Bridging Risk Measures and Classical Risk Processes

Wenjun Jiang

A Thesis
for The Department of
Mathematics and Statistics

Presented in Partial Fulfillment of the Requirements
for the Degree of Master of Science (Mathematics) at
Concordia University
Montréal, Québec, Canada

August 2015

CONCORDIA UNIVERSITY

School of Graduate Studies

This is to certify that the thesis prepared

By: **Wenjun Jiang**

Entitled: **Bridging Risk Measures and Classical Risk Processes**

and submitted in partial fulfillment of the requirements for the degree of

Master of Science (Mathematics)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

____ Examiner
Dr. Y.P. Chaubey

____ Examiner
Dr. M. Mailhot

____ Thesis Supervisor
Dr. J. Garrido

Approved by ____
Chair of Department or Graduate Program Director

____
Dean of Faculty

Date ____

ABSTRACT

Bridging Risk Measures and Classical Risk Processes

by Wenjun Jiang

The Cramér-Lundberg's risk model has been studied for a long time. It describes the basic risk process of an insurance company. Many interesting practical problems have been solved with this model and under its simplifying assumptions. In particular, the uncertainty of the risk process comes from several elements: the intensity parameter, the claim severity and the premium rate. Establishing an efficient method to measure the risk of such process is meaningful to insurance companies.

Although several methods have been proposed, none of these can fully reflect the influence of each element of the risk process. In this thesis, we try to analyze this risk from different perspectives. First, we analyze the survival probability for an infinitesimal period, we derive a risk measure which only relies on the distribution of the claim severity. A second way is to compare the adjustment coefficient graphically. After that, we extend the method proposed by Loisel and Trufin (2014). And last, inspired by the concept of the shareholders' deficit, we construct a new risk measure based on solvency criteria that include all the above risk elements.

In Chapter 5, we make use of the risk measures derived in this thesis to solve the classical problem of optimal capital allocation. We see that the optimal allocation strategy can be set out by use of the Lagrange method. Some recent findings on such problems are also presented.

Acknowledgments

First of all, I would like to express my sincere gratitude to my supervisor Dr. José Garrido for all his help and guidance throughout my Master's degree. Thank you, as well, to all the wonderful professors I had at Concordia University who helped me to further my knowledge and passion in mathematics. I am so grateful to the financial support provided by Concordia University, with which I can initiate my study in Canada and let me pursue my dream.

I would like to take this opportunity to thank my parents for their careful nurturing, their support and encouragement are the motivations that made me move forward. I also want to thank all my colleagues in the Department for their help through these two years both in academy and in life.

Contents

List of Figures

List of Symbols

u the initial capital of the surplus process.

c the premium rate of the surplus process.

θ the safety loading of the surplus process.

λ the intensity parameter of the counting process.

L_M the maximal deficit or the maximum severity of ruin.

τ the moment of ruin.

R the adjustment coefficient or Lundberg coefficient.

D^X the ladder height corresponds to the claim severity X.

L the sum of the ladder heights or the aggregate loss.

S_t the aggregate claims by the time t.

R_t the value of the risk process at time t.

U_t the value of the surplus process at time t.

W_t the Wiener process or standard Brownian motion.

$U_{\tau-}$ the value of the surplus process before ruin.

U_τ the deficit of the surplus process at ruin.

$\psi(u)$ the ruin probability given the initial capital u.

$\tilde{\psi}(s)$ the Laplace transform of the $\psi(u)$.

$\phi(u)$ the survival probability given the initial capital u.

$\varphi(u)$ the Gerber-Shiu function given the initial capital u.

$\rho[X]$ the risk measure on single random variables.

$\rho[R_t]$ the risk measure on risk processes.

$I^S_{T,c}(u)$ the area in the red for the surplus process $u + ct - S_t$.

Chapter 1

Risk Theory and Its Development

1.1 Introduction to Cramér-Lundberg's Model

In actuarial science and applied probability, risk theory uses mathematical models to describe the insurers' vulnerability to insolvency or ruin. The key problem of this theory is how to find the ruin probability, and analyze the distribution of the aggregrate loss and the fluctations of the surplus process.

The basic insurance risk model goes back to the early works of Filip Lundberg. He realized that Poisson processes are at the heart of non-life insurance models and therefore restricted his analysis to the homogeneous Poisson process. Lundberg's discovery was later incorporated in the theory of stochastic processes by Harald Cramér and his work laid the foundation of the classical risk theory model. The most basic results in classical risk theory can be found in Klugman, Panjer and Willmot (2008), Embrechts, Klüppelberg and Mikosch (1997) and Grandell (2001).

Definition 1.1.1 (Cramér-Lundberg Model)
The model established by Filip Lundberg and Harald Cramér makes the following underling assumption:

(a) The claim sizes $\{x_k\}_{k\in N}$ *are positive i.i.d random variables having common distribution* F*, finite mean* $\mu = \mathbb{E}[X_1]$ *and variance* $\sigma^2 = \mathbb{V}[X_1] \leq \infty$.

(b) Claims occur following an homogeneous Poisson process (counting process)

with intensity λ.

(c) The net premium rate includes a fixed positive safety loading: $c = (1+\theta)\lambda\mu$.

(d) The claim sizes and the Poisson process are independent.

Combining the assumptions above, one can write out the surplus model of an insurance company: $U_t = u + ct - S_t$, where u is the initial capital and S_t is the aggregate claims by time t: $S_t = \sum_{i=1}^{N(t)} X_i$. For convenience of later discussions, if considering the surplus process only at the time of each claim, the surplus at nth claim can be written as: $U_n = u - Z_n$, where $Z_n = \sum_{i=1}^{n} Y_i$, where $Y_i = X_i - c\Delta T_i$ and ΔT_i is the waiting time between the $(i-1)$th claim and the ith claim. There are many questions raised by this model, the most interesting one is the estimation of the ruin probability which is defined as follows:

Definition 1.1.2 (Ruin Moment)

(a) The moment of ruin over an infinite horizon is defined by

$$\tau = \inf\{t : t \geq 0, U_t < 0\}. \tag{1.1}$$

(b) The ruin probability with finite horizon:

$$\psi(u,T) = \mathbb{P}(\tau \leq T \mid U_0 = u), \qquad 0 < T < \infty, u > 0. \tag{1.2}$$

Similarly, the ruin probability with infinite horizon is defined as: $\psi(u) = \mathbb{P}(\tau < \infty \mid U_0 = u)$.

The problem introduced above with respect to the classical model, i.e. estimation of the ruin probability, has been studied by many scholars and some efficient techniques are proposed. Two popular methods deserve to be introduced here, since later some core ideas for these two techniques will still be applied to solve similiar problems, but in a more general case. One technique is by defining the ladder heights, and by analyzing the distribution of the sum of these ladder heights.

Definition 1.1.3 (Ladder Height)

The ladder height is defined as magnitude of the loss D, given that there is a claim incurs loss. Its density function is $f_D(x) = \frac{1-F(x)}{\mu}$ where $F(x)$ is the distribution function for claim severities, and $\mu = \mathbb{E}[x_1]$.

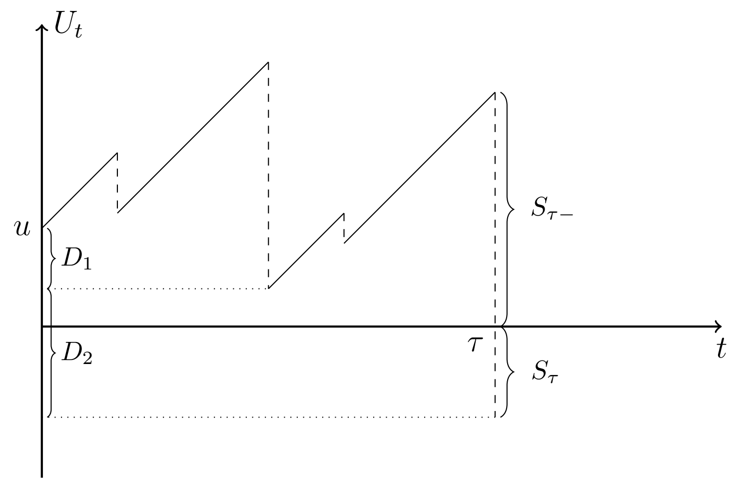

Figure 1.1: The surplus process of the classical model

In the classical model, usually reseachers are interested in the risk process $R_t = S_t - ct$. It has independent and stationary increments, thus one can find that the probability that a drop occurs is independent of the initial level. Further, it can be verified that the occurrence of k drops is equivalent to the occurrence of k failures and the success appears at the $(k+1)$th trial. Hence in the infinite horizon problem, the probability that a total of k drops occur is $\mathbb{P}(N = k) = \frac{\theta}{1+\theta}(\frac{1}{1+\theta})^k$, where $\frac{1}{1+\theta}$ is the probability for occurrence of one drop. Using $L = \sum_{i=1}^{M} D_i$ to represent the maximum loss in the infinite horizon problem, the ruin probability can be rewritten as $\mathbb{P}(L > u)$. It is not difficult to evaluate this probability

$$\mathbb{P}(L > u) = \psi(u) = 1 - \sum_{n=0}^{\infty} \frac{\theta}{1+\theta}(\frac{1}{1+\theta})^n F_D^{*(n)}(u), \qquad u > 0,$$

where $F_D^{*(n)}(x) = \mathbb{P}(\sum_{i=1}^{n} D_i \leq x)$ is the n-fold convolution of F_D, and $F_D^{*(0)}(u) = 1_{\{u>0\}}$, $F_D^{*(1)}(u) = F_D(u)$.

The second popular way is to condition on the first claim, then an integro-differential equation about the survival probability $\phi(u) = 1 - \psi(u)$ can be derived:

$$\phi(u) = \int_0^{\infty} \lambda e^{-\lambda t} \int_0^{u+ct} \phi(u + ct - x) dF_X(x) dt, \qquad u \geq 0. \tag{1.3}$$

After some algebraic manipulations, one can get the following defective renewal e-quation

$$\phi(u) = \frac{\theta}{1+\theta} + \frac{1}{1+\theta} \int_0^u \phi(u - x) dF_D(x), \qquad u \geq 0.$$

The survival probability in the defective renewal equation can be obtained by a Laplace transform. Clearly, the ruin probability $\psi(u) = 1 - \phi(u)$ found from (1.1) should be the same as that derived by the ladder height method.

Other results about the classical model are listed as follows.

Theorem 1.1.1 *Consider the Cramér-Lundberg model with a positive safety loading, if R is the smallest positive root of the equation*

$$1 + (1+\theta)\mu r = M_X(r), \qquad r < \delta_X,$$

where $M_X(r)$ is the moment generating function of the claims, δ_X is the convergence radius of $M_X(r)$, then the following relations hold:

(a) (Lundberg's inequality) For all $u \geq 0$, $\psi(u) \leq e^{-Ru}$.

(b) (Cramér-Lundberg's asymptotic approximation) When $u \to \infty$, then $\psi(u) \sim Ce^{-Ru}$, where

$$C = \frac{\mu\theta}{M_X'(r) - \mu(1+\theta)}.$$

(c) If the claims follow an exponential distribution with mean β, then the ruin probability is

$$\psi(u) = \frac{1}{1+\theta} \exp\{-\frac{\theta}{\beta(1+\theta)} u\}, \qquad u \geq 0.$$

The quantity R in the above theorem is called the adjustment coefficient (AC) or Lundberg's coefficient, and the corresponding equation is called Lundberg's equation. The adjustment coefficient will serve as a key quantity in the discussions later of this thesis.

1.2 Extended Risk Theory Models

1.2.1 Generalized Claim Counting Process

The Cramér-Lundberg model is a very simple model in risk theory, and it imposes several simplifying conditions. However, these conditions make the model deviate substantially from reality in most cases. In recent years, many modifications have been made about the classical model, especially for the underling claim counting process, to make it more realistic.

The renewal risk model was first introduced by Sparre Andersen (1957); he extended the classical model by allowing claim inter-arrival times to follow an arbitrary distribution. His idea is inherited and further analyzed by Thorin (1974,1982), and is fully illustrated in Grandell (1991). Both, Lundberg's inequality and Cramér-Lundberg's asymptotic approximation for the ultimate ruin probability under this case are derived and we will see later that these results are very similar to the classical ones.

There are two categories of renewal process which need introductions here, the ordinary renewal process and the corresponding delayed renewal process. In order to serve the discussion later, only definitions of these two renewal processes are given. For the details, see Karlin and Taylor (1975).

Definition 1.2.1 (Renewal Process)

1. Ordinary renewal process: A sequence of i.i.d. random variables $\{X_1, X_2, \dots, \}$*, where* X_i *is intepreted as the time interval between the* $(i-1)$*th event and the ith event.*

2. Delayed renewal process: A sequence of random variables $\{Y_1, X_2, X_3, \dots\}$*, where* $\{X_i, i = 2, 3, \dots, \}$ *are i.i.d. random variables,* Y_1 *follows a different distribution.*

For the ordinary renewal process, a more technical method is used to derive the upper bound of the ultimate ruin probability in Grandell (1991), that is the martingale method. Suppose that $\{Z_n\}_{n=1.2....}$ follows a risk process under this case, where the subscript indicates the occurrence of the nth claim, then the martingale is constructed as follows.

$$M_u(n) = \frac{e^{-r(u-Z_n)}}{g(r)^n}, \quad g(r) = \mathbb{E}[e^{rZ_1}], \qquad r < \delta_Y, \tag{1.4}$$

where δ_Y is the convergence radius of the moment generating function. It is easy to verify that the expression above defines a martingale with respect to the filtration $\mathscr{F}_n = \sigma\{Z_k; k = 0, 1, 2, \dots, n\}$. Consider the stopping time $N_u = \min\{n \mid Z_n > u\}$ and an arbitrary positive integer $n_0 < \infty$, combined with the optional stopping theorem, one can get

$$e^{-ru} = M_u(0) = \mathbb{E}[M_u(n_0 \wedge N_u)] \geq \mathbb{E}[M_u(N_u) \mid N_u \leq n_0]\mathbb{P}(N_u \leq n_0). \qquad (1.5)$$

Since $u - Z_{N_u} \leq 0$ on $\{N_u \leq \infty\}$, then

$$\mathbb{P}(N_u \leq n_0) \leq \frac{e^{-ru}}{\mathbb{E}[M_u(N_u) \mid N_u \leq n_0]} \leq \frac{e^{-ru}}{\mathbb{E}[g(r)^{-N_u} \mid N_u \leq n_0]} \leq e^{-ru} \max_{0 \leq n \leq n_0} g(r)^n,$$

when $n_0 \to \infty$, the inequality becomes $\psi(u) = \mathbb{P}(n_0 < \infty) \leq e^{-ru} \sup_{n \geq 0} g(r)^n$. As expected, the best choice for r here is the adjustment coefficient R, which satisfies Lundberg's equation for the renewal claim counting case. Hence, Lundberg's inequality is obtained: $\psi(u) \leq e^{-Ru}$. The second important result is the Cramér-Lundberg approximation for the ultimate ruin probability. It follows the same steps as for the classical model, and the result can be obtained by use of the key renewal theorem. For a detailed introduction about the renewal process, interested readers can refer to Karlin and Taylor (1975) or Ross (2010).

For the delayed renewal process, there is a particular case which interests many scholars – the stationary renewal process, in which the waiting time for the first claim follows the equilibrium distribution $G(y)$,

$$G(y) = \frac{1}{\mu} \int_0^y \big[1 - F(x)\big] dx, \qquad y \geq 0, \qquad (1.6)$$

where $F(x)$ is the distribution function for waiting time between successive claims. The following theorem tells us that the equilibrium distribution of the initial waiting time can be used to characterize the stationary renewal process.

Theorem 1.2.1 *If a modified renewal process $N(t)$ has the equilibrium distribution as the first waiting time, and $\sum_n$ is the total time spent for the occurrence of the nth event, then this process has the following properties.*

(a) $\mathbb{E}[N(t)] = \frac{t}{\mu}$ *for all* $t > 0$.

(b) $\mathbb{P}(B_t \leq y) = \frac{1}{\mu} \int_0^y [1 - F(x)]dx$, *where* $B_t = \sum_{N(t)+1} - t$ *is the residual time.*

(c) $N(t)$ *has stationary increments.*

Denoting by $\phi^*(u)$ the survival probability of the Sparre Andersen model with stationary renewal claim counting process, condition on the first claim amount and using the change of variable techniques, the defective renewal equation can be derived:

$$\phi^*(u) = \int_0^\infty G(s) \int_0^{u+cs} \phi(u+cs-x)dF(x)ds,$$
$$\Rightarrow \quad \phi^*(u) = \phi^*(0) + \frac{1}{c\mu}\int_0^u \phi(u-x)[1-F(x)]dx, \qquad u \geq 0.$$

Using Lundberg's inequality for the ordinary renewal case, one can get the upper bound of the ultimate ruin probability for stationary renewal case.

$$\psi^*(u) \leq \frac{1}{c\mu}\int_u^\infty 1-F(x)dx + \frac{1}{c\mu}\int_0^u e^{-R(u-x)}(1-F(x))dx, \tag{1.7}$$
$$\leq \frac{1}{c\mu}\int_0^\infty e^{-R(u-x)}(1-F(x))dx = \frac{1}{c\mu R}h(R)e^{-Ru}, \tag{1.8}$$

where $h(r) = \mathbb{E}[e^{rx}] - 1$, and R is the Lundberg coefficient. This gives Lundberg's inequality for stationary renewal process.

Besides discussing the ruin probability for general renewal claim counting processes, some specific renewal processes are studied, and among these the Erlang process attracts the most attention. Tijms (1994) shows how any positive continuous distribution can be approximated by mixed Erlang distributions to any arbitrary level of precision. Dickson (1997) studies a particular case, the Erlang(2) process, and gets an explicit result for the ruin probability. In fact, Dickson improves the idea of evaluating the ruin probability for the Erlang(n) process, but it is a little bit complicated because his approach relies heavily on the Laplace transform. It is time consuming if one wants to recover the ultimate ruin probability from its Laplace transform since it requires inverting the Laplace transform. A more convenient way is developed by Dickson and Hipp (2001). They study the moments of the ruin time for the Erlang(2) case, based on the analysis of the Gerber-Shiu function. Defining the "Dickson-Hipp" operator, they solve the integro-differential equation for the ruin probability and illustrate the application of this method with mixed exponential claims. Their technique is extended to the Erlang(n) process by Li and Garrido (2004), and more general analysis for the structure of the Gerber-Shiu function is given for Erlang(n) claim arrivals. Related topics such as the Gerber-Shiu function will be introduced later in detail.

A more general case is when claim arrivals follow a Cox process, which is initially discussed in an actuarial context by Reinhard (1984). A deeper analysis is given by Björk and Grandell (1988), their works are collected in Grandell (1991). A slight generalization of the Cox process is now called the Markov-modulated model. In Cox processes, the intensity parameter of the Poisson process in the ruin model is controlled by an external Markovian environment, usually the state of the external Markovian system is denoted by $I(t)$. In Grandell (1991), the waiting time for the next transition in the state process $I(t)$ follows an exponential distribution with intensity parameter $\eta(x)$, given that $I(t) = x$, and the transition probability from state x to a subset of the whole state space B is $p_L(x, B)$. If the state process satisfies these conditions, then this state process is usually called a jump process or piecewise Markov process. Under these assumptions, the ruin probability $\phi_x(u)$ for the Cox claim counting process can be written in the integro-differential form, for $u > 0$,

$$c\phi_x'(u) = x\phi_x(u) + \eta(x)\phi_x(u) - x\int_0^u \phi_x(u-z)dF(z) - \eta(x)\int_S p_L(x, dy)\phi_x(u), \quad (1.9)$$

where the ruin probability $\phi_x(u)$ depends on the initial state of the external system. The case of particular interest to researchers is when the jump process in the Cox model is stationary. The characterization of the stationary jump process is given in the next theorem.

Theorem 1.2.2 *Let Y be a Markovian jump process. If the initial distribution q_L of the states satisfies*

$$\int_B q_L(dy)\eta(y) = \int_S q_L(dy)\eta(y)p_L(y, B), \quad (1.10)$$

then Y is a stationary jump process.

In Reinhard (1984), only the non-stationary discrete case is considered. A matrix notation is used through his paper since the external Markov process is represented by a finite transition matrix. Its element h_{ij} indicates the transition probability from state i to state j in the external environment. By conditioning on the first jump epoch, Reinhard derives the ordinary differential equation system for finite horizontal ruin probability.

$$\frac{\partial}{\partial t}\phi_{ij}(u,t) - c_i\frac{\partial}{\partial u}R_{ij}(u,t) = -(\lambda_i+\eta_i)R_{ij}(u,t) + \lambda_i\int_0^u R_{ij}(u-x,t)dF_i(x) + \eta_i\sum_{k=1}^{\infty} h_{ik}R_{kj}(u,t), \qquad u,t>0,$$

where λ_i is the intensity parameter for the risk process when it is in the state i, and η_i is the intensity parameter for the external Poisson process when its current state is i. This integro-differential equation system can be solved following several steps that will be frequently used in the next subsection. Taking a Laplace transform on both sides of the equation and then isolating the transformed function, taking the inverse Laplace transform is the last but also the most complicated step.

Let $t \to \infty$, the function $\phi_{ij}(u,t)$ converges to be the ruin probability over an infinite horizon. Under simpler assumptions, as when the external environment only has two states, the transition matrix is

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

and claim amount follows an exponential distribution, then the non-ruin probabilities can be obtained.

1.2.2 Surplus Process Perturbed by a Diffusion

With the rapid expansion of the financial products offered on the markets, intensive study has been conducted about the Brownian motion, which is the building block of financial engineering. The close relation between the insurance industry and the financial industry has lead to the evolution of the models used in both areas. Dufresne and Gerber (1991) add a diffusion process to the classical risk model, and they define the surplus process as

$$U_t = u + ct + W_t - \sum_{i=1}^{N(t)} X_i, \qquad t \geq 0,$$

where $\{W_t\}$ is a Wiener process. This diffusion process can be interpreted as a model for additional uncertainty on the compound Poisson claims or on the premium income.

To find out the exact expression for the ruin or survival probability, they use a special method which decomposes the maximum aggregate loss into two parts, this allows them to inherit the ladder height method from the classical model. Denote by L_1 the first drop incurred by the diffusion process and L_2 the subsequent drop incurred by the jump process (or claims). The expression for the survival probability under this model is then similar to that for the classical one in (1.1):

$$\phi(u) = \sum_{n=0}^{\infty} \frac{\theta}{1+\theta}(\frac{1}{1+\theta})^n G_1^{*(n+1)} * G_2^{*(n)}(u), \qquad u > 0,$$

where G_1 is the cumulative distribution function of L_1, and G_2 is the cumulative distribution function of L_2. For consistencies in the notation, here $G_2^{*(0)}(u) = 1_{\{u>0\}}$, $G_2^{*(1)}(u) = G_2(u)$ and $G_1^{*(1)}(u) = G_1(u)$.

Schmidli (1995) further studies the perturbed model. He first defines a new probability measure and verifies that under this new measure the stationary and independent increments of the surplus process are preserved. The result he obtains is the following asymptotic expressions for the ruin probability when the initial capital goes to infinity.

$$\mathbb{P}(\tau < \infty, U_\tau = 0 \mid U_0 = u) \sim C^{(1)} e^{-ru}, \qquad C^{(1)} = \frac{\int_0^\infty (1 - G_1(x))dx}{\int_0^\infty (1 - H(x))dx},$$

$$\mathbb{P}(\tau < \infty, U_\tau < 0 \mid U_0 = u) \sim C^{(2)} e^{-ru}, \qquad C^{(2)} = \frac{\int_0^\infty (1 - G_2(x))e^{-rx}dx}{\int_0^\infty (1 - H(x))dx},$$

where H is the joint distribution of $L_1 + L_2$, the total drop at a single time.

1.2.3 Dependence Between Premium, Claims and Frequency

Besides considering more general claim counting processes, other authors relax the independence assumption between the premium rates, the claim sizes and the underling counting processes. Different dependence structures are imposed and added into the model to make it more realistic. One instance is the time-dependent premium risk model proposed by Asmussen (2000) in which premium rates are adjusted continuously based on the current insurer surplus level. Albrecher and Boxma (2004) studies the generalization of the classical model in another dependence setting, where the distribution of the inter-arrival times depends on the previous claim size. They derive an exact analytical expression for the Laplace transform of the ruin probability function. Boudreault *et al.* (2006) considers a particular dependence structure and derive the defective renewal equation satisfied by the expected discounted penalty function.

Cai and Zhou (2009) use more general assumptions for a perturbed risk model with dependence between premium rates and claim sizes,

$$c(x) = \begin{cases} c_1, & x \le b \\ c_2, & x > b \end{cases}, \qquad \sigma(x) = \begin{cases} \sigma_1, & x \le b \\ \sigma_2, & x > b \end{cases},$$

$$U_t = U_{\tau_n} + c(X_n)(t - \tau_n) + \sigma(X_n)(W_t - W_{\tau_n}), \qquad t > 0,$$
$$U_{\tau_{n+1}} = U_{\tau_n} + c(X_n)T_{n+1} - X_{n+1} + \sigma(X_n)(W_{\tau_{n+1}} - X_{\tau_n}),$$

where τ_n is the moment of occurrence of the nth claim, and W_t is the Wiener process. In this model, Cai and Zhou assume that the premium rates depend on the previous claims and the diffusion coefficient changes according to the changes in the premium rates. Their assumptions are meaningful since in auto-insurance, premium rates are usually adjusted after claims occur. To find out the survival probability for such a model, they derive the following integro-differential equations for $u > 0$,

$$\frac{1}{2}\sigma_1^2\phi_1^{''}(u) + c_1\phi_1^{'}(u) = \lambda\phi_1(u) - \lambda\int_0^u \Big[\phi_1(u-x)f_1(x) + \phi_2(u-x)f_2(x)\Big]dx, \quad (1.11)$$

$$\frac{1}{2}\sigma_2^2\phi_2^{''}(u) + c_2\phi_2^{'}(u) = \lambda\phi_2(u) - \lambda\int_0^u \Big[\phi_1(u-x)f_1(x) + \phi_2(u-x)f_2(x)\Big]dx. \quad (1.12)$$

To find explicit expressions for $\phi_1(u)$ and $\phi_2(u)$, Laplace transform are again used and the authors get two corresponding equations for the Laplace transforme of the ruin probabilities. Later, they use Rouché's theorem, combined with limit conditions, to derive two equations which can be used to find out the boundary conditions $\phi_1^{'}(0)$ and $\phi_2^{'}(0)$. Then the inverse Laplace transform can be used to evaluate the survival probabilities $\phi_1(u)$ and $\phi_2(u)$. The techniques used in their paper are common to the previous papers, which also shows the richness of the methods in risk theory, such as the transform techniques and complex analysis.

The ruin probability is a central quantity in risk theory research. Therefore it is worth reviewing the details about ruin probabilities under different settings, since these can serve as the most natural direct criteria to measure the risk of a risk process. Through the comparison of the ruin probabilities, actuaries in insurance companies can make better assessment of the initial capital or the premiums to cover the random loss incurred by the paid claims.

1.3 More Topics in Risk Theory

1.3.1 Gerber-Shiu Function

With the increased attention put on risk theory, and its wide applications in both the insurance industry and the financial industry, more interesting research tools are found by scholars. Gerber and Shiu (1998) study the joint distribution of the time of ruin τ, the surplus immediately before ruin $U_{\tau-}$ and the deficit at ruin U_τ. Including the force of interest, they define the following expected penalty function given the initial capital u. This function is also called Gerber-Shiu function in the actuarial literature:

$$\varphi(u) = \mathbb{E}[w(U_{\tau-}, |U_\tau|)e^{-\xi T} 1_{\{\tau<\infty\}} \mid U_0 = u], \qquad u \geq 0,$$

where $w(x, y)$ is interpreted as a penalty function. For the classical model, this function has been deeply analyzed by Gerber and Shiu (1998). Their study shows that this function satisfies a specific defective renewal equation: $\varphi = \varphi * g + h$, where $\varphi * g$ represents the convolution of φ and g, where g and h are two related functions. The solution to this defective equation can be written as a Neumann series:

$$\varphi = h + g * h + g * g * h + \dots.$$

This function is widely used in current studies in risk theory, since it can help derive several interesting results for risk processes. In the original paper of Gerber and Shiu (1998), many nice results are already given, such as an explicit ruin probabilitiew when the initial capital is zero, and asymptotic formula for ruin probabilities when the initial capital is very large. For arbitrary initial capital, the authors only consider a particular case, where the claim amount distribution is exponential or mixed exponential.

Dickson and Hipp (2001) consider the moments of the ruin time for Erlang(2) claim arrivals, and they use a Gerber-Shiu function of the following form,

$$\varphi(u) = \mathbb{E}[e^{-\delta\tau} 1_{\{\tau<\infty\}} \mid U_0 = u], \qquad u \geq 0,$$

By noting that

$$(-1)^k \frac{d^k}{d\delta^k}\varphi(u)|_{\delta=0} = \mathbb{E}[\tau^k 1_{\{\tau<\infty\}} \mid U_0 = u],$$

the moments of the time to ruin can be found. This method was extended to Erlang(n) case by Li and Garrido (2004). In fact, by suitable selection of the penalty function $w(x, y)$, the Gerber-Shiu function can be used to evaluate different quantities in risk theory. Since the distribution of time to ruin is very difficult to find, through comparison of the moments of the ruin time, the risks of the risk processes can be compared; this is considered in Kolkovska (2011) to be a risk measure. Later, it will be found that to compare the risks of different risk processes, many quantities in risk theory can be used as risk measures. However, these risk measures have different properties and sometimes cannot show a precise overall comparison of the risks in different risk processes.

1.3.2 Distribution of The Time to Ruin

The defective distribution of the time to ruin is defined as $\psi(u,t) = \mathbb{P}(\tau \leq t)$ where τ is the ruin time, so the distribution of the time to ruin is in fact the finite horizon ruin probability in (1.2). To get a proper distribution, one can divide $\psi(u,t)$ by $\psi(u)$, then the corresponding ruin time is defined as $\tau^* \triangleq \tau \mid \tau < \infty$, and the distribution is denoted by $\psi_c(u,t) = \frac{\psi(u,t)}{\psi(u)}$. The moments of the time to ruin have been found from the Gerber-Shiu function by setting the penalty function $w(U_{t-}, |U_t|) = 1$, which in fact leads to the Laplace transform of the random variable τ with respect to δ:

$$\begin{aligned} \varphi(u) &= \mathbb{E}[e^{-\delta\tau}1_{\{\tau<\infty\}} \mid U_0 = u], && (1.13)\\ &= \int_0^{\infty} (e^{-\delta t}\psi(u,t))dt, && (1.14)\\ &= \psi(u)\int_0^{\infty} (e^{-\delta t}\psi_c(u,t))dt. && (1.15)\end{aligned}$$

It is of interest to recover the expression of $\psi(u,t)$ for (1.2), or $\frac{\partial}{\partial t}\psi(u,t)$.

Since the general Gerber-Shiu function satisfies the defective renewal equation

$$\varphi(u) = \frac{1}{1+\theta}\int_0^u \varphi(u-x)dG(x) + \frac{1}{1+\theta}H(u), \qquad u \geq 0, \tag{1.16}$$

Willmot and Lin (1998) study its solution. Denoting the following compound geometric distribution by $K(u) = 1 - \bar{K}(u)$ by

$$\bar{K}(u) = \sum_{n=1}^{\infty} \frac{\theta}{1+\theta}(\frac{1}{1+\theta})^n \bar{G}^{*(n)}(u), \qquad u \geq 0.$$

Willmot and Lin (1998) show that from renewal theory the solution of (1.11) can be written as

$$\varphi(u) = \frac{1}{\theta}\int_0^u H(u-x)dK(x) + \frac{1}{1+\theta}H(u),$$

where the penalty function $w(U_{\tau-}, |U_\tau|) = 1$, the corresponding $H(u) = \bar{G}(u)$, and $\varphi(u) = \bar{K}(u)$. This shows that the distribution of the time to ruin can be represented as a compound geometric distribution.

There are helpful techniques to find the distribution of the time to ruin. Dickson and Waters (2002) apply numerical methods, especially the translated gamma approximation, to find the distribution of the time to ruin. Garcia (2005) uses a complex inversion formula, or inverse Laplace transform to find this distribution. From his discussion it is easy to find the inverse Laplace transform with respect to δ, that more work is required for the analysis of the residues of the different singularities of $e^{\delta s}\varphi(u)$. Dickson, Hughes and Zhang (2005) continue to use this method to recover the distribution of the time to ruin, but they focus their work on the Sparre Andersen process with Erlang arrivals and exponential claims.

For the classical model, Dickson and Willmot (2005) initiate their study with the compound geometric distribution derived from Willmot and Lin (1998). However, without applying the inverse Laplace transformation directly, they apply a more technical method which is innovative but somewhat time consuming. Their important discovery is that $\varphi(u)$ is the Laplace transform with respect to both δ and ρ, the latter being the unique positive root of the generalized Lundberg equation. Then Lagrange's implicit function theorem comes into play, it helps recover the structure of $\psi_c(0,t)$, and derive a representation of $\psi(u,t)$ with respect to $\xi(u,t)$, the inverse Laplace transform of $\varphi(u)$ with respect to ρ. Their final work is to revert the long representation form, term by term, by use of the comparison method. For details, see Dickson and Willmot (2005).

1.3.3 Deficit at Ruin

The surplus before the ruin and the deficit at the ruin are two other interesting quantities in the development of risk theory. If an insurance company does not immediately stop running when the surplus drops to be negative, then the deficit at the ruin is important to evaluate the duration of the negative surplus and indicates how long it will take for an insurance company to recover. Hence, this quantity can also partially reflect the risk of the surplus process because it can tell how severe a brankruptcy the insurance company can recover from.

The details of how to calculate the exact distribution of the deficit at the ruin $G(u,y)$ is first given in Gerber, Goovaerts and Kass (1987). It is not difficult to find out that the distribution $G(u,y)$ satisfies a defective renewal equation:

$$G(u,y) = \frac{\lambda}{c}\int_0^u G(u-x,y)\Big[1-F(x)\Big]dx+\frac{\lambda}{c}\int_u^{u+y}\Big[1-F(x)\Big]dx, \qquad u,y \geq 0. \quad (1.17)$$

The technique they use to solve for $G(u,y)$ is again the Laplace transform and its inverse. Later, a more efficient technique to find the distribution $G(u,y)$ was given by Lin and Willmot (1999). By full use of the results in Dufresne and Gerber (1988), the distribution $G(u,y)$ can be rewritten as

$$G(u,y) = \psi(u) - \frac{1}{\theta}\int_0^u \bar{F}_D(y+u-x)d\phi(x). \quad (1.18)$$

The conditional distribution of the deficit $G_u(y)$ is discussed in detail by Willmot (2000). Here, he points out the significant role that Erlang distributions plays in risk theory, which is further discussed in Willmot and Woo (2007).

As mentioned in the Section 2.1, Tijms (1994) finds that an arbitrary positive continuous distribution can be approximated to any level of precision by mixtures of Erlang distributions, so it is reasonable to assume that the claims follow a mixed Erlang distribution. Given that condition, the residual lifetime distribution, the compound aggregate claims distribution, the conditional distribution of the deficit at ruin and the infinite horizon ruin probability follow different mixed Erlang distributions, but all with the same scale parameter. Also, Willmot and Woo (2007) show that many well-known distributions can be expressed by mixed Erlang distributions. This is a very important result, since it can simplify the analysis without considering the most general case, and some nice analytical results can be derived under the assumption that the claims follow a mixed Erlang distribution.

1.3.4 Distribution of the Time to Recovery

The assumption of a positive safety loading leads to a positive drift for the surplus process, which indicates that the surplus will go to infinity with probability one. If ruin occurs, and the insurance company can let the process continue, then theoretically the surplus will only stay below the zero level temporarily. For convenience in the subsequent discussion, we say that the company is in the red if its surplus is negative. If the company can successfully refinance and operate when it is in the red, then it can recover. The recovery time is defined as

$$\tau' = \inf\left\{t \;\middle|\; t > \tau, U_t \geq 0\right\}, \qquad \tau > 0,$$

where τ is the ruin time.

Naturally, the quantity $\tau' - \tau$ can be treated as the time needed for the company to recover, and this quantity is certainly closely related to the deficit at the ruin U_τ, which in turn depends on the severity of a single claim, and the intensity of the claim counting process. So this quantity is an important index to evaluate the risk of

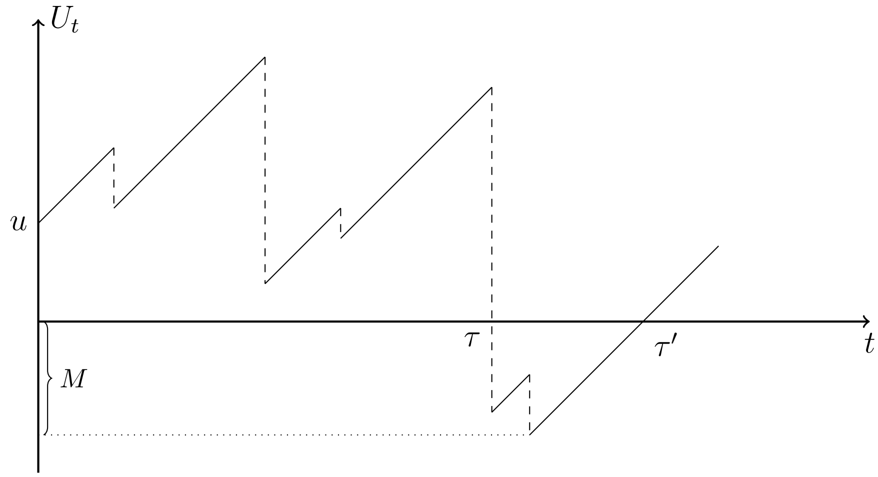

Figure 1.2: Maximal deficit during the first period of deficit

surplus process. The distribution of $\tau' - \tau$ for the classical Cramér-Lundberg model has been exhaustively studied by Egídio dos Reis (1993). Inspired by the method used in Gerber (1990), by regarding $\tau' - \tau$ as the time needed to pass the level $|U_\tau|$ from the initial capital level $U_0 = 0$, a martingale argument is smartly used by Egídio dos Reis to give the moment generating function of the time period $\tau' - \tau$. The expectation and variance of $\tau' - \tau$ can be derived from its moment generating function. Besides, Egídio dos Reis (1993) gives the moment generating function of the total duration of a negative surplus. It is not difficult to see that the moment generating function of $\tau' - \tau$ is a function of the initial capital $U_0 = u$ and the negative root of the generalized Lundberg equation. Thus, if the regulator wants to control the expected time to recovery below some certain level A, then the needed initial capital can be found from the expression $\mathbb{E}[\tau' - \tau \mid U_0 = u] \leq A$.

1.3.5 Maximal Severity of Ruin

The previous section shows that if the insurance company stays in the red for some time, but not very deeply, then there is still hope for the company to recover. However, if the claims come too frequently or a large claim occurs, then it will push the company to a worse situation, and it will no longer be able to borrow money to cover its heavy liabilities. In such case there is no hope for the company to recover.

The maximal severity of ruin is defined to measure the severity of a negative surplus process of the classical compound Poisson model:

$$L_M = \max_{\tau \le t \le \tau'} |U_t|. \tag{1.19}$$

This quantity has been fully studied by Picard (1994), the stationarity of the surplus process is heavily used in his paper. The main results can be summarized by the following theorem.

Theorem 1.3.1 *For L_M defined in (1.19), τ is the first time of ruin and τ' is the first time for recovery, then*

$$\mathbb{P}(L_M \le z \mid U_0 = u, \tau < \infty) = \frac{\psi(u) - \psi(u+z)}{\psi(u)(1-\psi(z))}, \qquad z \ge 0, \tag{1.20}$$

$$\mathbb{P}(L_M \le z \mid \tau < \infty, |U_\tau| = y) = \frac{1-\psi(z-y)}{1-\psi(z)}, \qquad z \ge y > 0. \tag{1.21}$$

where $\psi(u)$ is the ruin probability for the classical model with the initial capital u.

1.3.6 Expected Area in the Red (EAR)

Picard (1994) studies another quantity, which is called the cost of recovery in his paper, or the expected area in the red in Loisel (2005),

$$I = \int_\tau^{\tau'} |U_t| dt.$$

This quantity captures the main characteristics of the surplus process. If the cost exceeds the deficit, then by utility theory there exists a convex function g that can serve as the cost function

$$I_g = \int_\tau^{\tau'} g(|U_t|) dt.$$

The tool Picard (1994) uses is to study I_g as a martingale, but in a complicated way. To construct the martingale, he verifies the following theorem which connects the function g to another function f.

Theorem 1.3.2 *Let f and g be two real functions (g continuous and f in the class C^1) that are connected by the relation*

$$g(x) = -\lambda - cf'(x) + \lambda\mathbb{E}\Big[\exp\big\{f(x) - f(x - X_i)\big\}\Big], \qquad x \in \mathbb{R}. \tag{1.22}$$

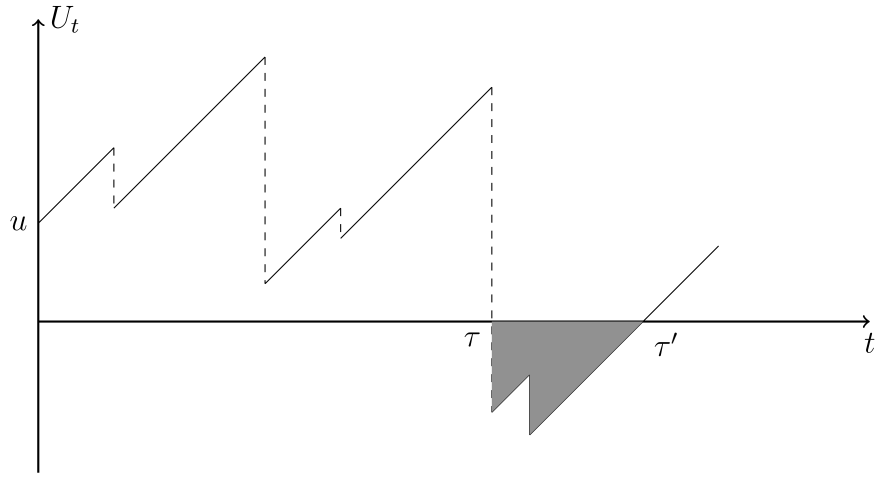

Figure 1.3: Area in the red when the surplus falls below zero

Thereby, for any $a \in \mathbb{R}$, $-f$, $-g$ and $|e^{-f}f'|$ admit an upper bound on $(-\infty, a]$. If selecting

$$W_t = \exp\Big\{-f(U_t) - \int_0^t g\big(U_s\big)ds\Big\}, \qquad t \geq 0, \tag{1.23}$$

then $(W_t, \mathscr{F}_t)_{t\geq 0}$ is a martingale.

To evaluate the expectation of I_g, specific conditions are imposed on the function f, and the connection between the I_g and f is revealed in the next theorem.

Theorem 1.3.3 *In addition to the foregoing hypotheses on f and g, suppose that $f(0) = 0$ and $f \geq 0$, $g \geq 0$ on $\mathbb{R}$, then*

$$\mathbb{E}\Big[\exp\Big\{\int_\tau^{\tau'} g(|U_t|)dt\Big\} \mid \tau < \infty, U_\tau\Big] = \exp\big\{-f(|U_\tau|)\big\}. \tag{1.24}$$

In order to use the equation (1.24), it is intuitive to choose the function g first and use (1.22) to solve for function f. However, this is imposible, since no efficient algorithm is proposed to solve the functional equation. So Picard (1994) selects the function f first, and focuses his discussion on polynomial functions, that is $f(x) = \sum_{i=1}^n b_i x^i$. By selecting $g(x) = 1$, he derives the expectation of the time to recovery given the deficit at ruin.

$$\mathbb{E}[\tau' - \tau \mid \tau < \infty, |U_\tau|] = \frac{|U_\tau|}{c - \lambda\mu}, \qquad c > \lambda\mu.$$

Taking expectation on both sides with respect to $|U_\tau|$, the expectation of the time to recovery coincides with the result given by Egídio dos Reis (1993).

Loisel (2005) studies this topic from another angle, and discusses the differentiation properties of the expected area in the red, and the differentation theorems he obtains are useful. They help him complete the exploration of a new kind of risk measure for surplus processes. In Loisel (2005), he first considers the simplest case:

$$\mathbb{E}[I_T(u)] = \mathbb{E}\Big[\int_0^T 1_{\{U_t<0\}}|U_t|dt\Big]. \tag{1.25}$$

Later he generalizes his results to the more complicated case.

$$\mathbb{E}[I_{g,h}(u)] = \mathbb{E}\Big\{\int_0^T \big[1_{\{U_t\geq 0\}}g(U_t) - 1_{\{U_t<0\}}h(|U_t|)\big]dt\Big\}, \tag{1.26}$$

where $U_0 = u$ is the initial capital. Denoting the time needed for recovery by $\tau(u,T) = \int_0^T 1_{\{U_t<0\}}dt$ and the time spent in state zero by $\tau_0(u,T) = \int_0^T 1_{\{U_t=0\}}dt$, then the results in Loisel (2005) can be summarized in the following theorems.

Theorem 1.3.4 *Assume that $T \in \mathbb{R}^+$. Let $(U_t)_{t\in[0,T)}$ be a stochastic process with almost surely time-integrable sample path. Denote by $\tau(u,T)$ the time needed for recovery of the surplus process, by $\tau_0(u,T)$ the time spent in zero by the surplus process., and let $f(u) = \mathbb{E}[I_T(u)]$. If $\mathbb{E}[\tau_0(u,T)] = 0$, then f is differentiable at u, and $f'(u) = -\mathbb{E}[\tau(u,T)]$.*

For the average time needed to recover, the differentiation theorem is stated as follows.

Theorem 1.3.5 *Let $R_t = ct - S_t$, where S_t is a jump process satisfying hypothesis: S_t has a finite expected number of nonegative jumps in every finite interval, and for each t, the distribution of S_t is absolutely continuous. For example, S_t is a compound Poisson process with a continuous jump size distribution. Consider $T < \infty$ and define h by $h(u) = \mathbb{E}[\tau(u,T)]$. Then h is differentiable on $\mathbb{R}^+$, and for $u > 0$, $h'(u) = -\frac{1}{c}\mathbb{E}[N^0(u,T)]$, where $N^0(u,T) = Card(\{t \in [0,T], u + ct - S_t = 0\})$.*

We can see that the expected area in the red can reflect the risk of a surplus process of an insurance company. It combines the severity of a single claim, the frequency of the claims and the premium rate, which are the main characteristics of a surplus process. This means reflects that it can perhaps serve as a new kind of risk measure. The detailed study of such a possible risk measure will be introduced in Chapter 3.

Chapter 2

Basic Risk Measure Theory

Risk measure theory is a broad topic. Although risk measures have been studied for a long time, the theory is not fully mature. Mathematically, a risk measure is a mapping from a class of random variables to the real line. Its properties, and how to apply them in different contexts and make them consistent with the observations, have recently received considerable attention in the financial and actuarial literature. Usually, the requirements of the decision-makers determine the selection of the risk measure. In this section, two main classes of risk measures will be introduced, the relationship between them and more pure mathematical results will be studied.

2.1 Coherent Risk Measure

What is a good risk measure? That's a frequently asked question. A lot of articles prove that no risk measures can meet all the requirements of the market and the decision-makers. Several risk measures are proposed in different contexts. Artzner *et al.* (1999) first introduce the concept of "coherence" when they study market risks and nonmarket risks. As usual, we denote by $\rho[X]$ the risk of X, where X is the random loss incurred by the claims or market changes.

Definition 2.1.1 (Coherent Risk Measure)

A risk measure $\rho[X]$ is coherent when it is equipped with the following four basic properties:

1. *Positive homogeneity:* $\rho[aX] = a\rho[X]$ *if* $a > 0$.

2. *Translation invariance:* $\rho[X + a] = \rho[X] + a$.

3. *Monotonicity: if* $\mathbb{P}(X \leq Y) = 1$, *then* $\rho[X] \leq \rho[Y]$.

4. *Subadditivity:* $\rho[X + Y] \leq \rho[X] + \rho[Y]$.

There are practical intepretations for these four basic properties. Positive homogeneity is often associated with the independence of the monetary units used. Translation invariance can be intepreted as follows: if there is a fixed loss a added on a risky position, then the extra capital needed to make this position acceptable should be a. Subadditivity indicates that a merger of risks does not create extra risk. This is an essential property for a risk measure since later we can see it is closely related to the portfolio optimization problem. These practical intepetations make coherent risk measures consistent with risk management on the market. Besides these basic properties, Denuit *et al.* (2005) supplements more desirable properties which are needed to make a risk measure as "good" as possible.

Property 2.1.1 (Supplements)

1. *Non-excessive loading:* $\rho[X] \leq \max[X]$.

2. *Non-negative loading:* $\rho[X] > \mathbb{E}[X]$.

3. *Continuity with respect to convergence in distribution: if* $\lim_{n\to\infty} X_n =_d X$, *then* $\lim_{n\to\infty} \rho[X_n] = \rho[X]$.

The non-excessive loading indicates that the needed extra capital should not exceed the maximal loss of the risky position; the non-negative loading means the needed extra capital should exceed the average loss, otherwise by the law of large numbers ruin will occur. For the third supplementary property, we can find that for two identically distributed random variables, their risks should be the same from the mathematical viewpoint. This property is later called "law-invariant" property, and is widely applied as the basic assumption for the risk measure. The law invariant coherent risk measure is systematically studied by Kusuoka (2001).

In fact, from Rockafeller (1970), we know that the combination of the first and fourth properties in the definition of coherent risk measures is equivalent to convexity.

Then the natural extension of coherent risk measures are the convex risk measures, which are systematically studied by Föllmer and Schied (2002).

Many coherent risk measures are being used in our daily life, both for theoretical research and practical management. Acerbi (2002) points out that convex combinations of these are still coherent, so infinitely many coherent risk measures can be generated from known ones. Next, we are going to study several common coherent risk measures.

2.1.1 Tail Value at Risk

To talk about the tail value-at-risk (TVaR), we need to study the value-at-risk (VaR) first. In fact, VaR is the most direct risk measure which is the benchmark of today's financial world, since it tells people how much loss they can afford within a certain period. The definition of VaR is very simple, which explains its wide recognition.

Definition 2.1.2 (Value at Risk)
Given a random variable X and a probability level $\alpha \in (0,1)$, then $VaR_\alpha[X]$ is defined to be the α-percentile of the distribution of X:

$$VaR_\alpha[X] = F_X^{-1}(\alpha) = \inf_{t \in R}\{t : F(t) \geq \alpha\}. \quad (2.1)$$

Even though VaR is easy to understand and widely used in real life, it is not a coherent risk measure. Counter-examples are given in most introductory level textbooks. Daníelsson *et al.* (2005) gives out some conditions to make the VaR risk measure subadditive in the tail region, but for the discussion in this section, we skip it and only put emphasis on the TVaR.

It is not hard to find that VaR can only reflect partial information of the random variable X, it cannot tell the decision-makers how thick the tail is. TVaR offsets this weakness of VaR by integrating losses in the tail.

Definition 2.1.3 (Tail Value at Risk)
Given a random variable X and a probability level $\alpha \in (0,1)$, the $TVaR_\alpha[X]$ is defined

as:

$$TVaR_\alpha[X] = \frac{1}{1-\alpha}\int_\alpha^1 VaR_\xi[X]d\xi, \tag{2.2}$$

provided the integral is finite.

There exists a risk measure which is very close to TVaR, called conditional tail expectation (CTE). It represents the conditional expected loss if finite, when the loss exceeds the given VaR value:

$$CTE_\alpha[X] = \mathbb{E}\Big[X \mid X > VaR_\alpha[X]\Big],$$

Acerbi and Tasche (2001) has shown that for discrete random variables, the CTE measure is not coherent, counter examples are also given in their paper. However, for continuous random variables, the CTE measure coincides with TVaR measure. More precisely, in Denuit *et al.* (2005), the $TVaR_\alpha$ and CTE_α are written as:

$$TVaR_\alpha[X] = VaR_\alpha + \frac{1}{1-\alpha}\mathbb{E}\Big[X - VaR_\alpha[X]\Big]_+,$$

$$CTE_\alpha[X] = VaR_\alpha[X] + \frac{1}{1-F_X(VaR_\alpha[X])}\mathbb{E}\Big[X - VaR_\alpha[X]\Big]_+,$$

then the relationship between TVaR and CTE can easily be found out.

There are other representations of TVaR, the most important one is given by Rockafeller and Uryasev (2000), and it is further studied and summarized by Pichler (2013):

$$\begin{aligned} TVaR_\alpha[X] &= \inf_{q\in R}\{q + \frac{1}{1-\alpha}\mathbb{E}[X-q]_+\}, \\ &= \sup\{\mathbb{E}[XZ] \mid 0 \le Z \le \frac{1}{1-\alpha}, \mathbb{E}[Z]=1\}, \\ &= \sup\{\mathbb{E}_{\tilde{P}}[X] \mid \tilde{P} \in \mathscr{P}\},\ \mathscr{P} = \{\tilde{P} \mid \frac{d\tilde{P}}{dP} \le \frac{1}{1-\alpha}\}. \end{aligned}$$

Now with the help of these representation forms, it is not difficult to verify the basic properties for TVaR; see Denuit *et al.* (2005) for the details of the proof.

Theorem 2.1.1 *The TVaR risk measure is coherent and satisfies all the supplementary properties.*

Besides being coherent, TVaR is in fact a particular case of distortion risk measures, which will be introduced in Section 2.2. However, for random variables with heavy tailed distributions, TVaR cannot be used to compare risks since the integration over the tail part diverges. This problem will be reconsidered in the Section 2.1.3.

2.1.2 Entropic Value at Risk

Motivated by Chernoff (1952), Ahmadi-Javid (2012) proposes to use the least upper bound of the VaR to serve as a new risk measure. The Chernoff inequality is given as follows: suppose the moment generating function of X exists, then for any constant a,

$$\mathbb{P}(X \geq a) \leq e^{-za} M_X(z), \qquad z > 0. \tag{2.3}$$

By solving the equation $e^{-za} M_X(z) = \alpha$ with respect to $\alpha \in (0, 1)$, one can obtain

$$a_X(\alpha, z) = z^{-1} \ln(\frac{M_X(z)}{\alpha}). \tag{2.4}$$

If such an a can be properly defined, then obviously $\mathbb{P}(X \geq a_X(\alpha, z)) \leq \alpha$. From this inequality, we can easily find that, for any $z > 0$, $a_X(\alpha, z)$ is an upper bound for $VaR_\alpha[X]$. Naturally, the new risk measure can be defined as the least upper bound for $VaR_\alpha[X]$, which is exactly the definition of the entropic value-at-risk.

Definition 2.1.4 (Entropic Value at Risk)
Given that the moment generating function of X exists, the entropic value-at-risk (EVaR) of X is defined to be

$$EVaR_\alpha[X] = \inf_{z>0}\{a_X(\alpha, z)\} = \inf_{z>0}\{z^{-1} \ln(\frac{M_X(z)}{\alpha})\}. \tag{2.5}$$

Ahmadi-Javid (2012) proves that this risk measure is also a coherent risk measure. In fact, the first three properties are easy to verify for EVaR, but to prove its subadditivity, we need the following lemma proved by Ahmadi-Javid (2012).

Lemma 2.1.1 *Given that the moment generating functions of X and Y exist, for $t_1, t_2 > 0$ and any $\alpha, \lambda \in (0, 1]$, the function $H_\alpha(X, t) = a_X(\alpha, t^{-1})$ is convex in the* *sense that*

$$H_\alpha\Big(\lambda X + (1-\lambda)Y, \lambda t_1 + (1-\lambda)t_2\Big) \le \lambda H_\alpha(X, t_1) + (1-\lambda)H_\alpha(Y, t_2). \qquad (2.6)$$

This lemma tells us that EVaR is a convex risk measure because it can be written as $EVaR_\alpha[X] = \inf_{t>0}\{H_\alpha(X,t)\}$. It is not difficult to verify that EVaR is positive homogeneous, if $b > 0$, then

$$\begin{aligned} EVaR_\alpha[bX] &= \inf_{z>0}\{z^{-1}\ln(\frac{M_{bX}(z)}{\alpha})\} \\ &= \inf_{z>0}\{z^{-1}\ln(\frac{M_X(bz)}{\alpha})\} \\ &= b\inf_{z>0}\{(bz)^{-1}\ln(\frac{M_X(bz)}{\alpha})\} \\ &= b\ EVaR_\alpha[X]. \end{aligned}$$

Once the positive homogeneity is verified, using the following theorem in Rockafeller (1970), it is easy to show that the EVaR is a coherent risk measure.

Theorem 2.1.2 *A positive homogeneous function f on $\mathbb{R}^n$ is convex if and only if*

$$f(X+Y) \le f(X) + f(Y)$$

for every $X, Y \in \mathbb{R}^n$.

Different decision-makers have different attitudes towards risks, those who are risk averse prefer using more conservative risk measurement tools. This section ends with a comparison of the risk measures introduced above.

In comparing the VaR to TVaR, it is easy to see that TVaR is more conservative than a VaR at the same confidence level α:

$$TVaR_\alpha[X] = \frac{1}{1-\alpha}\int_\alpha^1 VaR_\xi[X]d\xi \ge \frac{1}{1-\alpha}\int_\alpha^1 VaR_\alpha[X]d\xi = VaR_\alpha[X].$$

The following theorem from Ahmadi-Javid (2012) compares EVaR to a TVaR with the same α.

Theorem 2.1.3 *The EVaR is an upper bound for both VaR and TVaR with the same confidence level α, i.e., if the moment generating funtion of X exists, then for every $\alpha \in (0, 1]$,*

$$TVaR_\alpha[X] \le EVaR_\alpha[X].$$

Thus, EVaR is more conservative than VaR and TVaR, which suggests that a financial or insurance company using EVaR will allocate more initial capital to hedge the potential loss.

Related topics can be found in Föllmer and Knispel (2011), where they discuss a coherent version of the entropic risk measure, which is a generalized version of EVaR. Their method is quite theoretical and requires a deeper understanding of convex and real analysis. Applications of EVaR can be found in Firouzi and Luong (2014), where they use EVaR to solve a portfolio optimization problem, especially when the underling distribution of asset returns is non-elliptical.

2.1.3 General Representation

Some particular risk measures are reviewed above. However, it is impossible to enumerate all risk measures and discuss their properties and specific applications. A more effective approach is to explore a general representation of coherent risk measures. Studying the properties of the general form is more convenient and gives deeper insight about the structure of the coherent risk measures.

In the seminal paper on coherent risk measures, Artzner *et al.* (1999) give out the general form of these risk measures based on a discrete probability space.

Theorem 2.1.4 *A risk measure ρ is coherent if and only if there exits a family $\mathscr{P}$ of probability measures on the set of states, such that*

$$\rho[X] = \sup\left\{\mathbb{E}_{\mathbb{P}}[X] \;\middle|\; \mathbb{P} \in \mathscr{P}\right\}, \tag{2.7}$$

provided the expectation $\mathbb{E}_{\mathbb{P}}[X]$ exists.

It is also very easy to prove that if a risk measure is defined as in (2.7), then it satisfies the four basic properties for coherence. The two special coherent risk measures introduced in previous sections, TVaR and EVaR, both can be expressed in this form, the only difference is the underling space $\mathscr{P}$ of probability measures.

Delbaen (2002) later extended the result to a continuous probability space. His paper discusses the characterization of coherent risk measures through Fatou's property. Föllmer and Schied (2001) discuss convex risk measures and give their general representation. Kusuoka (2001) studies law invariant coherent risk measures and gives their general representation. Lastly, Frittelli and Gianin (2005) discuss law invariant convex risk measures and provide a representation.

Artzner *et al.* (1999) call the elements in $\mathscr{P}$ scenarios, so from (2.7) we can interpret the coherent risk measure as an expectation under the worst scenario.

It is easy to see that discussions on general representations are based on the assumption that their expectations exist under different scenarios. If the random variable X has a heavy-tail distribution, like Cauchy or Pareto distributions, then coherent risk measures are not applicable. Some methods which do not rely on integrals can partially solve this problem, at the cost of some desirable properties. Balbás, Blanco and Garrido (2014) solve this problem by extending coherent risk measures, such as TVaR, continuously to a larger space that contains some risks with infinite expectations. This is a very efficient way of extending risk measures, since in practical situations most risks will have finite expectations.

2.2 Distortion Risk Measure

There are two approaches to define a risk measure, and they are systemically discussed by Denuit *et al.* (2006). The first one is the axiomatic approach: a set of reasonable axioms for risk management are listed, then one tries to find risk measures that satisfy these axioms. However, risk measures defined this way can be inappropriate if the underling axioms do not agree with the situation at hand.

The second approach is to define a risk measure from the economic perspective. This is because economically, a risk measure should capture the preferences of decision-makers. Utility theory plays a significant role in the development of economics. The milestone of this theory is the expected utility theory proposed by von Neumann and Morgenstern (1947). Many risk measures can be defined from expected utility theory by the no difference principle.

Let us explore the effect of using utility theory on the final results. A person's preference should not change much on a fixed period, so the utility function u of this person can be regarded as deterministic. For a random variable X with some economic interpretation, like asset returns, its expectation under a utility function is

$$\mathbb{E}[u(x)] = \int_{-\infty}^{\infty} u(x)dF(x) = \int_{0}^{1} u(VaR_{\xi}[X])d\xi.$$

By comparison with the expectation without utility $\mathbb{E}[X] = \int_0^1 VaR_{\xi}[X]d\xi$, we can see that the utility function adjusts the values of the $VaR_{\xi}[X]$ in this person's preference. With the development of utility theory, many people started to criticize expected utility theory, since many paradoxes have been found which violate its principle. These paradoxes stimulated the development of other utility theories, many alternative approaches have been proposed. Kahneman and Tversky (1979) establish the prospect theory, which explores the certainty effect, the reflection effect and the isolation effect in people's choices. Quiggin (1982) proposes the anticipated theory, in which he analyzes the effect on the decision brought by the distortion of people's subjective probabilities towards the random loss. His theory later developed to be the dual theory of choice by Yaari (1987). Chateauneuf, Cohen and Meilijson (1997) combine the ideas from previous theories, and their rank-dependent expected utility theory can cover the explanation of the "irrational" phenomena in people's choices.

The distortion risk measure inherits the ideas of Quiggin (1982) and Yaari (1987). For convenience, we only talk about non-negative random variables in the first part of this section. The expectation of such random variables X can be written as

$$\mathbb{E}[X] = \int_{0}^{\infty} \Big[1 - F_X(y)\Big] dy = \int_{0}^{\infty} \bar{F}_X(y)dy.$$

A distortion risk measure distorts the tail probabilities of this random variable, to define a distorted expectation:

$$\rho_g[X] = \int_{0}^{\infty} g\Big(1 - F_X(y)\Big) dy = \int_{0}^{\infty} g\Big(\bar{F}_X(y)\Big) dy. \tag{2.8}$$

Here, the function g is called the distortion operator. Furthermore, the distorted expectation can be rewritten as:

$$\begin{aligned}
\rho_g[X] &= \int_0^{\infty} g(\bar{F}_X(y))dy \\
&= \int_0^{\infty}\int_0^{\bar{F}_X(y)} dg(p)dy \\
&= \int_0^1 VaR_p[X]dg(1-p), \quad g(x) \geq 0.
\end{aligned}$$

From this expression, we can see that the idea behind distorted risk measures agrees with the theories proposed by Quiggin (1982) and Yaari (1987).

For an exhaustive discussion on distortion risk measures see Denuit *et al.* (2005). The properties of this risk measure can be summarized as follows.

Theorem 2.2.1 *The distortion risk measure defined by (2.8) is positively homogeneous, translation invariant, monotone and comonotonic additive. Further, this risk measure is subadditive if, and only if, the distortion operator g is concave.*

2.2.1 Proportional Hazards Transform

The proportional hazards (PH) transform was first proposed by Wang (1995) in order to construct a new risk-adjusted premium principle. Since all the premium principles can be regarded as risk measures in the insurance industry, for convenience we do not distinguish between premium principles and risk measures in the following discussion.

The hazard rate function for a non-negative random variable X can be written as

$$\mu_X(t) = \frac{f_X(t)}{1-F_X(t)} = -\frac{d}{dt}\ln \bar{F}_X(t), \qquad t \geq 0. \tag{2.9}$$

Just as its name implies, the PH transform is multiplying the hazard rate function $\mu_X(t)$ by a constant $\frac{1}{\alpha}$, where $\alpha > 0$. Through this way a new random variable Y, which is characterized by $\mu_Y(t)$, can be generated:

$$\mu_Y(t) = \frac{1}{\alpha}\mu_X(t), \qquad t \geq 0. \tag{2.10}$$

One can find the survival function $\bar{F}_Y(t)$ for the random variable Y by using the equation (2.9):

$$\mu_Y(t) = \frac{1}{\alpha}\mu_X(t) \Rightarrow \bar{F}_Y(t) = \bar{F}_X(t)^{\frac{1}{\alpha}}, \qquad \alpha > 0. \tag{2.11}$$

If the random variable X stands for the potential loss from a risky position, then Y shares the same interpretation but with different tail probabilities. Since $0 \leq \bar{F}_X(t) \leq 1$, then if $\alpha > 1$, $\bar{F}_Y(t) \geq \bar{F}_X(t)$, which means that the decision-maker puts more weight on the tail part. If $0 < \alpha \leq 1$, then $\bar{F}_Y(t) \leq \bar{F}_X(t)$, which means the decision-maker puts less weight on the tail part. The value of α determines the attitude of the decision-maker to uncertainty.

Definition 2.2.1 (PH Transform)
For a risk X with survival function $\bar{F}_X(t)$, Y is defined through the PH transform with respect to X, and the risk-adjusted premium is defined as

$$\rho_\alpha[X] = \mathbb{E}[Y] = \int_0^\infty \bar{F}_X(t)^{\frac{1}{\alpha}} dt, \qquad \alpha \geq 0, \tag{2.12}$$

where α is called the risk-averse index. When $\alpha = 1$, $\rho_\alpha[X] = \mathbb{E}[X]$ which is the net expected loss.

The properties of this risk measure are easily established due to its simple form. This risk measure has the non-excessive loading property. When $\alpha > 1$, the PH transform leads to a non-negative loading. Furthermore, it is positively homogeneous and translation invariant. The proof of these properties can be found in Wang (1995). Subadditivity can be obtained if extra conditions are added.

Theorem 2.2.2 *For any two non-negative random variables X and Y, without assuming independence, the following inequality holds*

$$\rho_\alpha[X + Y] \leq \rho_\alpha[X] + \rho_\alpha[Y], \qquad \alpha \geq 1. \tag{2.13}$$

Proof. See Wang (1995). □

A very important property which makes this risk measure different from coherent risk measures is its layer additivity. Later we will see that this property can be generalized to the additivity of the distortion risk measures for comonotonic risks.

Theorem 2.2.3 *When a risk X is divided into layers $\{(x_i, x_{i+1}], i = 0, 1, \ldots\}$: $X = 1_{(0,x_1]} + 1_{(x_1,x_2]} + \ldots$, where $0 = x_0 < x_1 < x_2 < \ldots$, its risk-adjusted premium is the* *summation of the risk-adjusted premiums of all the layers:*

$$\rho_\alpha[X] = \sum_{i=0}^{\infty} \rho_\alpha[1_{(x_i, x_{i+1}]}]. \tag{2.14}$$

Given last property of this risk measure deserves mention: that it preserves s-tochastic order, which is denoted by $\leq_{st}$. If $X \leq_{st} Y$, this means $\bar{F}_X(t) \leq \bar{F}_Y(t)$, then from the PH transform, it is easy to get the rank of the risk-adjusted premiums for these two risks: $\rho_\alpha[X] \leq \rho_\alpha[Y]$. We will see that this property is satisfied by all distortion risk measures, if the distortion operator g is monotonely increasing.

2.2.2 Wang Transform

From the above discussion of the PH transform, we can see in fact that it is characterized by the distortion operator $g_\alpha(x) = x^{\frac{1}{\alpha}}$. From the introduction we also see that the distortion operator g determines the properties of this risk measure. In this section, a new distortion operator studied by Wang (2000) will be introduced. The corresponding transform $g(\bar{F}_X(t))$ is called Wang transform, and the risk measure characterized by this transform has many good properties.

If the standard normal distribution function is denoted by $u = \Phi(x)$, then its inverse function is $x = \Phi^{-1}(u)$, and the corresponding density function is $f(x) = \frac{d\Phi(x)}{dx} = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$. The Wang transform is characterized by the following distortion operator:

$$g_\alpha^*(u) = \Phi(\Phi^{-1}(u) + \alpha), \qquad \alpha \in (-\infty, \infty). \tag{2.15}$$

Discussing the properties of the risk measure derived from the Wang transform is equivalent to discussing the properties of the distortion operator.

In order to define a proper distribution after the transform, the necessary condition for the distortion operator is that is should be monotone increasing. This is easily verified for $g_\alpha^*(u)$, since the first-order derivative of this function is:

$$g_\alpha^{*\prime}(u) = \frac{f(x+\alpha)}{f(x)} = e^{-\alpha x - \alpha^2} > 0.$$

For the boundary conditions, it is also easily verified that:

$$g_\alpha^*(0) = \lim_{u \to 0^+} g_\alpha^*(u) = 0, \qquad g_\alpha^*(1) = \lim_{u \to 1^-} g_\alpha^*(u) = 1.$$

In order to let the risk measure defined from this distortion operator be subadditive, the sufficient and necessary condition is that the operator g be concave. The proof of this relationship can be found in Wang (2000) and Denuit *et al.* (2005). The Wang transform is subadditive since the second-order derivative of $g_\alpha(u)$ is:

$$g_\alpha^{*\,''}(u) = \frac{-\alpha f(x+\alpha)}{f(x)^2}.$$

So if $\alpha > 0$, $g_\alpha^{*\,''}(u) < 0$, then the distortion operator g_α^* meets all the requirements for the desirable properties. Thus the properties of the risk measure derived from the Wang transform can be summarized as follows.

Property 2.2.1 *If a distortion risk measure is characterized by the Wang transform, then it has the following properties:*

1. *Non-excessive loading:* $\rho_\alpha[X] = \int_0^\infty g_\alpha(S_X(t))dt \leq \max[X]$.
2. *Translation invariance:* $\rho_\alpha[X + a] = \rho_\alpha[X] + a$.
3. *Non-negative loading: If* $\alpha > 0$, $\rho_\alpha[X] > \mathbb{E}[X]$.
4. *Positive homogeneity:* $\rho_\alpha[bX] = b\rho_\alpha[X]$, *if* $b > 0$.
5. *Comonotonic additivity: if* X *and* Y *are comonotonic risks, then* $\rho_\alpha[X + Y] = \rho_\alpha[X] + \rho_\alpha[Y]$.
6. *Subadditivity:* $\rho_\alpha[X + Y] \leq \rho_\alpha[X] + \rho_\alpha[Y]$, *if* $\alpha > 0$.

Comonotonic additivity is explained in more detail in the next section.

2.2.3 General Representation

In fact, the general representation of distortion risk measures is given at the beginning of this chapter. However, that kind of representation is on the space of non-negative random variables. A more general representation is given by Wang, Young and Panjer (1997) based on the following four axioms:

Axiom 1: For a given market condition, the price of an insurance risk X depends only on its distribution.

Axiom 2: For two risks X and Y, if $X \leq Y$ a.s, then $\rho[X] \leq \rho[Y]$.

Axiom 3: If X and Y are comonotonic, then $\rho[X + Y] = \rho[X] + \rho[Y]$.

Axiom 4: For risk X and $d \geq 0$, the risk measure ρ satisfies

$$\lim_{b \to 0^+} \rho[(X-d)_+] = \rho[X], \qquad \lim_{d \to \infty} \rho[\min(X,d)] = \rho[X]. \tag{2.16}$$

If a risk measure satisfies the four axioms above, then it can be represented by the use of the Choquet integral, which is described in detail in Denneberg (1994). We skip the proof provided by Wang, Young and Panjer (1997) and give a representation for distortion risk measures for all real-valued risks:

$$\rho_g[X] = \int_0^\infty g[S_X(t)]dt + \int_{-\infty}^0 \Big(g[S_X(t)] - 1\Big)dt, \tag{2.17}$$

where the distortion operator g is increasing with $g(0) = 0$ and $g(1) = 1$.

Here we see that comonotonic additivity is a special property for the distortion risk measure. The comonotonicity is a very important concept which has been studied for a long time in actuarial science. Next, we give its definition and some related results.

Definition 2.2.2 (Comonotonic Variables)
A subset S in $\mathbb{R}^n$ is said to be comonotonic if for $(x_1, x_2, \ldots, x_n) \in S$ and $(y_1, y_2, \ldots, y_n) \in S$, $x_i \leq y_i$ for some $i \in \{1, 2, \ldots, n\}$ implies that $x_j \leq y_j$ for all $j \in \{1, 2, \ldots, n\}$.

Definition 2.2.3 (Comonotonic Random Vector)
A $\mathbb{R}^n$-valued random vector is comonotonic if it can be expressed as

$$(X_1, X_2, \ldots, X_n) =_d (F_{X_1}^{-1}(U), F_{X_2}^{-1}(U), \ldots, F_{X_n}^{-1}(U)), \tag{2.18}$$

where $=_d$ stands for the equality in distribution, $F_{X_i}(x)$ is the distribution function for X_i and U is a uniformly distributed random variable on $(0, 1)$.

Besides the definition given above, there is an equivalent condition to check whether a random vector is comonotonic or not.

Definition 2.2.4 (Test for Comonotonicity)
A $\mathbb{R}^n$-valued random vector is comonotonic if

$$\mathbb{P}(X_1 \leq x_1, X_2 \leq x_2, \ldots, X_n \leq x_n) = \min_{i \in \{1, \ldots, n\}} \mathbb{P}(X_i \leq x_i). \tag{2.19}$$

The details of the discussion of the comontonicity can be found in Denneberg (1994), Wang and Dhaene (1998) and Dhaene *et al.* (2006). A very important application of comonotonicity is that it can be used to derive upper bounds for the sum of random variables.

Theorem 2.2.4 *Let U be a uniform $(0,1)$ random variable. For any random vector $(X_1, X_2, \ldots, X_n)$ with marginal distributions $F_{X_1}, F_{X_2}, \ldots, F_{X_n}$, we have*

$$X_1 + X_2 + \cdots + X_n \leq_{cx} F_{X_1}^{-1}(U) + F_{X_2}^{-1}(U) + \cdots + F_{X_n}^{-1}(U). \tag{2.20}$$

Proof. See Kass, Dhaene and Goovaerts (2000). □

The inequality above is based on the convex order, which will be introduced in the next chapter.

The VaR risk measure introduced at the beginning of this chapter is not coherent, but it is a distortion risk measure. Its comonotonic additivity is easily verified; since a representation of the distortion risk measure based on VaR is given above, then it follows that VaR is a distortion risk measure, and hence is comonotonically additive.

A particular case is TVaR, it is not only a coherent risk measure, but also a distortion risk measure. The distortion operator of TVaR can be found to be

$$g_p(x) = \min\{\frac{1}{1-p}, 1\}, \qquad p \in (0,1), \quad x \in [0,1].$$

As discussed at the beginning of this chapter, there is no risk measure perfectly consistent with all the observations, so distortion risk measures are "flawed" in some sense. Balbás, Garrido and Mayoral (2009) use several examples to illustrate the inconsistencies one can reach using VaR and TVaR. Their focus is on distortion risk measures and they propose two new properties, completeness and adaptability, to help make the distortion risk measure as rational as possible and eliminate inconsistencies.

We end this chapter by establishing a connection between the two classes of risk measures discussed above. This connection was first explored by Pichler (2013). We know that coherent risk measures can be represented as the optimal solution of a scenario analysis. It is same for distortion risk measures. Pichler (2013) gives supremum and infimum representations for distortion risk measures, and most importantly, he establishes the connection between distortion risk measures and law-invariant coherent risk measures by use of a scenario analysis.

Theorem 2.2.5 *If $\rho[X]$ is a law-invariant coherent risk measure of risk X, and $\rho_g[X]$ is a distortion risk measure of risk X, then*

$$\rho[X] = \sup_{g \in S} \rho_g[X], \tag{2.21}$$

where S is the set of distortion operators, and is usually restricted on the space of continuous and strictly increasing distortion operators.

Proof. See Pichler (2013). □

The theorems and properties mentioned in this chapter are all based on theoretical studies. For practical applications, Heyde *et al.* (2010) study an empirical risk measure, which explains how to use weighted combinations of real data to measure risks. However, even though empirical representations of different risk measures are given in their paper, the coefficients weighting the data are based on a high dimensional supporting hyperplane. Additionally, the robustness of a risk measure is proposed as another important property for external decision-makers, due to the difficulty to frequently make substantial changes to reserve levels. Another contribution of Heyde *et al.* (2010) is that they verify that the empirical representation of coherent risk measure is not robust, therefore a new risk measure is propsoed– conditional tail median, which is easily verified to be robust.

Chapter 3

Stochastic Order

A stochastic order is defined to allow for the ordering of random variables in some probabilistic sense. Traditionally, to rank two random variables, their associated means were singled out for comparison. However, means of random variables are not so informative as they can only reflect partial information about the distributions. Therefore, other ways to rank the random variables that make a more comprehensive use of the information were sought. This information cannot be obtained only from tail values, the variance, the failure rate or other simple indexes. In this chapter, we will review some basic stochastic orders which will serve to connect the partial ordering between different processes and their corresponding risk measures.

3.1 First Order Stochastic Dominance

The concepts and the derivations on stochastic orders reviewed here can be found in many introductory level textbooks, such as Denuit *et al.* (2005) and Shaked and Shanthikumar (2007).

The first stochastic order we define is the most commonly used, it is called the first order stochastic dominance, or sometimes just usual stochastic order.

Definition 3.1.1 (First Order Stochastic Dominance)
Let X and Y be two random variables, X is said to be smaller than Y in the sense

of first order stochastic dominance $X \leq_{st} Y$ if

$$\mathbb{P}(X > x) \leq \mathbb{P}(Y > x), \qquad x \in (-\infty, \infty). \tag{3.1}$$

From above definition we can see that to rank two random variables using first order stochastic dominance, we only need to compare their survival functions or tail probabilities. There are different practical interpretations for first order stochastic dominance: in reliability theory, if $X \leq_{st} Y$, then Y is more reliable than X, but in the risk measure theory, it means that Y is more dangerous than X. There is an important equivalent condition to check whether $X \leq_{st} Y$.

Theorem 3.1.1 *$X \leq_{st} Y$ is true if, and only if, for any increasing function f,*

$$\mathbb{E}[f(X)] \leq \mathbb{E}[f(Y)]. \tag{3.2}$$

Proof. See Shaked and Shanthikumar (2007). □

Some questions arise from the definition of first order stochastic dominance, such as whether this order can be preserved under convolution? What conditions are needed to rank two compound random variables with first order stochastic dominance? These questions are very important in risk theory, and in our application to risk measures. They are answered by the following theorem.

Theorem 3.1.2 (Closure Property)

1. If $X \leq_{st} Y$, and g is any increasing function, then $g(X) \leq_{st} g(Y)$.

2. If $\{X_1, X_2, \ldots, X_m\}$ and $\{Y_1, Y_2, \ldots, Y_m\}$ are two sets of independent random variables, and $X_i \leq_{st} Y_i$ for each $i \in \{1, 2, \ldots, m\}$, then for any increasing function $\phi : \mathbb{R}^m \to \mathbb{R}$,

$$\phi(X_1, X_2, \ldots, X_m) \leq_{st} \phi(Y_1, Y_2, \ldots, Y_m). \tag{3.3}$$

Particularly, we can get the closure property of first order stochastic dominance under convolutions: $\sum_{i=1}^{m} X_i \leq_{st} \sum_{i=1}^{m} Y_i$.

3. If $\{X_1, X_2, \ldots, X_m\}$ and $\{Y_1, Y_2, \ldots, Y_m\}$ are two sets of independent random variables such that $X_i \leq_{st} Y_i$, and $X_i \xrightarrow{d} X$, $Y_i \xrightarrow{d} Y$, for each $i \in \{1, 2, \ldots, m\}$, then $X \leq_{st} Y$.

4. Let $\{X_i, i = 1, 2, \dots\}$ and $\{Y_i, i = 1, 2, \dots\}$ be two sets of independent random variables, and $X_i \leq_{st} Y_i$ for each $i \in \{1, 2, \dots, m\}$, if N and M are two other random variables that satistfy $N \leq_{st} M$, then

$$\sum_{i=1}^{N} X_i \leq_{st} \sum_{i=1}^{M} Y_i. \tag{3.4}$$

Proof. See Shaked and Shanthikumar (2007). □

3.1.1 Hazard Rate Order

In reliability theory, a frequently asked question is how reliable a device is after being used for several years. This question can be answered by evaluating its hazard rate function, which is defined as follows:

Definition 3.1.2 *If random variable X denotes the life time of a device, then its hazard rate function is given by*

$$r(t) = \lim_{\Delta t \to 0} \frac{\mathbb{P}(t < X \leq t + \Delta t \mid X > t)}{\Delta t} = \frac{f_X(t)}{\bar{F}_X(t)}, \qquad t \geq 0, \tag{3.5}$$

provided that $\bar{F}_X(t) \neq 0$.

From (3.5), we can see that the hazard rate function describes the failure intensity of the device, which explains why it is also commonly called failure rate. Multiplying both sides of (3.5) by dt, gives the failure probability within the next short period $(t, t+dt]$, after reaching age t. By comparing the hazard rate functions of two random variables, we can see which one is more likely to fail at each time epoch. This order is called the hazard rate order.

Definition 3.1.3 (Hazard Rate Oder)
If r_X and r_Y stand for the hazard rate functions of random variables X and Y, then $X \leq_{hr} Y$ means X is smaller than Y in the hazard rate order if

$$r_X(t) \leq r_Y(t), \qquad t \geq 0. \tag{3.6}$$

The first order stochastic dominance and the hazard rate order are closely related by use of the following expression:

$$\begin{aligned} r_X(t) \leq r_Y(t) \;\Rightarrow\; & \frac{\bar{F}_Y(t)}{\bar{F}_X(t)} \text{ is increasing in t} \\ \Rightarrow\; & \bar{F}_X(t_1)\bar{F}_Y(t_2) \geq \bar{F}_X(t_2)\bar{F}_Y(t_1), \qquad t_1 \leq t_2. \end{aligned}$$

Let $t_1 \to -\infty$, then one gets $\bar{F}_X(t_2) \leq \bar{F}_Y(t_2)$, which is $X \leq_{st} Y$. If the hazard rate function $r_X(t)$ is increasing in t, then we say X is an increasing failure rate (IFR) random variable. Similarly, we can define the decreasing failure rate (DFR) random variable. Some closure properties and characterization theorems with respect to this order are explored and the details are given in Denuit *et al.* (2005), or Shaked and Shanthikumar (2007).

3.2 Mean Residual Life Order

The mean residual life function $m(t)$ is another quantity which gives the information about the tail of the distribution. It is defined as the conditional expectation of the residual life given that the random variable reaches some age level.

Definition 3.2.1 (Mean Residual Life Function)
For random variable X, the mean residual life function m_X is defined as

$$m_X(t) = \mathbb{E}(X - t \mid X > t), \qquad t \geq 0, \tag{3.7}$$

provided the expectation is finite.

Like the TVaR measure, the mean residual life function describes how thick the tail of X is. In the context of reliability theory, it represents the expected residual life for a device, and larger $m_X(t)$ values lead to more durable devices. But in the context of the financial or insurance industry, if the random variable X stands for a potential loss, then the mean residual function represents the expected excess potential losses, then the larger $m_X(t)$ the more dangerous the position. The mean residual life function is thus used to rank random variables, and this order is called the mean residual life order.

Definition 3.2.2 (Mean Residual Life Order)
Let X and Y be two random variables for which their mean residual life functions m_X and m_Y exist. If $m_X(t) \leq m_Y(t)$ for all $t \geq 0$, then X is smaller than Y in the mean residual life order, which is denoted as $X \leq_{mrl} Y$.

There are some equivalent conditions which can help more quickly find this order:

Theorem 3.2.1 *Let X and Y be two random variables as in Definition 3.2.2., then $X \leq_{mrl} Y$ if, and only if, the following ratio of the tail integrals*

$$\frac{\int_t^\infty \bar{F}_Y(u)du}{\int_t^\infty \bar{F}_X(u)du} = \frac{\mathbb{E}[(Y-t)_+]}{\mathbb{E}[(X-t)_+]}, \qquad t \geq 0, \tag{3.8}$$

increases in t over the set $\{t \mid t \geq 0,\ \int_t^\infty \bar{F}_X(u)du > 0\}$.

Proof. Taking derivatives of (3.8) in t, then the mean residual order is obtained. □

Naturally, the connection between this order of random variables and other orders have been studied. Since the mean residual life function $m_X(t)$ can be written as

$$m_X(t) = \int_t^\infty \exp\{-\int_t^x r_X(u)du\}dx, \qquad t \in \text{support of } X, \tag{3.9}$$

where $r_X(u)$ is the hazard rate function. From (3.8), we can first explore the relations between the mean residual life order and the hazard rate order.

Theorem 3.2.2 *If X and Y are two random variables such that $X \leq_{hr} Y$, then $X \leq_{mrl} Y$.*

Proof. See Shaked and Shanthikumar (2007). □

If the mean residual life function $m_X(t)$ for random variable X is increasing, then X is said to be an increasing mean residual life (IMRL) random variable. The DMRL random variable can be defined similarly. For details about the closure properties of the IMRL and DMRL families of random variables, see Shaked and Shanthikumar (2007).

3.2.1 Harmonic Mean Residual Life Order

In Mathematics, the harmonic mean is one of several different kinds of averages. For some real numbers $x_1, x_2, \ldots, x_n$, the harmonic mean is defined as

$$H(x_1, x_2, \ldots, x_n) = \frac{n}{\sum_{i=1}^{n} \frac{1}{x_i}} = \Big[\frac{1}{n}(\sum_{i=1}^{n} \frac{1}{x_i})\Big]^{-1}.$$

Generalizing this idea gives what is the called harmonic mean residual life function which will be reviewed in this section. For practical financial and insurance applications, in this section, we only talk about nonnegative random variables. The harmonic mean residual life function for a nonnegative random variable X is defined as:

$$h_X(t) = \Big[\frac{1}{t}\int_0^t \frac{1}{m(u)}du\Big]^{-1}, \qquad t > 0, \tag{3.10}$$

where $m_X(u)$ is the mean residual life function. By use of (3.9), we can introduce a new order for nonnegative random variables, which is called harmonic mean residual life order.

Definition 3.2.3 (Harmonic Mean Residual Life Order)
Let X and Y be two nonnegative random variables with harmonic mean residual life functions $h_X(t)$ and $h_Y(t)$, then X is smaller than Y in the harmonic mean residual life order if

$$h_X(t) \le h_Y(t), \qquad t > 0, \tag{3.11}$$

and is denoted by $X \le_{hmrl} Y$.

From (3.9), we see that computing the harmonic mean residual life function can be complicated. Thus, Shaked and Shanthikumar (2007) simplify the comparison by use of the following relationship for $t > 0$:

$$[\frac{1}{t}\int_0^t \frac{1}{m_X(u)}du]^{-1} \le [\frac{1}{t}\int_0^t \frac{1}{m_Y(u)}du]^{-1} \iff \frac{\int_t^\infty \bar{F}_X(u)du}{\mathbb{E}[X]} \le \frac{\int_t^\infty \bar{F}_Y(u)du}{\mathbb{E}[Y]}. \tag{3.12}$$

Furthermore, the integral $\int_t^\infty \bar{F}_X(u)du$ can be rewritten as $\mathbb{E}[(X-t)_+]$, then we can get the following equivalent condition of the harmonic mean residual life order:

$$h_X(t) \le h_Y(t) \iff \frac{\mathbb{E}[(X-t)_+]}{\mathbb{E}[X]} \le \frac{\mathbb{E}[(Y-t)_+]}{\mathbb{E}[Y]}, \qquad t > 0. \tag{3.13}$$

We will see later that the truncated expectation $\mathbb{E}[(X-t)_+]$ can be used to define a new order between random variables.

We end this section by exploring the relationship between the harmonic mean residual life order and the mean residual life order. In fact, from the harmonic mean residual life function in (3.9), it is not difficult to get the following result.

Theorem 3.2.3 *Let X and Y be two nonnegative variables, if $X \leq_{mrl} Y$, then $X \leq_{hmrl} Y$.*

3.3 Convex Order

By now, several stochastic orders have been introduced, which are based on the comparison of the location and scale of the random variables. Starting from this section, some other stochastic orders will be introduced, where the variability of the random variables is compared.

The next stochastic order is worth studying and is frequently applied in actuarial science; it is called the convex order. The definition is as follows.

Definition 3.3.1 (Convex Order)
Let X and Y be two random variables. X is smaller than Y in the convex order if

$$\mathbb{E}[f(X)] \leq \mathbb{E}[f(Y)], \qquad f \textit{ is any convex function}, \tag{3.14}$$

and is denoted by $X \leq_{cx} Y$.

A rigorous definition of a convex function can be found in Rochafeller (1970), but here for convenience, we can use the following inequality to characterize a convex function:

$$f\Big(\lambda X + (1-\lambda)Y\Big) \leq \lambda f(X) + (1-\lambda)f(Y), \qquad \lambda \in (0,1). \tag{3.15}$$

Particularly, if the function f is continuous, then its convexity can be checked by its second-order derivative. From (3.13), we get the following information quickly if $X \leq_{cx} Y$: firstly, since $f(x) = x$ and $f(x) = -x$ are both convex functions, then we have that $\mathbb{E}[X] \leq \mathbb{E}[Y]$ and $\mathbb{E}[-X] \leq \mathbb{E}[-Y]$, which gives $\mathbb{E}[X] = \mathbb{E}[Y]$. Secondly, if we consider the function $f(x) = x^2$, also a convex function, so combining $\mathbb{E}[X^2] \leq \mathbb{E}[Y^2]$ with the previous results, one gets $\mathbb{V}[X] \leq \mathbb{V}[Y]$. From these information, we can roughly compare the variability of X and Y.

However, it is impossible to check whether the inequality in (3.13) is true for all convex functions, so some equivalent conditions are developed to compare two random variables in the convex order. To get equivalent conditions, we need to consider some other simpler stochastic orders.

3.3.1 Increasing Convex Order

The increasing convex order ($X \leq_{icx} Y$) is a special case of the convex order, with an additional restriction, that is the functions inserted in the expectation should be monotone increasing. This stochastic order has been proved to be equivalent with the stop loss order, which is widely used in the actuarial literature.

Definition 3.3.2 (Stop Loss Order)
Let X and Y be two random variables, given that their first order moments exist, then X is smaller than Y in the stop loss order if

$$\mathbb{E}[(X-t)_+] \leq \mathbb{E}[(Y-t)_+], \qquad t \in \mathbb{R}, \tag{3.16}$$

and is denoted by $X \leq_{sl} Y$.

The equivalence between these two stochastic orders can be verified as follows: similarly to the “standard mechanic” proof in Roydan and Fitzpatrick (2010), any Lebesgue integrable function can be approximated by a combination of simple functions. Any increasing convex function can be approximated by a combination of functions $\{f_i(x) = (x - t_i)_+\}_{i=1,2\ldots}$, and the latter belongs to the class of increasing convex functions. Then the equivalence between these two stochastic orders is obvious:

$$X \leq_{icx} Y \iff X \leq_{sl} Y. \tag{3.17}$$

Further, from the integral expression of $\mathbb{E}[(X-t)_+]$ we have:

$$\mathbb{E}[(X-t)_+] = \int_t^{\infty} \bar{F}_X(u)du.$$

Hence it is easy to derive the relation between first order stochastic dominance and the stop loss order:

$$X \leq_{st} Y \;\Rightarrow\; X \leq_{sl} Y. \tag{3.18}$$

Besides first order stochastic dominance, the mean residual order can also imply the stop loss order. Suppose $X \leq_{mrl} Y$, then

$$\frac{\int_t^{\infty} \bar{F}_X(u)du}{\bar{F}_X(t)} \leq \frac{\int_t^{\infty} \bar{F}_Y(u)du}{\bar{F}_Y(t)}, \qquad t \geq 0, \tag{3.19}$$

provided that $\bar{F}_X(t) \neq 0$ and $\bar{F}_Y(t) \neq 0$. Let $t \to -\infty$, then one gets $\mathbb{E}[X] \leq \mathbb{E}[Y]$. Combining with Theorem 3.2.1., then

$$\frac{\mathbb{E}[(Y-t)_+]}{\mathbb{E}[(X-t)_+]} \geq \frac{\mathbb{E}[Y]}{\mathbb{E}[X]} \geq 1 \;\Rightarrow\; \mathbb{E}[(X-t)_+] \leq \mathbb{E}[(Y-t)_+], \qquad \text{for all } t. \tag{3.20}$$

Now coming back to the discussion on the equivalent conditions for the convex order, the approximation method again comes into play here. The "increasing part" of any convex function can be approximated by a combination of functions $\{f_i(x) = (x-t_i)_+\}_{i=1,2\ldots}$, while the "decreasing part" can be approximated by a combination of functions $\{f_i(x) = (x-t_i)_+\}_{i=1,2\ldots}$ and the function $g(x) = -ax$, where $a > 0$. Thus, if X and Y are two random variables satisfying $\mathbb{E}[X] = \mathbb{E}[Y]$ and $X \leq_{sl} Y$, then $\mathbb{E}[g(X)] = \mathbb{E}[g(Y)]$ and $\mathbb{E}[(X-t)_+] \leq \mathbb{E}[(Y-t)_+]$, which results in $X \leq_{cx} Y$. To conclude, an equivalent condition to the convex order is:

$$X \leq_{cx} Y \iff X \leq_{sl} Y \text{with } \mathbb{E}[X] = \mathbb{E}[Y]. \tag{3.21}$$

For more properties, such as closure properties, see Denuit *et al.* (2005) and Shaked and Shanthikumar (2007).

3.4 Other Stochastic Orders

From the discussion in the Section 3.4, we find if two random variables X and Y have different expectations, then comparing them through the convex order is invalid. To rank the variablity of different random variables without considering their locations, some location-free stochastic orders are proposed, such as the dilation order.

Definition 3.4.1 (Dilation Order)
Let X and Y be two random variables with finite means. X is smaller than Y in the sense of the dilation order if

$$X - \mathbb{E}[X] \leq_{cx} Y - \mathbb{E}[Y], \tag{3.22}$$

and is denoted by $X \leq_{dil} Y$.

From this definition, we see that the dilation order dose not depend on the location of the random variables. A very important characterization theorem for this dilation order is proposed by Fagiuloi *et al.* (1999):

Theorem 3.4.1 *Let X and Y be two random variables with distributions F_X and F_Y, and each with finite expectation, then $X \leq_{dil} Y$ if, and only if,*

$$\frac{1}{1-p}\int_p^1 (F_X^{-1}(u) - F_Y^{-1}(u))du \leq \int_0^1 (F_X^{-1}(u) - F_Y^{-1}(u))du, \quad p \in [0,1). \tag{3.23}$$

Proof. See Fagiuloi *et al.* (1999) or Shaked and Shanthikumar (2007). □

Particularly, Belzunce *et al.* (1997) verify that, for nonnegative random variables X and Y, $X \leq_{dil} Y$ implies $X \leq_{icx} Y$. Since the claims are all nonnegative, if a risk measure agrees with the dilation order, then it must agree with the stop loss order.

Another special stochastic order is called the Lorenz order, which is defined as follows:

Definition 3.4.2 (Lorenz Order)
Let X and Y be two random variables, and X be smaller than Y in the sense of the Lorenz order if

$$\frac{X}{\mathbb{E}[X]} \leq_{cx} \frac{Y}{\mathbb{E}[Y]}, \tag{3.24}$$

and is denoted by $X \leq_{Lorenz} Y$.

For more details about this stochastic order, see Lefèvre and Utev (2001) and Kochar (2006).

Chapter 4

Risk Measure on Risk Processes

4.1 Risk Measures Derived From Risk Theory

Most risk measures theories are intended for losses represented by random variables. In this chapter, we consider the risk management problem for risk processes. Some efficient methods have been proposed, but all capture only part of the properties of the risk processes, rather than the overall characteristics.

4.1.1 Risk Measures Based on the Premium Rate

Dhaene *et al.* (2003) study a discrete surplus process

$$U_t = U(t-1) + ct - S_t, \quad t = 1, 2, \ldots, \tag{4.1}$$

and find that the initial capital is fixed then the premium rate (PR) can serve as a kind of risk measure for the aggregate claims S_t. For convenience, we write S_t as S for short in this subsection. In risk theory, the adjustment coefficient R is the smallest positive root of the Lundberg's equation $e^{Rc} = \mathbb{E}[e^{RS}]$, if it exists, and then the premium rate can be written as $c = \frac{1}{R} \ln \mathbb{E}[e^{RS}]$. Lundberg's inequality tells us that $\psi(u) \le e^{-Ru}$ if R exists. Then if the regulator wants to control the ruin probability at less than some fixed level, for example ϵ, then R can be selected suitably, as for instance $R = \frac{1}{u}|\ln \epsilon|$, such that the corresponding premium rate is:

$$\rho_{PR}(S) = c = \frac{1}{R} \ln \mathbb{E}[e^{RS}]. \tag{4.2}$$

This kind of risk measure coincides with the exponential premium formula which provides a suitable intepretation. In this case, the adjustment coefficient R can be treated as a risk-aversion parameter. The risk measure $\rho_{PR}(S)$ is equipped with some good properties that can be easily verified.

Theorem 4.1.1 *If the risk measure for the aggregate claims is defined by (4.2), then it has the following basic properties:*

1. Additivity for independent risk processes: in the case where S and T are independent aggregate claims, then $\rho_{PR}(S+T) = \rho_{PR}(S) + \rho_{PR}(T)$.

2. Preservation of the convex order for aggregate claims: if $S \leq_{cx} T$, then $\rho_{PR}(S) \leq \rho_{PR}(T)$.

3. ρ_{PR} is invariant for proportional changes in monetary units: $\rho_{PR}(aS) = a\rho_{PR}(S)$ for constant a.

4. If (S^*, T^*) are "more related" than (S, T), then $\rho_{PR}(S+T) \leq \rho_{PR}(S^* + T^*)$, with equality only if (S^*, T^*) and (S, T) have the same joint distribution.

A pair of random variables are defined to be "more related" than another with the same marginal distributions if their joint probability is larger. Moreover, one can find that when S and T are positively quadratic dependent, $\rho_{PR}(S)+\rho_{PR}(T) \leq \rho_{PR}(S+T)$; when S and T are negatively quadratic dependent, $\rho_{PR}(S) + \rho_{PR}(T) \geq \rho_{PR}(S+T)$. Thereby, this risk measure is neither subadditive nor superadditive. To find the definition of positive and negative quadratic dependence, see Dhaene *et al.* (2003).

However, if one wants to extend this risk measure to continuous surplus processes, the time variable may make the expectation diverge. The treatment of heavy-tail distributions is also needed to ensure that the adjustment coefficient is within the convergence radius of the moment generating function.

4.1.2 Risk Measures Based on the Ruin Probability

Trufin *et al.* (2011) improve the above method of Dhaene *et al.* (2003) and extend the risk measure to the continuous case. They consider a VaR-type risk measure defined as the smallest initial capital needed to ensure that the ruin probability is less than a given level. The ruin probability in the classical model in Definition 1.1.1, is given by:

$$\psi(u) = 1 - \frac{\theta}{1+\theta} \sum_{n=0}^{\infty} \left(\frac{1}{1+\theta}\right)^n F_D^{*(n)}(u), \tag{4.3}$$

where D is the ladder height introduced in the first chapter, just before (1.1). The VaR-type risk measure is expressed as the following:

$$\rho_\epsilon[X] = \inf\{v \geq 0 \mid \psi(v) \leq \epsilon\} = \psi^{-1}(\epsilon). \tag{4.4}$$

For a better understanding this risk measure, Trufin *et al.* (2011) establish the connection between the distribution of the aggregate loss L and (3.4) by the following equation:

$$\rho_\epsilon[X] = F_L^{-1}(1-\epsilon) = \inf\{v \mid F_L(v) \geq 1-\epsilon\}. \tag{4.5}$$

This kind of risk measure has an advantage over the previous one in Section 4.1.1: it just needs to evaluate the ruin probability without finding the adjustment coefficient. The latter can be a problem in some cases, as the adjustment coefficient does not exist for some heavy-tail distributions. Embrechts *et al.* (1997) investigate further the details of the ruin probability when the single claim severity follows a heavy-tail distribution; they also derive a Cramér-Lundberg approximation for the ruin probability when the initial capital approaches to infinity. Michel (1987) verifies that if D and E are two ladder height variables with respect to two claim variables X and Y, $D \leq_{st} E \Rightarrow \psi_X(u) \geq \psi_Y(u)$. Based on this result, the risk measure defined above has the following good properties:

Theorem 4.1.2 *If the risk measure for the surplus process is defined by (4.4), then it has the following basic properties:*

1. *Positive homogeneity:* $\rho_\epsilon[aX] = a\rho_\epsilon[X]$.
2. *Preserves the stop-loss order of* X *and* Y*: if* $X \leq_{sl} Y$*, then* $\rho_\epsilon[X] \leq \rho_\epsilon[Y]$.

3. If Z *is more positively dependent on* X *than* Y *in the sense that* $\mathbb{P}(X \leq x, Y \leq y) \leq \mathbb{P}(X \leq x, Z \leq y)$*, then* $\rho_\epsilon[X+Y] \leq \rho_\epsilon[X+Z]$*.*

4. Let Z *is the mixture of* X *and* Y *such as* $F_Z = p_1F_X + p_2F_Y + p_3F_{X+Y}$ *and* $p_1 + p_2 + p_3 = 1$*. If* $\alpha Z \leq_{hmrl} Y$*, where* $\alpha = \frac{\rho_\epsilon[Y]}{\rho_\epsilon[X]+\rho_\epsilon[Y]}$*, then* $\rho_\epsilon[Z] \leq \rho_\epsilon[X] + \rho_\epsilon[Y]$*.*

Trufin *et al.* (2011) further discuss the diversification effect of this kind of risk measure under conditions of exchangeable risks, negative quadrant dependent risks with equivalent sizes and risks that are members of scale family distributions. These results can be applied in diversifying the risks of portfolios. However, from the analysis done by Trufin *et al.* (2011), the intensity parameter of the counting process plays no role in the risk measure, since the VaR-type measure completely relies on the distribution of the aggregate loss L, and the latter only relies on the distribution of the claim severity X.

4.1.3 Risk Measures Based on the Ruin Probability and the Deficit at Ruin

Trufin and Mitric (2014) improve the VaR type method by including more information into consideration. To better describe the influence of the severity of the claim causing ruin, they add the expected deficit at ruin into the previous risk measure:

$$\xi_\epsilon[X] = \rho_\epsilon[X] + \mathbb{E}\big[|U_\tau| \;\big|\; \tau < \infty\big]. \tag{4.6}$$

This change makes the new risk measure less static than ρ_ϵ. Furthermore, the new risk measure combines the consideration of the ultimate ruin probability and the deficit at the time of ruin, which means the needed initial capital should guarantee the company the ability to recover if ruin occurs. With some basic change of variable techniques, Trufin and Mitric (2014) find the way to connect the aggregate loss and this new risk measure:

$$\xi_\epsilon[X] = TVaR[L; 1-\epsilon] - \mathbb{E}[L], \tag{4.7}$$

where L is the maximum aggregate loss defined in Chapter 1.

To compare the risks of two different risk processes, Trufin and Mitric (2014) derive several conditions after exploring the stochastic order of the claim severity distributions. Skipping the details of the proof, the following theorem summarizes the main results.

Theorem 4.1.3 *If the risk measure is defined as in (4.6) or (4.7), then for the comparison of different risk processes, the following relationships exist:*

1. $\xi_\epsilon[X] \leq \xi_\epsilon[Y]$ *if and only if* $L_x \leq_{dil} L_Y$.

2. Let D^X *and* D^Y *are the corresponding ladder heights of claim severity* X *and* Y*, if* $D^X \leq_{dil} D^Y$*, then* $\xi_\epsilon[X] \leq \xi_\epsilon[Y]$.

3. Let X *and* Y *be IMRL, where IMRL means increasing mean residual life, then this risk measure agrees with the mean residual life order, that is if* $X \leq_{mrl} Y$*, then* $\xi_\epsilon[X] \leq \xi_\epsilon[Y]$.

4. Let X *be DMRL and* Y *be IMRL, then this risk measure agrees with the harmonic mean residual life order, that is if* $X \leq_{hmrl} Y$*, then* $\xi_\epsilon[X] \leq \xi_\epsilon[Y]$.

Besides discussing the order of the risk processes under different conditions, the basic properties of this risk measure are given by Trufin and Mitric (2014). Except for the translation invariance, the risk measure $\xi_\epsilon[X]$ preserves the other properties of $\rho_\epsilon[X]$.

Theorem 4.1.4 *The risk measure* $\xi_\epsilon[X]$ *defined in (4.7) has the following basic properties:*

1. $\xi_\epsilon[X]$ *is positively homogeneous:* $\xi_\epsilon[aX] = a\xi_\epsilon[X]$.

2. Suppose X *and* Y *are identically distributed, then* $\xi_\epsilon[X+Y] \leq \xi_\epsilon[X] + \xi_\epsilon[Y]$ *if and only if* $L_{\frac{X+Y}{2}} \leq_{dil} L_Y$.

3. $\xi_\epsilon[X]$ *satisfies the following inequality:* $\xi_\epsilon[X+a] \leq (\geq) \xi_\epsilon[X] + a$ *for any positive constant* a *which satisfies* $a \geq (\leq) \frac{VaR[X]}{\mathbb{E}[X]} - \mathbb{E}[X]$.

For the diversification effect of this risk measure, Trufin and Mitric (2014) give out similar results as those in Trufin *et al.* (2011). If there exists a specific dependence structure between the claim variables X and Y, this risk measure will be subadditive.

It simplifies the second result in the previous theorem, since we do not need to verify the dilation order between $L_{\frac{X+Y}{2}}$ and L_Y.

Theorem 4.1.5 (More about subadditivity) *The risk measure ξ_ϵ is subadditive if there exists following dependence structure between the claim variables X and Y:*

1. X and Y are exchangeable risks, that is $\mathbb{P}(X \leq t_1, Y \leq t_1) = \mathbb{P}(X \leq t_2, Y \leq t_1)$.

2. X and Y are negatively quadrant dependent and identically distributed, that is the inequality $\mathbb{P}(X \leq t_1, Y \leq t_2) \leq \mathbb{P}(X \leq t_1)\mathbb{P}(Y \leq t_2)$ holds for all t_1 and t_2.

3. $X = \beta V_1$ and $Y = \gamma V_2$, where V_1 and V_2 are identically distributed.

4.1.4 Risk Measures Based on the EAR

In the first chapter, many quantities are mentioned, some that can reflect the influence of the claim severity and the intensity of the counting process, such as the expected area in the red (EAR). Henceforth, Loisel and Trufin (2014) propose to use this quantity as a new risk measure for the risk process. Similar to the methods used by Trufin and Mitric (2014), this risk measure is based on the needed initial capital to control the expected area in the red below a certain level.

Before presenting this risk measure of Loisel and Trufin (2014), we need to be clear about the notation used in this section. Denoting the aggregate claims $\sum_{i=1}^{N(t)} X_i$ by S_t, the premium rate c and initial capital u, the area in red within the time period $[0, T]$ is defined as:

$$I_{T,c}^S(u) = \int_0^T |U_t| 1_{\{U_t<0\}} dt, \qquad u \geq 0,$$

where $u_t = u + ct - S_t$, $t \geq 0$ (see Figure 4.1). For the expected area in red, it is easy to prove the following equation by use of Fubini's theorem:

$$\begin{aligned} \mathbb{E}\big[I_{T,c}^S(u)\big] &= \mathbb{E}\big[\int_0^T |U_t| 1_{\{U_t<0\}} dt\big] \\ &= \int_0^T \mathbb{E}\big[|U_t| 1_{\{U_t<0\}}\big] dt \\ &= \int_0^T \mathbb{E}\big[(S_t - ct - u)_+\big] dt. \end{aligned} \tag{4.8}$$

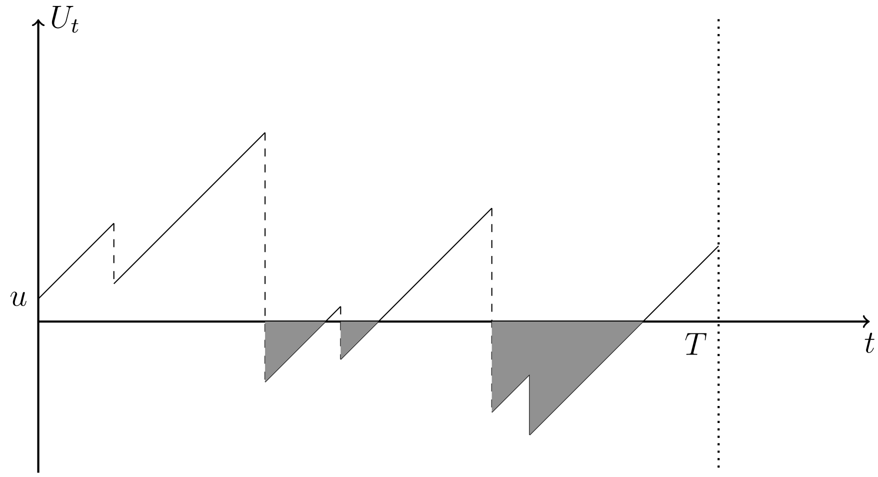

Figure 4.1: The area in the red during $[0, T]$

Fixing the time period, then Loisel and Trufin (2014) define their new risk measure as follows:

Definition 4.1.1 *For a surplus process with aggregate claims S_t and the premium rate c, the initial capital needed to control the expected area in the red within a given time period $[0, T]$ below level A is:*

$$\rho^T_{A,c}[S_t] = \inf\{v \geq 0 \mid \mathbb{E}[I^S_{T,c}(v)] \leq A\}. \tag{4.9}$$

With this definition, we can easily see the advantage of using this risk measure. Since the ruin probability within a finite time period is complicated to compute, no explicit form with respect to it can be used easily to evaluate the risk. We will see here that the infimum expression in (4.9) has many good properties, which are presented in the following several theorems.

Theorem 4.1.6 *Let S_t and $\tilde{S}_t$ be two different aggregate claim processes. If the risk measure $\rho^T_{A,c}$ on the surplus process is defined as (4.9), then it has the following properties:*

1. The risk measure $\rho^T_{A,c}$ agrees with the stop-loss order, that is $S_t \leq_{icx} \tilde{S}_t \Rightarrow \rho^T_{A,c}[S_t] \leq \rho^T_{A,c}[\tilde{S}_t]$.

2. The risk measure $\rho^T_{A,c}$ is monotone, that is $\mathbb{P}(S_t \leq \tilde{S}_t) = 1$ for all $t > 0$ leads to $\rho^T_{A,c}[S_t] \leq \rho^T_{A,c}[\tilde{S}_t]$.

3. The risk measure $\rho^T_{A,c}$ is translation invariant, that is $\rho^T_{A,c}[S_t + a] = \rho^T_{A,c}[S_t] + a$ for any constant $a > 0$.

4. The risk measure $\rho^T_{A,c}$ is positively homogeneous, that is $\rho^T_{A,c}[aS_t] = a\rho^T_{\frac{A}{a},\frac{c}{a}}[S_t]$ for any constant $a > 0$.

5. The risk measure $\rho^T_{A,c}$ is subadditive, in the sense that the inequality $\rho^T_{A,c}[S_t + \tilde{S}_t] \le \rho^T_{\beta A,\alpha c}[S_t] + \rho^T_{(1-\beta)A,(1-\alpha)c}[\tilde{S}_t]$ holds true for all $\alpha, \beta \in (0,1)$, whatever the dependence structure between the aggregate claims S_t and $\tilde{S}_t$.

As the risk measure $\rho^T_{A,c}$ satisfies the four basic properties of a coherent risk measure first introduced in Artzner *et al.* (1999) (see Definition 2.1.1), one can say that $\rho^T_{A,c}$ is coherent. However, when discussing positive homogeneity and subadditivity, we can find risk measures based on different criteria, which makes “coherence” debatable. In the next section, we will go back to the assumption of the classical model: $c = (1+\theta)\lambda\mathbb{E}[X]$ to discuss the properties of the risk measure for this case.

4.2 Some New Measures on Risk Processes

From the initial study in the previous section we see that, to compare the risks of different risk processes, two basic approaches are usually applied. The first way is by fixing the initial capital, then selecting the key quantity in the risk process which is comparable and easy to deal with to serve as the risk measure. The second way is to find out the necessary initial capital to control the adverse index, such as the ruin probability, expected deficit or expected area in the red.

In this section, other new and useful comparable quantities are introduced to serve as possible risk measures. The properties of these new risk measures are explored, to see if these results agree with the analysis done before.

4.2.1 Risk Measures Based on the Safety Loading

Under the original assumptions from the classical model, we can see that the premium is proportional to the expectation of the aggregate claims:

$$c = (1+\theta)\lambda\mathbb{E}[X]. \tag{4.10}$$

The classical model assumes that the safety loading is fixed and kept the same for all the risk processes. On the other hand, from the discussion at the beginning of this section, we see that fixing the initial capital, then the safety loading can serve as another kind of risk measure. In selecting a certain criteria, we can make this quantity comparable from one process to another. Intuitively, higher safety loadings lead to more dangerous risk processes. Let us then further study the safety loading.

Adminstrators of insurance companies want to see high levels of the surplus process. Mathematically, this could mean that at each time point, or during the following very short time period $(0, \Delta t]$, the cumulated premiums should exceed the claim amounts, or at least, the probability of this event should be large. Then consider this probability and make it greater than some given level:

$$\mathbb{P}(X\lambda\Delta t < c\Delta t) \geq 1-\epsilon, \tag{4.11}$$

where X is the claim severity random variable, and $\lambda\Delta t$ is the probability of occurrence of one claim. Since $c = (1+\theta)\lambda\mathbb{E}[X]$, then the inequality above becomes

$$\begin{aligned}
\mathbb{P}(X\lambda\Delta t < c\Delta t) \geq 1-\epsilon \quad &\Rightarrow \quad \mathbb{P}(X\lambda\Delta t < (1+\theta)\lambda\mathbb{E}[X]\Delta t) \geq 1-\epsilon \\
&\Rightarrow \quad \mathbb{P}(X < (1+\theta)\mathbb{E}[X]) \geq 1-\epsilon \\
&\Rightarrow \quad (1+\theta)\mathbb{E}[X] \geq F^{-1}_{1-\epsilon}(X) = VaR_{1-\epsilon}(X) \\
&\Rightarrow \quad \theta \geq \frac{VaR_{1-\epsilon}(X)}{\mathbb{E}[X]} - 1. \qquad (4.12)
\end{aligned}$$

From the expression above, we can define a proper risk measure based on the safety loading.

Definition 4.2.1 (Safety Loading) *For the surplus process with a given initial capital level, the safety loading can be defined as in (4.12) to control the probability in (4.11) to be greater than α. This safety loading argument hence generates risk measure*

$$\rho_\alpha^*[X] = \inf\{\theta \mid \theta \geq \frac{VaR_\alpha(X)}{\mathbb{E}[X]} - 1\} = \frac{VaR_\alpha(X)}{\mathbb{E}[X]} - 1. \tag{4.13}$$

To distinguish this risk measure and the VaR type risk measure in Trufin *et al.* (2011), we use ρ_α^* here.

Mostly, risk measures are defined over a finite or infinite horizon. Here, we reconsider the problem from an infinitesimal horizon for the practical convenience. With the above definition, we can derive the properties of this new risk measure, starting the discussion with coherence. In order to better explore the subadditivity, we first introduce a concept.

Definition 4.2.2 *A random vector (X, Y) has regularly varying right tails with tail index η if there is a function $a(t) > 0$ that is regularly varying at infinity with exponent $\frac{1}{\eta}$ and a nonzero measure μ on $(0, \infty)^2/\{0\}$ such that*

$$t\mathbb{P}\Big((X, Y) \in a(t)\Big) \to \mu, \tag{4.14}$$

as $t \to \infty$ vaguely in $(0, \infty)^2/\{0\}$.

Proposition 4.2.1 *Provided the first moment of the claim severity distribution exists, the risk measure based on the safety loading has the following properties:*

1. If $\frac{X}{\mathbb{E}[X]} \leq_{st} \frac{Y}{\mathbb{E}[Y]}$, then $\rho_\alpha^*[X] \leq \rho_\alpha^*[Y]$.

2. For any positive constant a, $\rho_\alpha^*[aX] = \rho_\alpha^*[X]$.

3. For any positive constant a, $\rho_\alpha^*[X + a] < \rho_\alpha^*[X]$.

4. If X and Y have jointly regularly varying non-degenerate tails with tail index $\beta > 1$, then $\rho_\alpha^*[X + Y] \leq \rho_\alpha^*[X] + \rho_\alpha^*[Y]$ for large α.

Proof. 1. Since the definition of θ is equivalent to:

$$\begin{aligned}\mathbb{P}(X\lambda\Delta t < c\Delta t) \geq \alpha &\Rightarrow \mathbb{P}(X < (1+\theta)\mathbb{E}[X]) \geq \alpha \\ &\Rightarrow \mathbb{P}(\frac{X - \mathbb{E}[X]}{\mathbb{E}[X]} < \theta) \geq \alpha \\ &\Rightarrow \theta = VaR_\alpha(\frac{X - \mathbb{E}[X]}{\mathbb{E}[X]}),\end{aligned}$$

so if $\frac{X}{\mathbb{E}[X]} \leq_{st} \frac{Y}{\mathbb{E}[Y]}$, then $\frac{X}{\mathbb{E}[X]} - 1 \leq_{st} \frac{Y}{\mathbb{E}[Y]} - 1$, which indicates that $VaR_\alpha(\frac{X-\mathbb{E}[X]}{\mathbb{E}[X]}) \leq VaR_\alpha(\frac{Y-\mathbb{E}[Y]}{\mathbb{E}[Y]})$. Then the first statement is true.

2. It is easy to see that $\frac{VaR_\alpha(aX)}{\mathbb{E}[aX]} - 1 = \frac{VaR_\alpha(X)}{\mathbb{E}[X]} - 1$, so $\rho_\alpha[aX] = \rho_\alpha[X]$. This property means that this risk measure is scale free. Moreover, similar to the risk measure presented by Trufin *et al.* (2011), the intensity parameter does not play any role here.

3. Since $\rho^*_\alpha[X] = \frac{VaR_\alpha(X)}{\mathbb{E}[X]} - 1 = \frac{VaR_\alpha(X)-\mathbb{E}[X]}{\mathbb{E}[X]}$, then it is easy to find that

$$\begin{aligned}\rho^*_\alpha[X+a] &= \frac{VaR_\alpha(X+a) - \mathbb{E}[X+a]}{\mathbb{E}[X+a]} = \frac{VaR_\alpha(X) - \mathbb{E}[X]}{\mathbb{E}[X]+a} \\ &< \frac{VaR_\alpha(X) - \mathbb{E}[X]}{\mathbb{E}[X]} = \rho^*_\alpha[X].\end{aligned}$$

In fact, we can find an intepretation for the inequality above. Once there is a fixed loss added to each claim, then the insurers will charge the insureds additional fees $(1+\theta)\lambda a$, that accounts for the uncertainty of the random claims. That is why the risk is reduced in this case.

4. Based on the results of Daníelsson *et al.* (2005), VaR has a diversification effect in the tail region, when X and Y have jointly regularly varying non-degenerate tails with tail index $\beta > 1$. So for large α, $VaR_\alpha(X+Y) \leq VaR_\alpha(X) + VaR_\alpha(Y)$. With this inequality, it is easy to prove the next inequality:

$$\begin{aligned}\rho^*_\alpha[X+Y] &= \frac{VaR_\alpha(X+Y)}{\mathbb{E}[X]+\mathbb{E}[Y]} - 1 \\ &\leq \frac{VaR_\alpha(X) + VaR_\alpha(Y)}{\mathbb{E}[X]+\mathbb{E}[Y]} - 1 \\ &= \frac{\mathbb{E}[X]}{\mathbb{E}[X]+\mathbb{E}[Y]}(\frac{VaR_\alpha(X)}{\mathbb{E}[X]} - 1) + \frac{\mathbb{E}[Y]}{\mathbb{E}[X]+\mathbb{E}[Y]}(\frac{VaR_\alpha(Y)}{\mathbb{E}[Y]} - 1) \\ &\leq \frac{VaR_\alpha(X)}{\mathbb{E}[X]} - 1 + \frac{VaR_\alpha(Y)}{\mathbb{E}[Y]} - 1 \\ &= \rho^*_\alpha[X] + \rho^*_\alpha[Y].\end{aligned}$$

Sometimes, the claims follow a heavy-tail distribution, such as Pareto-type distributions, which here always have the tail indexes that exceed 1. Then the regulators want a high α to avoid insolvency. Therefore, the subadditivity of this risk measure makes some sense for heavy-tail risks. □

In fact, the above property implies that the two risk processes must share the same intensity parameter, which may not be practical. Next we see that even for the most general case, subadditivity still holds. Consider two correlated aggregate claims $S_t^{(1)} = \sum_{i=1}^{N_1(t)+N(t)} X_i$ and $S_t^{(2)} = \sum_{i=1}^{N_2(t)+N(t)} Y_i$, then the corresponding risk measure or safety loading defined through (4.12) is

$$\mathbb{P}(\lambda_1 X + \lambda_2 Y + \lambda(X+Y) \leq (1+\theta)((\lambda_1+\lambda)\mu_1 + (\lambda_2+\lambda)\mu_2)) \geq \alpha$$
$$\Rightarrow \ \theta \geq \frac{VaR_\alpha((\lambda_1+\lambda)X + (\lambda_2+\lambda)Y)}{(\lambda_1+\lambda)\mu_1 + (\lambda_2+\lambda)\mu_2} - 1,$$

where λ_1, λ_2 and λ are the intensity parameters for the counting processes $N_1(t)$, $N_2(t)$ and $N(t)$, while μ_1 and μ_2 are the expectations of X and Y. We see that if the random variables X and Y meet the above conditions, then the risk measure for $S_t^{(1)} + S_t^{(2)}$ still satisfies subadditivity. The proof is similar to that of the last property.

The special case when the two risk processes share the same intensity parameter, and their claim variables are comonotonic, then we get the following result.

Proposition 4.2.2 *If X and Y are comonotonic, then the risk measure $\rho_\alpha^*[X]$ is still subadditive for large α.*

Proof. If X and Y are co-monotonic, from the basic properties about VaR established by Denuit *et al.* (2005), VaR is additive for co-monotonic risks. So when α is large enough, $VaR_\alpha(X) > \mathbb{E}[X]$, $VaR_\alpha(Y) > \mathbb{E}[Y]$ and $VaR_\alpha(X+Y) > \mathbb{E}[X+Y]$. Following the steps of the proof for Property 4 in the previous proposition, easily shows that $\rho_\alpha^*[X+Y] \leq \rho_\alpha^*[X] + \rho_\alpha^*[Y]$. □

Example 4.2.1 (Pareto claim severity) Suppose the claim severity follows the Pareto distribution with parameters α and θ, then the corresponding quantile and expectation are

$$VaR_p(X) = \theta[(1-p)^{-\frac{1}{\alpha}} - 1], \tag{4.15}$$
$$\mathbb{E}[X] = \frac{\theta}{\alpha - 1}, \qquad \alpha > 1. \tag{4.16}$$

Consider the rescaled random variable $\frac{X}{E[X]}$, its distribution is still of Pareto type:

$$\mathbb{P}(\frac{X}{\mathbb{E}[X]} \leq x) = \mathbb{P}(X \leq \frac{\theta}{\alpha-1}x) = 1 - (\frac{\alpha-1}{x+\alpha-1})^\alpha, \qquad x > 0.$$

Therefore, by use of the risk measure defined in this section, we find that

$$\rho_p^*[X] = \frac{VaR_p(X)}{\mathbb{E}[X]} - 1 = (\alpha - 1)[(1-p)^{-\frac{1}{\alpha}} - 1] - 1. \tag{4.17}$$

If X and Y are two such random variables, and with scale parameters α_1 and α_2, $\alpha_1 \leq \alpha_2$, then

$$\Big(\frac{\alpha_1 - 1}{x + \alpha_1 - 1}\Big)^{\alpha_1} \leq \Big(\frac{\alpha_2 - 1}{x + \alpha_2 - 1}\Big)^{\alpha_2},$$

which indicates that $\frac{X}{\mathbb{E}[X]} \leq_{st} \frac{Y}{\mathbb{E}[Y]}$. By simple computations of $\rho_p^*[X]$ and $\rho_p^*[Y]$, we see that $\rho_p^*[X] \leq \rho_p^*[Y]$.

Additionally, from Daníelsson *et al.* (2005), we know that the Pareto distribution belongs to the class of regularly varying distributions with tail index α. If $p \in (0, 1)$ and p is very close to 1, the VaR measure is subadditive for Pareto type random variables. Thus, the last property of the Proposition 4.2.1 holds true for the random variables in the above example. For more details about the subadditivity of the VaR measure, please refer to Daníelsson *et al.* (2005).

4.2.2 Risk Measures Based on the Adjustment Coefficient

In the previous section, the risk measure defined by Dhaene *et al.* (2003) relies on the Lundberg's inequality $\psi(u) \leq e^{-Ru}$, where $\psi(u)$ is the ruin probability (see Definition 1.1.2). Also, recall from the introduction of the first chapter, that in the classical model, the Cramér-Lundberg's approximation of terminal ruin probability when initial capital goes to infinity: $\psi(u) \sim Ce^{-Ru}$. These results show that larger adjustment coefficients R lead to lower ruin probabilities. Thus the adjustment coefficient can also serve as a new risk measure for risk processes. Hald and Schmidli (2004) use the adjustment coefficient as a measure for the risk processes and discuss the optimal strategy for proportional reinsurance in order to maximize the adjustment coefficient. Motivated by their work, we define a new risk measure for risk processes on the basis of adjustment coefficient.

The adjustment coefficient R is derived from the Lundberg fundamental equation. Gerber and Shiu (1998) give a general form of this equation;

$$\mathbb{E}[e^{-\xi t} e^{R(S_t - ct)}] = 1, \tag{4.18}$$

which is equivalent to the following equation:

$$\mathbb{E}[e^{RS_t}] = e^{(\xi+Rc)t}. \tag{4.19}$$

The left hand side expression of (4.18) is the moment generating function (mgf) of the compound renewal processes; it is difficult to find out its general analytical expression. There are many articles about finding the moments of the compound renewal process, such as Léveillé and Garrido (2001a, 2001b), in which they discuss the first and second moments of renewal sums with discounted claims. Jang (2004) uses the martingale approach to find the mgf of renewal sums with discounted claims when the counting process is Poisson and find the moments by differentiating the mgf and evaluating it at the original point.

Léveillé, Garrido and Wang (2010) make a breakthrough in this problem. They give integral equations of the moment generating function of renewal sums with discounted claims:

$$M_{Z(t)}(x) = \bar{F}(t) + \int_0^t M_X(se^{-\xi v})M_{Z(t-v)}(se^{-\xi v})dF(v), \qquad x > 0,$$

where $t > 0$, $\xi \geq 0$ is the force of interest, and $\bar{F}(t) = 1 - F(t)$. Then they use it to derive the analytical expression for the moment generating function recursively:

$$M_{Z(t)}(x) = \sum_{n=0}^{\infty} H_n(t,s), \qquad x > 0, \tag{4.20}$$

$$H_n(t,s) = \int_0^t M_X(se^{-\xi v})H_{n-1}(t-v, se^{-\xi v})dF(v), \quad H_0(t,s) = \bar{F}(t). \tag{4.21}$$

It is difficult to initiate the study of the roots of (4.17) by using the expression above, so in this thesis we only restrict our discussion to the context of the classical risk model.

Since our discussion is based on the adjustment coefficient, first suppose that the mgf for the severity distribution exists. When the force of interest $\xi = 0$, and the underling counting process is Poisson, it easy to get the corresponding equation for (4.17):

$$1 + \frac{c}{\lambda}r = M_X(r), \tag{4.22}$$

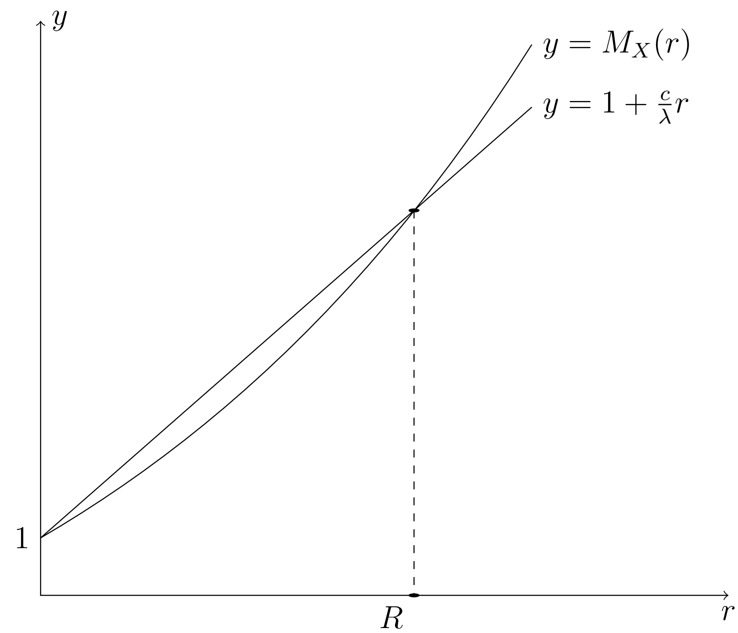

Figure 4.2: Adjustment coefficient of the classical model

in which $M_X(r)$ is the claim severity mgf, c is the premium rate and λ is the frequency parameter of the Poisson process. Since the mgf of a nonnegative random variable is always convex, and the right side of the equation (4.22) is a linear function, then the moment generating function $M_X(r)$ will intersect with the linear function $g(r) = 1 + \frac{c}{\lambda}r$ at mostly two points. When taking $r = 0$, left side equals to right side in the equation (4.22), so there is an unique positive root to (4.22), if $r = R$ is within the convergence radius of the moment generating function.

Definition 4.2.3 (Adjustment Coefficient) *If the positive root of equation (4.22) exists, then this root is the adjustment coefficient for the Lundberg equation, and the risk measure for the whole risk process is defined as:*

$$\rho_{AC}[R_t] = \rho\Big[\sum_{i=0}^{N(t)} X_i - ct\Big] = f(R), \tag{4.23}$$

where

$$R = \inf\Big\{r > 0 \;\Big|\; 1 + \frac{c}{\lambda}r = M_X(r)\Big\}, \tag{4.24}$$

and the function f is a decreasing function. From the analysis at the beginning of this chapter, we have that more dangerous risk processes lead to smaller adjustment coefficients, explaining why a decreasing function is applied to the root in (4.23). This way, if R_t is more dangerous, $\rho_{AC}[R_t]$ is larger. In fact, to compare the risks of different surplus processes is equivalent to compare the adjustment coefficients, but the properties of this risk measure heavily rely on the choice of the function f. Motivated by Lundberg's inequality $\psi(u) \le e^{-Ru} \le \epsilon$, one can get the relationship between the adjustment coefficient and the needed initial capital $u \ge \frac{-\ln \epsilon}{R}$, so one suitable choice for f is reciprocal function $f(R) = \frac{1}{R}$. Here we relax the assumption of the classical Cramér-Lundberg model that the premium rate is $c = (1+\theta)\lambda\mathbb{E}[X]$, since the equation (4.23) would be degenerate under this assumption.

This risk measure has several interesting properties, the first and also the most important is that it can diversify risks for two different risk processes.

Proposition 4.2.3 *If $R_t^{(1)}$ and $R_t^{(2)}$ are two classical Poisson risk processes, then the risk measure defined by (4.23) satisfies the following inequality: $\rho_{AC}[R_t^{(1)} + R_t^{(2)}] \le \rho_{AC}[R_t^{(1)}] + \rho_{AC}[R_t^{(2)}]$.*

Proof. Suppose that the risk processes $R_t^{(1)}$ and $R_t^{(2)}$ are defined as:

$$R_t^{(1)} = \sum_{i=1}^{N_1(t)} X_i - c_1 t, \quad R_t^{(2)} = \sum_{i=1}^{N_2(t)} Y_i - c_2 t,$$

$$1 + \frac{c_1}{\lambda_1} R = M_X(R), \tag{4.25}$$

$$1 + \frac{c_2}{\lambda_2} R = M_X(R), \tag{4.26}$$

where λ_1 and λ_2 are the corresponding intensity parameters of the counting processes $N_1(t)$ and $N_2(t)$. Now let R_1 be the positive root of equation (4.24) and R_2 be the positive root of equation (4.25), such that $R_1 \le R_2$. Then consider the process $R_t^{(1)} + R_t^{(2)}$. Since $N_1(t)$ and $N_2(t)$ are two Poisson processes with parameters λ_1 and λ_2, then $N_1(t) + N_2(t)$ is still a Poisson process with parameter $\lambda_1 + \lambda_2$, so $R_t^{(1)} + R_t^{(2)}$ can be treated as a new risk process:

$$R_t^{(1)} + R_t^{(2)} = \sum_{i=1}^{N_3(t)} Z_i - c_3 t,$$

where $N_3(t)$ is a Poisson process with parameter $\lambda_3 = \lambda_1 + \lambda_2$, the premium rate $c_3 = c_1 + c_2$, and the density function for Z_i is:

$$f_Z(x) = \frac{\lambda_1}{\lambda_1 + \lambda_2} f_X(x) + \frac{\lambda_2}{\lambda_1 + \lambda_2} f_Y(x), \qquad x \geq 0.$$

It is easy to write out Lundberg's equation about this new risk process:

$$1 + \frac{c_3}{\lambda_3} R = M_Z(R), \tag{4.27}$$

which is equivalent to:

$$1 + \frac{c_1 + c_2}{\lambda_1 + \lambda_2} R = \frac{\lambda_1}{\lambda_1 + \lambda_2} M_X(R) + \frac{\lambda_2}{\lambda_1 + \lambda_2} M_Y(R). \tag{4.28}$$

We can rearrange the order of the terms in the equation above and rewrite it as:

$$\frac{\lambda_1}{\lambda_1 + \lambda_2}(1 + \frac{c_1}{\lambda_1} R) + \frac{\lambda_2}{\lambda_1 + \lambda_2}(1 + \frac{c_2}{\lambda_2} R) = \frac{\lambda_1}{\lambda_1 + \lambda_2} M_X(R) + \frac{\lambda_2}{\lambda_1 + \lambda_2} M_Y(R). \tag{4.29}$$

Suppose the positive root of (4.27) is R_3, we can rank R_1, R_2 and R_3 by comparing equations (4.27), (4.25) and (4.26). Since $R_1 \leq R_2$, then the following equality and inequality hold:

$$1 + \frac{c_1}{\lambda_1} R_1 = M_X(R_1), \tag{4.30}$$

$$1 + \frac{c_2}{\lambda_2} R_1 \geq M_Y(R_1). \tag{4.31}$$

Thus we can get the following inequality:

$$1 + \frac{c_3}{\lambda_3} R_1 \geq M_Z(R_1). \tag{4.32}$$

Furthermore, we can get a similar inequality substituting in R_2:

$$1 + \frac{c_3}{\lambda_3} R_2 \leq M_Z(R_2). \tag{4.33}$$

Thereby $R_1 \leq R_3 \leq R_2$, which is equivalent to $\frac{1}{R_1} \geq \frac{1}{R_3} \geq \frac{1}{R_2}$, so we get $\rho_{AC}[R_t^{(1)}] \geq \rho_{AC}[R_t^{(3)}] \geq \rho_{AC}[R_t^{(2)}]$, then the following inequality holds:

$$\rho_{AC}[R_t^{(3)}] = \rho_{AC}[R_t^{(1)} + R_t^{(2)}] \leq \rho_{AC}[R_t^{(1)}] + \rho_{AC}[R_t^{(2)}].$$

$\square$

Here it makes no sense to consider translation invariance since the initial capital is not included into the definition. For positive homogeneity, we get the following result.

Proposition 4.2.4 *If R_t is a classical Poisson risk process with premium rate c and frequency parameter λ, then the following equality is true for any constant $a > 0$ with $f(R) = \frac{1}{R}$:*

$$\rho_{AC}[aR_t] = a\rho_{AC}[R_t].$$

Proof. The risk process is $R_t = \sum_{i=1}^{N(t)} X_i - ct$, and $N(t)$ is a Poisson process with parameter λ, then the corresponding Lundberg equation is:

$$1 + \frac{c}{\lambda}R = M_X(R). \tag{4.34}$$

The Lundberg equation for the scaled risk process $aR_t = \sum_{i=1}^{N(t)} aX_i - act$ is:

$$1 + \frac{ac}{\lambda}R = 1 + \frac{c}{\lambda}aR = M_X(aR). \tag{4.35}$$

Denoting the positive root of equation (4.34) by R_1 and that of equation (4.35) by R_2, compare equations (4.34) and (4.35). Since the positive root of Lundberg's equation in the Poisson case is unique, it is easy to find that $aR_2 = R_1$. So the relationship between $\rho_{AC}[R_t]$ and $\rho_{AC}[aR_t]$ is:

$$\rho_{AC}[aR_t] = \frac{1}{R_2} = a\frac{1}{aR_2} = a\frac{1}{R_1} = a\rho_{AC}[R_t].$$

□

For coherent risk measures or distortion risk measures, monotonicity is easily verified since these measures are based on a single variable. However, now the risk object is the whole risk process, and the uncertainty of the process not only comes from the claim severity, but also from the underling counting process and the premium principle. In order to compare the risks of two risk processes, a crude approach is provided by the following theorem.

Proposition 4.2.5 *Let $R_t^{(1)} = \sum_{i=1}^{N_1(t)} X_i - c_1 t$ and $R_t^{(2)} = \sum_{i=1}^{N_2(t)} Y_i - c_2 t$ be two classial Poisson risk processes, λ_1 and λ_2 are parameters for $N_1(t)$ and $N_2(t)$. If* *these two risk processes satisfy* $\frac{\lambda_1}{c_1} \leq \frac{\lambda_2}{c_2}$*, and their claim severities satisfy* $X \leq_{icx} Y$*, then* $\rho_{AC}[R_t^{(1)}] \leq \rho_{AC}[R_t^{(2)}]$.

Proof. For the risk processes $R_t^{(1)}$ and $R_t^{(2)}$, their corresponding Lundberg equations are:

$$1 + \frac{c_1}{\lambda_1}R = M_X(R) \Rightarrow \frac{\lambda_1}{c_1} + R = \frac{\lambda_1}{c_1}M_X(R), \tag{4.36}$$

$$1 + \frac{c_2}{\lambda_2}R = M_Y(R) \Rightarrow \frac{\lambda_2}{c_2} + R = \frac{\lambda_2}{c_2}M_Y(R). \tag{4.37}$$

The left hand side of the above equations is linear with the same slope but different intercepts. The right hand side are two convex functions about R. For convenience let $m_1(R) = \frac{\lambda_1}{c_1}M_X(R)$ and $m_2(R) = \frac{\lambda_2}{c_2}M_Y(R)$; it is easy to see that $m_1(0) = \frac{\lambda_1}{c_1}$ and $m_2(0) = \frac{\lambda_2}{c_2}$. Since $X \leq_{icx} Y$, the two following inequalities can be verified by using the properties of increasing convex order:

$$0 < \frac{\lambda_1 \mathbb{E}[X]}{c_1} = m_1^{'}(0) \leq \frac{\lambda_2 \mathbb{E}[Y]}{c_2} = m_2^{'}(0) < 1, \tag{4.38}$$

$$\frac{\lambda_1}{c_1}\mathbb{E}[X^2 e^{RX}] = m_1^{''}(R) \leq \frac{\lambda_2}{c_2}\mathbb{E}[Y^2 e^{RY}] = m_2^{''}(R). \tag{4.39}$$

If R_1 is the positive root of equation (4.36) and R_2 is the positive root of (4.37), then combing the inequalities above we see that $R_1 \geq R_2$, then $\rho_{AC}[R_t^{(1)}] \leq \rho_{AC}[R_t^{(2)}]$ follows. □

Example 4.2.2 (Exponential claim severity) If the claim severities are i.i.d. exponentially distributed with mean $\frac{1}{\beta}$, then the corresponding moment generating function is $M_x(r) = \frac{\beta}{\beta - r}$ for $r < \beta$, and then the smallest positive solution of Lundberg's equation is $R = \beta - \frac{\lambda}{c}$. Clearly, when λ decreases, then the claims will occur less frequently, and this will make the whole process less dangerous. If the premium rate c is higher, then the safety loading is increased, which may decrease the risk of the whole process. If instead the parameter β of the exponential distribution is increased, considering the tail distribution of the exponential distribution $\bar{F}_{X_i}(x) = 1 - F_{X_i}(x) = e^{-\beta_i x}$, then:

$$X_1 \leq_{st} X_2 \;\Leftrightarrow\; \bar{F}_{X_1}(x) \leq \bar{F}_{X_2}(x), \qquad \beta_1 \geq \beta_2, \tag{4.40}$$

which indicates that X_2 is more dangerous than X_1, so increasing the parameter of exponential distribution is equivalent to decreasing the risk of the whole process. All these operations will increase the adjustment coefficient, so the adjustment coefficient can clearly be used to measure the risk of the risk processes.

4.2.3 Generalized Risk Measures Based on the EAR

In the previous section, we discuss a drawback in the proposal by Loisel and Trufin (2014). The fixed premium rate makes this risk measure totally characterized by the aggregate claim process rather than the risk process. In order to make the risk measure more dynamic, in this section we generalize the method of Loisel and Trufin (2014) under the classical assumption $c = (1+\theta)\lambda\mathbb{E}[X]$, where X is the claim severity, λ is the intensity of the Poisson counting process and θ is the safety loading.

Under these classical risk assumptions, a new risk measure is based on the following definition of the expected area in the red.

Definition 4.2.4 *Let a classical risk process be defined as $R_t = S_t - (1+\theta)\lambda\mu t$, where $\mu = \mathbb{E}[X]$, and S_t is the aggregate claims. The expected area in red in the time interval $[0, T]$ with initial capital u is given by*

$$\mathbb{E}[I_T^{R_t}(u)] = \mathbb{E}\Big[\int_0^T [|R_t - u| 1_{\{R_t > u\}}] dt\Big]. \tag{4.41}$$

Using similar techniques as in Loisel and Trufin (2014), by Fubini's theorem we can rewrite (4.37) as follows:

$$\begin{aligned} \mathbb{E}[I_T^{R_t}(u)] &= \mathbb{E}\Big[\int_0^T [|R_t - u| 1_{\{R_t > u\}}] dt\Big], & (4.42)\\ &= \int_0^T \mathbb{E}[|R_t - u| 1_{\{R_t > u\}}] dt, & (4.43)\\ &= \int_0^T \mathbb{E}[(S_t - (1+\theta)\lambda\mu t - u)_+] dt. & (4.44) \end{aligned}$$

Now, we can define a new risk measure based on the risk process R_t, which is similar to the definition given by Loisel and Trufin (2014) for S_t.

Definition 4.2.5 (Limited expected area) *For a limited expected area A, the minimum initial capital needed to control the expected area less than A within $[0, T]$ is*

$$\rho_A^T[R_t] = \inf\{v \geq 0 \mid \mathbb{E}[I_T^R(v)] \leq A\}. \tag{4.45}$$

Now consider the effect of this small modification in the definition. In Loisel and Trufin (2014), the risk measure $\rho_{A,c}^T$ agrees with the stop loss order of the aggregate losses. The modified risk measure ρ_A^T should obviously agree with the stop loss order of the risk processes, which is

$$R_t^{(1)} \leq_{icx} R_t^{(2)}. \tag{4.46}$$

Under the stop loss order, one can get $\mathbb{E}[R_t^{(1)}] \leq \mathbb{E}[R_t^{(2)}]$, which is equivalent to $\lambda_1\mu_1 \geq \lambda_2\mu_2$. Moreover, once the inequality above is satisfied, we can focus on the stochastic order of the aggregrate losses, and derive the following result.

Proposition 4.2.6 *If $\mathbb{E}[S_t^{(1)}] \geq \mathbb{E}[S_t^{(2)}]$, the risk measure ρ_A^T agrees with the dilation order of the aggregate loss, that is $S_t^{(1)} \leq_{dil} S_t^{(2)}$ for all $t \in (0, T]$, then $\rho_A^T[R_t^{(1)}] \leq \rho_A^T[R_t^{(2)}]$.*

Proof. The proof is similar to that given by Belzunce *et al* (1997). According to the definition of the dilation order,

$$S_t^{(1)} \leq_{dil} S_t^{(2)} \;\Rightarrow\; S_t^{(1)} - \mathbb{E}[S_t^{(1)}] \leq_{cx} S_t^{(2)} - \mathbb{E}[S_t^{(1)}]. \tag{4.47}$$

Since $\mathbb{E}[S_t^{(1)}] \geq \mathbb{E}[S_t^{(2)}]$, then one can get that

$$S_t^{(1)} \leq_{cx} S_t^{(2)} + (\mathbb{E}[S_t^{(1)}] - \mathbb{E}[S_t^{(2)}]). \tag{4.48}$$

Further, considering the positive safety loading θ gives

$$S_t^{(1)} \leq_{cx} S_t^{(2)} + (1+\theta)(\mathbb{E}[S_t^{(1)}] - \mathbb{E}[S_t^{(2)}]), \tag{4.49}$$

which is exactly

$$S_t^{(1)} - (1+\theta)\mathbb{E}[S_t^{(1)}] \leq_{cx} S_t^{(2)} - (1+\theta)\mathbb{E}[S_t^{(2)}]. \tag{4.50}$$

Consequently, $R_t^{(1)} \leq_{icx} R_t^{(2)}$ and the result $\rho_A^T[R^{(1)}] \leq \rho_A^T[R^{(2)}]$ follows. □

Next, consider the analysis in Loisel and Trufin (2014) of the properties of their risk measure $\rho_{A,c}^T$ but for our modified risk measure ρ_A^T. For convenience, we adopt the notations used in Loisel and Trufin (2014), that is $R + a = \{R_t + a, t > 0\}$ and $aR = \{aR_t, t > 0\}$.

Property 4.2.1 *If a risk measure is defined for risk processes as in (4.45), then it possesses the following properties:*

1. *Monotonicity: if* $\mathbb{P}(R_t^{(1)} \leq R_t^{(2)}) = 1$ *for all* $t \in (0, T]$*, then* $\rho_A^T[R_t^{(1)}] \leq \rho_A^T[R_t^{(2)}]$.
2. *Translation invariance: for any positive constant* a*,* $\rho_A^T[R_t^{(1)} + a] = \rho_A^T[R_t^{(1)}] + a$.
3. *Positive homogeneity: for any positive constant* a*,* $\rho_A^T[aR_t^{(1)}] = a\rho_{\frac{A}{a}}^T[R_t^{(1)}]$.
4. *Subadditivity: for all* $\beta \in (0, 1)$*,* $\rho_A^T[R_t^{(1)} + R_t^{(2)}] \leq \rho_{\beta A}^T[R_t^{(1)}] + \rho_{(1-\beta)A}^T[R_t^{(2)}]$.

Proof. The proof is similar to that in Loisel and Trufin (2014). From the definition of the risk measure ρ_A^T, the first property is easily verified. To verify the second and the third properties, it is sufficient to note that

$$\int_0^T \mathbb{E}[(aS_t - a(1+\theta)\lambda\mu t - u)_+]dt = a\int_0^T \mathbb{E}[(S_t - (1+\theta)\lambda\mu t - \frac{u}{a})_+]dt,$$

$$\int_0^T \mathbb{E}[(S_t - (1+\theta)\lambda\mu t + a - u)_+]dt = a\int_0^T \mathbb{E}[(S_t - (1+\theta)\lambda\mu t - (u-a)_+]dt,$$

then by definition of ρ_A^T, the second and the third properties can be derived.

For the last property, we need to use the following inequality,

$$\begin{aligned}
&\int_0^T \mathbb{E}[(S_t^{(1)} + S_t^{(2)} - (1+\theta)(\lambda_1\mu_1 + \lambda_2\mu_2)t - u)_+]dt \\
\leq &\int_0^T \mathbb{E}[(S_t^{(1)} - (1+\theta)\lambda_1\mu_1 t + a - \beta u)_+]dt \\
&+ \int_0^T \mathbb{E}[(S_t^{(2)} - (1+\theta)\lambda_2\mu_2 t + a - (1-\beta)u)_+]dt, \qquad \beta \in (0,1).
\end{aligned}$$

Then, for $\beta \in (0, 1)$,

$$\mathbb{E}[I_T^{R_t^{(1)}+R_t^{(2)}}(\rho_{\beta A}^T[R_t^{(1)}] + \rho_{(1-\beta)A}^T[R_t^{(2)}])] \leq \mathbb{E}[I_T^{R_t^{(1)}}(\rho_{\beta A}^T[R_t^{(1)}])] + \mathbb{E}[I_T^{R_t^{(2)}}(\rho_{(1-\beta)A}^T[R_t^{(2)}])], \tag{4.51}$$

which implies $\rho_A^T[R_t^{(1)} + R_t^{(2)}] \leq \rho_{\beta A}^T[R_t^{(1)}] + \rho_{(1-\beta)A}^T[R_t^{(2)}]$ as needed. But one should note that the risk measures on the right hand side are no longer based on the same area A. □

To better understand the implementation of this risk measure, we study a specfic example. First, it is necessary to introduce a theorem in Loisel and Trufin (2014), which is derived from the differentation theorem introduced at the end of Chapter 1.

Theorem 4.2.1 *For the compound Poisson model with a positive safety loading, the expected area in red* $\mathbb{E}[I_{\infty,c}(u)]$ *and the ultimate ruin probability* $\psi(u)$ *are linked by the following equation:*

$$\frac{d^2}{du^2}\mathbb{E}[I_{\infty,c}(u)] = \frac{1}{c}\frac{\psi(u)}{1-\psi(0)}, \qquad u \geq 0. \tag{4.52}$$

Proof. See Loisel and Trufin (2014), mainly relies on the differentation theorem introduced in Loisel (2005). □

For the model discussed in this section, equation (4.52) should be modified to be

$$\frac{d^2}{du^2}\mathbb{E}[I_{\infty}(u)] = \frac{1}{(1+\theta)\lambda\mu}\frac{\psi(u)}{1-\psi(0)}, \qquad u > 0. \tag{4.53}$$

Example 4.2.3 Suppose that the claims severity follows an exponential distribution with mean β. From the introduction in Chapter 1 we know that the ruin probability has a closed form: $\psi(u) = (1-\beta R)e^{-Ru}$, where R is Lundberg's adjustment coefficient. After some computations, we get $R = \frac{\theta}{\beta(1+\theta)}$. Then by use of the differential equation in (4.52), we get the following result:

$$\mathbb{E}[I_{\infty}(u)] = \frac{\beta(1+\theta)}{\lambda\theta^3}\exp\{-\frac{\theta}{\beta(1+\theta)}u\}, \qquad u > 0. \tag{4.54}$$

Thereby, the risk measure $\rho_A^{\infty}(R_t^{(1)})$ can be written as $\rho_A^{\infty}(R_t^{(1)}) = \inf\{u \geq 0 \mid \mathbb{E}[I_{\infty}(u)] \leq A\}$, which is the following expression:

$$\rho_A^{\infty}[R_t^{(1)}] = \begin{cases} \frac{\beta(1+\theta)}{\theta}\ln\frac{\beta(1+\theta)}{A\lambda\theta^3}, & A < \frac{\beta(1+\theta)}{\lambda\theta^3}, \\ 0, & A \geq \frac{\beta(1+\theta)}{\lambda\theta^3}. \end{cases} \tag{4.55}$$

From (4.55), a basic analysis can be done: increasing intensity parameters λ lead to smaller measures of risk; increasing the parameter β, or increasing the expectation of the claim severity for this example, leads to a larger measure of risk.

Trufin and Mitric (2014) combines the ruin probability and the expected deficit at ruin to generate a new type of risk measure. Inspired by it, we argue that considering only one criterion may not fully reflect the dangerousness of the risk processes. Take the risk measure introduced in this section for example, if two risk processes have close expected areas in red, then we say that these two risk processes are undistinguishable. However, we know once the deficit reaches some level, the company may run out of credit and the severe insolvency will lead the company to bankruptcy. Therefore, we consider the combination of the expected area in red and the quantile of the maximal severity of ruin (see definition in (1.17)) to define a new kind of risk measure.

Recall the distribution of the maximal severity of ruin in (1.18) given the initial capital in Chapter 1, for convenience we continue the above example for further illustration. The claim severity now follows a exponential distribution with mean β, the ruin probability in this case is $\psi(u) = (1 - \beta R)e^{-Ru}$. Thus, the distribution of the maximal severity of ruin given the initial capital level is given by:

$$F_{L_M}(z) = \frac{\psi(u) - \psi(u+z)}{\psi(u)(1 - \psi(z))} = \frac{1 - e^{-Rz}}{1 - e^{-Rz} + \beta R e^{-Rz}}, \qquad z > 0. \tag{4.56}$$

Surprisingly, we can find that the simplified expression above is not related with the initial capital. Then the $p-$quantile of this distribution can be obtained:

$$VaR_p(L_M) = \frac{\beta(1+\theta)}{\theta} \ln \frac{(1 - p + \theta)}{(1-p)(1+\theta)}. \tag{4.57}$$

Finaly, we compute $VaR_p(L_M) + \rho_A^{\infty}(R_t^{(1)})$:

$$VaR_p(L_M) + \rho_A^{\infty}[R_t^{(1)}] = \frac{\beta(1+\theta)}{\theta} \ln \frac{\beta(1 + \frac{\theta}{1-p})}{A\lambda\theta^3}, \qquad A < \frac{\beta(1+\theta)}{\lambda\theta^3}. \tag{4.58}$$

The final expression of the new risk measure has a similar form to (4.55), which indicates the same results can be obtained based on this new risk measure.

To evaluate the insolvency risk, the information needed by the adminstrators differs from that of external regulators, i.e. the government. Thereby, different risk measures can be constructed using the idea illustrated above.

4.2.4 Risk Measures Based on the Expected Loss Ratio

From the analyses in Trufin *et al.* (2011) and Trufin and Mitric (2014), the risk processes are evaluated through VaR or TVaR measures of the aggregate loss $L = \sum_{i=1}^{M} D_i$, which was introduced in the Section 1.1. The variable L represents the real loss faced by the insurance company with regards to the insolvency. Hence, comparing the risks of different risk processes is equivalent to comparing the variables L in the sense of the stochastic order.

However, as seen in Chapter 1, under the assumptions of the classical risk model, the distribution of the aggregate loss L does not depend on the intensity parameter λ. This means that the risk of risk processes may not be distinguishable if their claim severity follows the same distribution. In this section, we construct a risk measure which takes into account the distribution of L and the intensity parameter λ.

Suppose that each shareholder of an insurance company only takes a limited liability, that is, if the insurance company runs into a bankruptcy the shareholders will not pay any losses which exceed the initial capital u. If L is the aggregate loss, then $(L-u)_+$ is the shareholders' deficit or insurer's default option. This quantity has been fully studied in recent years for the allocation problem. Sherris (2006) studies how to distribute the expected shareholders' deficit to each business line under a complete market condition. Kim and Hardy (2009) studies how to establish a new allocation principle on the basis of the expected shareholders' deficit. Dhaene *et al.* (2009) studies the effect of merging and diversifying the shareholders' deficit when some correlation orders are imposed. We will see in the next section that the allocation problem based on our new risk measure, defined below, can be perfectly solved by use of the unified principle proposed in Dhaene *et al.* (2010).

In fact, one could use directly the expected shareholders deficit (ESD) $\mathbb{E}[(L-u)_+]$ as a risk measure, but this measure does not depend on the intensity parameter either. Since the solvency of an insurance company heavily depends on its premium principle, we hence divide the expected shareholders' deficit by the premium charged by the insurance company. Under the assumptions of the classical model, the new quantity considered would be of the following loss ratio:

$$l(u) = \frac{\mathbb{E}[(L-u)_+]}{(1+\theta)\lambda\mu}, \tag{4.59}$$

where $\mu = \mathbb{E}[X]$ is the expectation of the claim severity. Here $l(u)$ is a scaled expected shareholders' deficit, and we call it the expected loss ratio (ELR) in this section. We can find that it is more likely that the insurance company will be able to repay its liabilities with a smaller $l(u)$. Thereby, $l(u)$ reflects the solvency of an insurance company.

To protect the shareholders's and shareholders alike, the insurance company will set sufficient reserves u in order to control $l(u)$ and keep it under a certain level. The needed reserve, or initial capital, based on this new criteria defines a new risk measure, based on the risk process with the following expression,

$$\rho_{ELR}[R_t] = \inf\{u \mid l(u) \leq \epsilon\}, \qquad \epsilon > 0. \tag{4.60}$$

As usual, we now analyze the basic properties of this expected loss ratio from the perspective of coherence. Consider two classical risk processes $R_t^{(1)} = \sum_{i=1}^{N_1(t)} X_i - (1+\theta)\lambda_1\mu_1$ and $R_t^{(2)} = \sum_{i=1}^{N_2(t)} Y_i - (1+\theta)\lambda_2\mu_2$, where $\{X_i\}_{i=1,2,\ldots}$ are positive i.i.d. random variables with distribution F_X, and $\{Y_i\}_{i=1,2,\ldots}$ are positive i.i.d. random variables with distribution F_Y.

Proposition 4.2.7 *The risk measure defined by (4.59) has the following properties:*

1. *If $\lambda_1 \geq \lambda_2$, $X_1 \leq_{cx} Y_1$, then $\rho_{ELR}[R_t^{(1)}] \leq \rho_{ELR}[R_t^{(2)}]$.*
2. *For any positive constant a, $\rho_{ELR}(aR_t^{(1)}) = a\rho_{ELR}[R_t^{(1)}]$.*
3. *For any positive constant a, $\rho_{ELR}(R_t^{(1)} + a) = \rho_{ELR}[R_t^{(1)}] + a$.*
4. *If $R_t^{(1)}$ and $R_t^{(2)}$ are indepedent, and $Y \leq_{cx} X$, then $\rho_{ELR}[R_t^{(1)} + R_t^{(2)}] \leq \rho_{ELR}[R_t^{(1)}] + \rho_{ELR}[R_t^{(2)}]$.*

Proof. 1. Using the results of Trufin *et al.* (2011), we can conclude that:

$$X \leq_{cx} Y \Rightarrow D^X \leq_{st} D^Y \Rightarrow D^X \leq_{sl} D^Y,$$

where D^X and D^Y are corresponding ladder heights. Since the stop-loss order is preserved under the compound sum, then $L_1 = \sum_{i=1}^{G_1} D_i^X \leq_{sl} L_2 = \sum_{i=1}^{G_2} D_i^Y$, which leads to $\mathbb{E}[(L_1 - u)_+] \leq \mathbb{E}[(L_2 - u)_+]$. The condition $X \leq_{cx} Y$ also implies that $\mathbb{E}[X] = \mathbb{E}[Y]$. Finally, since $\lambda_1 \geq \lambda_2$, the first property is verified.

2. If the risk process is multiplied by a positive constant a, then the claim severity will be aX_1. The corresponding ladder height D^{X*} follows the distribution

$$\begin{aligned} F_{D^{X*}}(y) &= \int_0^y \frac{1 - F_{aX}(x)}{\mathbb{E}[aX]} dx \\ &= \int_0^y \frac{1 - F_X(x/a)}{\mathbb{E}[X]} d(x/a) \\ &= \int_0^{\frac{y}{a}} \frac{1 - F_X(x)}{\mathbb{E}[X]} dx \\ &= F_{aD^X}(y). \end{aligned}$$

The derivation above shows that the $L' = aL$. Therefore

$$\begin{aligned} \rho_{ELR}[aR_t^{(1)}] &= \inf\left\{u \mid \frac{\mathbb{E}[(aL - u)_+]}{(1+\theta)\lambda_1\mathbb{E}[aX]} \leq \epsilon\right\} \\ &= \inf\left\{u \mid \frac{\mathbb{E}[(L - \frac{u}{a})_+]}{(1+\theta)\lambda_1\mathbb{E}[X]} \leq \epsilon\right\} \\ &= a\rho_{ELR}[R_t^{(1)}]. \end{aligned}$$

3. Note that adding a positive constant a on the risk process defined above, implies that the same constant gets added onto the aggregate loss L. Then the expected loss ratio for translated process is

$$\begin{aligned} \rho_{ELR}[R_t^{(1)} + a] &= \inf\left\{u \mid \frac{\mathbb{E}[(L + a - u)_+]}{(1+\theta)\lambda_1\mathbb{E}[X]}\right\} \\ &= \inf\left\{u \mid \frac{\mathbb{E}[(L + (u - a))_+]}{(1+\theta)\lambda_1\mathbb{E}[X]}\right\} \\ &= a + \rho_{ELR}[R_t^{(1)}]. \end{aligned}$$

4. If $R_t^{(1)}$ and $R_t^{(2)}$ are independent, then $R_t^{(3)} = R_t^{(1)} + R_t^{(2)}$ will be a new risk process with intensity parameter $\lambda_3 = \lambda_1 + \lambda_2$ and claims severity Z with the following mixed distribution

$$F_Z(x) = \frac{\lambda_1}{\lambda_1 + \lambda_2} F_X(x) + \frac{\lambda_2}{\lambda_1 + \lambda_2} F_Y(x), \qquad x > 0.$$

Using the definition, one can find the distribution of the ladder height D^Z:

$$F_{D^Z}(y) = \alpha F_{D^X}(y) + (1-\alpha) F_{D^Y}(y), \qquad \alpha = \frac{\lambda_1\mathbb{E}[X]}{\lambda_1\mathbb{E}[X] + \lambda_2\mathbb{E}[Y]}.$$

This means that the new ladder height D^Z also follows a mixed distribution. Since $Y \leq_{cx} X$, then $D^Y \leq_{st} D^X$, then it leads to

$$F_{D^Y}(y) \geq F_{D^X}(y) \Rightarrow F_{D^Z}(y) \geq F_{D^X}(y) \Rightarrow D^Z \leq_{st} D^X.$$

The first order of dominance can be preserved under the compound sum, hence we get the following ordering:

$$\sum_{i=1}^{M_3} D_i^Z \leq_{st} \sum_{i=1}^{M_3} D_i^X \leq_{cx} \sum_{i=1}^{M_1} D_i^X + \sum_{i=1}^{M_2} D_i^Y,$$

where M_1, M_2 and M_3 are i.i.d. geometric random variables. Therefore, the following ordering obviously holds.

$$\sum_{i=1}^{M_3} D_i^Z \leq_{sl} \sum_{i=1}^{M_1} D_i^X + \sum_{i=1}^{M_2} D_i^Y,$$

Let $L_3 = \sum_{i=1}^{M_3} D_i^Z$, then

$$\begin{aligned}
&\frac{\mathbb{E}[(L_3 - (\rho_{ELR}(R_t^{(1)}) + \rho_{ELR}(R_t^{(2)}))_+]}{(1+\theta)(\lambda_1\mathbb{E}[X] + \lambda_2\mathbb{E}[Y])} \leq \frac{\mathbb{E}[(L_1 + L_2 - (\rho_{ELR}(R_t^{(1)}) + \rho_{ELR}(R_t^{(2)})))_+]}{(1+\theta)(\lambda_1\mathbb{E}[X] + \lambda_2\mathbb{E}[Y])} \\
&\leq \frac{\mathbb{E}[(L_1 - \rho_{ELR}(R_t^{(1)}))_+]}{(1+\theta)(\lambda_1\mathbb{E}[X] + \lambda_2\mathbb{E}[Y])} + \frac{\mathbb{E}[(L_2 - \rho_{ELR}(R_t^{(2)}))_+]}{(1+\theta)(\lambda_1\mathbb{E}[X] + \lambda_2\mathbb{E}[Y])} \\
&= \alpha\frac{\mathbb{E}[(L_1 - \rho_{ELR}(R_t^{(1)}))_+]}{(1+\theta)\lambda_1\mathbb{E}[X]} + (1-\alpha)\frac{\mathbb{E}[(L_2 - \rho_{ELR}(R_t^{(2)}))_+]}{(1+\theta)\lambda_2\mathbb{E}[Y]} \\
&\leq \alpha\epsilon + (1-\alpha)\epsilon = \epsilon.
\end{aligned}$$

The inequality above implies that $\rho_{ELR}[R_t^{(1)} + R_t^{(2)}] \leq \rho_{ELR}[R_t^{(1)}] + \rho_{ELR}[R_t^{(2)}]$. □

Finally, we give an illustrative example. First note that the expected shareholders deficit $\mathbb{E}[(L-u)_+]$ can be rewritten as:

$$\mathbb{E}[(L-u)_+] = \int_u^\infty \psi(x)dx, \tag{4.61}$$

where $\psi(x)$ is the ruin probability with initial capital x. For convenience, we again let the claims severity follows the exponential distribution, because the ruin probability of this case is available analytically.

Example 4.2.4 If the claim severity $X_i \sim \text{Exp}(\frac{1}{\beta})$, then the ruin probability $\psi(x)$ is given as

$$\psi(x) = (1 - \beta R)e^{-Rx}, \qquad x > 0,$$

where R is the adjustment coefficient. Hence

$$\frac{\mathbb{E}[(L-u)_+]}{(1+\theta)\lambda \mathbb{E}[X]} = \frac{\int_u^{\infty}(1-\beta R)e^{-Rx}dx}{(1+\theta)\lambda\beta} = \frac{(1-\beta R)e^{-Ru}}{(1+\theta)\lambda\beta R}.$$

Thereby, the needed initial capital to control the expected loss ratio under level ϵ is

$$\rho_{ELR}[R_t] = \inf\left\{\frac{(1-\beta R)e^{-Ru}}{(1+\theta)\lambda\beta R} \leq \epsilon\right\} = \frac{1}{R}\ln\frac{1-\beta R}{(1+\theta)\lambda\beta R\epsilon}.$$

Since the adjustment coefficient $R = \frac{\theta}{\beta(1+\theta)}$ under the assumptions of the classical model, the final expression for the initial capital needed is

$$\rho_{ELR}[R_t] = \frac{\beta(1+\theta)}{\theta}\ln\frac{1}{(1+\theta)\theta\lambda\epsilon}.$$

Remark 4.2.1 If we let $\epsilon < \frac{1}{(1+\theta)\theta\lambda}$, then one can select $A = \frac{\beta(1+\theta)^2\epsilon}{\theta^2}$ in (4.55), and hence $\rho_{ELR}[R_t] = \rho_A^{\infty}[R_t]$, which means that the ELR risk measure is equivalent to the EAR risk measure in some special cases.

Chapter 5

Applications

In this chapter, we apply the risk measures introduced in the Chapter 4 to solve some insurance risk management problems. One is called the optimal allocation problem, which is intensively studied in the financial and actuarial literature. It contains two subproblems: the capital allocation problem and the risk allocation problem. Another interesting problem is the optimal reinsurance problem, which is more difficult to deal with, especially when multi-period decisions are involved. These problems are all particular cases of the optimization problem, which means some well known techniques can be applied.

5.1 Optimal Allocation Problem

The optimal allocation problem is always a central theme in applying a new risk measure. Shareholders and investors are concerned with the risk of their capital investment and the return it will generate. Usually, two subproblems are considered. In the capital allocation problem, it is assumed that the total amount of the initial capital is fixed, one needs to find a way to distribute this capital to the sub-lines of business in order to minimize the sum of their risks. The other subproblem is called the risk allocation problem, which is similar to the capital allocation problem, but from a different point of view. Assuming the total risks are fixed, the question is how to distribute them to the sublines of business in order to minimize the needed initial capital. Both problems have been discussed extensiely in recent years.

Many principles have been used to carry out solutions to the allocation problem. Dhaene *et al.* (2010) summarize various allocation principles which are popular in use, and find that they can be regarded as the proportional allocation principle below with different risk measures:

$$K_i = \frac{K}{\sum_{j=1}^{n} \rho[X_j]} \rho[X_i], \qquad i = 1, \dots, n, \tag{5.1}$$

where K is the total initial capital, and ρ is the risk measure. For convenience, denoting the sum of the risks as $S = \sum_{i=1}^{n} X_i$, and the correspnding comonotonic sum of the risks as $S^c = \sum_{i=1}^{n} F_{X_i}^{-1}(U)$, where $U \sim \text{Unif}(0,1)$, the popular risk measures ρ used in (5.1) can be summarized as follows:

$$\begin{aligned}
\text{Haircut allocation}: &\quad \rho[X_i] = F_{X_i}^{-1}(p), \qquad p \in (0,1), \\
\text{Quantile allocation}: &\quad \rho[X_i] = F_{X_i}^{-1}(F_{S^c}(K)), \\
\text{Covariance allocation}: &\quad \rho[X_i] = \mathbb{C}\text{ov}[X_i, S], \\
\text{CTE allocation}: &\quad \rho[X_i] = \mathbb{E}[X_i \mid S > F_S^{-1}(p)], \quad p \in (0,1).
\end{aligned}$$

However, allocating capital based on the proportions in (5.1) is debated, since there is no unique criterion. Therefore, Dhaene *et al.* (2010) propose the following criterion: minimizing the distance between the uncertain claims and the allocated capitals associated with the corresponding different business lines, given the total amount of initial capital. The corresponding allocation problem is transformed into the following optimization problem:

$$\min_{K_1,\dots,K_n} \qquad \sum_{i=1}^{n} v_i \mathbb{E}\Big[\xi_i \tilde{D}\Big(\frac{X_i - K_i}{v_i}\Big)\Big], \tag{5.2}$$

$$\text{s.t} \qquad \sum_{i=1}^{n} K_i = K. \tag{5.3}$$

In the optimization problem above, the v_i are measures of exposure or business volume, the ξ_i are random variables with $\mathbb{E}[\xi_i] = 1$, $\tilde{D}$ is the function that measures the distance between the uncertain loss and the allocated capital in each business line. Dhaene *et al.* (2010) solve problem (5.2) with the distance functions $\tilde{D}(x) = x^2$ and

$\tilde{D}(x) = |x|$. In their analysis, an important assumption is that the first two moments of the claims severity be finite. Their approach can be used to solve the allocation problem based on the aggregate loss $L = \sum_{i=1}^{M} D_i$.

Although many risk measures have been proposed for risk processes, for the optimal allocation problem, only those that are a function of the initial capital can be applied. The risk measures based on other characteristic quantities of the risk process, but not on the initial capital, cannot be applied to solve the above optimization problems.

5.2 Capital Allocation

In this section, we first see an example from Dhaene *et al.* (2003). Later we apply the same technique to study the optimal strategy for the allocation problem with different risk measures. A nice result is derived when using the expected area in the red.

Example 5.2.1 *(Risk Measure Based on the Premium Rate)*

The risk measure studied by Dhaene *et al.* (2003) introduced in Section 4.1.1 depends on the initial capital, since the adjustment coefficient is chosen as $R = \frac{1}{u} |\ln \epsilon|$. With this risk measure applied to the aggregate loss, Dhaene *et al.* (2003) consider the following allocation problem:

$$\min_{u_1,\ldots,u_n} \quad \sum_{i=1}^{n} \frac{u_i}{|\ln \epsilon|} \ln \mathbb{E}\Big[\exp\Big(\frac{|\ln \epsilon|}{u_i} X_i \Big) \Big], \tag{5.4}$$

$$\text{s.t} \quad \sum_{i=1}^{n} u_i = u. \tag{5.5}$$

The solution is not difficult to derive. To show the allocation strategy, it is necessary to introduce the exponential and the Esscher premium principles for X_i with parameter $\frac{|\ln \epsilon|}{u_i}$:

$$\text{Exponential premium} \quad : \quad \rho^i_{exp}(X_i) = \frac{u_i}{|\ln \epsilon|} \ln \mathbb{E}\Big[\exp\Big(\frac{|\ln \epsilon|}{u_i} X_i \Big) \Big], \tag{5.6}$$

$$\text{Esscher premium} \quad : \quad \rho^i_{Ess}(X_i) = \frac{\mathbb{E}\big[X_i e^{(|\ln \epsilon|/u_i) X_i} \big]}{\mathbb{E}\big[e^{(|\ln \epsilon|/u_i) X_i} \big]} \tag{5.7}$$

Using Lagrange multipliers, we can get the optimal allocation strategy: choosing u_i to be the solution of the following equation:

$$\frac{1}{u_j}\left[\rho^j_{exp}(X_j) - \rho^j_{Ess}(X_j)\right] = \frac{1}{u}\sum_{i=1}^{n}\left[\rho^i_{exp}(X_i) - \rho^i_{Ess}(X_i)\right]. \tag{5.8}$$

The equation above is somewhat complicated to solve, thereby some approximation techniques are proposed in Dhaene *et al.* (2003). Considering the cumulant generating function $K(t) = \ln M_X(t)$, and taking $t = \frac{|\ln \epsilon|}{u_i}$, we get

$$\rho^i_{exp}(X_i) = K'(t),$$
$$\rho^i_{Ess}(X_i) = \frac{1}{t}K(t).$$

Taylor's expansion for the cumulant generating function $K(t)$ is given by

$$K(t) = \ln M_X(t) = \mathbb{E}[X]t + \mathbb{V}[X]\frac{t^2}{2!} + O(t^3). \tag{5.9}$$

Henceforth, taking ϵ small enough, $t = \frac{|\ln \epsilon|}{u_i}$ will also be very small, making the error term negligible. Hence, the following equation can be regarded as the approximation to (5.8):

$$\frac{u_j}{u} = \frac{\mathbb{V}[X_j]/(2u_j)}{\sum_{i=1}^{n}\mathbb{V}[X_i]/(2u_i)}. \tag{5.10}$$

With this approximation in (5.10), the optimal allocation strategy will be easily to implement.

We see that, applying Lagrange multipliers is an efficient method to solve constrained optimization problems. Next, we consider another allocation problem but use the expected area in the red as the risk measure.

Example 5.2.2 *(Risk Measures Based on EAR)*
For the risk measure based on the expected area in the red, the new optimal allocation problem can be written in the following form:

$$\min_{u_1,\dots,u_n} \quad \sum_{i=1}^{n}\mathbb{E}[I^i_\infty(u_i)], \tag{5.11}$$

$$\text{s.t} \quad \sum_{i=1}^{n} u_i = u. \tag{5.12}$$

In this problem $\mathbb{E}[I^i_\infty(u_i)]$ represents the expected area in the red for the ith business line, or ith subcompany, given an initial capital u_i. With the differentation theorems derived in Loisel (2005), we still can use the Lagrange multiplier method to solve this problem. Let $f_i(u_i) = \mathbb{E}[I^i_\infty(u_i)]$, then the augmented Lagrange function is

$$\sum_{i=1}^{n} f_i(u_i) + \gamma(\sum_{i=1}^{n} u_i - u). \tag{5.13}$$

The necessary condition for reaching the optimal solution of this problem should be

$$f_i'(u_i) = -\mathbb{E}[\tau_i(u_i)] = -\gamma, \tag{5.14}$$

where $\mathbb{E}[\tau_i(u_i)]$, the expectation of the total time spent below zero, is defined in the Section 1.3.6. This quantity is fully studied by Egídio dos Reis (1993), which is introduced in the Section 1.3.4. With the help of the moment generating function derived by Egídio dos Reis (1993), the expected duration $\mathbb{E}[\tilde{\tau} \mid U_0]$ of the negative surplus is

$$\mathbb{E}[\tilde{\tau} \mid U_0] = \psi(U_0)(\mathbb{E}[\tau' - \tau| \, U_0] + \mathbb{E}[\tilde{\tau} \mid U_0 = 0]), \tag{5.15}$$

where $\mathbb{E}[\tau' - \tau \mid U_0]$ is the expected duration of the first recovery period, and Egídio dos Reis (1993) also gives out its expression:

$$\mathbb{E}[\tau' - \tau \mid U_0] = \frac{\mathbb{E}[|U_\tau| \mid U_0]}{c\phi(0)}. \tag{5.16}$$

The expression for $\mathbb{E}[|U_\tau| \mid U_0]$, the expected deficit at ruin time τ, is given by Truffin and Mitric (2014):

$$\mathbb{E}[|U_\tau| \mid U_0 = u] = \int_u^\infty \frac{\psi(x)}{\psi(u)} dx - \frac{\mathbb{E}[X^2]}{2\theta\mathbb{E}[X]}. \tag{5.17}$$

Combining all the equations above, we can get an explicit expression for $\mathbb{E}[\tau_i(u_i)]$:

$$\begin{aligned}
\mathbb{E}[\tau_i(u_i)] &= \mathbb{E}[\tilde{\tau} \mid U_0 = u_i] && (5.18)\\
&= \psi_i(u_i)\Big(\frac{\mathbb{E}[|U_\tau| \mid U_0]}{c\phi_i(0)} + \mathbb{E}[\tilde{\tau} \mid U_0 = 0]\Big) && (5.19)\\
&= \frac{1}{c_i\phi_i(0)} \int_{u_i}^\infty \psi_i(x)dx + \psi_i(u_i)\frac{\mathbb{E}[X^2]}{2c_i\phi_i(0)}\Big(\frac{\lambda_i}{c_i\phi_i(0)} - \frac{1}{\theta\mathbb{E}[X]}\Big) && (5.20)\\
&= \frac{1}{c_i\phi_i(0)} \int_{u_i}^\infty \psi_i(x)dx. && (5.21)
\end{aligned}$$

Since the safety loading is the same for all the business lines, so $\phi_i(0) = \frac{\theta}{1+\theta}$, which makes (5.14) simplify to

$$\frac{1}{c_i}\int_{u_i}^{\infty} \psi_i(x)dx = \frac{\theta}{1+\theta}\gamma. \tag{5.22}$$

With the introduction of the aggregate loss $L = \sum_{i=1}^{M} D_i$, where D_i is the ladder height, Trufin and Mitric (2014) have shown that the integral in the equation above can be rewritten as

$$\int_{u_i}^{\infty} \psi_i(x)dx = \mathbb{E}\big[(L_i - u_i)_+\big], \tag{5.23}$$

where L_i is the aggregate loss for the ith business line. Then we find the optimal allocation strategy, choosing $\{u_i, i = 1, \dots, n\}$ which satisfy the following equations:

$$\frac{1}{c_i}\mathbb{E}\big[(L_i - u_i)_+\big] = \cdots = \frac{1}{c_n}\mathbb{E}\big[(L_n - u_n)_+\big], \tag{5.24}$$

$$\sum_{i=1}^{n} u_i = u. \tag{5.25}$$

In reality, the decision-makers should care more about the aggregate loss $L_j = \sum_{i=1}^{M} D_{ji}$, because each ladder D_{ji} represents the real loss the company suffers from the jth business line. To avoid insolvency, each business line should have enough reserves to offset L_j. This generates a new allocation problem, which is explained in the next example.

Example 5.2.3 *(Risk Measures based on the ESD)*

Denoting the aggregate loss of the ith subcompany by L_i, consider the following problem:

$$\min_{u_1,\dots,u_n} \quad \sum_{i=1}^{n} \mathbb{E}\big[(L_i - u_i)_+\big], \tag{5.26}$$

$$\text{s.t} \quad \sum_{i=1}^{n} u_i = u. \tag{5.27}$$

From the introduction in Section 1.1, we know the aggregate loss L_i follows a compound geometric distribution. Some articles deal with this problem, with simpler loss variables. The expectation $\mathbb{E}[(L - u)_+]$ is deeply studied by Sherris (2006). Based on this value, Kim and Hardy (2009) establish a new strategy for capital allocation.

For the convenience of a later illustration, we now introduce the following new notation:

$$F_X^{-1}(p) = \inf\{x \in \mathbb{R} \mid F_X(x) \geq p\}, \qquad p \in [0,1], \tag{5.28}$$

$$F_X^{-1+}(p) = \sup\{x \in \mathbb{R} \mid F_X(x) \leq p\}, \qquad p \in [0,1], \tag{5.29}$$

$$F_{X,\alpha}^{-1}(p) = \alpha F_X^{-1}(p) + (1-\alpha) F_X^{-1+}(p), \qquad p \in (0,1) \quad \alpha \in [0,1]. \tag{5.30}$$

Dhaene *et al.* (2010) discuss this problem as a particular case of the optimization problem (5.2)-(5.3), and get the following allocation strategy which can help solve the problem in this example.

Theorem 5.2.1 *Assuming that* $F_{S^c}^{-1+}(0) < u < F_{S^c}^{-1}(1)$*, the optimal allocation problem*

$$\min_{u_1,\ldots,u_n} \quad \sum_{i=1}^n \mathbb{E}\big[(X_i - u_i)_+\big], \qquad \text{s.t} \quad \sum_{i=1}^n u_i = u, \tag{5.31}$$

has the following solution:

$$u_i = F_{X_i,\alpha}^{-1}\big(F_{S^c}(u)\big), \qquad i = 1, \ldots, n, \tag{5.32}$$

where S^c*, the sum of the comonotonic risks, is defined after (5.1) and* α *is determined by the following equation:*

$$u = F_{S^c,\alpha}^{-1}\big(F_{S^c}(u)\big). \tag{5.33}$$

Proof. See Dhaene *et al.* (2010). □

Replacing X_i by L_i, if the condition $F_{S^c}^{-1+}(0) < u < F_{S^c}^{-1}(1)$ still holds, then it indicates that this theorem provides a allocation strategy for the problem in (5.26)-(5.27).

The last example we study here is similar to Example 5.1.3, the difference being that the objective function is now of the form (4.59). Dhaene *et al.* (2010) also provide a strategy for this more general problem.

Example 5.2.4 *(Risk Measure based on ELR)*

If the objective function is of the form (4.59), then we should consider the following optimization problem:

$$\min_{u_1,\dots,u_n} \quad \sum_{i=1}^{n} \frac{\mathbb{E}[(L_i - u_i)_+]}{(1+\theta)\lambda_i\mu_i}, \tag{5.34}$$

$$\text{s.t} \quad \sum_{i=1}^{n} u_i = u. \tag{5.35}$$

This problem is also a particular case of the problem (5.2)-(5.3). To avoid tedious notation, we replace $(1+\theta)\lambda_i\mu_i$ by c_i, which is the premium rate charged for the ith line of business. Since now we have different weights $\frac{1}{c_i}$ in front of each expected shareholders deficit $\mathbb{E}\Big[(L_i - u)_+\Big]$, the method introduced in Dhaene *et al.* (2010) can not be applied. We use Lanrange multipliers to see what strategy should it be, and also if the explicit solution can be found.

Let $g_i(u_i) = \frac{\mathbb{E}[(L_i - u_i)_+]}{c_i}$, the augmented function in this example is

$$\sum_{i=1}^{n} g_i(u_i) + \gamma(\sum_{i=1}^{n} u_i - u). \tag{5.36}$$

Taking derivative of the function (5.36) with respect to each u_i and γ, and equate the derivatives with zeros, we can get the following equations:

$$F_{L_i}(u_i) = 1 - c_i\gamma, \qquad i = 1, 2, \dots, n, \tag{5.37}$$

$$\textstyle\sum_{i=1}^{n} u_i = u. \tag{5.38}$$

To find each u_i, the key point is to find suitable γ which satisfies equations (5.37)-(5.38), which leads to the next equation about γ:

$$\sum_{i=1}^{n} F_{L_i}^{-1}(1 - c_i\gamma) = u. \tag{5.39}$$

Note that the left hand side of equation (5.39) is a decreasing function of γ, some numerical methods can be applied here to find γ, like bisetion algorithm. Once γ is found, each u_i is determined by the equations in (5.37).

Conclusion

This thesis studies various risk measures to be applied on the risk processes of insurance companies. The models and results derived in this thesis are based on the assumptions of the Cramér-Lundberg classical risk model. In Chapter 1 we summarize the results and recent developments in risk theory, to first better understand the characteristics of risk processes, that can serve in risk management. A brief introduction of the basic theory of risk measures is also given, which is later used to establish the right criteria in evaluating and selecting the methods for our proposed risk measures.

Risk measures based on risk processes are designed to evaluate the uncertainty in the Cramér-Lundberg model. Chapter 2 lists the methods found in the literature and analyze their strengths and weaknesses. None of these risk measures reflects the influence of all the risk characteristics of risk processes. In Chapter 4 we propose extensions by modifying the conditions in the definitions of these measures. At the same time, we try to consider new ways to create risk measures, based on the other risk characteristics such as safety loading, the adjustment coefficient and the expected loss ratio. Some desirable properties of the classical risk measures are preserved and simple examples are provided. Even though not all of our new risk measures can capture the overall characteristics of the risk processes, easy implementation may make them efficient from a practical viewpoint.

In the last chapter, we apply the risk measures proposed in Chapter 4 to solve the optimal allocation problem. Some nice forms of the optimal strategy are obtained which help reveal connections between different risk measures.

Bibliography

[1] Acerbi, C. (2002) *Spectral measures of risk: a coherent representation of subjective risk aversion.* Journal of Banking & Finance, **Vol. 2002(26)**: 1505-1518.

[2] Ahmadi-Javid, A. (2012) *Entropic Value-at-Risk: a new coherent risk measure.* Journal of Optimization Theory and Applications, **Vol. 2012(155)**: 1105-1123.

[3] Albrecher, H. and Boxma, O.J. (2004) *A ruin model with dependence between claim sizes and claim intervals.* Insurance: Mathematics and Economics, **Vol. 35(2)**: 245-254.

[4] Andersen, E.S. (1957) *On the collective theory of risk in case of contagion between claims.* Bulletin of the Institute of Mathematics and its Applications, **Vol. 12**: 275-279.

[5] Artzner, P., Delbaen, F., Eber, J.M. and Heath, D. (1999) *Coherent measures of risk.* Mathematical Finance, **Vol. 9(3)**: 203-228.

[6] Asmussen, S. (2000) *Ruin Probability.* World Science, River Edge, N.J.

[7] Balbás, A., Garrido, J. and Mayoral, S. (2009) *Properties of distortion risk measures.* Methodology and Computing in Applied Probability, **Vol. 11(3)**: 385-399.

[8] Balbás, A., Blanco, I. and Garrido J. (2014) *Measuring risk when expected losses are unbounded.* Risks, **Vol. 2014(4)**: 411-424.

[9] Belzunce, F., Pellerey, F., Ruiz, J.M. and Shaked, M. (1997) *The dilation order, the dispersion order, and orderings of residual lives.* Statistics & Probability Letters, **Vol. 33(3)**: 263-275.

[10] Boudreault, M., Cossette, H., Landriault, D. and Marceau. E. (2006) *On a risk model with dependence between interclaim arrivals and claim sizes*. Scandinavian Actuarial Journal, **Vol. 2006(5)**: 265-285.

[11] Daníelsson, J., Jorgensen, B.N., Samorodnitsky, G., Sarama, M. and de Vries, C.G. (2005) *Subadditivity re-examined: the case for Value-at-Risk*. Working paper, London School of Economics.

[12] Delbaen, F. (2002) *Coherent risk measures on general probability spaces*. Advances in Finance and Stochastics, 1-37.

[13] Denuit, M., Dhaene, J., Goovaerts, M. and Kaas, R. (2005) *Actuarial Theory for Dependent Risks: Measures, Orders and Models*. Wiley, Chichester, N.Y.

[14] Denuit, M., Dhaene, J., Goovaerts, M. and Kass, R. (2006) *Risk measurement with equivalent uitlity principles*. Statistics & Decisions, **Vol. 2006(24)**: 1-25.

[15] Dhaene, J., Goovaerts, M. and Kass, R. (2003) *Economic capital allocation derived from risk measures*. North American Actuarial Journal, **Vol. 7(2)**: 44-56.

[16] Dhaene, J., Denuit, M. and Vanduffel, S. (2009) *Correlation order, merging and diversification*. Insurance: Mathematics and Economics, **Vol. 45(3)**: 325-332.

[17] Dhaene, J., Tsanakas, A., Valdez, E.A. and Vanduffel,S. (2010) *Optimal capital allocation principle*. 9th International Congress on Insurance: Mathematics and Economics, Laval.

[18] Dickson, C.M. (1997) *On a class of renewal risk processes*. North American Actuarial Journal, **Vol. 2(3)**: 60-68.

[19] Dickson, C.M. and Hipp, C. (2001) *On the time to ruin for Erlang(2) risk processes*. Insurance: Mathematics and Economics, **Vol. 29(3)**: 333-344.

[20] Dickson, C.M., Hughes, B.D. and Zhang, L.Z. (2005) *The density of the time to ruin for a Sparre Andersen process with Erlang arrivals and exponential claims*. Scandinavian Actuarial Journal, **Vol. 2005(5)**: 358-376.

[21] Dickson, C.M. and Waters, H.R. (2002) *The distribution of the time to ruin in the classical risk model*. ASTIN Bulletin, **Vol. 32(2)**: 299-313.

[22] Dickson, C.M. and Willmot, G.E. (2005) *The density of the time to ruin in the classical Poisson risk model*. ASTIN Bulletin, **Vol. 35(1)**: 45-60.

[23] Dufresne, F. and Gerber, H.U. (1988) *The surpluses immediately before and at ruin, and the amount of the claim causing ruin*. Insurance: Mathematics and Economics, **Vol. 7(3)**: 193-199.

[24] Dufresne, F. and Gerber, H.U. (1991) *Risk theory for the compound Poisson process that is perturbed by diffusion*. Insurance: Mathematics and Economics, **Vol. 10(1)**: 51-59.

[25] Egídio dos Reis, A. (1993) *How long is the surplus below zero*. Insurance: Mathematics and Economics, **Vol. 12(1)**: 23-38.

[26] Embrechts, P., Klüppelberg, C. and Mikosch, T. (1997) *Modelling Extremal Events*. Springer, Berlin.

[27] Fagiuoli, E., Pellerey, F. and Shaked, M. (1999) *A characterization of the dilation order and its applications*. Statistical Papers, **Vol. 40(4)**: 393-406.

[28] Firouzi, H.O. and Luong, A. (1999) *Optimal Portfolio Problem Using Entropic Value at Risk: When the Underlying Distribution is Non-Elliptical*. Working paper, University of Montreal.

[29] Föllmer, H. and Schied, A. (2001) *Convex measures of risk and trading constraints*. Finance and Stochastics, **Vol. 6(4)**: 429-447.

[30] Gerber, H.U. (1990) *When does the surplus reach a given target*. Insurance: Mathematics and Economics, **Vol. 9(2-3)**: 115119.

[31] Gerber, H.U., Goovaerts, M.J. and Kass, R. (1987) *On the probability and severity of ruin*. ASTIN Bulletin, **Vol. 17(2)**: 151-163.

[32] Gerber, H.U. and Shiu, S.W. (1998) *On the time value of ruin.* North American Actuarial Journal, **Vol. 2(1)**: 48-72.

[33] Grandell, J. (2001) *Aspects of Risk Theory.* Springer, New York.

[34] Hald, M. and Schmidli, H. (2004) *On the maximisation of the adjustment coefficient under proportional reinsurance.* ASTIN Bulletin, **Vol. 34(1)**: 75-83.

[35] Jang, J.W. (2004) *Martingale approach for moments of discounted aggregate claims.* The Journal of Risk and Insurance, **Vol. 71(2)**: 201-211.

[36] Kahneman, D. and Tversky, A. (1979) *Prospect theory: an analysis of decision under risk.* Econometrica, **Vol. 47(2)**: 263-292.

[37] Karlin, S. and Taylor, H.M. (1975) *A First Course in Stochastic Processes.* Academic Press, Burlington, N.Y.

[38] Kass, R., Dhaene, J. and Goovaerts, M. (2000) *Upper and lower bounds for sums of random variables.* Insurance: Mathematics and Economics, **Vol. 27(2)**: 151-168.

[39] Kim, H.T. and Hardy, M.R. (2009) *A capital allocation based on a solvency exchange option.* Insurance: Mathematics and Economics, **Vol. 44(3)**: 357-366.

[40] Klugman, S.A., Panjer, H.H. and Willmot, G. (2008) *Loss Model: From Data to Decisions.* Wiley, Hoboken, N.J.

[41] Kochar, S. (2011) *Lorenz ordering of order statistics.* Statistics & Probability Letters, **Vol. 2006(76)**: 1855-1860.

[42] Kolkovska, E.T. (2011) *Risk measures for classical and perturbed risk processes.* Pliska Stud. Math. Bulgar., **Vol. 2011(20)**: 121-134.

[43] Kou, S., Xinghua, P., Heyde, C.C. (2010) *External risk measures and Basel accords.* Mathematics of Operational Research, **Vol. 38(3)**: 393-417.

[44] Kusuoka, S. (2001). *On law invariant coherent risk measures.* Advances in Mathematical Economics, **Vol. 3**: 8395.

[45] Léveillé, G. and Garrido, J. (2001a) *Moments of compound renewal sums with discounted claims.* Insurance: Mathematics and Economics, **Vol. 28(2)**: 217-231.

[46] Léveillé, G. and Garrido, J. (2001b) *Recursive moments of compound renewal sums with discounted claims.* Scandinavian Actuarial Journal, **Vol. 2001(2)**: 98-110.

[47] Léveillé, G., Garrido, J. and Wang, Y.F. (2010) *Moment generating function of compound renewal sums with discounted claims.* Scandinavian Actuarial Journal, **Vol. 2010(3)**: 165-184.

[48] Li, S. and Garrido, J. (2004) *On ruin for the Erlang(n) risk process.* Insurance: Mathematics and Economics, **Vol. 34(3)**: 391-408.

[49] Lin, X.S. and Willmot, G.E. (1999) *Analysis of a defective renewal equation arising in ruin theory.* Insurance: Mathematics and Economics, **Vol. 25(1)**: 6384.

[50] Loisel, S. (2005) *Differentiation of some functionals of risk processes, and optimal reserve allocation.* Journal of Applied Probability, **Vol. 42(2)**: 379-392.

[51] Loisel, S. and Trufin, J. (2014) *Properties of a risk measure derived from the expected area in red.* Insurance: Mathematics and Economics, **Vol. 55**: 191-199.

[52] Michel, R. (1987) *A partial ordering of claim amount distributions and its relation to ruin probabilities in the Poisson model.* Bulletin of the Swiss Association of Actuaries, **1987**: 75-80.

[53] Picard, P. (1994) *On some measures of the severity of ruin in the classical Poisson model.* Insurance: Mathematics and Economics, **Vol. 14(2)**: 107-115.

[54] Pichler, A. (2013) *Insurance pricing under ambiguity.* European Actuarial Journal, **Vol. 2014(4)**: 335-364.

[55] Pichler, A. (2014) *Premiums and reserves, adjusted by distortions*. Scandinavian Actuarial Journal, **Vol. 2015(4)**: 332-351.

[56] Quiggin, J. (1982) *A theory of anticipated utility*. Journal of Economic Behavior and Organization, **Vol. 1982(3)**: 323-343.

[57] Reinhard, J.M. (1984) *On a class of semi-Markov risk models obtained as classical risk models in a Markovian environment*. ASTIN Bulletin, **Vol. 14(1)**: 23-43.

[58] Rockafeller, R.T. (1970) *Convex Analysis*. Princeton University Press, Princeton, N.J.

[59] Ross, S.M. (2010) *Introduction to Probability Models*. Academic Press, Burlington, N.Y.

[60] Royden, H. and Fitzpatrick, P. (2010) *Real Analysis*. Pearson, New York.

[61] Schmidli, H. (1995) *Cramér-Lundberg approximations for ruin probabilities of risk processes perturbed by diffusion*. Insurance: Mathematics and Economics, **Vol. 16(2)**: 135-149.

[62] Shaked, M. and Shanthikumar, J.G. (2007) *Stochastic Orders*. Springer, New York.

[63] Sherris, M. (2006) *Solvency, capital allocation, and fair rate of return in insurance*. The Journal of Risk and Insurance, **Vol. 73(1)**: 71-96.

[64] Thorin, O. (1974) *On the asymptotic behavior of the ruin probability for an infinite period when the epochs of claims form a renewal process*. Scandinavian Actuarial Journal, **Vol. 1974(2)**: 81-99.

[65] Thorin, O. (1982) *Probability of ruin*. Scandinavian Actuarial Journal, **Vol. 1982(2)**: 65-103.

[66] Tijms, H.C. (1994) *Stochastic Models: An Algorithmic Approach*. Wiley, New York.

[67] Trufin, J., Albrecher, H. and Denuit, M. (2011) *Properties of a risk measure derived from ruin theory*. The Geneva Risk and Insurance Review, **Vol. 2011(36)**: 174-188.

[68] Trufin, J. and Mitric, L. (2014) *On a risk measure inspired from the ruin probability and the expected deficit at ruin*. 3rd Workshop on Insurance Mathematics, Université Laval.

[69] Wang, S.S. (1995) *Insurance pricing and increased limits ratemaking by propotional hazards transforms*. Insurance: Mathematics and Economics, **Vol. 17(1)**: 43-54.

[70] Wang, S.S. (2000) *A class of distortion operators for pricing financial and insurance risks*. Journal of Risk and Insurance, **Vol. 67(1)**: 15-36.

[71] Wang, S.S. and Dhaene, J. (1998) *Comonotonicity, correlation order and premium principles*. Insurance: Mathematics and Economics, **Vol. 22(3)**: 235-242.

[72] Wang, S.S., Young, V.R. and Panjer, H.H. (1997) *Axiomatic characterization of insurance prices*. Insurance: Mathematics and Economics, **Vol. 21(2)**: 173-183.

[73] Willmot, G.E. and Lin, X.S. (1998) *Exact and approximation properties of the distribution of surplus before and after ruin*. Insurance: Mathematics and Economics, **Vol. 23(1)**: 91-110.

[74] Willmot, G.E. (2000) *On evaluation of the conditional distribution of the deficit at the time of ruin*. Scandinavian Actuarial Journal, **Vol. 2000(1)**: 63-79.

[75] Willmot, G.E. and Woo, J.K. (2007) *On the class of Erlang mixtures with risk theoretic applications*. North American Actuarial Journal, **Vol. 11(2)**: 99-115.

[76] Yaari, M.E. (1987) *The dual theory of choice under risk*. Econometrica, **Vol. 55(1)**: 95-115.

[77] Zhou, M. and Cai, J. (2009) *A perturbed risk model with dependence between premium rates and claim sizes*. Insurance: Mathematics and Economics, **Vol. 45(3)**: 382-392.