. We give explicit computation methods for the Black-Scholes and Binomial models in Section 2.4.

The solution for Expected Shortfall Minimization is studied in a semimartingale model in Föllmer and Leukert [9] and Xu [25]. Apply Proposition 4.1 in [9] to the above shortfall problem, we get the following result.

Theorem 2.1 (Solution to Expected Shortfall Minimization Problem)**.** *For any constant a, define sets based on the size of the Radon Nikodým derivative between the risk neutral* *probability measure* $\tilde{P}$ *and the physical probability measure* P*:* $A = \left\{\omega \in \Omega : \frac{d\tilde{P}}{dP}(\omega) > a\right\}$, $B = \left\{\omega \in \Omega : \frac{d\tilde{P}}{dP}(\omega) < a\right\}$*, and* $C = \left\{\omega \in \Omega : \frac{d\tilde{P}}{dP}(\omega) = a\right\}$*. The optimal solutions* X^* *and the corresponding value function* $v(x)$ *to the Expected Shortfall Minimization Problem in* **Step 1** *are given as the following:*

Case 1 $x \leq x_d$*:*

$X^* =$ *any random variable* X *with values in* $[x_d, x_u]$ *satisfying* $\tilde{E}[X^*] = x_r$*.*

$v(x) = 0$*.*

Case 2 $x_d \leq x \leq x_r < x_u$*:*

$X^* =$ *any random variable* X *with values in* $[x, x_u]$ *satisfying* $\tilde{E}[X^*] = x_r$*.*

$v(x) = 0$*.*

Case 3 $x_d < x_r \leq x \leq x_u$*:*

$X^* = x_d \mathbf{I}_{A_x} + k_x \mathbf{I}_{C_x} + x \mathbf{I}_{B_x}$*, where sets* A_x, B_x, C_x *are decided by level* a_x *defined as*

$$a_x = \sup\left\{a \,:\, \tilde{P}(B) \leq \frac{x_r - x_d}{x - x_d}\right\},$$

and k_x *is chosen so that the constraint*

$$x_r = \tilde{E}[X^*] = x_d \tilde{P}(A_x) + k_x \tilde{P}(C_x) + x \tilde{P}(B_x)$$

is satisfied, i.e.,

$$k_x = \frac{x_r - x_d \tilde{P}(A_x) - x \tilde{P}(B_x)}{\tilde{P}(C_x)} \mathbf{I}_{\{\tilde{P}(C_x) > 0\}}.$$

$v(x) = (x - x_d) P(A_x) + (x - k_x) P(C_x)$*.*

Case 4 $x \geq x_u$ *(when* $x_u < \infty$*):*

$X^* = x_d \mathbf{I}_{\bar{A}} + \bar{k} \mathbf{I}_{\bar{C}} + x_u \mathbf{I}_{\bar{B}}$*, where* $\bar{A}, \bar{B}, \bar{C}$ *are decided by level* $\bar{a}$ *defined as*

$$\bar{a} = \sup\left\{a \,:\, \tilde{P}(B) \leq \frac{x_r - x_d}{x_u - x_d}\right\},$$

and $\bar{k}$ is chosen so that the constraint

$$x_r = \tilde{E}[X^*] = x_d\tilde{P}(\bar{A}) + \bar{k}\tilde{P}(\bar{C}) + x_u\tilde{P}(\bar{B})$$

is satisfied, i.e.,

$$\bar{k} = \frac{x_r - x_d\tilde{P}(\bar{A}) - x_u\tilde{P}(\bar{B})}{\tilde{P}(\bar{C})}\, \mathbf{I}_{\{\tilde{P}(\bar{C})>0\}}.$$

$v(x) = (x - x_d)P(\bar{A}) + (x - \bar{k})P(\bar{C}) + (x - x_u)P(\bar{B})$.

Remark. Notice that the numbers a, x, k and sets A, B, C are all related. We call the collection x_u, $\bar{a}$, $\bar{k}$ and $\bar{A}$, $\bar{B}$, $\bar{C}$ that corresponds to x_u the 'bar-system'. Later in the paper we will also have 'r-system' and 'star-system'. We reserve the non-indexed system x, a, k and A, B, C for general definitions, and we use a_x, k_x and A_x, B_x, C_x to describe a system for fixed x.

Remark. The global minimum for function $v(x)$ is 0. For the first two cases where $x \leq x_r$, the minimal value of 0 can be easily achieved by an admissible $X^* \geq x$, including the special example of $X^* \equiv x_r$ that naturally satisfies the constraint of $\tilde{E}[X^*] = x_r$. For the latter two cases where $x > x_r$, the solution comes from Neyman-Pearson Lemma. A part of the X^* should be as large as possible to minimize $v(x)$ on the good set 'B', while the other part should be taken at the lower bound to offset this large number so that the risk-neutral expectation of X^* is guaranteed to stay at x_r.

In **Case 3**, we can equivalently define a_x as

$$a_x = \sup\left\{a \,:\, \tilde{P}(A) \geq \frac{x - x_r}{x - x_d}\right\},$$

because for fixed x level, A_x is the smallest set satisfying $\tilde{P}(A_x) \geq \frac{x-x_r}{x-x_d}$, and B_x is the largest set satisfying $\tilde{P}(B_x) \leq \frac{x_r-x_d}{x-x_d}$. When there is point mass at a_x, set C_x has non-zero probability and k_x has to be chosen to satisfy the constraint of $\tilde{E}[X^*] = x_r$. When there is no point mass at a_x, set C_x has zero probability under both physical and risk-neutral probability measures, and we have exact equalities $\tilde{P}(A_x) = \frac{x-x_r}{x-x_d}$ and $\tilde{P}(B_x) = \frac{x_r-x_d}{x-x_d}$. Note that the sets A, B, C and the number k in **Case 3** are functions of x, while in **Case**

4 they are not.

Remark. To use Theorem 2.1 to solve the CVaR Minimization Problem in **Step 2**, we need to find the global minimum among four cases when $x_u < \infty$:

$$\begin{aligned}
\frac{1}{\lambda}\inf_{x\le x_d}(v(x)-\lambda x) &= \frac{1}{\lambda}\inf_{x\le x_d}(0-\lambda x) = -x_d,\\
\frac{1}{\lambda}\inf_{x_d\le x\le x_r}(v(x)-\lambda x) &= \frac{1}{\lambda}\inf_{x_d\le x\le x_r}(0-\lambda x) = -x_r \le -x_d,\\
\frac{1}{\lambda}\inf_{x_r\le x\le x_u}(v(x)-\lambda x) &= \frac{1}{\lambda}\inf_{x_r\le x\le x_u}\big((x-x_d)P(A_x)+(x-k_x)P(C_x)-\lambda x\big),\\
\frac{1}{\lambda}\inf_{x\ge x_u}(v(x)-\lambda x) &= \frac{1}{\lambda}\inf_{x\ge x_u}\big((x-x_d)P(\bar{A})+(x-\bar{k})P(\bar{C})+(x-x_u)P(\bar{B})-\lambda x\big).
\end{aligned}$$

When $x_u = \infty$, only the first three cases need to be considered. We rewrite the third case as

$$\begin{aligned}
\frac{1}{\lambda}\inf_{x_r\le x\le x_u}(v(x)-\lambda x) &= \frac{1}{\lambda}\inf_{x_r\le x\le x_u}((x-x_d)P(A_x)+(x-k_x)P(C_x)-\lambda x)\\
&= -x_r + \frac{1}{\lambda}\inf_{x_r\le x\le x_u}((x-x_d)P(A_x)+(x-k_x)P(C_x)-\lambda x+\lambda x_r)\\
&= -x_r + \frac{1}{\lambda}\inf_{x_r\le x\le x_u}((x-x_d)(P(A_x)-\lambda\tilde{P}(A_x))+(x-k_x)(P(C_x)-\lambda\tilde{P}(C_x)))\\
&= -x_r + \frac{1}{\lambda}\inf_{x_r\le x\le x_u} h(x)
\end{aligned}$$

where we define

$$h(x) = (x-x_d)(P(A_x)-\lambda\tilde{P}(A_x))+(x-k_x)(P(C_x)-\lambda\tilde{P}(C_x)),$$

and solve the problem

$$\inf_{x_r\le x\le x_u} h(x) \tag{2.8}$$

in Lemma 2.2. In case four when we have $x_u < \infty$, the minimization is simpler because $\bar{A}, \bar{B}, \bar{C}$ and $\bar{k}$ are irrelevant to x. The function is linear in x with positive slope so the minimum is obtained at $x = x_u$:

$$\begin{aligned}
&\frac{1}{\lambda}\inf_{x\geq x_u}\big((x-x_d)P(\bar{A})+(x-\bar{k})P(\bar{C})+(x-x_u)P(\bar{B})-\lambda x\big)\\
&\quad=\frac{1}{\lambda}\big((x_u-x_d)P(\bar{A})+(x_u-\bar{k})P(\bar{C})-\lambda x_u\big)\\
&\quad\geq\frac{1}{\lambda}\inf_{x_r\leq x\leq x_u}\big((x-x_d)P(A_x)+(x-k_x)P(C_x)-\lambda x\big).
\end{aligned}$$

We have shown here that the minimum obtained in the fourth case will not provide the global minimum because it is dominated by the result from the third case. Note that the solutions for the first two cases are simple where we observe the second case dominates the first case. It is easy to see that case two is also dominated by case three because it coincides with the result in case three when $x = x_r$. Therefore, once we solve (2.8) in Lemma 2.2, we arrive naturally at the result of **Step 2** in Theorem 2.4.

Lemma 2.2. *Recall from Theorem 2.1, sets A, B, C are defined according to the number a, namely $A = \Big\{\omega \in \Omega : \frac{d\tilde{P}}{dP}(\omega) > a\Big\}$, $B = \Big\{\omega \in \Omega : \frac{d\tilde{P}}{dP}(\omega) < a\Big\}$, and $C = \Big\{\omega \in \Omega : \frac{d\tilde{P}}{dP}(\omega) = a\Big\}$. Also recall for any fixed x, we define*

$$a_x = \sup\left\{a\ :\ \tilde{P}(B) \leq \frac{x_r - x_d}{x - x_d}\right\}, \quad k_x = \frac{x_r - x_d\tilde{P}(A_x) - x\tilde{P}(B_x)}{\tilde{P}(C_x)}\mathbf{I}_{\{\tilde{P}(C_x)>0\}},$$

where the sets A_x, B_x and C_x are related to x as $A_x = \Big\{\omega \in \Omega : \frac{d\tilde{P}}{dP}(\omega) > a_x\Big\}$, etc. Denote the parameters $\bar{a}$, $\bar{k}$, $\bar{A}$, $\bar{B}$, $\bar{C}$ corresponding to $x = x_u$ as the 'bar-system'; parameters a_r, k_r, A_r, B_r, C_r corresponding to $x = x_r$ as the 'r-system', parameters a^*, k^*, A^*, B^*, C^* corresponding to $x = x^*$ as the 'star-system'. The solution to the minimization problem

$$\inf_{x_r\leq x\leq x_u} h(x),$$

where

$$h(x) = (x - x_d)(P(A_x) - \lambda\tilde{P}(A_x)) + (x - k_x)(P(C_x) - \lambda\tilde{P}(C_x)),$$

is

- *If* $\frac{d\tilde{P}}{dP} \leq \frac{1}{\lambda}$, *$P-$a.s., then the minimum is achieved by the 'r-system' and*

$$\inf_{x_r \leq x \leq x_u} h(x) = h(x_r) = 0.$$

- *Otherwise, if* $\frac{1}{\bar{a}} \leq \frac{\lambda - P(\bar{A})}{1-\tilde{P}(\bar{A})}$, *then the minimum is achieved by the 'bar-system' and*

$$\inf_{x_r \leq x \leq x_u} h(x) = h(x_u).$$

If $\frac{1}{\bar{a}} > \frac{\lambda - P(\bar{A})}{1-\tilde{P}(\bar{A})}$, *then the minimum is achieved by the 'star-system' and*

$$\inf_{x_r \leq x \leq x_u} h(x) = h(x^*) = (x^* - x_d)(P(A^*) - \lambda\tilde{P}(A^*)).$$

Here $a^* = \sup\left\{a \,:\, \frac{1}{a} \geq \frac{\lambda - P(A)}{1-\tilde{P}(A)}\right\}$, $A^* = \left\{\omega \in \Omega : \frac{d\tilde{P}}{dP}(\omega) > a^*\right\}$, $k^* = x^* = \frac{x_r - x_d\tilde{P}(A^*)}{1-\tilde{P}(A^*)}$, *are the parameters that defines the 'star-system'.*

Remark. The 'r-system' corresponds to parameters: $a_r = ess\sup\frac{d\tilde{P}}{dP}$, $\tilde{P}(B_r) = P(B_r) = 1$, $\tilde{P}(A_r) = \tilde{P}(C_r) = 0$, and $k_r = 0$. When $x_u < \infty$, the definition for the 'bar-system' is straightforward. When $x_u = \infty$, the 'bar-system' corresponds to the set of parameters satisfying $\bar{a} = ess\inf\frac{d\tilde{P}}{dP}$, $\tilde{P}(B) = P(B) = 0$, $\tilde{P}(A) + \tilde{P}(C) = 1$. With this definition, we do not need to differentiate the cases of $x_u < \infty$ and $x_u = \infty$ in the above lemma. In particular, when $x_u = \infty$, $\frac{1}{\bar{a}} > \frac{\lambda - P(\bar{A})}{1-\tilde{P}(\bar{A})}$ is automatically satisfied under the condition $P(\frac{d\tilde{P}}{dP} > \frac{1}{\lambda}) > 0$ thus the optimal is always achieved by the 'star-system'.

Corollary 2.3. *In the case where the probability space is atomless and the Radon Nikodým derivative* $\frac{d\tilde{P}}{dP}(\omega)$ *has continuous distribution, we have* $\tilde{P}(C) = P(C) = 0$ *and* $\tilde{P}(B) = 1 - \tilde{P}(A)$, *so set C will become irrelavant. The definition* $a_x = \sup\left\{a \,:\, \tilde{P}(B) \leq \frac{x_r - x_d}{x - x_d}\right\}$ *yields the precise equalities* $\tilde{P}(A_x) = \frac{x - x_r}{x - x_d}$ *and* $\tilde{P}(B_x) = \frac{x_r - x_d}{x - x_d}$. *The solution to the minimization problem*

$$\inf_{x_r \leq x \leq x_u} h(x)$$

where

$$h(x) = (x - x_d)(P(A_x) - \lambda\tilde{P}(A_x))$$

is

- *If* $\frac{d\tilde{P}}{dP} \leq \frac{1}{\lambda}$, $P-$*a.s., then the minimum is achieved by the 'r-system' and*

$$\inf_{x_r \leq x \leq x_u} h(x) = h(x_r) = 0.$$

- *Otherwise, if* $\frac{1}{\bar{a}} \leq \frac{\lambda - P(\bar{A})}{1-\tilde{P}(\bar{A})}$, *then the minimum is achieved by the 'bar-system' and*

$$\inf_{x_r \leq x \leq x_u} h(x) = h(x_u).$$

  If $\frac{1}{\bar{a}} > \frac{\lambda - P(\bar{A})}{1-\tilde{P}(\bar{A})}$, *then the minimum is achieved by the 'star-system' and*

$$\inf_{x_r \leq x \leq x_u} h(x) = h(x^*),$$

  where $x^* = \frac{x_r - x_d \tilde{P}(A^*)}{1-\tilde{P}(A^*)}$ *and* $A^* = \left\{\omega \in \Omega : \frac{d\tilde{P}}{dP}(\omega) > a^*\right\}$ *satisfies* $\frac{1}{a^*} = \frac{\lambda - P(A^*)}{1-\tilde{P}(A^*)}$.

Remark. Recall from the definitions in Remark 2.3, when $x_u = \infty$, the 'bar-system' corresponds to the set of parameters satisfying $\bar{a} = ess\inf \frac{d\tilde{P}}{dP}$, $\tilde{P}(B) = P(B) = 0$, $\tilde{P}(A) = P(A) = 1$. As in Lemma 2.2, we do not need to differentiate the cases of $x_u < \infty$ and $x_u = \infty$ in the above corollary. In particular, when $x_u = \infty$, $\frac{1}{\bar{a}} > \frac{\lambda - P(\bar{A})}{1-\tilde{P}(\bar{A})}$ is automatically satisfied under the condition $P(\frac{d\tilde{P}}{dP} > \frac{1}{\lambda}) > 0$ thus the optimal is always achieved by the 'star-system'.

Proof for Corollary 2.3. Let us first prove Corollary 2.3 in the continuous distribution case. Suppose $\frac{d\tilde{P}}{dP} \leq \frac{1}{\lambda}$, $P-$a.s. Then for any $x \in [x_r, x_u]$,

$$\tilde{P}(A_x) = \int_{A_x} \frac{d\tilde{P}}{dP}(\omega) dP(\omega) \leq \frac{1}{\lambda} P(A_x).$$

Thus $h(x) = (x - x_d)(P(A_x) - \lambda \tilde{P}(A_x)) \geq 0$. When $x = x_r$, $\tilde{P}(B_r) = \frac{x_r - x_d}{x_r - x_d} = 1$ and $P(A_r) = \tilde{P}(A_r) = 0$, therefore $h(x_r) = 0$. We conclude,

$$\inf_{x_r \leq x \leq x_u} h(x) = h(x_r) = 0.$$

Now suppose $P(\frac{d\tilde{P}}{dP} > \frac{1}{\lambda}) > 0$. Notice that when $\frac{d\tilde{P}}{dP}$ has a continuous distribution, we have the exact equalities $\tilde{P}(A_x) = \frac{x-x_r}{x-x_d}$ and $\tilde{P}(B_x) = \frac{x_r-x_d}{x-x_d}$, and we observe the following:

- A_x increases as x increases; a_x decreases as x increases.

  Define function $f(x) = \frac{x-x_r}{x-x_d}$. We see that $f(x)$ is an increasing function since $f'(x) = \frac{x_r-x_d}{(x-x_d)^2} > 0$. Notice that the probability function $\tilde{P}(A_x)$ is an increasing function of A_x, so $x \nearrow \Leftrightarrow f(x) \nearrow \Leftrightarrow \tilde{P}(A_x) \nearrow \Leftrightarrow A_x \nearrow$. In this special case where the Radon-Nikodým derivative $\frac{d\tilde{P}}{dP}(\omega)$ has continuous distribution but could skip values, a_x is a decreasing function of x where at times it can jump downward.

- $\tilde{P}(A_r) = 0$ and

$$\tilde{P}(\bar{A}) = \begin{cases} \frac{x_u-x_r}{x_u-x_d}, & x_u < \infty, \\ 1, & x_u = \infty. \end{cases}$$

  See Remark 2.3 and note $\tilde{P}(C) = 0$ in this case.

- $\frac{d\tilde{P}(A_x)}{dx} = \frac{x_r-x_d}{(x-x_d)^2}$; $D^-P(A_x) = \frac{1}{a_{x-}}\frac{d\tilde{P}(A_x)}{dx}$, $D^+P(A_x) = \frac{1}{a_{x+}}\frac{d\tilde{P}(A_x)}{dx}$.

  Use the definition of $f(x)$, $\frac{d\tilde{P}(A_x)}{dx} = f'(x) = \frac{x_r-x_d}{(x-x_d)^2}$. Notice that a_x may not be a continuous function of x, but the left-hand and right-hand limit a_{x-} and a_{x+} exist for all x because it is a decreasing function. In fact, we have $\tilde{P}(a_{x+} < \frac{d\tilde{P}}{dP}(\omega) < a_{x-}) = 0$, see Fig. 2.1. Since P and $\tilde{P}$ are equivalent, we also have $P(a_{x+} < \frac{d\tilde{P}}{dP}(\omega) < a_{x-}) = 0$ and $P(A_x) = P(\frac{d\tilde{P}}{dP}(\omega) > a_{x-}) = P(\frac{d\tilde{P}}{dP}(\omega) > a_{x+})$. If we denote the left-hand and right-hand derivatives as

$$D^-P(A_x) = \lim_{\epsilon \nearrow 0} \frac{P(A_{x+\epsilon}) - P(A_x)}{\epsilon},$$
$$D^+P(A_x) = \lim_{\epsilon \searrow 0} \frac{P(A_{x+\epsilon}) - P(A_x)}{\epsilon},$$

  we have

$$\begin{aligned} D^-P(A_x) &= \lim_{\epsilon \nearrow 0} \frac{P(A_{x+\epsilon}) - P(A_x)}{\epsilon} = \lim_{\epsilon \nearrow 0} \frac{E[1_{A_{x+\epsilon}}] - E[1_{A_x}]}{\epsilon} \\ &= \lim_{\epsilon \nearrow 0} \frac{\tilde{E}[\frac{dP}{d\tilde{P}} 1_{\frac{d\tilde{P}}{dP} > a_{x+\epsilon}}] - \tilde{E}[\frac{dP}{d\tilde{P}} 1_{\frac{d\tilde{P}}{dP} > a_{x-}}]}{\epsilon} = \lim_{\epsilon \nearrow 0} \frac{\tilde{E}[\frac{dP}{d\tilde{P}} 1_{a_{x-} < \frac{d\tilde{P}}{dP} \le a_{x+\epsilon}}]}{-\epsilon}. \end{aligned}$$

Since

$$\frac{\tilde{E}[\frac{dP}{d\tilde{P}}1_{a_{x-}<\frac{d\tilde{P}}{dP}\le a_{x+\epsilon}}]}{-\epsilon} \ge \frac{1}{a_{x+\epsilon}}\frac{\tilde{E}[1_{a_{x-}<\frac{d\tilde{P}}{dP}\le a_{x+\epsilon}}]}{-\epsilon}$$
$$= \frac{1}{a_{x+\epsilon}}\frac{\tilde{P}(A_{x+\epsilon})-\tilde{P}(A_x)}{\epsilon} \to \frac{1}{a_{x-}}\frac{d\tilde{P}(A_x)}{dx},$$

and

$$\frac{\tilde{E}[\frac{dP}{d\tilde{P}}1_{a_{x-}<\frac{d\tilde{P}}{dP}\le a_{x+\epsilon}}]}{-\epsilon} < \frac{1}{a_{x-}}\frac{\tilde{E}[1_{a_{x-}<\frac{d\tilde{P}}{dP}\le a_{x+\epsilon}}]}{-\epsilon}$$
$$= \frac{1}{a_{x-}}\frac{\tilde{P}(A_{x+\epsilon})-\tilde{P}(A_x)}{\epsilon} \to \frac{1}{a_{x-}}\frac{d\tilde{P}(A_x)}{dx},$$

as $\epsilon \nearrow 0$. We conclude that the left derivative is

$$D^{-}P(A_x) = \frac{1}{a_{x-}}\frac{d\tilde{P}(A_x)}{dx}.$$

Similarly, the right derivative is

$$D^{+}P(A_x) = \frac{1}{a_{x+}}\frac{d\tilde{P}(A_x)}{dx}.$$

If a_x is continuous at x, i.e., $a_{x+} = a_{x-} = a_x$, then the derivative exists

$$\frac{dP(A_x)}{dx} = \frac{1}{a_x}\frac{d\tilde{P}(A_x)}{dx} = \frac{x_r - x_d}{a_x(x-x_d)^2}.$$

Now let us turn to the first and second derivatives of the function we would like to minimize:

$$h(x) = (x - x_d)(P(A_x) - \lambda\tilde{P}(A_x)),$$

and show it to be a convex function. On $x \in (x_r, x_u)$, when a_x is continuous at x, we have

$$\begin{aligned}
h'(x) &= (P(A_x) - \lambda \tilde{P}(A_x)) + (x - x_d)\left(\frac{dP(A_x)}{dx} - \lambda \frac{d\tilde{P}(A_x)}{dx}\right) \\
&= \left(P(A_x) - \lambda \frac{x - x_r}{x - x_d}\right) + (x - x_d)\left(\frac{x_r - x_d}{a_x (x - x_d)^2} - \lambda \frac{x_r - x_d}{(x - x_d)^2}\right) \\
&= P(A_x) - \lambda \frac{x - x_r}{x - x_d} + \left(\frac{1}{a_x} - \lambda\right)\frac{x_r - x_d}{x - x_d} \\
&= P(A_x) - \lambda + \frac{1}{a_x}(1 - \tilde{P}(A_x)).
\end{aligned}$$

When a_x is discontinuous at x, we can define the left- and right-derivatives

$$\begin{aligned}
D^- h(x) &= \lim_{\epsilon \nearrow 0} \frac{h(x+\epsilon) - h(x)}{\epsilon}, \\
D^+ h(x) &= \lim_{\epsilon \searrow 0} \frac{h(x+\epsilon) - h(x)}{\epsilon}.
\end{aligned}$$

Similar to the above calculation, we get

$$\begin{aligned}
D^- h(x) &= P(A_x) - \lambda + \frac{1}{a_{x-}}(1 - \tilde{P}(A_x)), \\
D^+ h(x) &= P(A_x) - \lambda + \frac{1}{a_{x+}}(1 - \tilde{P}(A_x)).
\end{aligned}$$

When a_x is continuous at x, $\tilde{P}(A_x) = \tilde{P}(\frac{d\tilde{P}}{dP}(\omega) > a_x) = 1 - \tilde{P}(\frac{d\tilde{P}}{dP}(\omega) \leq a_x) = 1 - \tilde{F}(a_x)$, where $\tilde{F}(\cdot)$ is the cumulative distribution function of the Radon Nikodým derivative $\frac{d\tilde{P}}{dP}$. Since $\tilde{P}(A_x) = \frac{x - x_r}{x - x_d}$, $\tilde{F}(a_x) = \frac{x_r - x_d}{x - x_d}$. We have also started by assuming $\frac{d\tilde{P}}{dP}$ has a continuous distribution, therefore the derivative of $\tilde{F}(\cdot)$ exists and is the probability density function $\tilde{f}(\cdot)$. When $a_{x-} = a_{x+}$, $\tilde{P}(A_x)$ is strictly increasing as x increases, thus $\tilde{f}(a_x) > 0$. By Inverse Differentiation Theorem, the derivative of a_x exists and can be computed as

$$(a_x)' = -\frac{x_r - x_d}{\tilde{f}(a_x)(x - x_d)^2} < 0.$$

By Chain Rule, we know

$$\left(\frac{1}{a_x}\right)' = -\frac{a_x'}{a_x^2} > 0.$$

Now we can compute the second derivative of $h(x)$:

$$\begin{aligned} h''(x) &= \frac{dP(A_x)}{dx} + \left(\frac{1}{a_x}\right)'(1-\tilde{P}(A_x)) - \frac{1}{a_x}\frac{d\tilde{P}(A_x)}{dx} \\ &= \left(\frac{1}{a_x}\right)'(1-\tilde{P}(A_x)) > 0. \end{aligned}$$

Here $1-\tilde{P}(A_x) = \tilde{P}(B_x) = \frac{x_r - x_d}{x - x_d}$ is strictly positive on the set $x \in (x_r, x_u)$. Clearly, the second derivative indicates that $h'(x)$ is strictly increasing at those points $x \in (x_r, x_u)$ where a_x is continuous.

When a_x is discontinuous, we have

$$D^- h(x) = P(A_x) - \lambda + \frac{1}{a_{x-}}(1-\tilde{P}(A_x)) < D^+ h(x) = P(A_x) - \lambda + \frac{1}{a_{x+}}(1-\tilde{P}(A_x)).$$

We recognize that this is a kink point for $h(x)$. Finally, we conclude $h(x)$ is convex on (x_r, x_u).

When $x_u < \infty$, it is easy to see that $h(x)$ is continuous at both left and right end points with the definition in Remark 2.3. Therefore, it is convex on the closed interval $[x_r, x_u]$. If we can find $x^* \in [x_r, x_u]$, where $0 \in [D^- h(x^*), D^+ h(x^*)]$, then it is the minimum. Otherwise if $D^+ h(x_r) \geq 0$, then the infimum is obtained at $x = x_r$; if $D^- h(x_u) \leq 0$, then the infimum is obtained at $x = x_u$. If the derivative of $h(x)$ exists at $x = x^*$, then the condition $0 \in [D^- h(x^*), D^+ h(x^*)]$ collapses to $h'(x^*) = 0$, or equivalently, $\frac{1}{a^*} = \frac{\lambda - P(A^*)}{1 - \tilde{P}(A^*)}$. When the derivative does not exist, the condition that $0 \in [D^- h(x^*), D^+ h(x^*)]$ corresponds to $\frac{1}{a^{*-}} < \frac{\lambda - P(A^*)}{1-\tilde{P}(A^*)}$, and $\frac{1}{a^{*+}} > \frac{\lambda - P(A^*)}{1-\tilde{P}(A^*)}$. In this case, we can always find an $\frac{1}{a^*} \in [\frac{1}{a_{x^*-}}, \frac{1}{a_{x^*+}}]$ where $\frac{1}{a^*} = \frac{\lambda - P(A^*)}{1-\tilde{P}(A^*)}$, and the corresponding x^* can be computed from the equation $\tilde{P}(A^*) = \frac{x^* - x_r}{x^* - x_d}$, i.e., $x^* = \frac{x_r - x_d \tilde{P}(A)}{1 - \tilde{P}(A)}$.

Recall that $P(\frac{d\tilde{P}}{dP} > \frac{1}{\lambda}) > 0$. Recall from Remark 2.3, $\tilde{P}(A_r) = P(A_r) = 0$, $a_r = ess\sup \frac{d\tilde{P}}{dP} > \frac{1}{\lambda}$. So $D^+ h(x_r) = P(A_r) - \lambda + \frac{1}{a_r}(1 - \tilde{P}(A_r)) = -\lambda + \frac{1}{a_r} < 0$. On the other end, $D^- h(x_u) = P(\bar{A}) - \lambda + \frac{1}{\bar{a}}(1 - \tilde{P}(\bar{A}))$. We have $D^- h(x_u) > 0$ if and only if $\frac{1}{\bar{a}} > \frac{\lambda - P(\bar{A})}{1 - \tilde{P}(\bar{A})}$. In this case, the minimum occurs at $x^* \in (x_r, x_u)$ where $\frac{1}{a^*} = \frac{\lambda - P(A^*)}{1-\tilde{P}(A^*)}$. If $\frac{1}{\bar{a}} \leq \frac{\lambda - P(\bar{A})}{1-\tilde{P}(\bar{A})}$, $h(x)$ decreases on $[x_r,\, x_u]$, and the minimum is achieved at the right end point with the value $h(x_u)$.

When $x_u = \infty$, $h(x)$ is convex on $[x_d, \infty)$. As $x \to \infty$, $P(A_x) \to 1$, $\tilde{P}(A_x) \to 1$, $a_x \to ess\inf \frac{d\tilde{P}}{dP}$. $D^+h(x)$ becomes positive sooner or later, and the minimum is obtained in the interior where we define the 'star-system'. *Q.E.D.*

Proof for Lemma 2.2. As in the proof for Corollary 2.3, let $\tilde{F}(\cdot)$ be the cumulative distribution function of the Radon Nikodým derivative $\frac{d\tilde{P}}{dP}$. Then for fixed x, we have $\tilde{F}(a_x) = 1 - \tilde{P}(A_x)$. In the proof for Corollary 2.3, we have assumed that $\frac{d\tilde{P}}{dP}$ has a continuous distribution. This essentially dealt with case where $\tilde{F}(\cdot)$ is continuous: it could either be strictly increasing or flat. Now to deal with the general case, we only need to discuss the remaining case where $\tilde{F}(\cdot)$ has a jump, i.e., there is a point mass at $\frac{d\tilde{P}}{dP} = a_x$, see Fig. 2.1.

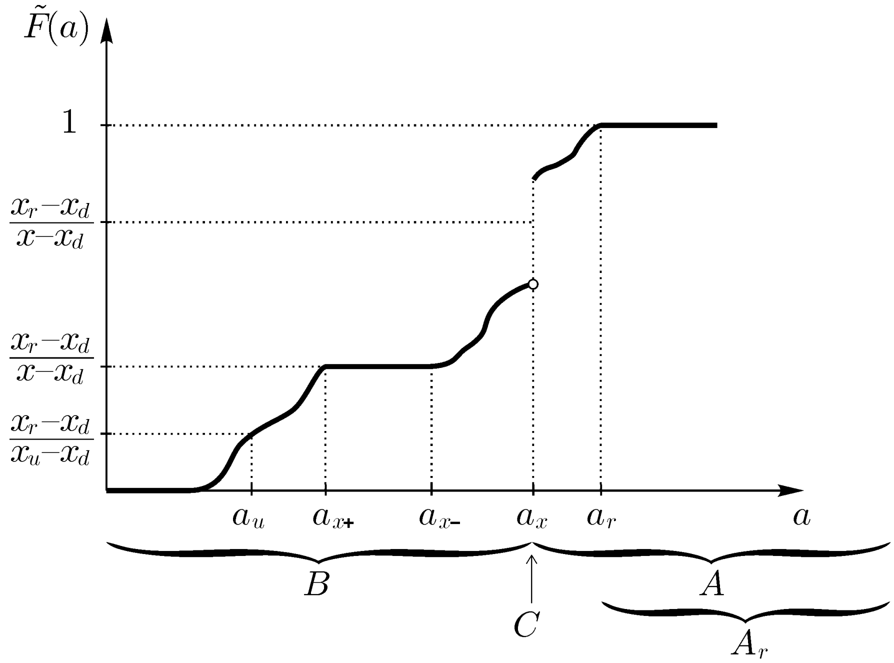

Figure 2.1: $\tilde{F}(a)$ is the cumulative distribution function of the Radon-Nikodým derivative $\frac{d\tilde{P}}{dP}$.

Recall the definitions

$$
\begin{aligned}
a_x &= \sup\left\{a \,:\, \tilde{P}(B) \le \frac{x_r - x_d}{x - x_d}\right\},\\
A_x &= \left\{\omega \in \Omega : \frac{d\tilde{P}}{dP}(\omega) > a_x\right\}, C_x = \left\{\omega \in \Omega : \frac{d\tilde{P}}{dP}(\omega) = a_x\right\}, B_x = \left\{\omega \in \Omega : \frac{d\tilde{P}}{dP}(\omega) < a_x\right\},\\
k_x &= \frac{x_r - x_d\tilde{P}(A_x) - x\tilde{P}(B_x)}{\tilde{P}(C_x)} 1_{\{\tilde{P}(C_x)>0\}},\\
h(x) &= (x - x_d)(P(A_x) - \lambda\tilde{P}(A_x)) + (x - k_x)(P(C_x) - \lambda\tilde{P}(C_x)),
\end{aligned}
$$

and we would like to find

$$\inf_{x_r \le x \le x_u} h(x).$$

When $\frac{d\tilde{P}}{dP}$ has a point mass at a_x, i.e., $\tilde{P}(\frac{d\tilde{P}}{dP} = a_x) = \tilde{P}(C_x) > 0$, the distribution function $\tilde{F}$ has a jump at a_x: $\tilde{F}(a_x) - \tilde{F}(a_{x^-}) = \tilde{P}(C_x)$. As in the proof for Corollary 2.3, we first discuss the case $\frac{d\tilde{P}}{dP} \le \frac{1}{\lambda}$, $P-$a.s. Similarly we can show $\tilde{P}(A_x) \le \frac{1}{\lambda}P(A_x)$ and $\tilde{P}(C_x) \le \frac{1}{\lambda}P(C_x)$ for $x \in [x_r, x_u]$. It is easy to check that $x - k_x \ge 0$ when $x_r \le x \le x_u$, so $h(x) \ge 0$ on $[x_r,\, x_u]$. Also notice that $h(x_r) = 0$, we conclude,

$$\inf_{x_r \le x \le x_u} h(x) = h(x_r) = 0.$$

Now suppose $P(\frac{d\tilde{P}}{dP} > \frac{1}{\lambda}) > 0$. If we can find an a such that $\tilde{P}(A) = \frac{x - x_r}{x - x_d}$ exactly and $\tilde{P}(B) = \frac{x_r - x_d}{x - x_d}$ exactly, then $\tilde{P}(C) = 0$. This is the situation where the distribution of $\frac{d\tilde{P}}{dP}$ is continuous. If such an a cannot be found then we have the situation where $\tilde{P}(C) > 0$, and this corresponds to the situation where there is a point mass that we need to work out. Therefore, we need to discuss three cases:

1 $\tilde{P}(A_x) = \frac{x - x_r}{x - x_d}$, $\tilde{P}(B_x) < \frac{x_r - x_d}{x - x_d}$ and $\tilde{P}(C_x) > 0$,

2 $\tilde{P}(A_x) < \frac{x - x_r}{x - x_d}$, $\tilde{P}(B_x) = \frac{x_r - x_d}{x - x_d}$ and $\tilde{P}(C_x) > 0$,

3 $\tilde{P}(A_x) < \frac{x - x_r}{x - x_d}$, $\tilde{P}(B_x) < \frac{x_r - x_d}{x - x_d}$ and $\tilde{P}(C_x) > 0$.

We first deal with the last case where we fix a_1, A_1, B_1 and C_1 where $\tilde{P}(A_1) < \frac{x - x_r}{x - x_d}$, $\tilde{P}(B_1) < \frac{x_r - x_d}{x - x_d}$ and $\tilde{P}(C_1) = \tilde{P}(\frac{d\tilde{P}}{dP} = a_1) > 0$ and $x_r = x_d\tilde{P}(A_1) + x\tilde{P}(B_1) + k_x\tilde{P}(C_1)$ is satisfied. As

k_x decreases from x to x_d, x increases from $x_1 = \frac{x_r - x_d \tilde{P}(A_1)}{\tilde{P}(B_1)+\tilde{P}(C_1)}$ to $x_2 = \frac{x_r - x_d(\tilde{P}(A_1)+\tilde{P}(C_1))}{\tilde{P}(B_1)}$, while at the same time A_1, B_1, C_1 and a_1 remain unchanged. The derivative of $h(x)$ on the interval $x \in (x_1,\, x_2)$ is easily calculated as

$$\begin{aligned}
h'(x) &= (P(A_1) - \lambda \tilde{P}(A_1)) + (1 - \frac{dk_x}{dx})(P(C_1) - \lambda \tilde{P}(C_1)) \\
&= (P(A_1) - \lambda \tilde{P}(A_1)) + (1 + \frac{\tilde{P}(B_1)}{\tilde{P}(C_1)})(P(C_1) - \lambda \tilde{P}(C_1)) \\
&= (P(A_1) - \lambda \tilde{P}(A_1)) + (1 - \tilde{P}(A_1))(\frac{P(C_1)}{\tilde{P}(C_1)} - \lambda) \\
&= (P(A_1) - \lambda \tilde{P}(A_1)) + (1 - \tilde{P}(A_1))(\frac{1}{a_1} - \lambda) \\
&= P(A_1) - \lambda + \frac{1}{a_1}(1 - \tilde{P}(A_1)).
\end{aligned}$$

The formula reads exactly the same as the one in the continuous case except that $h'(x)$ is constant now on this open interval, and the originally curved $h(x)$ degenerates to a straight line. At the end point $x = x_2$, k_x dropped to x_d and we have $\tilde{P}(B_1) = \frac{x_r - x_d}{x - x_d}$. Still we have $\tilde{P}(C_1) = \tilde{P}(\frac{d\tilde{P}}{dP} = a_1) > 0$ and $\tilde{P}(A_1) < \frac{x - x_r}{x - x_d}$. This corresponds to the second case in the above list. There are three possibilities at this point.

(a) There is a point $a_2 < a_1$ where $\tilde{F}(a)$ is constant on the interval (a_2, a_1) and has a jump at a_2, i.e., $\tilde{F}(a_2-) < \tilde{F}(a_2)$.

(b) There is a point $a_{1+} < a_1$ and a_{1+} is the smallest number such that $\tilde{F}(a)$ is constant on the interval (a_{1+}, a_1) and has no jump at a_{1+}.

(c) $\tilde{F}(a)$ is strictly increasing to the left of a_1.

These three cases correspond to how the function $h(x)$ at $x = x_2$ is connected to its right-hand side:

(a) A kink connection to another line with different slope.

(b) A kink connection to a curve.

(c) A smooth connection to a curve.

If $\tilde{F}(a)$ is flat until it encounters another point mass at a_2 as in case (a), then the old sets A_1 and C_1 combine to produce the new set $A_2 = A_1 \bigcup C_1$ and $C_2 = \{\omega : \frac{d\tilde{P}}{dP}(\omega) = a_2\}$. The left derivative at this point is computed above

$$D^- h(x_2) = P(A_1) - \lambda + \frac{1}{a_1}(1 - \tilde{P}(A_1)).$$

The right derivative is the same formula applied to the new sets:

$$D^+ h(x_2) = P(A_2) - \lambda + \frac{1}{a_2}(1 - \tilde{P}(A_2)).$$

The difference

$$\begin{aligned}
&D^+ h(x_2) - D^- h(x_2) \\
&= P(A_2) - \lambda + \frac{1}{a_2}(1 - \tilde{P}(A_2)) - \left(P(A_1) - \lambda + \frac{1}{a_1}(1 - \tilde{P}(A_1))\right) \\
&= P(C_1) + (\frac{1}{a_2} - \frac{1}{a_1})(1 - \tilde{P}(A_2)) - \frac{1}{a_1}\tilde{P}(C_1) \\
&= (\frac{1}{a_2} - \frac{1}{a_1})\tilde{P}(B_1) \geq 0.
\end{aligned}$$

Here we used the definition of set $C_1 = \left\{\frac{d\tilde{P}}{dP} = a_1\right\}$ to yield $P(C_1) - \frac{1}{a_1}\tilde{P}(C_1) = 0$, and $\frac{1}{a_2} \geq \frac{1}{a_1}$ since $a_2 < a_1$. Therefore, the convexity of $h(x)$ at $x = x_2$ is kept. In case (c), $\tilde{F}(a)$ is increasing on the left of a_1 and we shall now return to the continuous case in the proof for Corollary 2.3 to conclude that

$$D^+ h(x_2) = P(A_1) - \lambda + \frac{1}{a_{1+}}(1 - \tilde{P}(A_1)) \geq h'(x), \quad \text{for } x \in (x_1, x_2),$$

because $a_{1+} \leq a_1$. In case (b),

$$D^+ h(x_2) = P(A_2) - \lambda + \frac{1}{a_{1+}}(1 - \tilde{P}(A_2)),$$

and the proof for

$$D^+ h(x_2) \geq D^- h(x_2)$$

is similar to that of case (a). Thus $h(x)$ is convex on $x \in (x_1,\, x_2]$.

Now consider the other end point $x = x_1$. Here $k_x = x$ and we have $\tilde{P}(A_1) = \frac{x - x_r}{x - x_d}$, $\tilde{P}(C_1) = \tilde{P}(\frac{d\tilde{P}}{dP} = a_1) > 0$ and $\tilde{P}(B_1) < \frac{x_r - x_d}{x - x_d}$. This corresponds to the first case in the above list. We can carry out similar discussion as in the second case and conclude that

$$D^- h(x_1) \leq h'(x), \quad \text{for } x \in (x_1,\, x_2),$$

thus we have the convexity of function $h(x)$ on the closed interval $[x_1,\, x_2]$. In summary: when there is a point mass at $\frac{d\tilde{P}}{dP} = a_x$, i.e., $\tilde{F}(a_x-) < \tilde{F}(a_x)$ the convex function $h(x)$ becomes linear; in contrast to the fact that when $\tilde{F}(a)$ is flat, $h(x)$ will have a kink point where its derivative jumps. As shown in Fig. 2.2, in a case like the Binomial model where there are only point masses, $h(x)$ is a piecewise constant convex function; in a case like the Black-Scholes model where the distribution is continuous and spans the whole positive part of the real line, $h(x)$ is a continuously differentiable convex function. In general, these two pictures can be mixed. In any case, combining the results we have just shown and those in the proof of Corollary 2.3, we know that $h(x)$ is convex all the time on $x \in [x_r, x_u]$.

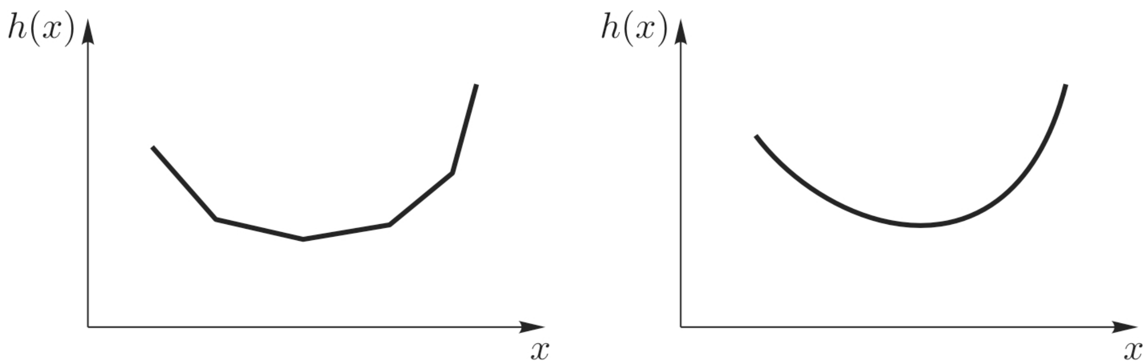

Figure 2.2: The left picture is how $h(x)$ look like in the Binomial model; the right pictures is for the Black-Scholes model.

The discussion in the proof of Corollary 2.3 dealt with minimizing $h(x)$ when $h(x)$ is curved and contains kink points. The optimal condition is the existence of an a^* such that

$$\frac{1}{a^*} = \frac{\lambda - P(A^*)}{1 - \tilde{P}(A^*)}. \tag{2.9}$$

Now we deal with the situation where $h(x)$ is a straight line on $[x_1,\, x_2]$ where the minimum can only occur at end points. For downward slopping case where $h'(x) < 0$ on (x_1, x_2), the minimum occurs at the right-end point x_2 where either a kink or a smooth connection to a curved situation can happen, or a kink to another line can happen. When it is a smooth connection, x_2 can not be a global minimum because $h'(x_2)$ exists and is strictly negative. When it is a kink to a smooth curve, then $0 \in [D^- h(x_2),\, D^+ h(x_2)]$ corresponds to $\frac{1}{a_1} < \frac{\lambda - P(A_1)}{1-\tilde{P}(A_1)}$ and $\frac{1}{a_{1+}} \geq \frac{\lambda - P(A_2)}{1-\tilde{P}(A_2)}$. When it is connected with a kink to another line, a observes a jump from a_1 to a_2, the set A jumps from A_1 to A_2 and k jumps from x_d to x_2. The optimal condition $0 \in [D^- h(x_2),\, D^+ h(x_2)]$ corresponds to $\frac{1}{a_1} < \frac{\lambda - P(A_1)}{1-\tilde{P}(A_1)}$ and $\frac{1}{a_2} \geq \frac{\lambda - P(A_2)}{1-\tilde{P}(A_2)}$. In both cases, the optimal a can be expressed as

$$a^* = \sup\left\{a \,:\, \frac{1}{a} \geq \frac{\lambda - P(A)}{1-\tilde{P}(A)}\right\}, \tag{2.10}$$

and $x_2 = \frac{x_r - x_d \tilde{P}(A_2)}{1-\tilde{P}(A_2)} = \frac{x_r - x_d \tilde{P}(A^*)}{1-\tilde{P}(A^*)} = k$. The case of upward sloping can be similarly analyzed. It is easy to check the conditions when the slope is zero. Recognizing (2.10) is a generalization of (2.9) we had for the continuous case, we arrive at the optimal condition

$$a^* = \sup\left\{a \,:\, \frac{1}{a} \geq \frac{\lambda - P(A)}{1-\tilde{P}(A)}\right\},$$

and $x^* = \frac{x_r - x_d \tilde{P}(A^*)}{1-\tilde{P}(A^*)} = k^*$. The only remaining issue to be checked is the condition for 'bar-system' to be optimal when x_u corresponds to a point mass at $\bar{a}$. The arguments given in the proof of Corollary 2.3 for the continuous case work here too both for finite and infinite x_u. For example, in the case $x_u = \infty$, we have defined in Remark 2.3 that $\bar{a} = ess\inf \frac{d\tilde{P}}{dP}$. Therefore, we have $\tilde{P}(\bar{B}) = P(\bar{B}) = 0$, and $\tilde{P}(\bar{C}) + \tilde{P}(\bar{A}) = P(\bar{C}) + P(\bar{A}) = 1$ where $\bar{C} = \left\{\omega \in \Omega \,:\, \frac{d\tilde{P}}{dP}(\omega) = \bar{a}\right\}$. Since this is a line segment for $h(x)$, we have already calculated its slope

$$\begin{aligned} h'(x) &= P(\bar{A}) - \lambda + \frac{1}{\bar{a}}(1 - \tilde{P}(\bar{A})) = P(\bar{A}) - \lambda + \frac{1}{\bar{a}}\tilde{P}(\bar{C}) \\ &= P(\bar{A}) - \lambda + \frac{1}{\bar{a}}\bar{a}P(\bar{C}) = 1 - \lambda > 0. \end{aligned}$$

So the optimal will be obtained by the 'star-system' in the interior. *Q.E.D.*

Theorem 2.4 (Solution to CVaR Minimization Problem)**.** *Define the sets A, B, C and the numbers a_x, k_x and the sets A_x, B_x, C_x for fixed number x the same way as in Lemma 2.2 and Theorem 2.1. Denote the 'r-system', 'bar-system' and 'star-system' as in Lemma 2.2. The solution to problem (2.5) and consequently our main problem (1.5) is as follows:*

- *If $\frac{d\tilde{P}}{dP} \leq \frac{1}{\lambda}$, $P-$a.s., then $X^* = x_r$ is the optimal final portfolio value, and the minimal risk is $CVaR_\lambda(X^*) = -x_r$.*

- *Otherwise, find the 'bar-system' using definitions*

$$\bar{a} = \begin{cases} \sup\left\{a : \tilde{P}(B) \leq \frac{x_r - x_d}{x_u - x_d}\right\}, & x_u < \infty, \\ \operatorname{ess\,inf} \frac{d\tilde{P}}{dP}, & x_u = \infty. \end{cases}$$

$$\bar{A} = \left\{\omega \in \Omega : \frac{d\tilde{P}}{dP}(\omega) > \bar{a}\right\}, \quad \bar{B} = \left\{\omega \in \Omega : \frac{d\tilde{P}}{dP}(\omega) < \bar{a}\right\},$$

$$\bar{C} = \left\{\omega \in \Omega : \frac{d\tilde{P}}{dP}(\omega) = \bar{a}\right\}, \quad \bar{k} = \frac{x_r - x_d\tilde{P}(\bar{A}) - x_u\tilde{P}(\bar{B})}{\tilde{P}(\bar{C})}\mathbf{I}_{\{\tilde{P}(\bar{C})>0\}}.$$

  - *If $\frac{1}{\bar{a}} \leq \frac{\lambda - P(\bar{A})}{1-\tilde{P}(\bar{A})}$, then the optimal risk is achieved by this 'bar-system': the optimal final portfolio value and the associated minimal risk are*

$$X^* = x_d\mathbf{I}_{\bar{A}} + \bar{k}\mathbf{I}_{\bar{C}} + x_u\mathbf{I}_{\bar{B}},$$
$$CVaR_\lambda(X^*) = -x_r + \frac{1}{\lambda}[(x_u - x_d)(P(\bar{A}) - \lambda\tilde{P}(\bar{A})) + (x_u - \bar{k})(P(\bar{C}) - \lambda\tilde{P}(\bar{C}))].$$

  - *If $\frac{1}{\bar{a}} > \frac{\lambda - P(\bar{A})}{1-\tilde{P}(\bar{A})}$, then the optimal risk is achieved by the 'star-system' obtained by $a^* = \sup\left\{a : \frac{1}{a} \geq \frac{\lambda - P(A)}{1-\tilde{P}(A)}\right\}$, $A^* = \left\{\omega \in \Omega : \frac{d\tilde{P}}{dP}(\omega) > a^*\right\}$, $x^* = \frac{x_r - x_d\tilde{P}(A^*)}{1-\tilde{P}(A^*)}$. The optimal final portfolio value and the associated minimal risk are*

$$X^* = x_d\mathbf{I}_{A^*} + x^*\mathbf{I}_{A^{*c}},$$
$$CVaR_\lambda(X^*) = -x_r + \frac{1}{\lambda}(x^* - x_d)(P(A^*) - \lambda\tilde{P}(A^*)).$$

Remark. For the continuous case as in Remark 2.3 we can simplify the results as following

- If $\frac{d\tilde{P}}{dP} \leq \frac{1}{\lambda}$, $P-$a.s., then $X^* = x_r$ is the optimal final portfolio value, and the minimal risk is $CVaR_\lambda(X^*) = -x_r$.

- Otherwise, find the 'bar-system' using definitions

$$\bar{a} = \begin{cases} \sup\left\{a : \tilde{P}(B) \le \frac{x_r - x_d}{x_u - x_d}\right\}, & x_u < \infty, \\ ess\inf \frac{d\tilde{P}}{dP}, & x_u = \infty. \end{cases}$$
$$\bar{A} = \left\{\omega \in \Omega : \frac{d\tilde{P}}{dP}(\omega) > \bar{a}\right\}.$$

  - If $\frac{1}{\bar{a}} \le \frac{\lambda - P(\bar{A})}{1 - \tilde{P}(\bar{A})}$, then the optimal portfolio is achieved by this 'bar-system':the optimal final portfolio value and the associated minimal risk are

$$X^* = x_d \mathbf{I}_{\bar{A}} + x_u \mathbf{I}_{\bar{A}^c},$$
$$CVaR_\lambda(X^*) = -x_r + \frac{1}{\lambda}(x_u - x_d)(P(\bar{A}) - \lambda\tilde{P}(\bar{A})).$$

  - If $\frac{1}{\bar{a}} > \frac{\lambda - P(\bar{A})}{1 - \tilde{P}(\bar{A})}$, then the optimal risk is achieved by the 'star-system' obtained by $a^* = \{a : \frac{1}{a} = \frac{\lambda - P(A)}{1 - \tilde{P}(A)}\}$, $A^* = \left\{\omega \in \Omega : \frac{d\tilde{P}}{dP}(\omega) > a^*\right\}$. $x^* = \frac{x_r - x_d\tilde{P}(A^*)}{1 - \tilde{P}(A^*)}$. The optimal final portfolio value and the associated minimal risk are

$$X^* = x_d \mathbf{I}_{A^*} + x^* \mathbf{I}_{A^{*c}},$$
$$CVaR_\lambda(X^*) = -x_r + \frac{1}{\lambda}(x^* - x_d)(P(A^*) - \lambda\tilde{P}(A^*)).$$

Proof. Theorem 2.1 gives the solution to the shortfall problem in **Step 1**. Leveraging these results, we have discussed in Remark 2.3 that the solution for **step 2** is achieved by finding the solution to the third case of

$$\frac{1}{\lambda} \inf_{x_r \le x \le x_u} (v(x) - \lambda x) = -x_r + \frac{1}{\lambda} \inf_{x_r < x \le x_u} h(x),$$

where $h(x) = (x - x_d)(P(A_x) - \lambda\tilde{P}(A_x)) + (x - k_x)(P(C_x) - \lambda\tilde{P}(C_x))$. Now combine the solution to

$$\inf_{x_r \le x \le x_u} h(x),$$

found in Lemma 2.2, we quickly arrive at the conclusion. *Q.E.D.*

2.4 Application to Some Complete Market Examples

2.4.1 Binomial Model

Consider a recombining binomial tree, we have the following dynamics for the stock S_n and the self-financing portfolio X_n:

$$\begin{aligned} S_{n+1}(H) &= uS_n, \text{ with } P(\omega_n = H) = p \text{ and } \tilde{P}(\omega_n = H) = \tilde{p}, \\ S_{n+1}(T) &= dS_n, \text{ with } P(\omega_n = T) = q \text{ and } \tilde{P}(\omega_n = T) = \tilde{q}, \\ X_{n+1} &= \xi_n S_{n+1} + (X_n - \xi_n S_n)(1+r), \end{aligned}$$

where $\tilde{p}$, $\tilde{q}$ are risk-neutral probabilities, p, q are physical probabilities, u, d are the step sizes for up move and down move respectively, and r is the risk-free interest earned for one time step. Given initial stock price S_0 and initial portfolio value X_0, our main goal is first to find

$$CVaR_\lambda(X_N^*) := \inf_{\xi_n} CVaR_\lambda(X_N) \text{ s.t. } \tilde{E}[X_N] = x_r,\ x_d \le X_N \le x_u, \tag{2.11}$$

where the constants satisfy $-\infty < x_d < x < x_u \le \infty$, and then to find the corresponding dynamic hedging ξ_n.

Denote the final states of an N-step binomial tree as $\Omega = \{\omega_1, \omega_2, ..., \omega_{2^N}\}$, where we require the Radon-Nikodým derivatives to be arranged in descending order, i.e., $\frac{\tilde{P}(\omega_i)}{P(\omega_i)} \ge \frac{\tilde{P}(\omega_j)}{P(\omega_j)}$, $\forall i < j$. Note that in this case where the tree is recombining, the Radon-Nikodým derivative $\frac{\tilde{P}(\omega_i)}{P(\omega_i)} = \frac{\tilde{p}^m \tilde{q}^{N-m}}{p^m q^{N-m}}$ depends on the total number of up moves m in state ω_i, and is monotonic in m depending on whether $\frac{\tilde{p}}{p} > 1$ or $\frac{\tilde{p}}{p} < 1$. We group the distinct final nodes into sets $\Omega = \Omega_0 \cup \Omega_1 \cup ...\Omega_N$. Suppose $\frac{\tilde{p}}{p} < 1$, $\Omega_0 = \{\omega_1\}$ contains the only state where there are N down moves, $\Omega_1 = \{\omega_2, \omega_3, \ldots \omega_{N+1}\}$ contains those states where there are $N-1$ down moves and one up move. In general, $\Omega_k = \left\{\omega \in \Omega : \frac{\tilde{P}(\omega_i)}{P(\omega_i)} = \frac{\tilde{p}^{N-k}\tilde{q}^k}{p^{N-k}q^k}\right\}$ contains $\frac{N!}{k!(N-k)!}$ states. If instead $\frac{\tilde{p}}{p} > 1$, the order will be reversed and w_1 is the state with N up moves.

Proposition 2.5. *Following the definitions of 'r-system', 'bar-system' and 'star-system' as in Theorem 2.4. We can compute the solution to problem (2.11) and the corresponding dynamic hedging strategy ξ_n in an N-step Binomial Model as below:*

- *If either $1 < \frac{\tilde{p}}{p} \leq \sqrt[N]{\frac{1}{\lambda}}$ or $1 < \frac{\tilde{q}}{q} \leq \sqrt[N]{\frac{1}{\lambda}}$ holds, then the optimal portfolio is $X_N^* = x_r$ and the optimal strategy is $\xi_n = 0$, for all $n = 0, 1, ..., N-1$. The corresponding minimal risk is $CVaR_\lambda(X_N^*) = -x_r$.*

- *Otherwise, if $x_u < \infty$, find the 'bar-system' using definitions*

$$i_u = \begin{cases} N+1, & \tilde{P}(\Omega_N) > \frac{x_r - x_d}{x_u - x_d}, \\ \min\left\{k : \sum_{i=k}^N \tilde{P}(\Omega_i) \leq \frac{x_r - x_d}{x_u - x_d}\right\}, & o.w. \end{cases}$$

$$\bar{B} = \bigcup_{i : i \geq i_u} \Omega_i, \quad \bar{C} = \Omega_{i_u - 1}, \quad \bar{A} = \bigcup_{i : i \geq i_u - 2} \Omega_i,$$

$$\bar{a} = \frac{\tilde{P}(\Omega_{i_u-1})}{P(\Omega_{i_u-1})}, \quad \bar{k} = \frac{x_r - x_d \tilde{P}(\bar{A}) - x_u \tilde{P}(\bar{B})}{\tilde{P}(\bar{C})}.$$

  - *If $\frac{1}{\bar{a}} \leq \frac{\lambda - P(\bar{A})}{1 - \tilde{P}(\bar{A})}$, then the optimal is achieved by this 'bar-system': the optimal final portfolio value and the associated minimal risk are*

$$X_N^* = x_d \mathbf{I}_{\bar{A}} + \bar{k}\mathbf{I}_{\bar{C}} + x_u \mathbf{I}_{\bar{B}},$$

$$CVaR_\lambda(X_N^*) = -x_r + \frac{1}{\lambda}[(x_u - x_d)(P(\bar{A}) - \lambda\tilde{P}(\bar{A})) + (x_u - \bar{k})(P(\bar{C}) - \lambda\tilde{P}(\bar{C}))].$$

  - *If $\frac{1}{\bar{a}} > \frac{\lambda - P(\bar{A})}{1 - \tilde{P}(\bar{A})}$, then find the 'star-system' obtained by*

$$a^* = \sup\left\{a : a > \frac{\tilde{P}(\Omega_k)}{P(\Omega_k)}, \frac{1}{a} \geq \frac{\lambda - \sum_{i=0}^k P(\Omega_i)}{1 - \sum_{i=0}^k \tilde{P}(\Omega_i)}\right\},$$

$$A^* = \bigcup_i \left\{\Omega_i : \frac{\tilde{P}(\Omega_i)}{P(\Omega_i)} > a^*\right\},$$

$$x^* = \frac{x_r - x_d \tilde{P}(A^*)}{1 - \tilde{P}(A^*)}.$$

  The minimal risk is $CVaR_\lambda(X_N^*) = -x_r + \frac{1}{\lambda}(x^* - x_d)(P(\bar{A}) - \lambda\tilde{P}(\bar{A}))$, where the optimal final portfolio value is $X_N^* = x_d \mathbf{I}_{A^*} + x^* \mathbf{I}_{A^{*c}}$.

  If $x_u = \infty$, then the optimal risk is achieved by the 'star-system' calculated above.

In either case, hedging is calculated as $\xi_n^* = \frac{X_{n+1}^*(H)-X_{n+1}^*(T)}{S_{n+1}(H)-S_{n+1}(T)}$ *from the Risk Neutral Pricing formula* $X_n^* = \frac{1}{(1+r)^{(N-n)}}\tilde{E}[X_N^*|\mathcal{F}_n]$.

Proof. The first condition in Theorem 2.4, $\frac{d\tilde{P}}{dP} \leq \frac{1}{\lambda}$, $P - a.s.$, is equivalent to $\frac{\tilde{P}(\omega_1)}{P(\omega_1)} \leq \frac{1}{\lambda}$ since the states are already arranged in descending order according to the size of the Radon Nikodým derivative. When $\frac{\tilde{p}}{p} > 1$, $\frac{\tilde{P}(\omega_1)}{P(\omega_1)} = \frac{\tilde{p}^N}{p^N} \leq \frac{1}{\lambda} \Leftrightarrow \frac{\tilde{p}}{p} \leq \sqrt[N]{\frac{1}{\lambda}}$. Similarly, we get the condition for the case $\frac{\tilde{p}}{p} > 1$, i.e., $\frac{\tilde{q}}{q} > 1$.

The second part of Theorem 2.4 can be translated easily by realizing that set B is the union of all the Ω_i's such that their Radon Nikodým derivatives are less than a threshold a. Now since the final states are ordered, once we find the state Ω_{i_u} for the 'bar-system', then all the states with a bigger index, i.e., Ω_{i_u}, Ω_{i_u+1}, ..., Ω_N, comprise the set $\bar{B}$. $\bar{A}$ and $\bar{C}$ are determined accordingly. Note that when $x_u < \infty$, $\sum_{i=0}^{N} \tilde{P}(\Omega_i) = 1 > \frac{x_r - x_d}{x_u - x_d}$, so i_u can only take values from $1, 2, ..., N+1$, so set $\bar{C}$ is non-empty. *Q.E.D.*

Algorithm 2.6. *CVaR Minimization for Binomial Model*

1. *If* $1 < \frac{\tilde{p}}{p} \leq \sqrt[N]{\frac{1}{\lambda}}$ *or* $1 < \frac{\tilde{q}}{q} \leq \sqrt[N]{\frac{1}{\lambda}}$, *then*

$$\inf_{\xi_n} CVaR_\lambda(X_N) = -x_r, \quad X_N^* = x_r, \quad \xi_n^* = 0 \textit{ for all } n.$$

   Stop.

2. *If* $x_u = \infty$, *go to Step* (4). *Otherwise, find 'bar-system', go to Step* (3).

3. *If* $\frac{1}{\bar{a}} \leq \frac{\lambda - P(\bar{A})}{1-\tilde{P}(\bar{A})}$, *then*

$$\inf_{\xi_n} CVaR_\lambda(X_N) = -x_r + \frac{1}{\lambda}[(x_u - x_d)(P(\bar{A}) - \lambda\tilde{P}(\bar{A})) + (x_u - \bar{k})(P(\bar{C}) - \lambda\tilde{P}(\bar{C}))],$$

$$X_N^* = x_d\mathbf{I}_{\bar{A}} + \bar{k}\mathbf{I}_{\bar{C}} + x_u\mathbf{I}_{\bar{B}}.$$

   Go to Step (5).

4. Find the 'star-system'.

$$\inf_{\xi_n} CVaR_\lambda(X_N) = -x_r + \frac{1}{\lambda}(x^* - x_d)(P(A^*) - \lambda\tilde{P}(A^*)),$$
$$X_N^* = x_d\mathbf{I}_{A^*} + x^*\mathbf{I}_{A^{*c}}.$$

Go to Step (5)*.*

5. Calculate $X_n^* = \frac{1}{(1+r)^{(N-n)}}\tilde{E}[X_N^*|\mathcal{F}_n]$ *for all* n*. Calculate* $\xi_n^* = \frac{X_{n+1}^*(H)-X_{n+1}^*(T)}{S_{n+1}(H)-S_{n+1}(T)}$*. Stop.*

2.4.2 Black-Scholes Model

Let us turn our attention to the Black-Scholes model, a complete market model with continuously distributed stock price. The dynamics of the stock price and the self-financing portfolio are as follows:

$$dS_t = S_t(\mu dt + \sigma dW_t),$$
$$dX_t = \xi_t dS_t + (X_t - \xi_t S_t)rdt.$$

Our main goal is once again first to find

$$\inf_{\xi_t} CVaR_\lambda(X_T) \quad \text{s.t.} \quad \tilde{E}[X_T] = x_r, \ x_d \leq X_T \leq x_u \, a.s., \tag{2.12}$$

and then to find the corresponding dynamic hedging ξ_t.

Definition 2.1. Let the market price of risk $\theta = \frac{\mu - r}{\sigma} > 0$ be as usual, and define the functions

$$d_-(a,s,t) = \frac{1}{\theta\sqrt{T-t}}[-\ln a + \frac{\theta}{\sigma}(\frac{\mu+r-\sigma^2}{2}t - \ln\frac{s}{S_0}) + \frac{\theta^2}{2}(T-t)], \quad d_+(a,s,t) = -d_-(a,s,t).$$

Denote $N(\cdot)$ as the cumulative distribution function for standard normal distribution.

Proposition 2.7. *Following the definitions of 'r-system', 'bar-system', 'star-system' as in Theorem 2.4, we can compute the solution to problem (2.12) and the corresponding dynamic* *hedging ξ_n in the Black-Scholes model as below:*

If $x_u < \infty$, find the 'bar-system' using equations

$$\begin{aligned} \tilde{P}(\bar{A}) &= \frac{x_u - x_r}{x_u - x_d}, \\ \bar{a} &= e^{\theta\sqrt{T}\left[\frac{\theta\sqrt{T}}{2} - N^{-1}(\tilde{P}(\bar{A}))\right]}, \\ P(\bar{A}) &= N(-\tfrac{\theta\sqrt{T}}{2} - \tfrac{\ln\bar{a}}{\theta\sqrt{T}}). \end{aligned}$$

- *If $\frac{1}{\bar{a}} \le \frac{\lambda - P(\bar{A})}{1-\tilde{P}(\bar{A})}$, then the optimal portfolio is achieved by this 'bar-system', in which case the minimum risk, the optimal portfolio and the hedging strategy are*

$$\begin{aligned} CVaR_\lambda(X_T^*) &= -x_r + \tfrac{1}{\lambda}(x_u - x_d)(P(\bar{A}) - \lambda\tilde{P}(\bar{A})), \\ X_t^* &= e^{-r(T-t)}\left[x_u N\left(d_+(\bar{a}, S_t, t)\right) + x_d N\left(d_-(\bar{a}, S_t, t)\right)\right], \\ \xi_t^* &= \frac{x_u - x_d}{\sigma S_t\sqrt{2\pi(T-t)}} e^{-r(T-t) - \frac{d_-^2(\bar{a}, S_t, t)}{2}}. \end{aligned}$$

- *If $\frac{1}{\bar{a}} > \frac{\lambda - P(\bar{A})}{1-\tilde{P}(\bar{A})}$, then the optimal risk is achieved by the 'star-system' calculated as*

$$\begin{aligned} a^* &= \left\{ a : \, a = \frac{1 - N(\frac{\theta\sqrt{T}}{2} - \frac{\ln a}{\theta\sqrt{T}})}{\lambda - N(-\frac{\theta\sqrt{T}}{2} - \frac{\ln a}{\theta\sqrt{T}})} \right\}, \\ P(A^*) &= N(-\tfrac{\theta\sqrt{T}}{2} - \tfrac{\ln a^*}{\theta\sqrt{T}}), \\ \tilde{P}(A^*) &= N(\tfrac{\theta\sqrt{T}}{2} - \tfrac{\ln a^*}{\theta\sqrt{T}}), \\ x^* &= \frac{x_r - x_d\tilde{P}(A^*)}{1 - \tilde{P}(A^*)}. \end{aligned}$$

  The minimum risk, the optimal portfolio and the hedging strategy are

$$\begin{aligned} CVaR_\lambda(X_T^*) &= -x_r + \tfrac{1}{\lambda}(x^* - x_d)(P(A^*) - \lambda\tilde{P}(A^*)), \\ X_t^* &= e^{-r(T-t)}\left[x^* N\left(d_+(a^*, S_t, t)\right) + x_d N\left(d_-(a^*, S_t, t)\right)\right], \\ \xi_t^* &= \frac{x^* - x_d}{\sigma S_t\sqrt{2\pi(T-t)}} e^{-r(T-t) - \frac{d_-^2(a^*, S_t, t)}{2}}. \end{aligned}$$

If $x_u = \infty$, then the optimal risk is achieved by the 'star-system' calculated above.

Remark. Note that the formulae for X_t^* and ξ_t^* resemble the Black-Scholes formulae for a European call option. The reason is that X_t^* is the Risk Neutral price of the optimal final value $X_T^* = x_d \mathbf{I}_{A^*} + x^* \mathbf{I}_{A^{*c}}$ from Theorem 2.4. We will see in the following proof that set $A^* = \left\{ \frac{d\tilde{P}}{dP} > a^* \right\} = \{ S_T < c^* \}$ where $c^* = S_0 e^{\frac{\mu + r - \sigma^2}{2} T - \frac{\sigma}{\theta} \ln a^*}$. Thus the optimal final value becomes piecewise constant depending on the final stock price below or above a constant threshold $X_T^* = x_d \mathbf{I}_{\{S_T < c^*\}} + x^* \mathbf{I}_{\{S_T \geq c^*\}}$.

Proof. According to Theorem 2.4 and Remark 2.3, we need to check whether $\frac{d\tilde{P}}{dP}(\omega)\big|_T \leq \frac{1}{\lambda}$ $P-$a.s., and if not, then whether $\frac{1}{\bar{a}} \leq \frac{\lambda - P(\bar{A})}{1 - \tilde{P}(\bar{A})}$ for the case $x_u < \infty$. The Radon Nikodým derivative process for geometric Brownian motion model is $Z_t := \frac{d\tilde{P}}{dP}|_t = e^{-\theta W_t - \frac{\theta^2}{2} t}$, where $\theta = \frac{\mu - r}{\sigma}$. Obviously, $P\left(\frac{d\tilde{P}}{dP}(\omega)\big|_T > \frac{1}{\lambda} \right) > 0$ since $\operatorname{ess\,sup} Z_T = \infty$. To check the second inequality, and possibly to find solution to the equation $a = \frac{1 - \tilde{P}(A)}{\lambda - P(A)}$, we need to find the explicit relation among the three elements $\tilde{P}(A)$, $P(A)$ and a. Notice that

$$A = \left\{ \frac{d\tilde{P}}{dP}\Big|_T > a \right\} = \left\{ \frac{W_T}{\sqrt{T}} < -\frac{\theta\sqrt{T}}{2} - \frac{\ln a}{\theta\sqrt{T}} \right\} = \left\{ \frac{\tilde{W}_T}{\sqrt{T}} < \frac{\theta\sqrt{T}}{2} - \frac{\ln a}{\theta\sqrt{T}} \right\},$$

we have

$$P(A) = N\left(-\frac{\theta\sqrt{T}}{2} - \frac{\ln a}{\theta\sqrt{T}}\right), \quad \tilde{P}(A) = N\left(\frac{\theta\sqrt{T}}{2} - \frac{\ln a}{\theta\sqrt{T}}\right).$$

Link the above two equations with the definition $\tilde{P}(A) = \frac{x - x_r}{x - x_d}$, we can solve for $\tilde{P}(\bar{A})$, $\bar{a}$, $P(\bar{A})$ sequentially:

$$\tilde{P}(\bar{A}) = \frac{x_u - x_r}{x_u - x_d}, \quad \bar{a} = e^{\theta\sqrt{T}\left[\frac{\theta\sqrt{T}}{2} - N^{-1}(\tilde{P}(\bar{A}))\right]}, \quad P(\bar{A}) = N\left(-\frac{\theta\sqrt{T}}{2} - \frac{\ln \bar{a}}{\theta\sqrt{T}}\right).$$

To obtain 'star-system', we only need to solve the equation $a = \frac{1 - \tilde{P}(A)}{\lambda - P(A)}$ through its explicit form:

$$a = \frac{1 - N\left(\frac{\theta\sqrt{T}}{2} - \frac{\ln a}{\theta\sqrt{T}}\right)}{\lambda - N\left(-\frac{\theta\sqrt{T}}{2} - \frac{\ln a}{\theta\sqrt{T}}\right)}.$$

A direct application of Theorem 2.4 gives the optimal final portfolio X_T^* and the minimal $CVaR$. The corresponding optimal hedging ξ_t^* is the usual 'Delta-hedge' in the Black-Scholes model where the derivative payoff is X_T^*, so we first calculate the Risk Neutral price process $X_t = v(S_t, t)$, and then differentiate with respect to the stock price S_t.

$$X_t = e^{r(T-t)}\tilde{E}[X_T^*|\mathcal{F}_t] = e^{-r(T-t)}\left[x^*\tilde{P}_t(A^{*c}) + x_d\tilde{P}_t(A^*)\right] = e^{-r(T-t)}\left[x^* + (x_d - x^*)\tilde{P}_t(A^*)\right],$$

where $\tilde{P}_t(A^*)$ is the conditional probability under the Risk Neutral measure. Since

$$A^* = \{Z_T > a^*\} = \left\{Z_t e^{-\theta(W_T-W_t)-\frac{\theta^2}{2}(T-t)} > a\right\} = \left\{\frac{\tilde{W}_T-\tilde{W}_t}{\sqrt{T-t}} < -\frac{\ln\frac{a^*}{Z_t}}{\theta\sqrt{T-t}} + \frac{\theta}{2}\sqrt{T-t}\right\},$$

we have then

$$\tilde{P}_t(A^*) = N\left(-\frac{\ln\frac{a^*}{Z_t}}{\theta\sqrt{T-t}} + \frac{\theta}{2}\sqrt{T-t}\right). \tag{2.13}$$

Note that Z_t can be represented by the stock price S_t:

$$S_t = S_0 e^{\sigma W_t + (\mu - \frac{1}{2}\sigma^2)t} \Rightarrow W_t = \frac{\ln\frac{S_t}{S_0} - (\mu-\frac{1}{2}\sigma^2)t}{\sigma},$$

$$Z_t = e^{-\theta W_t - \frac{1}{2}\theta^2 t} \Rightarrow Z_t = g(S_t, t), \text{ where } g(s,t) = e^{\frac{\theta}{\sigma}[\frac{\mu+r-\sigma^2}{2}t - \ln\frac{s}{S_0}]}.$$

Substitute $g(S_t, t)$ into (2.13) we have

$$\tilde{P}_t(A^*) = N\left(d_-(a^*, S_t, t)\right),$$

where

$$d_-(a,s,t) = \frac{1}{\theta\sqrt{T-t}}[-\ln a + \frac{\theta}{\sigma}(\frac{\mu+r-\sigma^2}{2}t - \ln\frac{s}{S_0}) + \frac{\theta^2}{2}(T-t)].$$

Hence

$$X_t = v(S_t, t),$$

where

$$v(s,t) = e^{-r(T-t)}\left[x^* N\left(d_+(a^*,s,t)\right) + x_d N\left(d_-(a^*,s,t)\right)\right], \quad d_+(a^*,s,t) = -d_-(a^*,s,t)$$

Rewrite $v(s,t) = e^{-r(T-t)}\left[x^* + (x_d - x^*)N\left(d_-(a^*,s,t)\right)\right]$, we find the partial derivative

$$v_s(s,t) = \frac{x^* - x_d}{\sigma s\sqrt{2\pi(T-t)}} e^{-r(T-t) - \frac{d_-^2(a^*,s,t)}{2}}.$$

Given the stock price S_t at time t, the optimal strategy ξ_t^* is:

$$\xi_t^* = v_s(S_t, t).$$

Q.E.D.

Algorithm 2.8. *CVaR Minimization for Black-Scholes Model*

1. *If* $x_u = \infty$, *go to Step* (3). *Otherwise, find 'bar-system', go to Step* (2).

2. *If* $\frac{1}{\bar{a}} \leq \frac{\lambda - P(\bar{A})}{1-\tilde{P}(\bar{A})}$, *then*

$$\inf_{\xi_t} CVaR_\lambda(X_T) = -x_r + \frac{1}{\lambda}(x_u - x_d)(P(\bar{A}) - \lambda\tilde{P}(\bar{A})),$$

$$X_t^* = e^{-r(T-t)}\left[x_u N\left(d_+(\bar{a}, S_t, t)\right) + x_d N\left(d_-(\bar{a}, S_t, t)\right)\right],$$

$$\xi_t = \frac{x_u - x_d}{\sigma S_t\sqrt{2\pi(T-t)}} e^{-r(T-t) - \frac{d_-^2(\bar{a}, S_t, t)}{2}}.$$

   Stop.

3. *Find the 'star-system'.*

$$\inf_{\xi_t} CVaR_\lambda(X_T) = -x_r + \frac{1}{\lambda}(x^* - x_d)(P(A^*) - \lambda\tilde{P}(A^*)),$$

$$X_t^* = e^{-r(T-t)}\left[x^* N\left(d_+(a^*, S_t, t)\right) + x_d N\left(d_-(a^*, S_t, t)\right)\right],$$

$$\xi_t^* = \frac{x^* - x_d}{\sigma S_t\sqrt{2\pi(T-t)}} e^{-r(T-t) - \frac{d_-^2(a^*, S_t, t)}{2}}.$$

   Stop.

2.5 Comparison: Optimal Portfolios with Dynamic & Static Hedging

2.5.1 Binomial Model

Dynamic Hedging

Take a 2-step binomial model as an example, where the sampling times are $t = t_0, t_1, t_2$. Also assume $p = \frac{7}{8}$, $q = \frac{1}{8}$, $u = 2$, $d = \frac{1}{2}$, $r = \frac{1}{4}$, $S_0 = 4$, $X_0 = 1$, $x_u = 2$, $x_d = 1$, $\lambda = 0.1$. We calculate the risk-neutral probabilities $\tilde{p} = \frac{1+r-d}{u-d} = \frac{1}{2}$, $\tilde{q} = \frac{u-1-r}{u-d} = \frac{1}{2}$, and $x_r = (1+r)^2 = \frac{25}{16}$. Following Algorithm 2.6, we have:

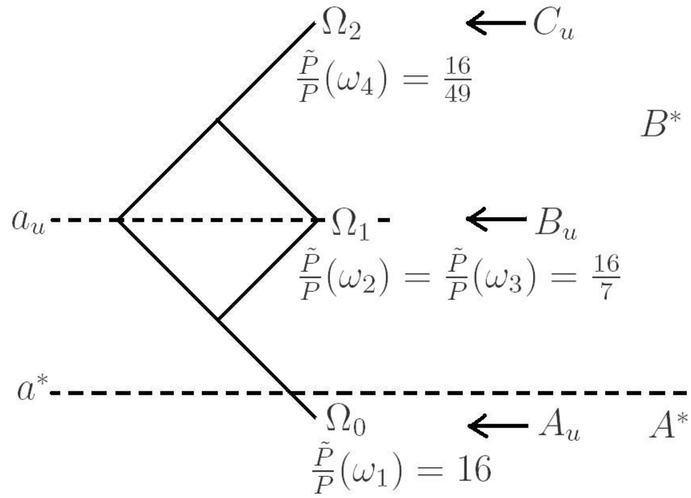

Figure 2.3: Binomial Model Example

1. Neither $1 < \frac{\tilde{p}}{p} \leq \sqrt[N]{\frac{1}{\lambda}}$ nor $1 < \frac{\tilde{q}}{q} \leq \sqrt[N]{\frac{1}{\lambda}}$ is satisfied.

2. Find 'bar-system':

   $\bar{a} = \frac{16}{7}$, $\bar{B} = \Omega_2$, $\bar{A} = \Omega_0$, $\bar{C} = \Omega_1$.

   So $\tilde{P}(\bar{A}) = \frac{1}{4}$, $P(\bar{A}) = \frac{1}{64}$.

3. Condition $\frac{1}{\bar{a}} \leq \frac{\lambda - P(\bar{A})}{1-\tilde{P}(\bar{A})}$ does not hold.

4. Find the 'star-system':

   $a^* = 8.8889$, $A^* = \Omega_0$, $B^* = \Omega_1 \bigcup \Omega_2$, $x^* = 1.75$.

   $\inf_{\xi_n} CVaR_\lambda(X_2) = -1.6328$, and $X_2^*(\Omega_0) = 1$, $X_2^*(\Omega_1) = X_2^*(\Omega_2) = 1.75$.

5. $X_1^*(H) = 1.4$, $X_1^*(T) = 1.1$. $\xi_1^*(H) = 0$, $\xi_1^*(T) = 0.25$, $\xi_0^* = 0.05$.

Static Hedging

If we can only determine the hedging at the beginning, i.e., ξ_0, then the portfolio values along the binomial tree are $X_2 = \xi_0 S_2 + (1+r)^2(X_0 - \xi_0 S_0)$. To constrain X_2 on the interval $[x_d, x_u]$, we require $\xi_0 \in [-0.0577, 0.0449]$. We notice that the optimal ξ_0^* in the dynamic case is outside of this range because ξ_1^* can still be adjusted for the final outcome to be admissible. The resulting $CVaR_\lambda(X_2)$ with $\lambda = 0.1$, is tabulated as follows:

$P(\omega)$	$X_2(\omega)$	ξ_0 = -0.0577	$\xi_0 \in (-0.0577, 0)$	$\xi_0 = 0$	$\xi_0 \in (0, 0.0449)$	$\xi_0 = 0.0449$
$\frac{49}{64}$	$X_2(HH)$	1	$\nearrow$	1.5625	$\nearrow$	2
$\frac{7}{64}$	$X_2(HT)$	1.6923	$\searrow$	1.5625	$\searrow$	1.4615
$\frac{7}{64}$	$X_2(TH)$	1.6923	$\searrow$	1.5625	$\searrow$	1.4615
$\frac{1}{64}$	$X_2(TT)$	1.8654	$\searrow$	1.5625	$\searrow$	1.3269
	$CVaR_{0.1}(X_2)$	-1	-1	-1.5625	$\nearrow$	-1.4405

Table 2.1: Binomial Model Example

Obviously the static hedging is not as good as the dynamic hedging since the optimal risk is $\inf_{\xi_0} CVaR_\lambda(X_2) = -1.5625$, achieved at $\xi_0 = 0$.

2.5.2 Black-Scholes Model

Assume $\lambda = 5\%$, $T = 2$, $r = 5\%$, $S_0 = 10$, $X_0 = 10$, $x_d = 5$, $x_u = 30$.

$(x_u < \infty)$	Example 1	Example 2	Example 3
μ	0.1	0.2	0.3
σ	0.2	0.1	0.1
Dynamic initial hedge ξ_0^*	0	2.6117	4.9958
Dynamic minimal $CVaR(X_T^*)$	-11.0517	-13.3297	-28.8575

Table 2.2: Black-Scholes Model Example with Finite Upper Bound

Here we use a finite upper bound $x_u = 30$ to first illustrate two cases where the optimal is achieved by the 'star-system' in Example 2 and 'bar-system' in Example 3. Although we choose a fairly reasonable set of parameters in Example 1, it turns out that the optimal 'star-system' gives value very close to the 'r-system' where no investment in the stock is made.

Next we let $x_u = \infty$ while keeping all other constants at the same level, and compare the results between static and dynamic hedging. To compute the results for the static hedging case, note that

$$X_T = x_r + \xi_0(S_T - S_0 e^{rT}), \quad x_r = X_0 e^{rT}.$$

For $X_T \in [x_d, \infty)$, we require $\xi_0 \in [0, \frac{x_r - x_d}{S_0 e^{rT}}]$.

$$\begin{aligned} CVaR_\lambda(X_T) &= -\tfrac{1}{\lambda} E[X_T 1_{\{X_T < q_\lambda\}}], \quad \text{where} \quad q_\lambda = -VaR_\lambda(X_T) \\ &= -x_r - \xi_0 S_0 (e^{\mu T} \tfrac{N\left(N^{-1}(\lambda) - \sigma\sqrt{T}\right)}{\lambda} - e^{rT}). \end{aligned}$$

Since the $CVaR_\lambda(X_T)$ is linear in ξ_0, the minimal is obtained at one of the end points.

$(x_u = \infty)$	Example 1	Example 2	Example 3
μ	0.1	0.2	0.3
σ	0.2	0.1	0.1
Static hedge ξ_0^*	0	0	0.5476
Static minimal $CVaR(X_T^*)$	-11.0517	-11.0517	-12.3889
Dynamic initial hedge ξ_0^*	0	2.6117	7.6179
Dynamic minimal $CVaR(X_T^*)$	-11.0517	-13.3297	-57.9182

Table 2.3: Black-Scholes Model Example with Infinite Upper Bound

We see that dynamic hedging provide quite different results both for the minimal $CVaR$ and the hedge in Example 2 and Example 3, while remaining almost indistinguishable in Example 1.

CHAPTER 3: AN OPTIMIZATION PROBLEM WITH EXPECTED RETURN REQUIREMENT

3.1 Two-Constraint Optimization Problem

In this Chapter, we will focus on solving the Main Problem (1.3) with additional assumption 1.2, restated below:

$$
\begin{aligned}
& \inf_{\xi_t} CVaR_\lambda(X_T) \\
\text{subject to} \quad & E[X_T] \geq z, \\
& x_d \leq X_t \leq x_u \, a.s., \quad \forall t \in [0,T].
\end{aligned}
$$

Similar to the approach of reformulating the Main Problem without expected return requirement (1.5) described in Chapter 2, we use the equivalence between conditional Value-at-Risk and the Fenchel-Legendre dual of the expected shortfall derived in Rockafellar and Uryasev ([16] and [17]), namely,

$$CVaR_\lambda(X) = \frac{1}{\lambda} \inf_{x \in \mathbb{R}} \left(E[(x-X)^+] - \lambda x \right), \quad \forall \lambda \in (0,1), \tag{3.1}$$

to reduce the Main Problem (1.3) to the following static convex optimization problem

$$
\begin{aligned}
& CVaR(X^*) = \inf_{X \in \mathcal{F}} \inf_{x \in \mathcal{R}} \frac{1}{\lambda}(E[(x-X)^+] - \lambda x), \qquad (3.2) \\
\text{subject to} \quad & E[X] \geq z, \quad (\textit{return constraint}) \\
& \tilde{E}[X] = x_r, \quad (\textit{capital constraint}) \\
& x_d \leq X \leq x_u \, a.s..
\end{aligned}
$$

Interchanging the order of infimum, we have the following two-step procedure for the Main Problem (1.3):

Two-Constraint Problem:

Step 1: Minimization of Expected Shortfall

$$
\begin{aligned}
& v(x) = \inf_{X \in \mathcal{F}} E[(x - X)^+] \\
\text{subject to} \quad & E[X] \geq z, \quad (\textit{return constraint}) \\
& \tilde{E}[X] = x_r, \quad (\textit{capital constraint}) \\
& x_d \leq X \leq x_u \, a.s.;
\end{aligned}
\tag{3.3}
$$

Step 2: Minimization of conditional Value-at-Risk

$$
\inf_{X \in \mathcal{F}} CVaR_\lambda(X) = \frac{1}{\lambda} \inf_{x \in \mathbb{R}} (v(x) - \lambda x) \tag{3.4}
$$

With the additional constraint on the expectation $E[X] \geq z$, Rockafellar and Uryasev [16] provides a linear programming solution for the Monte-Carlo simulation of the one-time step problem. The dynamic solution given in Ruszczyński and Shapiro [21] requires the modification of the CVaR into a dynamic version. The new results obtained in this thesis is to provide a solution to the problem of *Minimization of Expected Shortfall* in (3.3) with the constraint on the expectation $E[X] \geq z$, then to the problem of *Minimization of CVaR* in (3.4), thus solve the Main Problem in (1.3).

Recall the two-step procedure for the Main Problem without expected return requirement (2.5) solved in Chapter 2:

Step 1: Minimization of Expected Shortfall

$$
\begin{aligned}
& v(x) = \inf_{X \in \mathcal{F}} E[(x - X)^+] \\
\text{subject to} \quad & \tilde{E}[X] = x_r, \quad (\textit{capital constraint}) \\
& x_d \leq X \leq x_u \, a.s.;
\end{aligned}
\tag{3.5}
$$

Step 2: Minimization of conditional Value-at-Risk

$$\inf_{X\in\mathcal{F}} CVaR_\lambda(X) = \frac{1}{\lambda}\inf_{x\in\mathbb{R}}\left(v(x)-\lambda x\right) \tag{3.6}$$

Föllmer and Leukert [9] derived the optimal solution to **Step 1** of this One-Constraint Problem:

$$X(x) = x_d \mathbb{I}_{\left\{\frac{d\tilde{P}}{dP}>a\right\}} + x\mathbb{I}_{\left\{\frac{d\tilde{P}}{dP}\leq a\right\}}, \quad \text{for } x_d < x < x_u \, a.s.. \tag{3.7}$$

The above $X(x)$ is the solution under a special case when the Radon-Nikodým derivative $\frac{d\tilde{P}}{dP}|_T$ is restricted to have a continuous distribution to minimize the complication in its presentation. The optimality of $X(x)$ can be proved in various ways, but it is clearly a result of Neyman-Pearson lemma once the connection between the problem of Minimization of Expected Shortfall and that of hypothesis testing between P and $\tilde{P}$ is established. To view it as a solution from convex duality approach, see Theorem 1.19 in Xu [25]. Note that in (3.7), a is computed from the budget constraint $\tilde{E}[X] = x_r$ for fixed constant x. To proceed to **Step 2**, we varied the value of x and looked for the best x^*. Define set $A = \left\{\omega \in \Omega : \frac{d\tilde{P}}{dP}(\omega) > a\right\}$. Let a^* and A^* be the solution to the equation $\frac{1}{a} = \frac{\lambda - P(A)}{1-\tilde{P}(A)}$. Define $x^* = \frac{x_r - x_d\tilde{P}(A^*)}{1-\tilde{P}(A^*)}$. Under some technical conditions, the solution to **Step 2** of the Main Problem without expected return constraint is shown in Chapter 2 to be

$$X^* = x_d\mathbb{I}_{A^*} + x^*\mathbb{I}_{A^{*c}}, \quad \textbf{(Two-Line Configuration)} \tag{3.8}$$
$$CVaR_\lambda(X^*) = -x_r + \frac{1}{\lambda}(x^*-x_d)\left(P(A^*)-\lambda\tilde{P}(A^*)\right),$$

regardless whether $x_u < \infty$ or $x_u = \infty$. More general solutions in the case when the Radon-Nikodým derivative $\frac{d\tilde{P}}{dP}|_T$ is not restricted to have a continuous distribution is fully treated in Chapter 2. Note that the two-line solution in (3.8) is inherited from the Neyman-Pearson lemma. We will see later in this Chapter that when $x_u < \infty$, the solutions to both **Step 1** and **Step 2** of the static formulation (3.2), and thus the Main Problem (1.3) and (1.6), turn out to be a three-line solution of the form

$$X^{**} = x_d\mathbb{I}_{A^{**}} + x^{**}\mathbb{I}_{B^{**}} + x_u\mathbb{I}_{D^{**}}, \quad \textbf{(Three-Line Configuration)}$$

where

$$A = \left\{\omega \in \Omega \,:\, \tfrac{d\tilde{P}}{dP}(\omega) > a\right\}, B = \left\{\omega \in \Omega \,:\, b \leq \tfrac{d\tilde{P}}{dP}(\omega) \leq a\right\}$$
$$D = \left\{\omega \in \Omega \,:\, \tfrac{d\tilde{P}}{dP}(\omega) < b\right\},$$

and A^{**}, B^{**} and D^{**} are associated to the optimal choice of a^{**} and b^{**}. When $x_u = \infty$, the optimal solution X most likely will not exist, but the infimum of the CVaR can still be computed, some insight can be found in the Black-Scholes example in section 3.5.

The key to finding the exact solution to the Main Problem without expected return constraint (1.5), is to find the pair (a^*, x^*), and Theorem (2.4) states that (a^*, x^*) is the solution to the *capital constraint* ($\tilde{E}[X] = x_r$) and *Euler first order optimality condition* ($v'(x) = 0$ in **Step 2**):

$$x_d\tilde{P}(A) + x\tilde{P}(A^c) = x_r,$$
$$P(A) + \frac{\tilde{P}(A^c)}{a} - \lambda = 0.$$

Likewise, we will see later in Proposition 3.10 and Theorem 3.11 that (a^{**}, b^{**}, x^{**}) is the solution to the same two conditions plus the *return constraint* ($E[X] = z$):

$$x_dP(A) + xP(B) + x_uP(D) = z,$$
$$x_d\tilde{P}(A) + x\tilde{P}(B) + x_u\tilde{P}(D) = x_r,$$
$$P(A) + \frac{\tilde{P}(B) - bP(B)}{a - b} - \lambda = 0.$$

3.2 Case: $x_u < \infty$

Before establishing their existence, we first define some particular Two-Line solutions and the general Three-Line solution that satisfy their respective capital and expected return constraints. Recall the definitions of the sets:

Definition 3.1.

$$A = \left\{\omega \in \Omega \,:\, \tfrac{d\tilde{P}}{dP}(\omega) > a\right\}, \quad B = \left\{\omega \in \Omega \,:\, b \leq \tfrac{d\tilde{P}}{dP}(\omega) \leq a\right\}, \quad D = \left\{\omega \in \Omega \,:\, \tfrac{d\tilde{P}}{dP}(\omega) < b\right\}.$$

Definition 3.2. Three-Line Configuration has the structure $X = x_d \mathbb{I}_A + x\mathbb{I}_B + x_u \mathbb{I}_D$.

Two-Line Configuration $X = x\mathbb{I}_B + x_u \mathbb{I}_D$ is a degenerated form of Three-Line Configuration with $a = \infty$, $B = \left\{\omega \in \Omega : \frac{d\tilde{P}}{dP}(\omega) \geq b\right\}$ and $D = \left\{\omega \in \Omega : \frac{d\tilde{P}}{dP}(\omega) < b\right\}$.

Two-Line Configuration $X = x_d \mathbb{I}_A + x\mathbb{I}_B$ is a degenerated form of Three-Line Configuration with $b = 0$, $A = \left\{\omega \in \Omega : \frac{d\tilde{P}}{dP}(\omega) > a\right\}$, and $B = \left\{\omega \in \Omega : \frac{d\tilde{P}}{dP}(\omega) \leq a\right\}$.

Two-Line Configuration $X = x_d \mathbb{I}_A + x_u \mathbb{I}_D$ is a degenerated form of Three-Line Configuration with $a = b$, $A = \left\{\omega \in \Omega : \frac{d\tilde{P}}{dP}(\omega) > a\right\}$, and $D = \left\{\omega \in \Omega : \frac{d\tilde{P}}{dP}(\omega) < a\right\}$.

Remark. Under Assumption 1.2, the 'bar-system' ($\bar{a}$, $\bar{A}$, $\bar{B}$) in Chapter 2 is the same as the system achieved by the Two-Line Configuration $X = x_d \mathbb{I}_A + x_u \mathbb{I}_D$, here we treat the system as a shorthand from its general form ($\bar{a}$, $\bar{b}$, $\bar{x}$, $\bar{A}$, $\bar{B}$, $\bar{D}$), where $\bar{x} = x_u$, $\bar{b} = \bar{a}$ and $P(\bar{B}) = \tilde{P}(\bar{B}) = 0$. Similarly the 'star-system' in Chapter 2 corresponds to the Two-Line Configuration $X = x_d \mathbb{I}_A + x\mathbb{I}_B$, and we treat it as a shorthand from (a^*, b^*, x^*, A^*, B^*, D^*), where $b^* = 0$ and $P(D^*) = \tilde{P}(D^*) = 0$.

Definition 3.3. General Constraints are the capital constraint and the equality part of the expected return constraint for configuration $X = x_d \mathbb{I}_A + x\mathbb{I}_B + x_u \mathbb{I}_D$:

$$E[X] = x_d P(A) + xP(B) + x_u P(D) = z,$$
$$\tilde{E}[X] = x_d \tilde{P}(A) + x\tilde{P}(B) + x_u \tilde{P}(D) = x_r.$$

Degenerated Constraints 1 are the capital constraint and the equality part of the expected return constraint for configuration $X = x\mathbb{I}_B + x_u \mathbb{I}_D$:

$$E[X] = xP(B) + x_u P(D) = z,$$
$$\tilde{E}[X] = x\tilde{P}(B) + x_u \tilde{P}(D) = x_r.$$

Degenerated Constraints 2 are the capital constraint and the equality part of the expected return constraint for configuration $X = x_d \mathbb{I}_A + x \mathbb{I}_B$:

$$E[X] = x_d P(A) + x P(B) = z,$$

$$\tilde{E}[X] = x_d \tilde{P}(A) + x \tilde{P}(B) = x_r.$$

Degenerated Constraints 3 are the capital constraint and the equality part of the expected return constraint for configuration $X = x_d \mathbb{I}_A + x_u \mathbb{I}_D$:

$$E[X] = x_d P(A) + x_u P(D) = z,$$

$$\tilde{E}[X] = x_d \tilde{P}(A) + x_u \tilde{P}(D) = x_r.$$

Definition 3.4. For fixed $-\infty < x_d < x_r < x_u < \infty$, let $\bar{a} = \bar{b}$ be the constant that satisfies capital constraint $\tilde{E}[X] = x_d \tilde{P}(A) + x_u \tilde{P}(D) = x_r$ for configuration $X = x_d \mathbb{I}_A + x_u \mathbb{I}_D$ in **Degenerated Constraints 3**. Consequently, $\bar{A}$, $\bar{D}$ and $\bar{X}$ are associated to the constant $\bar{a} = \bar{b}$, i.e., $\bar{X} = x_d \mathbb{I}_{\bar{A}} + x_u \mathbb{I}_{\bar{D}}$ where $\bar{A} = \left\{ \omega \in \Omega : \frac{d\tilde{P}}{dP}(\omega) > \bar{a} \right\}$, and $\bar{D} = \left\{ \omega \in \Omega : \frac{d\tilde{P}}{dP}(\omega) < \bar{a} \right\}$. Define $\bar{z} = E[\bar{X}] = x_d P(\bar{A}) + x_u P(\bar{D})$.

Note that $\bar{z}$ is the expected return achieved by the 'bar-system' defined in Chapter 2; it is the unique expected value of a Two-Line configuration that satisfy **Degenerated Constraints 3**.

Lemma 3.1. *$\bar{z}$ is the highest return that can be obtained by a portfolio with initial capital x_0 and is bounded between x_d and x_u:*

$$\bar{z} = \max_{X \in \mathcal{F}} E[X] \quad s.t. \quad \tilde{E}[X] = x_r, \quad x_d \le X \le x_u \, a.s.$$

Proof. The problem of

$$\bar{z} = \max_{X \in \mathcal{F}} E[X] \quad s.t. \quad \tilde{E}[X] = x_r, \quad x_d \le X \le x_u \, a.s.$$

is equivalent to the Expected Shortfall Problem

$$\bar{z} = \min_{X \in \mathcal{F}} E[(x_u - X)^+] \quad s.t. \quad \tilde{E}[X] = x_r, \quad X \geq x_d \, a.s..$$

Therefore, the answer is immediate. *Q.E.D.*

From now on, we will concern ourselves with $z \in [x_r, \bar{z}]$. The lower bound can be interpreted that the investment will yield a higher return than the risk-free rate r, i.e., $z = E[X] \geq x_0 e^{rT} = x_r$. Mathematically, when $z \in (-\infty, x_r)$, the optimal solution X^* to the **One-Constraint Problem** satisfies the *return constraint* $E[X^*] \geq z$ automatically (see Lemma 3.2), thus it is also the optimal solution to the **Two-Constraint Problem**.

Lemma 3.2. *For fixed $-\infty < x_d < x_r < x_u < \infty$, and any $x \in [x_d, x_r]$, choose b so that configuration $X = x\mathbb{I}_B + x_u\mathbb{I}_D$ satisfies the capital constraint $\tilde{E}[X] = x\tilde{P}(B) + x_u\tilde{P}(D) = x_r$ in **Degenerated Constraints 1**. Let $z = E[X] = xP(B) + x_uP(D)$. Then z decreases continuously from $\bar{z}$ to x_r as x increases from x_d to x_r. For any $x \in [x_r, x_u]$, choose a so that configuration $X = x_d\mathbb{I}_A + x\mathbb{I}_B$ satisfies the capital constraint $\tilde{E}[X] = x_d\tilde{P}(A) + x\tilde{P}(B) = x_r$ in **Degenerated Constraints 2**. Let $z = E[X] = x_dP(A) + xP(B)$. Then z increases continuously from x_r to $\bar{z}$ as x increases from x_r to x_u.*

Proof. Choose $x_d \leq x_1 < x_2 \leq x_r$. Let $X_1 = x_1\mathbb{I}_{B_1} + x_u\mathbb{I}_{D_1}$ where $B_1 = \left\{\omega \in \Omega : \frac{d\tilde{P}}{dP}(\omega) \geq b_1\right\}$ and $D_1 = \left\{\omega \in \Omega : \frac{d\tilde{P}}{dP}(\omega) < b_1\right\}$. Choose b_1 such that $\tilde{E}[X_1] = x_r$. This capital constraint means $x_1\tilde{P}(B_1) + x_u\tilde{P}(D_1) = x_r$. Since $\tilde{P}(B_1) + \tilde{P}(D_1) = 1$, $\tilde{P}(B_1) = \frac{x_u - x_r}{x_u - x_1}$ and $\tilde{P}(D_1) = \frac{x_r - x_1}{x_u - x_1}$. Define $z_1 = E[X_1]$. Similarly, z_2, X_2, B_2, D_2, b_2 corresponds to x_2 where $b_1 > b_2$ and $\tilde{P}(B_2) = \frac{x_u - x_r}{x_u - x_2}$ and $\tilde{P}(D_2) = \frac{x_r - x_2}{x_u - x_2}$. Note that $D_2 \subset D_1$, $B_1 \subset B_2$ and

$D_1 \backslash D_2 = B_2 \backslash B_1$. We have

$$
\begin{aligned}
z_1 - z_2 &= x_1 P(B_1) + x_u P(D_1) - x_2 P(B_2) - x_u P(D_2) \\
&= (x_u - x_2) P(B_2 \backslash B_1) - (x_2 - x_1) P(B_1) \\
&= (x_u - x_2) P\left(b_2 < \tfrac{d\tilde{P}}{dP}(\omega) < b_1\right) - (x_2 - x_1) P\left(\tfrac{d\tilde{P}}{dP}(\omega) \geq b_1\right) \\
&= (x_u - x_2) \int_{\left\{b_2 < \frac{d\tilde{P}}{dP}(\omega) < b_1\right\}} \tfrac{dP}{d\tilde{P}}(\omega) d\tilde{P}(\omega) - (x_2 - x_1) \int_{\left\{\frac{d\tilde{P}}{dP}(\omega) \geq b_1\right\}} \tfrac{dP}{d\tilde{P}}(\omega) d\tilde{P}(\omega) \\
&> (x_u - x_2) \frac{1}{b_1} \tilde{P}(B_2 \backslash B_1) - (x_2 - x_1) \frac{1}{b_1} \tilde{P}(B_1) \\
&= (x_u - x_2) \frac{1}{b_1} \left(\frac{x_u - x_r}{x_u - x_2} - \frac{x_u - x_r}{x_u - x_1} \right) - (x_2 - x_1) \frac{1}{b_1} \frac{x_u - x_r}{x_u - x_1} = 0.
\end{aligned}
$$

For any given $\epsilon > 0$, choose $x_2 - x_1 \leq \epsilon$, then

$$
\begin{aligned}
z_1 - z_2 &= (x_u - x_1) P(B_2 \backslash B_1) - (x_2 - x_1) P(B_2) \\
&\leq (x_u - x_1) P(B_2 \backslash B_1) \\
&\leq (x_u - x_1) \left(\frac{x_u - x_r}{x_u - x_2} - \frac{x_u - x_r}{x_u - x_1} \right) \\
&\leq \frac{(x_2 - x_1)(x_u - x_r)}{x_u - x_2} \leq x_2 - x_1 \leq \epsilon.
\end{aligned}
$$

Therefore, z decreases continuously as x increases when $x \in [x_d, x_r]$. When $x = x_d$, $z = \bar{z}$ from Definition 3.4. When $x = x_r$, $X \equiv x_r$ and $z = E[X] = x_r$. Similarly, we can show that z increases continuously from x_r to $\bar{z}$ as x increases from x_r to x_u. *Q.E.D.*

From the above lemma, we can see that for given x value, we can compute the corresponding z value in **Degenerated Constraints 1** and **Degenerated Constraints 2**. Since their relationship is monotone and continuous in each situation, given z we can find the corresponding x value in both situations.

Definition 3.5. For fixed $-\infty < x_d < x_r < x_u < \infty$, and fixed $z \in [x_r, \bar{z}]$, define x_{z1} and x_{z2} to be the corresponding x values for configurations that satisfy **Degenerated Constraints 1** and **Degenerated Constraints 2** respectively.

Definition 3.5 means that when we fix z in a proper interval $[x_r, \bar{z}]$, we can find two feasible solutions: $X = x_{z1}\mathbb{I}_B + x_u\mathbb{I}_D$ satisfying $\tilde{E}[X] = x_{z1}\tilde{P}(B) + x_u\tilde{P}(D) = x_r$ and $E[X] = x_{z1}P(B) + x_uP(D) = z$; $X = x_d\mathbb{I}_A + x_{z2}\mathbb{I}_B$ satisfying $\tilde{E}[X] = x_d\tilde{P}(A) + x_{z2}\tilde{P}(B) = x_r$ and $E[X] = x_dP(A) + x_{z2}P(B) = z$.

Now we fix $x \in [x_d, x_{z1}]$, and as in Lemma 3.2, choose b so that configuration $X = x\mathbb{I}_B + x_u\mathbb{I}_D$ satisfies the capital constraint $\tilde{E}[X] = x\tilde{P}(B) + x_u\tilde{P}(D) = x_r$ in **Degenerated Constraints 1**. At the left end point $x = x_d$, we encounter $\bar{X}$ given by **Degenerated Constraints 3** in Definition 3.4 and corresponding $\bar{z} = E[\bar{X}] \geq z$. At the right end point $x = x_{z1}$, we encounter $X = x_{z1}\mathbb{I}_B + x_u\mathbb{I}_D$ such that $E[X] = z$. In between, $E[X]$, where $X = x\mathbb{I}_B + x_u\mathbb{I}_D$ and $\tilde{E}[X] = x_r$, is decreasing according to Lemma 3.2. We recognize that $E[X] = xP(B) + x_uP(D) \geq z$, for all $x \in [x_d, x_{z1}]$. Similar analysis can be applied to the interval $x \in [x_{z2}, x_u]$. We make this conclusion in the following lemma.

Lemma 3.3. *For fixed $-\infty < x_d < x_r < x_u < \infty$, and fixed $z \in [x_r, \bar{z}]$,*

1. *If we fix $x \in [x_d, x_{z1}]$, the Two-Line Configuration $X = x\mathbb{I}_B + x_u\mathbb{I}_D$ which satisfies the capital constraint $\tilde{E}[X] = x\tilde{P}(B) + x_u\tilde{P}(D) = x_r$ in Degenerated Constraints 1* ***satisfies*** *the expected return constraint: $E[X] = xP(B) + x_uP(D) \geq z$;*

2. *If we fix $x \in (x_{z1}, x_r]$, the Two-Line Configuration $X = x\mathbb{I}_B + x_u\mathbb{I}_D$ which satisfies the capital constraint $\tilde{E}[X] = x\tilde{P}(B) + x_u\tilde{P}(D) = x_r$ in Degenerated Constraints 1* ***fails*** *the expected return constraint: $E[X] = xP(B) + x_uP(D) < z$;*

3. *If we fix $x \in [x_r, x_{z2})$, the Two-Line Configuration $X = x_d\mathbb{I}_A + x\mathbb{I}_B$ which satisfies the capital constraint $\tilde{E}[X] = x_d\tilde{P}(A) + x\tilde{P}(B) = x_r$ in Degenerated Constraints 2* ***fails*** *the expected return constraint: $E[X] = xP(B) + x_uP(D) < z$;*

4. *If we fix $x \in [x_{z2}, x_u)$, the Two-Line Configuration $X = x_d\mathbb{I}_A + x\mathbb{I}_B$ which satisfies the capital constraint $\tilde{E}[X] = x_d\tilde{P}(A) + x\tilde{P}(B) = x_r$ in Degenerated Constraints 2* ***satisfies*** *the expected return constraint: $E[X] = xP(B) + x_uP(D) \geq z$.*

Proposition 3.4. *For fixed $-\infty < x_d < x_r < x_u < \infty$, and fixed $z \in [x_r, \bar{z}]$, if we fix $x \in [x_d, x_{z1}]$, then there exists a* ***Two-Line Configuration*** *$X = x\mathbb{I}_B + x_u\mathbb{I}_D$ which is the optimal solution to* ***Step 1*** *of the* ***Two-Constraint Problem****; if we fix $x \in [x_{z2}, x_u]$, then* *there exists a* ***Two-Line Configuration*** $X = x_d \mathbb{I}_A + x \mathbb{I}_B$ *which is the optimal solution to* ***Step 1*** *of the* ***Two-Constraint Problem****.*

Lemma 3.3 and Proposition 3.4 are natural logical consequences and their proofs will be skipped.

When $x \in (x_{z1}, x_{z2})$, the Two-Line Configurations that can be achieved with the right amount of initial capital ($\tilde{E}[X] = x_r$) do not generate high enough expected return ($E[X] < z$) to be feasible, so we have to look for a novel solution of Three-Line Configuration that is both feasible and optimal.

Lemma 3.5. *For fixed* $-\infty < x_d < x_r < x_u < \infty$*, fixed* $z \in [x_r, \bar{z}]$*, and fixed* $x \in (x_{z1}, x_{z2})$*, choose the pair of real numbers* $-\infty < b \leq a < \infty$ *so that configuration* $X = x_d \mathbb{I}_A + x \mathbb{I}_B + x_u \mathbb{I}_D$ *always satisfies the capital constraint* $\tilde{E}[X] = x_d \tilde{P}(A) + x \tilde{P}(B) + x_u \tilde{P}(D) = x_r$ *in* ***General Constraints****. When* $b = \bar{b} = \bar{a} = a$*, then* $X = \bar{X}$ *and* $E[\bar{X}] = \bar{z}$*. When* $b < \bar{b}$ *and* $a > \bar{a}$*, the expected value* $E[X] = x_d P(A) + x P(B) + x_u P(D)$ *decreases continuously as* b *decreases and* a *increases. In the extreme case* $b = 0$*, the Three-Line configuration becomes the Two-Line Configuration* $X = x \mathbb{I}_B + x_u \mathbb{I}_D$*; in the extreme* $a = \infty$*, the Three-Line configuration becomes the Two-Line Configuration* $X = x_d \mathbb{I}_A + x \mathbb{I}_B$*. In either extreme cases, the expected value is below* z *by Lemma 3.3.*

Proof. Choose $-\infty < b_1 < b_2 \leq \bar{b} = \bar{a} \leq a_2 < a_1 < \infty$. Let configuration $X_1 = x_d \mathbb{I}_{A_1} + x \mathbb{I}_{B_1} + x_u \mathbb{I}_{D_1}$ correspond to the pair (a_1, b_1) where $A_1 = \left\{ \omega \in \Omega : \frac{d\tilde{P}}{dP}(\omega) > a_1 \right\}, B_1 = \left\{ \omega \in \Omega : b_1 \leq \frac{d\tilde{P}}{dP}(\omega) \leq a_1 \right\}, D_1 = \left\{ \omega \in \Omega : \frac{d\tilde{P}}{dP}(\omega) < b_1 \right\}$. Similarly, let configuration $X_2 = x_d \mathbb{I}_{A_2} + x \mathbb{I}_{B_2} + x_u \mathbb{I}_{D_2}$ correspond to the pair (a_2, b_2). Define $z_1 = E[X_1]$ and $z_2 = E[X_2]$. Since both X_1 and X_2 satisfy the capital constraint, we have

$$x_d \tilde{P}(A_1) + x \tilde{P}(B_1) + x_u \tilde{P}(D_1) = x_r = x_d \tilde{P}(A_2) + x \tilde{P}(B_2) + x_u \tilde{P}(D_2).$$

This simplifies to the equation

$$(x - x_d) \tilde{P}(A_2 \backslash A_1) = (x_u - x) \tilde{P}(D_2 \backslash D_1). \tag{3.9}$$

Then

$$
\begin{aligned}
z_2 - z_1 &= x_d P(A_2) + x P(B_2) + x_u P(D_2) - x_d P(A_1) - x P(B_1) - x_u P(D_1) \\
&= (x_u - x) P(D_2 \backslash D_1) - (x - x_d) P(A_2 \backslash A_1) \\
&= (x_u - x) P(D_2 \backslash D_1) - (x_u - x) \frac{\tilde{P}(D_2 \backslash D_1)}{\tilde{P}(A_2 \backslash A_1)} P(A_2 \backslash A_1) \\
&= (x_u - x) \tilde{P}(D_2 \backslash D_1) \left(\frac{P(D_2 \backslash D_1)}{\tilde{P}(D_2 \backslash D_1)} - \frac{P(A_2 \backslash A_1)}{\tilde{P}(A_2 \backslash A_1)} \right) \\
&= (x_u - x) \tilde{P}(D_2 \backslash D_1) \left(\frac{\int_{\left\{ b_1 \le \frac{d\tilde{P}}{dP}(\omega) < b_2 \right\}} \frac{dP}{d\tilde{P}}(\omega) d\tilde{P}(\omega)}{\tilde{P}(D_2 \backslash D_1)} - \frac{\int_{\left\{ a_2 < \frac{d\tilde{P}}{dP}(\omega) \le a_1 \right\}} \frac{dP}{d\tilde{P}}(\omega) d\tilde{P}(\omega)}{\tilde{P}(A_2 \backslash A_1)} \right) \\
&\ge (x_u - x) \tilde{P}(D_2 \backslash D_1) \left(\frac{1}{b_1} - \frac{1}{a_1} \right) > 0.
\end{aligned}
$$

Suppose the pair (a_1, b_1) is chosen so that X_1 satisfies the budget constraint $\tilde{E}[X_1] = x_r$. For any given $\epsilon > 0$, choose $b_2 - b_1$ small enough such that $P(D_2 \backslash D_1) \le \frac{\epsilon}{x_u - x}$. Now choose a_2 such that $a_2 < a_1$ and equation (3.9) is satisfied. Then X_2 also satisfies the budget constraint $\tilde{E}[X_2] = x_r$, and

$$
z_2 - z_1 = (x_u - x) P(D_2 \backslash D_1) - (x - x_d) P(A_2 \backslash A_1) \le (x_u - x) P(D_2 \backslash D_1) \le \epsilon.
$$

We conclude that the expected value of the Three-Line configuration decreases continuously as b decreases and a increases. *Q.E.D.*

Proposition 3.6. *For fixed $-\infty < x_d < x_r < x_u < \infty$, and fixed $z \in [x_r, \bar{z}]$, if we fix $x \in (x_{z1}, x_{z2})$, then there exists a **Three-Line Configuration** $X = x_d \mathbb{I}_A + x \mathbb{I}_B + x_u \mathbb{I}_D$ that satisfies the **General Constraints** which is the optimal solution to **Step 1** of the **Two-Constraint Problem**.*

Proof. Denote $\rho = \frac{d\tilde{P}}{dP}$. According to Lemma 3.5, there exists a Three-Line configuration $\hat{X} = x_d \mathbb{I}_A + x \mathbb{I}_B + x_u \mathbb{I}_D$ that satisfies the General Constraints:

$$
\begin{aligned}
E[X] &= x_d P(A) + x P(B) + x_u P(D) = z, \\
\tilde{E}[X] &= x_d \tilde{P}(A) + x \tilde{P}(B) + x_u \tilde{P}(D) = x_r.
\end{aligned}
$$

where

$$A = \{\omega \in \Omega \,:\, \rho(\omega) > \hat{a}\}\,, \quad B = \left\{\omega \in \Omega \,:\, \hat{b} \le \rho(\omega) \le \hat{a}\right\}, \quad D = \left\{\omega \in \Omega \,:\, \rho(\omega) < \hat{b}\right\}.$$

As standard for convex optimization problems, if we can find a pair of Lagrange multipliers $\lambda \ge 0$ and $\mu \in \mathbb{R}$ such that $\hat{X}$ is the solution to the minimization problem

$$\inf_{X \in \mathcal{F},\; x_d \le X \le x_u} E[(x - X)^+ - \lambda X - \mu \rho X] = E[(x - \hat{X})^+ - \lambda \hat{X} - \mu \rho \hat{X}], \tag{3.10}$$

then $\hat{X}$ is the solution to the constrained problem

$$\inf_{X \in \mathcal{F},\; x_d \le X \le x_u} E[(x - X)^+], \quad s.t. \quad E[X] \ge z, \quad \tilde{E}[X] = x_r.$$

Define

$$\lambda = \frac{\hat{b}}{\hat{a} - \hat{b}}, \quad \mu = -\frac{1}{\hat{a} - \hat{b}}.$$

Then (3.10) becomes

$$\inf_{X \in \mathcal{F},\; x_d \le X \le x_u} E\left[(x - X)^+ + \tfrac{\rho - \hat{b}}{\hat{a} - \hat{b}} X\right].$$

Choose any $X \in \mathcal{F}$ where $x_d \le X \le x_u$, and denote $G = \{\omega \in \Omega \,:\, X(\omega) \ge x\}$ and $L = \{\omega \in \Omega \,:\, X(\omega) < x\}$. Note that $\frac{\rho - \hat{b}}{\hat{a} - \hat{b}} > 1$ on set A, $0 \le \frac{\rho - \hat{b}}{\hat{a} - \hat{b}} \le 1$ on set B, $\frac{\rho - \hat{b}}{\hat{a} - \hat{b}} < 0$ on set D. Then the difference

$$
\begin{aligned}
&E\left[(x-X)^{+}+\tfrac{\rho-\hat{b}}{\hat{a}-\hat{b}}X\right]-E\left[(x-\hat{X})^{+}+\tfrac{\rho-\hat{b}}{\hat{a}-\hat{b}}\hat{X}\right]\\
&=E\left[(x-X)\mathbb{I}_L+\tfrac{\rho-\hat{b}}{\hat{a}-\hat{b}}X\left(\mathbb{I}_A+\mathbb{I}_B+\mathbb{I}_D\right)\right]-E\left[\left(x-x_d\right)\mathbb{I}_A+\tfrac{\rho-\hat{b}}{\hat{a}-\hat{b}}(x_d\mathbb{I}_A+x\mathbb{I}_B+x_u\mathbb{I}_D)\right]\\
&=E\left[(x-X)\mathbb{I}_L+\left(\tfrac{\rho-\hat{b}}{\hat{a}-\hat{b}}(X-x_d)-(x-x_d)\right)\mathbb{I}_A+\tfrac{\rho-\hat{b}}{\hat{a}-\hat{b}}\left(X-x\right)\mathbb{I}_B+\tfrac{\rho-\hat{b}}{\hat{a}-\hat{b}}\left(X-x_u\right)\mathbb{I}_D\right]\\
&\geq E\left[(x-X)\mathbb{I}_L+\left(X-x\right)\mathbb{I}_A+\tfrac{\rho-\hat{b}}{\hat{a}-\hat{b}}\left(X-x\right)\mathbb{I}_B+\tfrac{\rho-\hat{b}}{\hat{a}-\hat{b}}\left(X-x_u\right)\mathbb{I}_D\right]\\
&=E\left[(x-X)\left(\mathbb{I}_{L\cap A}+\mathbb{I}_{L\cap B}+\mathbb{I}_{L\cap D}\right)+(X-x)\left(\mathbb{I}_{A\cap G}+\mathbb{I}_{A\cap L}\right)+\tfrac{\rho-\hat{b}}{\hat{a}-\hat{b}}\left(X-x\right)\mathbb{I}_B\right.\\
&\qquad\left.+\tfrac{\rho-\hat{b}}{\hat{a}-\hat{b}}\left(X-x_u\right)\mathbb{I}_D\right]\\
&=E\left[(x-X)\left(\mathbb{I}_{L\cap B}+\mathbb{I}_{L\cap D}\right)+(X-x)\,\mathbb{I}_{A\cap G}+\tfrac{\rho-\hat{b}}{\hat{a}-\hat{b}}\left(X-x\right)\mathbb{I}_B+\tfrac{\rho-\hat{b}}{\hat{a}-\hat{b}}\left(X-x_u\right)\mathbb{I}_D\right]\\
&=E\left[(x-X)\left(\mathbb{I}_{L\cap B}+\mathbb{I}_{L\cap D}\right)+(X-x)\,\mathbb{I}_{A\cap G}+\tfrac{\rho-\hat{b}}{\hat{a}-\hat{b}}\left(X-x\right)\left(\mathbb{I}_{B\cap G}+\mathbb{I}_{B\cap L}\right)\right.\\
&\qquad\left.+\tfrac{\rho-\hat{b}}{\hat{a}-\hat{b}}\left(X-x_u\right)\left(\mathbb{I}_{D\cap G}+\mathbb{I}_{D\cap L}\right)\right]\\
&=E\left[(x-X)\left(1-\tfrac{\rho-\hat{b}}{\hat{a}-\hat{b}}\right)\mathbb{I}_{B\cap L}+\left(x-X+\tfrac{\rho-\hat{b}}{\hat{a}-\hat{b}}\left(X-x_u\right)\right)\mathbb{I}_{D\cap L}+(X-x)\,\mathbb{I}_{A\cap G}\right.\\
&\qquad\left.+\tfrac{\rho-\hat{b}}{\hat{a}-\hat{b}}\left(X-x\right)\mathbb{I}_{B\cap G}+\tfrac{\rho-\hat{b}}{\hat{a}-\hat{b}}\left(X-x_u\right)\mathbb{I}_{D\cap G}\right]\\
&\geq 0.
\end{aligned}
$$

The last inequality holds because each term inside the expectation is greater than or equal to zero. *Q.E.D.*

Let us recall the first step for the Two-Constraint Problem:

Step 1: Minimization of Expected Shortfall

$$
\begin{aligned}
&v(x)=\inf_{X\in\mathcal{F}}E[(x-X)^{+}] &&(3.11)\\
\text{subject to}\quad &E[X]\geq z,\quad (\textit{return constraint})\\
&\tilde{E}[X]=x_r,\quad (\textit{capital constraint})\\
&x_d\leq X\leq x_u\ a.s.;
\end{aligned}
$$

Theorem 3.7 (Solution to **Step 1:** Minimization of Expected Shortfall)**.** *For fixed* $-\infty<x_d<x_r<x_u<\infty$*, and fixed* $z\in[x_r,\bar{z}]$*. The optimal* $X(x)$ *and the corresponding value* *function* $v(x)$ *to* ***Step 1: Minimization of Expected Shortfall*** *of the* ***Two-Constraint Problem*** *are as follows:*

- $x \in (-\infty, x_d]$*:*

$$\begin{aligned} X(x) = & \text{ any random variable with values in } [x_d, x_u] \text{ satisfying both} \\ & \tilde{E}[X(x)] = x_r \text{ and } E[X(x)] \geq z, \\ v(x) = & \ 0. \end{aligned}$$

- $x \in [x_d, x_{z1}]$*:*

$$\begin{aligned} X(x) = & \text{ any random variable with values in } [x, x_u] \text{ satisfying both} \\ & \tilde{E}[X(x)] = x_r \text{ and} E[X(x)] \geq z, \\ v(x) = & \ 0. \end{aligned}$$

- $x \in (x_{z1}, x_{z2})$*:*

$$\begin{aligned} X(x) = & \ x_d \mathbb{I}_{A_x} + x \mathbb{I}_{B_x} + x_u \mathbb{I}_{D_x} \\ & \text{where } A_x, B_x, D_x \text{ are determined by } a_x \text{ and } b_x \text{ through definitions 3.1} \\ & \text{satisfying the General Constraints:} \tilde{E}[X(x)] = x_r \text{ and } E[X(x)] = z, \\ v(x) = & \ (x - x_d) P(A_x). \end{aligned}$$

- $x \in [x_{z2}, x_u]$*:*

$$\begin{aligned} X(x) = & \ x_d \mathbb{I}_{A_x} + x \mathbb{I}_{B_x} \\ & \text{where } A_x, B_x \text{ are determined by } a_x \text{ as in Definition 3.2 satisfying both} \\ & \tilde{E}[X(x)] = x_r \text{ and } E[X(x)] \geq z, \\ v(x) = & \ (x - x_d) P(A_x). \end{aligned}$$

- *$x \in [x_u, \infty)$:*

$$X(x) = x_d \mathbb{I}_{\bar{A}} + x_u \mathbb{I}_{\bar{B}} = \bar{X}$$

where $\bar{A}, \bar{B}$ are associated to $\bar{a}$ as in Definition 3.4 satisfying both

$$\tilde{E}[X(x)] = x_r \text{ and } E[X(x)] = \bar{z} \geq z,$$

$$v(x) = (x - x_d)P(\bar{A}) + (x - x_u)P(\bar{B}).$$

Theorem 3.7 is a direct consequence of Lemma 3.3, Proposition 3.4, and Proposition 3.6.

To solve **Step 2** of the **Two-Constraint Problem**, we need to find

$$\frac{1}{\lambda} \inf_{x \in \mathbb{R}} (v(x) - \lambda x),$$

where we have already computed $v(x)$ in Theorem 3.7. It turns out that depending on the z level in the return constraint of the **Two-Constraint Problem**, sometimes the optimal is obtained by the Two-Line solution to the **One-Constraint Problem**, other times it is obtained by a true Three-Line solution. To accomplish this, we have to recall the results of Theorem 2.4 and the Remark on continuous case in Chapter (2). Recognizing the equivalence of the "bar-system" and "bar-system" defined in Definition 3.4, we restate the result as follows:

Theorem 3.8 (Theorem 2.4 and its Remark in Chapter 2 with Assumption 1.2)**.**

1. *Suppose* $\operatorname{ess\,sup} \frac{d\tilde{P}}{dP} \leq \frac{1}{\lambda}$. *$X = x_r$ is the optimal solution to **Step 2: Minimization of Conditional Value-at-Risk** of the **One-Constraint Problem** and the associated minimal risk is*

$$CVaR(X) = -x_r.$$

2. *Suppose* $\operatorname{ess\,sup} \frac{d\tilde{P}}{dP} > \frac{1}{\lambda}$.

   - *If $\frac{1}{\bar{a}} \leq \frac{\lambda - P(\bar{A})}{1 - \tilde{P}(\bar{A})}$, then $\bar{X} = x_d \mathbb{I}_{\bar{A}} + x_u \mathbb{I}_{\bar{D}}$ is the optimal solution to **Step 2: Minimization of Conditional Value-at-Risk** of the **One-Constraint Problem*** *and the associated minimal risk is*

$$CVaR(\bar{X}) = -x_r + \frac{1}{\lambda}(x_u - x_d)(P(\bar{A}) - \lambda\tilde{P}(\bar{A})).$$

- *Otherwise, let a^* be the solution to the equation $\frac{1}{a} = \frac{\lambda - P(A)}{1-\tilde{P}(A)}$. Then A^* and B^* defined in 3.1 and Definition 3.2 are the sets associated to the level a^*. Define $x^* = \frac{x_r - x_d\tilde{P}(A^*)}{1-\tilde{P}(A^*)}$ so that configuration*

$$X^* = x_d\mathbb{I}_{A^*} + x^*\mathbb{I}_{B^*}$$

*satisfies the capital constraint $\tilde{E}[X^*] = x_d\tilde{P}(A^*) + x^*\tilde{P}(B^*) = x_r$ in **Degenerated Constraints 2**. Then X^* is the optimal solution to **Step 2: Minimization of Conditional Value-at-Risk** of the **One-Constraint Problem** and the associated minimal risk is*

$$CVaR(X^*) = -x_r + \frac{1}{\lambda}(x^* - x_d)(P(A^*) - \lambda\tilde{P}(A^*)).$$

Definition 3.6. In the situation where $\operatorname{ess\,sup} \frac{d\tilde{P}}{dP} > \frac{1}{\lambda}$ of Theorem 3.8, define $z^* = \bar{z}$ in the first case; define $z^* = E[X^*]$ in the second case.

It is straightforward to see that when z is less than z^*, the Two-Line solution provided in Theorem 3.8 is indeed the optimal solution to **Step 2: Minimization of Conditional Value-at-Risk** of the **Two-Constraint Problem**. While when z is greater than z^* the Two-Line solutions are no longer feasible in the **Two-Constraint Problem** and we will show now that the Three-Line solutions are not only feasible but also optimal.

For $z \in (z^*, \bar{z}]$, **Step 2** of the **Two-Constraint Problem**

$$\frac{1}{\lambda}\inf_{x\in\mathbb{R}}(v(x) - \lambda x)$$

is the minimum of the following five sub-problems after applying Theorem 3.7:

Case 1

$$\frac{1}{\lambda}\inf_{(-\infty,x_d]}(v(x)-\lambda x)=\frac{1}{\lambda}\inf_{(-\infty,x_d]}(-\lambda x)=-x_d;$$

Case 2

$$\frac{1}{\lambda}\inf_{[x_d,x_{z1}]}(v(x)-\lambda x)=\frac{1}{\lambda}\inf_{[x_d,x_{z1}]}(-\lambda x)=-x_{z1}\leq -x_d;$$

Case 3

$$\frac{1}{\lambda}\inf_{(x_{z1},x_{z2})}(v(x)-\lambda x)=\frac{1}{\lambda}\inf_{(x_{z1},x_{z2})}((x-x_d)P(A_x)-\lambda x)\,;$$

Case 4

$$\frac{1}{\lambda}\inf_{[x_{z2},x_u]}(v(x)-\lambda x)=\frac{1}{\lambda}\inf_{[x_{z2},x_u]}((x-x_d)P(A_x)-\lambda x)\,;$$

Case 5

$$\frac{1}{\lambda}\inf_{[x_u,\infty)}(v(x)-\lambda x)=\frac{1}{\lambda}\inf_{[x_u,\infty)}\left((x-x_d)P(\bar{A})+(x-x_u)P(\bar{B})-\lambda x\right).$$

We first establish the convexity of the objective function and its continuity in Lemma 3.9, then we prove the Three-Line solution which is feasible and satisfies the first order condition is indeed optimal.

Lemma 3.9. *$v(x)$ is a convex function for $x\in\mathbb{R}$, and thus continuous.*

Proof. The convexity of $v(x)$ is a simple consequence of its definition (3.3). Real-valued convex functions on $\mathbb{R}$ are continuous on its interior of the domain, so $v(x)$ is continuous on $\mathbb{R}$. *Q.E.D.*

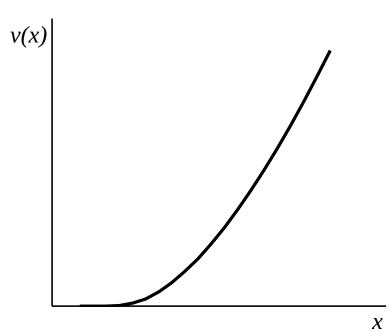

Figure 3.1: Sample $v(x)$ for Black-Scholes Model

Proposition 3.10. *For fixed $-\infty < x_d < x_r < x_u < \infty$, and fixed $z \in (z^*, \bar{z}]$. Suppose* $\operatorname{ess\,sup} \frac{d\tilde{P}}{dP} > \frac{1}{\lambda}$. *The solution a^{**}, b^{**} and x^{**} (and consequently, A^{**}, B^{**} and D^{**}) to the system*

$$\begin{aligned} x_d P(A) + x P(B) + x_u P(D) &= z, \quad (\textit{return constraint}) \\ x_d \tilde{P}(A) + x \tilde{P}(B) + x_u \tilde{P}(D) &= x_r, \quad (\textit{capital constraint}) \\ P(A) + \frac{\tilde{P}(B) - bP(B)}{a-b} - \lambda &= 0, \quad (\textit{first order Euler condition}) \end{aligned}$$

exists. $X^{**} = x_d \mathbb{I}_{A^{**}} + x^{**} \mathbb{I}_{B^{**}} + x_u \mathbb{I}_{D^{**}}$ *is the optimal solution to **Step 2: Minimization of Conditional Value-at-Risk** of the **Two-Constraint Problem** where*

$$\frac{1}{\lambda} \inf_{x \in \mathbb{R}} (v(x) - \lambda x) = \frac{1}{\lambda} \min_{(x_{z1}, x_{z2})} (v(x) - \lambda x),$$

and the associated minimal risk is

$$CVaR(X^{**}) = -x_r + \frac{1}{\lambda} \left((x^{**} - x_d) P(A^{**}) - \lambda x^{**} \right).$$

Proof. Obviously, **Case 2** dominates **Case 1** in the sense that its minimum is lower. In **Case 3**, by the continuity of $v(x)$, we have

$$\frac{1}{\lambda} \inf_{(x_{z1}, x_{z2})} ((x - x_d) P(A_x) - \lambda x) \leq \frac{1}{\lambda} ((x_{z1} - x_d) P(A_{x_{z1}}) - \lambda x_{z1}) = -x_{z1}.$$

The last equality comes from the fact $P(A_{x_{z1}}) = 0$: As in Lemma 3.5, we know that when $x = x_{z1}$, the Three-Line configuration $X = x_d \mathbb{I}_A + x \mathbb{I}_B + x_u \mathbb{I}_D$ degenerates to the Two-Line configuration $X = x_{z1} \mathbb{I}_B + x_u \mathbb{I}_D$ where $a_{x_{z1}} = \infty$. Therefore, **Case 3** dominates **Case 2**.

In **Case 5**,

$$\begin{aligned}
\frac{1}{\lambda}\inf_{[x_u,\infty)}(v(x)-\lambda x) &= \frac{1}{\lambda}\inf_{[x_u,\infty)}\big((x-x_d)P(\bar{A})+(x-x_u)P(\bar{B})-\lambda x\big)\\
&= \frac{1}{\lambda}\inf_{[x_u,\infty)}\big((1-\lambda)x-x_dP(\bar{A})-x_uP(\bar{B})\big)\\
&= \frac{1}{\lambda}\big((1-\lambda)x_u-x_dP(\bar{A})-x_uP(\bar{B})\big)\\
&= \frac{1}{\lambda}\big((x_u-x_d)P(\bar{A})-\lambda x_u\big)\\
&\geq \frac{1}{\lambda}\inf_{[x_{z2},x_u]}\left((x-x_d)P(A_x)-\lambda x\right).
\end{aligned}$$

Therefore, **Case 4** dominates **Case 5**. When $x \in [x_{z2}, x_u]$ and $\operatorname{ess\,sup} \frac{d\tilde{P}}{dP} > \frac{1}{\lambda}$, Theorem 3.7 and Theorem 3.8 imply that the infimum in **Case 4** is achieved either by $\bar{X}$ or X^*. Since we restrict $z \in (z^*, \bar{z}]$ where $z^* = \bar{z}$ by Definition 3.6 in the first case, we need not consider this case in the current proposition. In the second case, Lemma 3.2 implies that $x^* < x_{z2}$ (because $z > z^*$). By the convexity of $v(x)$, and then the continuity of $v(x)$,

$$\begin{aligned}
\frac{1}{\lambda}\inf_{[x_{z2},x_u]}\left((x-x_d)P(A_x)-\lambda x\right) &= \frac{1}{\lambda}\left((x_{z2}-x_d)P(A_{x_{z2}})-\lambda x_{z2}\right)\\
&\geq \frac{1}{\lambda}\inf_{(x_{z1},x_{z2})}\left((x-x_d)P(A_x)-\lambda x\right).
\end{aligned}$$

Therefore, **Case 3** dominates **Case 4**. We have shown that **Case 3** actually provides the globally infimum:

$$\frac{1}{\lambda}\inf_{x\in\mathbb{R}}(v(x)-\lambda x) = \frac{1}{\lambda}\inf_{(x_{z1},x_{z2})}(v(x)-\lambda x).$$

Now we focus on $x \in (x_{z1}, x_{z2})$, where $X(x) = x_d\mathbb{I}_{A_x} + x\mathbb{I}_{B_x} + x_u\mathbb{I}_{D_x}$ satisfy the general constraints:

$$\begin{aligned}
E[X(x)] &= x_dP(A_x)+xP(B_x)+x_uP(D_x) = z,\\
\tilde{E}[X(x)] &= x_d\tilde{P}(A_x)+x\tilde{P}(B_x)+x_u\tilde{P}(D_x) = x_r,
\end{aligned}$$

and the definition for sets A_x, B_x and D_x are

$$A_x = \left\{\omega \in \Omega \,:\, \tfrac{d\tilde{P}}{dP}(\omega) > a_x\right\}, \quad B_x = \left\{\omega \in \Omega \,:\, b_x \leq \tfrac{d\tilde{P}}{dP}(\omega) \leq a_x\right\},$$
$$D_x = \left\{\omega \in \Omega \,:\, \tfrac{d\tilde{P}}{dP}(\omega) < b_x\right\}.$$

Note that $v(x) = (x - x_d)P(A_x)$ (see Theorem 3.7). Since $P(A_x) + P(B_x) + P(D_x) = 1$ and $\tilde{P}(A_x) + \tilde{P}(B_x) + \tilde{P}(D_x) = 1$, we rewrite the capital and return constraints as

$$x - z = (x - x_d)P(A_x) + (x - x_u)P(D_x),$$
$$x - x_r = (x - x_d)\tilde{P}(A_x) + (x - x_u)\tilde{P}(D_x).$$

Differentiating both sides with respect to x, we get

$$P(B_x) = (x - x_d)\frac{dP(A_x)}{dx} + (x - x_u)\frac{dP(D_x)}{dx},$$
$$\tilde{P}(B_x) = (x - x_d)\frac{d\tilde{P}(A_x)}{dx} + (x - x_u)\frac{d\tilde{P}(D_x)}{dx}.$$

Since

$$\frac{d\tilde{P}(A_x)}{dx} = a_x\frac{dP(A_x)}{dx}, \quad \frac{d\tilde{P}(D_x)}{dx} = b_x\frac{dP(D_x)}{dx},$$

we get

$$\frac{dP(A_x)}{dx} = \frac{\tilde{P}(B_x) - bP(B_x)}{(x - x_d)(a - b)}.$$

Therefore,

$$(v(x) - \lambda x)' = P(A_x) + (x - x_d)\frac{dP(A_x)}{dx} - \lambda$$
$$= P(A_x) + \frac{\tilde{P}(B_x) - bP(B_x)}{a - b} - \lambda.$$

When the above derivative is zero, we arrive at the first order Euler condition

$$P(A_x) + \frac{\tilde{P}(B_x) - bP(B_x)}{a - b} - \lambda = 0.$$

To be precise, the above differentiation should be replaced by left-hand and right-hand derivatives as detailed in the Proof for Corollary (2.3) in Chapter 2. But the first order Euler condition will turn out to be the same because we have assumed that the Radon-Nikodým derivative $\frac{d\tilde{P}}{dP}$ has continuous distribution.

To finish this proof, we need to show that there exists an $x \in (x_{z1}, x_{z2})$ where the first order Euler condition is satisfied. From Lemma 3.5, we know that as $x \searrow x_{z1}$, $a_x \nearrow \infty$, and $P(A_x) \searrow 0$. Therefore,

$$\lim_{x \searrow x_{z1}} (v(x) - \lambda x)' = -\lambda < 0.$$

As $x \nearrow x_{z2}$, $b_x \searrow 0$, and $P(D_x) \searrow 0$. Therefore,

$$\lim_{x \searrow x_{z1}} (v(x) - \lambda x)' = P(A_{x_{z2}}) - \frac{\tilde{P}(A^c_{x_{z2}})}{a_{x_{z2}}} - \lambda.$$

This derivative coincide with the derivative of the value function of the Two-Line configuration that is optimal on the interval $x \in [x_{z2}, x_u]$ provided in Theorem 3.7 (see Proof for Corollary (2.3) in Chapter 2). Again when $x \in [x_{z2}, x_u]$ and $\operatorname{ess\,sup} \frac{d\tilde{P}}{dP} > \frac{1}{\lambda}$, Theorem 3.7 and Theorem 3.8 imply that the infimum of $v(x) - \lambda x$ is achieved either by $\bar{X}$ or X^*. Since we restrict $z \in (z^*, \bar{z}]$ where $z^* = \bar{z}$ by Definition 3.6 in the first case, we need not consider this case in the current proposition. In the second case, Lemma 3.2 implies that $x^* < x_{z2}$ (because $z > z^*$). This in turn implies

$$P(A_{x_{z2}}) - \frac{\tilde{P}(A^c_{x_{z2}})}{a_{x_{z2}}} - \lambda < 0.$$

We have just shown that there exist some $x^{**} \in (x_{z1}, x_{z2})$ such that $(v(x) - \lambda x)'|_{x=x^{**}} = 0$. By the convexity of $v(x) - \lambda x$, this is the point where it obtains the minimum value. Now

$$\begin{aligned} CVaR(X^{**}) &= \frac{1}{\lambda} \left(v(x^{**}) - \lambda x^{**} \right) \\ &= \frac{1}{\lambda} \left((x^{**} - x_d) P(A^{**}) - \lambda x^{**} \right). \end{aligned}$$

Q.E.D.

Theorem 3.11 (Solution to **Step 2:** Minimization of Conditional Value-at-Risk)**.** *For fixed*

$-\infty < x_d < x_r < x_u < \infty$.

1. *Suppose* $\operatorname{ess\,sup} \frac{d\tilde{P}}{dP} \leq \frac{1}{\lambda}$ *and* $z = x_r$. $X = x_r$ *is the optimal solution to* ***Step 2: Minimization of Conditional Value-at-Risk*** *of the* ***Two-Constraint Problem*** *and the associated minimal risk is*

$$CVaR(X) = -x_r.$$

2. *Suppose* $\operatorname{ess\,sup} \frac{d\tilde{P}}{dP} \leq \frac{1}{\lambda}$ *and* $z \in (x_r, \bar{z}]$. *The optimal solution to* ***Step 2: Minimization of Conditional Value-at-Risk*** *of the* ***Two-Constraint Problem*** *does not exist and the minimal risk is*

$$CVaR(X) = -x_r.$$

3. *Suppose* $\operatorname{ess\,sup} \frac{d\tilde{P}}{dP} > \frac{1}{\lambda}$ *and* $z \in [x_r, z^*]$.

   - *If* $\frac{1}{\bar{a}} \leq \frac{\lambda - P(\bar{A})}{1 - \tilde{P}(\bar{A})}$ *(see Definition 3.4), then* $\bar{X} = x_d \mathbb{I}_{\bar{A}} + x_u \mathbb{I}_{\bar{D}}$ *is the optimal solution to* ***Step 2: Minimization of Conditional Value-at-Risk*** *of the* ***Two-Constraint Problem*** *and the associated minimal risk is*

$$CVaR(\bar{X}) = -x_r + \frac{1}{\lambda}(x_u - x_d)(P(\bar{A}) - \lambda \tilde{P}(\bar{A})).$$

   - *Otherwise,* $X^* = x_d \mathbb{I}_{A^*} + x^* \mathbb{I}_{B^*}$ *defined in Theorem 3.8 is the optimal solution to* ***Step 2: Minimization of Conditional Value-at-Risk*** *of the* ***Two-Constraint Problem*** *and the associated minimal risk is*

$$CVaR(X^*) = -x_r + \frac{1}{\lambda}(x^* - x_d)(P(A^*) - \lambda \tilde{P}(A^*)).$$

4. *Suppose* $\operatorname{ess\,sup} \frac{d\tilde{P}}{dP} > \frac{1}{\lambda}$ *and* $z \in (z^*, \bar{z}]$. $X^{**} = x_d \mathbb{I}_{A^{**}} + x^{**} \mathbb{I}_{B^{**}} + x_u \mathbb{I}_{D^{**}}$ *defined in Proposition 3.10 is the optimal solution to* ***Step 2: Minimization of Conditional Value-at-Risk*** *of the* ***Two-Constraint Problem*** *and the associated minimal risk*

is

$$CVaR(X^{**}) = -x_r + \frac{1}{\lambda}\left((x^{**} - x_d)P(A^{**}) - \lambda x^{**}\right).$$

Proof. Case 3 and 4 are already proved in Theorem 3.8 and Proposition 3.10. In Case 1 where $\operatorname{ess\,sup} \frac{d\tilde{P}}{dP} \leq \frac{1}{\lambda}$ and $z = x_r$, $X = x_r$ is both feasible and optimal by Theorem 3.8. In Case 2, fix arbitrary $\epsilon > 0$. We will look for a Two-Line solution $X_\epsilon = x_\epsilon \mathbb{I}_{A_\epsilon} + \alpha_\epsilon \mathbb{I}_{B_\epsilon}$ with the right parameters $a_\epsilon, x_\epsilon, \alpha_\epsilon$ which satisfies both the capital constraint and return constraint:

$$E[X_\epsilon] = x_\epsilon P(A_\epsilon) + \alpha_\epsilon P(B_\epsilon) = z, \tag{3.12}$$

$$\tilde{E}[X_\epsilon] = x_\epsilon \tilde{P}(A_\epsilon) + \alpha_\epsilon \tilde{P}(B_\epsilon) = x_r, \tag{3.13}$$

where

$$A_\epsilon = \left\{\omega \in \Omega \,:\, \tfrac{d\tilde{P}}{dP}(\omega) > a_\epsilon\right\}, \quad B_\epsilon = \left\{\omega \in \Omega \,:\, \tfrac{d\tilde{P}}{dP}(\omega) \leq a_\epsilon\right\},$$

and produces a CVaR level close to the lower bound:

$$CVaR(X_\epsilon) \leq CVaR(x_r) + \epsilon = -x_r + \epsilon.$$

First, we choose $x_\epsilon = x_r - \epsilon$. To find the remaining two parameters a_ϵ and α_ϵ so that equations (3.12) and (3.13) are satisfies, we note

$$x_r P(A_\epsilon) + x_r P(B_\epsilon) = x_r,$$

$$x_r \tilde{P}(A_\epsilon) + x_r \tilde{P}(B_\epsilon) = x_r,$$

and conclude that it is equivalent to find a pair of a_ϵ and α_ϵ such that the following two equalities are satisfied:

$$-\epsilon P(A_\epsilon) + (\alpha_\epsilon - x_r)P(B_\epsilon) = \gamma,$$

$$-\epsilon \tilde{P}(A_\epsilon) + (\alpha_\epsilon - x_r)\tilde{P}(B_\epsilon) = 0,$$

where we denote $\gamma = z - x_r$. If we can find a solution a_ϵ to the equation

$$\frac{\tilde{P}(B_\epsilon)}{P(B_\epsilon)} = \frac{\epsilon}{\gamma + \epsilon}, \tag{3.14}$$

then

$$\alpha_\epsilon = x_r + \frac{\tilde{P}(A_\epsilon)}{\tilde{P}(B_\epsilon)}\epsilon,$$

and we have the solutions for equations (3.12) and (3.13). It is not difficult to prove that the fraction $\frac{\tilde{P}(B)}{P(B)}$ increases continuously from 0 to 1 as a increases from 0 to $\frac{1}{\lambda}$. Therefore, we can find a solution $a_\epsilon \in (0, \frac{1}{\lambda})$ where (3.14) is satisfied. By definition (1.2),

$$CVaR_\lambda(X_\epsilon) = \frac{1}{\lambda}\inf_{x\in\mathbb{R}}\left(E[(x - X_\epsilon)^+] - \lambda x\right) \le \frac{1}{\lambda}\left(E[(x_\epsilon - X_\epsilon)^+] - \lambda x_\epsilon\right) = -x_\epsilon.$$

The difference

$$CVaR_\lambda(X_\epsilon) - CVaR(x_r) \le -x_\epsilon + x_r = \epsilon.$$

Under Assumptions 1.1, 1.2, the solution in Case 2 is almost surely unique, the result is proved. *Q.E.D.*

3.3 Case: $x_u = \infty$

We first restate Theorem 3.8 in the current context. When $x_u = \infty$, we interpret $\bar{A} = \Omega$ and $\bar{z} = \infty$.

Theorem 3.12 (Theorem 2.4 and its Remark in Chapter 2 when $x_u = \infty$)**.**

1. *Suppose* $\operatorname{ess\,sup} \frac{d\tilde{P}}{dP} \le \frac{1}{\lambda}$. *$X = x_r$ is the optimal solution to **Step 2: Minimization of Conditional Value-at-Risk** of the **One-Constraint Problem** and the associated minimal risk is*

$$CVaR(X) = -x_r.$$

2. *Suppose* $\operatorname{ess\,sup} \frac{d\tilde{P}}{dP} > \frac{1}{\lambda}$. *Let a^* be the solution to the equation $\frac{1}{a} = \frac{\lambda - P(A)}{1 - \tilde{P}(A)}$. Associate sets $A^* = \left\{\omega \in \Omega : \frac{d\tilde{P}}{dP}(\omega) > a^*\right\}$ and $B^* = \left\{\omega \in \Omega : \frac{d\tilde{P}}{dP}(\omega) \le a^*\right\}$ to level a^*.*

Define $x^* = \frac{x_r - x_d \tilde{P}(A^*)}{1-\tilde{P}(A^*)}$ *so that configuration*

$$X^* = x_d \mathbb{I}_{A^*} + x^* \mathbb{I}_{B^*}$$

satisfies the capital constraint $\tilde{E}[X^*] = x_d \tilde{P}(A^*) + x^* \tilde{P}(B^*) = x_r$ *in* ***Degenerated Constraints 2****. Then* X^* *is the optimal solution to* ***Step 2: Minimization of Conditional Value-at-Risk*** *of the* ***One-Constraint Problem*** *and the associated minimal risk is*

$$CVaR(X^*) = -x_r + \frac{1}{\lambda}(x^* - x_d)(P(A^*) - \lambda \tilde{P}(A^*)).$$

Theorem 3.13 (Minimization of Conditional Value-at-Risk When $x_u = \infty$)**.** *For fixed* $-\infty < x_d < x_r < x_u < \infty$.

1. *Suppose* $\operatorname{ess\,sup} \frac{d\tilde{P}}{dP} \leq \frac{1}{\lambda}$ *and* $z = x_r$. $X = x_r$ *is the optimal solution to* ***Step 2: Minimization of Conditional Value-at-Risk*** *of the* ***Two-Constraint Problem*** *and the associated minimal risk is*

$$CVaR(X) = -x_r.$$

2. *Suppose* $\operatorname{ess\,sup} \frac{d\tilde{P}}{dP} \leq \frac{1}{\lambda}$ *and* $z \in (x_r, \infty)$. *The optimal solution to* ***Step 2: Minimization of Conditional Value-at-Risk*** *of the* ***Two-Constraint Problem*** *does not exist and the minimal risk is*

$$CVaR(X) = -x_r.$$

3. *Suppose* $\operatorname{ess\,sup} \frac{d\tilde{P}}{dP} > \frac{1}{\lambda}$ *and* $z \in [x_r, z^*]$. X^* *is the optimal solution to* ***Step 2: Minimization of Conditional Value-at-Risk*** *of the* ***Two-Constraint Problem*** *and the associated minimal risk is*

$$CVaR(X^*) = -x_r + \frac{1}{\lambda}(x^* - x_d)(P(A^*) - \lambda \tilde{P}(A^*)).$$

4. *Suppose* $\operatorname{ess\,sup} \frac{d\tilde{P}}{dP} > \frac{1}{\lambda}$ *and* $z \in (z^*, \infty)$*. The optimal solution to* ***Step 2: Minimization of Conditional Value-at-Risk*** *of the* ***Two-Constraint Problem*** *does not exist and the minimal risk is*

$$CVaR(X^*) = -x_r + \frac{1}{\lambda}(x^* - x_d)(P(A^*) - \lambda\tilde{P}(A^*)).$$

Proof. Case 1 and 3 are obviously true in light of Theorem 3.12. The proof for Case 2 is similar to that in the Proof of Theorem 3.11, so we will not repeat it here. Since $E[X^*] = z^* < z$ in case 4, $CVaR(X^*)$ is only a lower bound in this case. We first show that it is the true infimum obtained in Case 4. Fix arbitrary $\epsilon > 0$. We will look for a Three-Line solution $X_\epsilon = x_d\mathbb{I}_{A_\epsilon} + x_\epsilon\mathbb{I}_{B_\epsilon} + \alpha_\epsilon\mathbb{I}_{D_\epsilon}$ with the right parameters $a_\epsilon, b_\epsilon, x_\epsilon, \alpha_\epsilon$ which satisfies the general constraints:

$$E[X_\epsilon] = x_d P(A_\epsilon) + x_\epsilon P(B_\epsilon) + \alpha_\epsilon P(D_\epsilon) = z, \tag{3.15}$$

$$\tilde{E}[X_\epsilon] = x_d \tilde{P}(A_\epsilon) + x_\epsilon \tilde{P}(B_\epsilon) + \alpha_\epsilon \tilde{P}(D_\epsilon) = x_r, \tag{3.16}$$

where

$$A_\epsilon = \left\{\omega \in \Omega \,:\, \tfrac{d\tilde{P}}{dP}(\omega) > a_\epsilon\right\}, \quad B_\epsilon = \left\{\omega \in \Omega \,:\, b_\epsilon \le \tfrac{d\tilde{P}}{dP}(\omega) \le a_\epsilon\right\},$$
$$D_\epsilon = \left\{\omega \in \Omega \,:\, \tfrac{d\tilde{P}}{dP}(\omega) < b_\epsilon\right\},$$

and produces a CVaR level close to the lower bound:

$$CVaR(X_\epsilon) \le CVaR(X^*) + \epsilon.$$

First, we choose $a_\epsilon = a^*$, $A_\epsilon = A^*$, $x_\epsilon = x^* - \delta$, where we define $\delta = \frac{\lambda}{\lambda - P(A^*)}\epsilon$. To find the remaining two parameters b_ϵ and α_ϵ so that equations (3.15) and (3.16) are satisfies, we note

$$E[X^*] = x_d P(A^*) + x^* P(B^*) = z^*,$$
$$\tilde{E}[X^*] = x_d \tilde{P}(A^*) + x^* \tilde{P}(B^*) = x_r,$$

and conclude that it is equivalent to find a pair of b_ϵ and α_ϵ such that the following two equalities are satisfied:

$$-\delta(P(B^*) - P(D_\epsilon)) + (\alpha_\epsilon - x^*)P(D_\epsilon) = \gamma,$$
$$-\delta(\tilde{P}(B^*) - \tilde{P}(D_\epsilon)) + (\alpha_\epsilon - x^*)\tilde{P}(D_\epsilon) = 0,$$

where we denote $\gamma = z - z^*$. If we can find a solution b_ϵ to the equation

$$\frac{\tilde{P}(D_\epsilon)}{P(D_\epsilon)} = \frac{\tilde{P}(B^*)}{\frac{\gamma}{\delta} + P(B^*)}, \tag{3.17}$$

then

$$\alpha_\epsilon = x^* + \left(\frac{\tilde{P}(B^*)}{\tilde{P}(D_\epsilon)} - 1 \right) \delta,$$

and we have the solutions for equations (3.15) and (3.16). It is not difficult to prove that the fraction $\frac{\tilde{P}(D)}{P(D)}$ increases continuously from 0 to $\frac{\tilde{P}(B^*)}{P(B^*)}$ as b increases from 0 to a^*. Therefore, we can find a solution $b_\epsilon \in (0, a^*)$ where (3.17) is satisfied. By definition (1.2),

$$\begin{aligned} CVaR_\lambda(X_\epsilon) &= \frac{1}{\lambda} \inf_{x \in \mathbb{R}} \left(E[(x - X_\epsilon)^+] - \lambda x \right) \\ &\le \frac{1}{\lambda} \left(E[(x_\epsilon - X_\epsilon)^+] - \lambda x_\epsilon \right) \\ &= \frac{1}{\lambda}(x_\epsilon - x_d)P(A_\epsilon) - x_\epsilon. \end{aligned}$$

The difference

$$\begin{aligned} CVaR_\lambda(X_\epsilon) - CVaR(X^*) &\le \frac{1}{\lambda}(x_\epsilon - x_d)P(A_\epsilon) - x_\epsilon - \frac{1}{\lambda}(x^* - x_d)P(A^*) + x^* \\ &= \frac{1}{\lambda}(x^* - x_d)(P(A_\epsilon) - P(A^*)) + \left(1 - \frac{P(A_\epsilon)}{\lambda}\right)(x^* - x_\epsilon) = \epsilon. \end{aligned}$$

Under Assumptions 1.1, 1.2, the solution in Case 4 is almost surely unique, the result is proved. *Q.E.D.*

3.4 Application to Black-Scholes Model

The dynamics of the stock price and the self-financing portfolio in the Black-Scholes model are as follows:

$$dS_t = S_t(\mu dt + \sigma dW_t),$$
$$dX_t = \xi_t dS_t + (X_t - \xi_t S_t) r dt.$$

Our main goal is once again first to find

$$\inf_{\xi_t} CVaR_\lambda(X_T) \quad \text{s.t.} \quad \tilde{E}[X_T] = x_r,\ E[X_T] \geq z,\ x_d \leq X_T \leq x_u\ a.s., \tag{3.18}$$

and then to find the corresponding dynamic hedging ξ_t.

Proposition 3.14. *The calculations of 'bar-system', 'star-system' are as in Theorem 3.11, and the calculation of 'double-star-system' is calculated as in Proposition 3.10. Namely:*

bar-system

$$\tilde{P}(\bar{A}) = \frac{x_u - x_r}{x_u - x_d},$$
$$\bar{a} = e^{\theta\sqrt{T}\left[\frac{\theta\sqrt{T}}{2} - N^{-1}(\tilde{P}(\bar{A}))\right]},$$
$$P(\bar{A}) = N\left(-\tfrac{\theta\sqrt{T}}{2} - \tfrac{\ln \bar{a}}{\theta\sqrt{T}}\right).$$

star-system

$$
\begin{aligned}
a^* &= \left\{ a : \, a = \frac{1-N(\frac{\theta\sqrt{T}}{2} - \frac{\ln a}{\theta\sqrt{T}})}{\lambda - N(-\frac{\theta\sqrt{T}}{2} - \frac{\ln a}{\theta\sqrt{T}})} \right\}, \\
P(A^*) &= N(-\frac{\theta\sqrt{T}}{2} - \frac{\ln a^*}{\theta\sqrt{T}}), \\
\tilde{P}(A^*) &= N(\frac{\theta\sqrt{T}}{2} - \frac{\ln a^*}{\theta\sqrt{T}}), \\
x^* &= \frac{x_r - x_d\tilde{P}(A^*)}{1-\tilde{P}(A^*)}.
\end{aligned}
$$

double-star-system

The system (a^{**}, b^{**}, x^{**}) (and consequently, A^{**}, B^{**}, D^{**},) is the solution to the following equations:

$$
\begin{aligned}
&x_d P(A) + xP(B) + x_u P(D) = z, \quad (\textit{return constraint}) \\
&x_d \tilde{P}(A) + x\tilde{P}(B) + x_u \tilde{P}(D) = x_r, \quad (\textit{capital constraint}) \\
&P(A) + \frac{\tilde{P}(B) - bP(B)}{a-b} - \lambda = 0. \quad (\textit{first order Euler condition})
\end{aligned}
$$

$\bar{z}$ and z^* are defined in Definition 3.4 and 3.6:

$$
\bar{z} = x_d P(\bar{A}) + x_u(1 - P(\bar{A})),
$$

$$
z^* = \begin{cases} x_d P(\bar{A}) + x_u(1-P(\bar{A})) & : \ \textit{if } x_u < \infty \textit{ and } \frac{1}{\bar{a}} \leq \frac{\lambda - P(\bar{A})}{1-\tilde{P}(\bar{A})} \\ x_d P(A^*) + x^*(1-P(A^*)) & : \ \textit{otherwise.} \end{cases}
$$

We can compute the solution to problem (3.18) and the corresponding dynamic hedging ξ_t in the Black-Scholes model as below:

1. *Suppose $z \in [x_r, z^*]$.*

   - *If $x_u < \infty$ and $\frac{1}{\bar{a}} \leq \frac{\lambda - P(\bar{A})}{1-\tilde{P}(\bar{A})}$, then the minimal risk, the optimal portfolio and the* *hedging strategy are*

$$
\begin{aligned}
CVaR_\lambda(X_T^{**}) &= -x_r + \tfrac{1}{\lambda}(x_u - x_d)(P(\bar{A}) - \lambda\tilde{P}(\bar{A})),\\
X_t^{**} &= e^{-r(T-t)}\left[x_u N\left(d_+(\bar{a}, S_t, t)\right) + x_d N\left(d_-(\bar{a}, S_t, t)\right)\right],\\
\xi_t^{**} &= \frac{x_u - x_d}{\sigma S_t\sqrt{2\pi(T-t)}} e^{-r(T-t) - \frac{d_-^2(\bar{a}, S_t, t)}{2}}.
\end{aligned}
$$

- *Otherwise, the minimal risk, the optimal portfolio and the hedging strategy are*

$$
\begin{aligned}
CVaR_\lambda(X_T^{**}) &= -x_r + \tfrac{1}{\lambda}(x^* - x_d)(P(A^*) - \lambda\tilde{P}(A^*)),\\
X_t^{**} &= e^{-r(T-t)}\left[x^* N\left(d_+(a^*, S_t, t)\right) + x_d N\left(d_-(a^*, S_t, t)\right)\right],\\
\xi_t^{**} &= \frac{x^* - x_d}{\sigma S_t\sqrt{2\pi(T-t)}} e^{-r(T-t) - \frac{d_-^2(a^*, S_t, t)}{2}}.
\end{aligned}
$$

2. *Suppose* $z \in (z^*, \bar{z}]$.

- *If* $x_u < \infty$, *then the minimal risk, optimal portfolio and the hedging strategy are*

$$
\begin{aligned}
CVaR(X_T^{**}) &= -x_r + \frac{1}{\lambda}\left((x^{**} - x_d)P(A^{**}) - \lambda x^{**}\right),\\
X_t^{**} &= e^{-r(T-t)}[x^{**}N(d_+(a^{**}, S_t, t)) + x_d N(d_-(a^{**}, S_t, t))]\\
&\quad + e^{-r(T-t)}[x^{**}N(d_-(b^{**}, S_t, t)) + x_u N(d_+(b^{**}, S_t, t))]\\
&\quad - e^{r(T-t)}x^{**},\\
\xi_t^{**} &= \frac{x^{**} - x_d}{\sigma S_t\sqrt{2\pi(T-t)}} e^{-r(T-t) - \frac{d_-^2(a^{**}, S_t, t)}{2}}\\
&\quad + \frac{x^{**} - x_u}{\sigma S_t\sqrt{2\pi(T-t)}} e^{-r(T-t) - \frac{d_+^2(b^{**}, S_t, t)}{2}}.
\end{aligned}
$$

- *If* $x_u = \infty$, *then the optimal portfolio or the hedging strategy does not exist, but the minimal risk is*

$$
CVaR(X^{**}) = -x_r + \tfrac{1}{\lambda}(x^* - x_d)(P(A^*) - \lambda\tilde{P}(A^*)).
$$

Remark. Recall Definition 3.4, the bar-system ($\bar{a}$, $\bar{b}$, $\bar{x}$) (and consequently, $\bar{A}$, $\bar{B}$, $\bar{D}$), where

$\bar{a} = \bar{b}$ (and consequently $P(\bar{B}) = \tilde{P}(\bar{B}) = 0$), is the solution to the following equation:

$$x_d \tilde{P}(A) + x_u \tilde{P}(D) = x_r. \quad \text{(capital constraint)}$$

The star-system (a^*, b^*, x^*) (and consequently, A^*, B^*, D^*), where $b^* = 0$ (and consequently $P(\bar{D}) = \tilde{P}(\bar{D}) = 0$), is the solution to the following equations:

$$x_d \tilde{P}(A) + x \tilde{P}(D) = x_r, \quad \text{(capital constraint)}$$
$$P(A) + \frac{\tilde{P}(D)}{a} - \lambda = 0. \quad \text{(first order Euler condition)}$$

Proof. The Radon Nikodým derivative process for geometric Brownian motion model is $Z_t := \frac{d\tilde{P}}{dP}|_t = e^{-\theta W_t - \frac{\theta^2}{2} t}$, where $\theta = \frac{\mu - r}{\sigma}$. Obviously, $P\left(\frac{d\tilde{P}}{dP}(\omega)\big|_T > \frac{1}{\lambda}\right) > 0$ since $\operatorname{ess\,sup} Z_T = \infty$. According to Theorem 3.11 and Theorem 3.13, if $z \in [x_r, z^*]$, then the optimal solution is the same as that for the one-constraint problem described in Proposition 2.7, otherwise we need to see if $x_u < \infty$. When x_u is unbounded, there is no optimal final portfolio X_T^{**} or optimal strategy ξ_t^{**}, but the minimal CVaR is given. When x_u is bounded from above, the optimal is achieved by double-star-system (a^{**}, b^{**}, x^{**}) (and consequently, A^{**}, B^{**}, D^{**},), which can be found numerically. A direct application of Theorem 3.11 gives the optimal final portfolio X_T^{**} and the minimal CVaR. The corresponding optimal hedging ξ_t^{**} is the usual 'Delta-hedge' in the Black-Scholes model where the derivative payoff is X_T^{**}, so we first calculate the Risk Neutral price process $X_t = v(S_t, t)$, and then differentiate with respect to the stock price S_t.

$$X_t = e^{r(T-t)} \tilde{E}[X_T^{**} | \mathcal{F}_t] = e^{-r(T-t)} \left[x_d \tilde{P}_t(A^{**}) + x^{**} \tilde{P}_t(B^{**}) + x_u \tilde{P}_t(D^{**}) \right]$$

where $\tilde{P}_t(\bullet)$ is the conditional probability under the Risk Neutral measure. Since

$$A^{**} = \{Z_T > a^{**}\} = \left\{ Z_t e^{-\theta(W_T - W_t) - \frac{\theta^2}{2}(T-t)} > a^{**} \right\} = \left\{ \frac{\tilde{W}_T - \tilde{W}_t}{\sqrt{T-t}} < -\frac{\ln \frac{a^{**}}{Z_t}}{\theta\sqrt{T-t}} + \frac{\theta}{2}\sqrt{T-t} \right\},$$

we have then

$$\tilde{P}_t(A^{**}) = N\left(-\frac{\ln \frac{a^{**}}{Z_t}}{\theta\sqrt{T-t}} + \frac{\theta}{2}\sqrt{T-t}\right). \tag{3.19}$$

Similarly, we have

$$\tilde{P}_t(D^{**}) = 1 - N\left(-\frac{\ln \frac{b^{**}}{Z_t}}{\theta\sqrt{T-t}} + \frac{\theta}{2}\sqrt{T-t}\right). \tag{3.20}$$

Note that Z_t can be represented by the stock price S_t:

$$S_t = S_0 e^{\sigma W_t + (\mu - \frac{1}{2}\sigma^2)t} \Rightarrow W_t = \frac{\ln \frac{S_t}{S_0} - (\mu - \frac{1}{2}\sigma^2)t}{\sigma},$$
$$Z_t = e^{-\theta W_t - \frac{1}{2}\theta^2 t} \Rightarrow Z_t = g(S_t, t), \text{ where } g(s,t) = e^{\frac{\theta}{\sigma}[\frac{\mu + r - \sigma^2}{2}t - \ln \frac{s}{S_0}]}.$$

Substitute $g(S_t, t)$ into (3.19) and (3.20) we have

$$\tilde{P}_t(A^{**}) = N\left(d_-(a^{**}, S_t, t)\right),$$
$$\tilde{P}_t(D^{**}) = 1 - N\left(d_-(b^{**}, S_t, t)\right) = N\left(d_+(b^{**}, S_t, t)\right).$$

where

$$d_-(a,s,t) = \frac{1}{\theta\sqrt{T-t}}[-\ln a + \frac{\theta}{\sigma}(\frac{\mu + r - \sigma^2}{2}t - \ln \frac{s}{S_0}) + \frac{\theta^2}{2}(T-t)],$$
$$d_+(a,s,t) = \frac{1}{\theta\sqrt{T-t}}[\ln a + \frac{\theta}{\sigma}(\frac{\mu + r - \sigma^2}{2}t - \ln \frac{s}{S_0}) + \frac{\theta^2}{2}(T-t)],$$

Hence

$$X_t = v(S_t, t),$$

where

$$v(s,t) = e^{-r(T-t)} x_d N\left(d_-(a^{**}, s, t)\right) + e^{-r(T-t)} x^{**}[1 - N\left(d_-(a^{**}, s, t)\right) - N\left(d_+(b^{**}, s, t)\right)]$$
$$+ e^{-r(T-t)} x_u N\left(d_+(b^{**}, s, t)\right).$$

Rewrite

$$
\begin{aligned}
v(s,t) &= e^{-r(T-t)}\left[x^{**}N\left(d_+(a^{**},s,t)\right) + x_d N\left(d_-(a^{**},s,t)\right)\right] \\
&\quad + e^{-r(T-t)}\left[x^{**}N\left(d_-(b^{**},s,t)\right) + x_u N\left(d_+(b^{**},s,t)\right)\right] - e^{-r(T-t)}x^{**} \\
&= e^{-r(T-t)}\left[x^{**} + (x_d - x^{**})N\left(d_-(a^{**},s,t)\right)\right] \\
&\quad + e^{-r(T-t)}\left[x^{**} + (x_u - x^{**})N\left(d_+(b^{**},s,t)\right)\right] - e^{-r(T-t)}x^{**},
\end{aligned}
$$

we find the partial derivative

$$
v_s(s,t) = \frac{x^{**} - x_d}{\sigma s\sqrt{2\pi(T-t)}} e^{-r(T-t) - \frac{d_-^2(a^{**},s,t)}{2}} + \frac{x^{**} - x_u}{\sigma s\sqrt{2\pi(T-t)}} e^{-r(T-t) - \frac{d_+^2(b^{**},s,t)}{2}}.
$$

Given the stock price S_t at time t, the optimal strategy ξ_t^{**} is:

$$
\xi_t^{**} = v_s(S_t, t).
$$

Q.E.D.

Algorithm 3.15. *CVaR Minimization for Black-Scholes Model*

1. *Find 'bar-system', 'star-system', $\bar{z}$ and z^*.*

2. *If $z \in [x_r,\, z^*]$, then go to Step* (3). *Otherwise, go to Step* (4).

3. *If $x_u < \infty$ and $\frac{1}{\bar{a}} \leq \frac{\lambda - P(\bar{A})}{1 - \tilde{P}(\bar{A})}$, then*

$$
\begin{aligned}
\inf_{\xi_t} CVaR_\lambda(X_T) &= -x_r + \tfrac{1}{\lambda}(x_u - x_d)(P(\bar{A}) - \lambda\tilde{P}(\bar{A})), \\
X_t^* &= e^{-r(T-t)}\left[x_u N\left(d_+(\bar{a}, S_t, t)\right) + x_d N\left(d_-(\bar{a}, S_t, t)\right)\right], \\
\xi_t^* &= \frac{x_u - x_d}{\sigma S_t\sqrt{2\pi(T-t)}} e^{-r(T-t) - \frac{d_-^2(\bar{a}, S_t, t)}{2}}.
\end{aligned}
$$

Otherwise

$$\begin{aligned}\inf_{\xi_t} CVaR_\lambda(X_T) = -x_r + \tfrac{1}{\lambda}(x^* - x_d)(P(A^*) - \lambda\tilde{P}(A^*)),\\
X_t^* &= e^{-r(T-t)}\left[x^* N\left(d_+(a^*, S_t, t)\right) + x_d N\left(d_-(a^*, S_t, t)\right)\right],\\
\xi_t^* &= \frac{x^* - x_d}{\sigma S_t\sqrt{2\pi(T-t)}} e^{-r(T-t) - \frac{d_-^2(a^*, S_t, t)}{2}}.\end{aligned}$$

Stop.

4. *If $x_u < \infty$, then find 'double-star-system', and*

$$\begin{aligned}CVaR(X_T^{**}) &= -x_r + \frac{1}{\lambda}\left((x^{**} - x_d)P(A^{**}) - \lambda x^{**}\right),\\
X_t^{**} &= e^{-r(T-t)}[x^{**}N(d_+(a^{**}, S_t, t)) + x_d N(d_-(a^{**}, S_t, t))]\\
&\quad + e^{-r(T-t)}[x^{**}N(d_-(b^{**}, S_t, t)) + x_u N(d_+(b^{**}, S_t, t))]\\
&\quad - e^{r(T-t)}x^{**},\\
\xi_t^{**} &= \frac{x^{**} - x_d}{\sigma S_t\sqrt{2\pi(T-t)}} e^{-r(T-t) - \frac{d_-^2(a^{**}, S_t, t)}{2}}\\
&\quad + \frac{x^{**} - x_u}{\sigma S_t\sqrt{2\pi(T-t)}} e^{-r(T-t) - \frac{d_+^2(b^{**}, S_t, t)}{2}}.\end{aligned}$$

Otherwise, the optimal portfolio or the hedging strategy does not exist, and

$$CVaR(X^{**}) = -x_r + \tfrac{1}{\lambda}(x^* - x_d)(P(A^*) - \lambda\tilde{P}(A^*)).$$

Stop.

3.5 Comparison: Minimal CVaR with & without Expected Return Constraint

Let us take Example 2 in Chapter 2, where the optimal of the one-constraint problem is achieved by 'star-system'. Recall that the parameters are $\lambda = 5\%$, $T = 2$, $r = 5\%$, $S_0 = 10$, $X_0 = 10$, $x_d = 5$, $\mu = 0.2$, $\sigma = 0.1$. The levels of z is selected so that $z \in [x_r,\, z^*]$ since it is of the most interest to see cases where optimal of the two-constraint problem is achieved by 'double-star-system'.

One-Constraint Optimization			Two-Constraint Optimization			
x_u	30	50	x_u	30	30	50
			z	18	25	25
			$\bar{z}$	28.0609	28.0609	43.0320
			z^*	15.3352	15.3352	15.3352
x^*	15.4407	15.4407	x^{**}	15.4859	15.0151	15.3753
a^*	14.5304	14.5304	a^{**}	14.2057	11.2895	13.9225
			b^{**}	0.0152	0.2769	0.0308
$CVaR(X_T^*)$	-13.3297	-13.3297	$CVaR(X_T^{**})$	-13.2585	-12.2534	-13.1624

Table 3.1: Black-Scholes Model Example with & without Expected Return Constraint

Whether the upper bound x_u being 30 or 50 does not have any impact on the 'star-system' (x^*, a^*), and $\hat{z}$ calculated from this system is not impacted either. However, as the upper bound x_u increases, $\bar{z}$ increases, thus we can require a higher expected return z.

Compare the results of the three cases where $x_u = 30$ in the table, we see that the higher the required return, the harder it is to obtain a low CVaR. This is also true for the two cases where $x_u = 50$. Now let us compare the two columns to the right: the two cases have the same required return 25. When the upper bound x_u is higher (=50), the attainable return $\bar{z}$ is higher (=43.0320), the required return $z = 25$ is relatively easier to achieve, thus has less an impact in minimizing CVaR. (CVaR only decreases a little from -13.3297 to -13.1624.) An intuition we can get from this comparison is that when we let the upper bound be extremely large ($x_u \uparrow \infty$), the attainable return $\bar{z}$ will also be so large that any required return z will seem to be effortless to obtain, thus the value of minimal CVaR is almost not impacted.

Also have a look at the threshold b^{**}: when this number is small, the optimal of the two-constraint problem is very close to the optimal of the one-constraint problem. In the extreme that $b^{**} = 0$, the two problems coincide. With the same upper bound, the higher the required return, the more adjustment needs to be made on top of X_T^*, thus the larger the b value. With the same required return, the higher the upper bound, the less the effort in adjusting X_T^*, thus the less the b value. In the extreme when $x_u \uparrow \infty$, $b^{**} \downarrow 0$, thus $CVaR(X_T^{**}) \downarrow CVaR(X_T^*)$.

CHAPTER 4: CONCLUSION AND DISCUSSION

We have so far found exact solutions for CVaR minimization in complete market models. Without the expected return constraint, we get results with both Continuous and Discrete distribution of Radon-Nikodým derivative, cite Binomial model and Black-Scholes model as examples, and describe procedures of finding the optimals with exact formulae. With the additional constraint on expected return, we require the Radon-Nikodým derivative to have continuous distribution as in Assumption 1.2, and a theoretical solution is found. When this assumption is weakened, the results should still hold, albeit in a more complicated form. It will also be very interesting to extend this result for CVaR minimization to minimizing Law-Invariant Risk Measures in general.

REFERENCES

[1] Acerbi, C., D. Tasche (2002): "On the coherence of expected shortfall," *Journal of Banking and Finance*, **26**, 1487–1503.

[2] Bielecki, T., H. Jin, S. R. Pliska, X. Y. Zhou (2005): "Continuous-time mean-variance portfolio selection with bankruptcy prohibition", *Mathematical Finance*, **15**, 213–244.

[3] Alexander, S., T. F. Coleman, Y. Li (2004): "Derivative portfolio hedging based on CVaR", *Risk Measures for the 21st Century*, John Wiley & Sons, Hoboken, NJ, USA, 339–363.

[4] Artzner, P., F. Delbaen, J.-M. Eber, D. Heath (1997): "Thinking coherently," *Risk*, **10**, 68–71.

[5] Artzner, P., F. Delbaen, J.-M. Eber, D. Heath (1999): "Coherent measures of risk," *Mathematical Finance*, **9**, 203–228.

[6] Delbaen, F., W. Schachermayer (1994): "A general version of the fundamental theorem of asset pricing," *Mathematische Annalen*, **300**, 463–520.

[7] El Karoui, N., M. C. Quenez (1995): "Dynamic programming and pricing of contingent claims in an incomplete market," *SIAM Journal on Control and Optimization*, **33, 1**, 29–66.

[8] Föllmer, H., Yu. M. Kabanov (1998): "Optional decomposition and lagrange multipliers," *Finance and Stochastics*, **2**, 69–81.

[9] Föllmer, H., P. Leukert (2000): "Efficient hedging: cost versus shortfall risk," *Finance and Stochastics*, **4**, 117–146.

[10] Föllmer, H., A. Schied (2002): *Stochastic finance - an introduction in discrete time,* Walter de Gruyter, Berlin, Germany, Studies in Mathematics, **27**.

[11] Glasserman, P. (2004): *Monte Carlo methods in financial engineering,* Springer.

[12] Kramkov (1996): "Optional decomposition of supermartingales and hedging contingent claims in incomplete security markets," *Probability Theory and Related Fields*, **105**, 459–479.

[13] Li, J., M. Xu (2008): "Risk minimizing portfolio optimization and hedging with conditional Value-at-Risk", *Review of Futures Markets*, **16,4**, 471–506.

[14] Markowitz, H. (1952): "Portfolio Selection", *The Journal of Finance*, **7,1**, 77–91.

[15] Morgan Guaranty Trust Company (1994): "RiskMetrics - Technical Document", *Morgan Guaranty Trust Company, Global Research, New York.*

[16] Rockafellar, R. T., S. Uryasev (2000): "Optimization of Conditional Value-at-Risk," *The Journal of Risk*, **2, 4**, 21–51.

[17] Rockafellar, R. T., S. Uryasev (2002): "Conditional value-at-risk for general loss distributions," *Journal of Banking and Finance*, **26**, 1443–1471.

[18] Rudloff, B. (2007): “Convex hedging in incomplete markets”, *Applied Mathematical Finance*, **14**, 437–452.

[19] Ruszczyński, A., G. CH. Pflug (2004): “A risk measure for income proceses”, *Risk Measures for the 21st Century*, John Wiley & Sons, 249–269.

[20] Ruszczyński, A., G. CH. Pflug (2005): “Measuring risk for income streams”, *Computational Optimization and Applications*, **32**, 249–269.

[21] Ruszczyński, A., A. Shapiro (2006): “Conditional risk mapping,” *Mathematics of Operations Research*, **31, 3**, 544–561.

[22] Schied, A. (2004): “On the Neyman-Pearson problem for law-invariant risk measures and robust utility functionals,” *The Annals of Applied Probability*, **14, 3**, 1398–1423.

[23] Shreve, S. E. (2004): *Stochastic calculus for finance II - continuous-time models*, Springer.

[24] Sekine, J. (2004): “Dynamic minimization of worst conditional expectation of shortfall”, *Mathematical Finance*, **14**, 605–618.

[25] Xu, M. (2004): “Minimizing shortfall risk using duality approach - an application to partial hedging in incomplete markets”, *Ph.D. thesis*, Carnegie Mellon University.