August 2024

CTE Induced Premium Principles and Properties

Linjiao Wu
University of Wisconsin-Milwaukee

CTE INDUCED PREMIUM PRINCIPLES AND PROPERTIES

by

Linjiao Wu

A Dissertation Submitted in

Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

in Mathematics

at

The University of Wisconsin–Milwaukee

August 2024

ABSTRACT

CTE INDUCED PREMIUM PRINCIPLES AND PROPERTIES

by

Linjiao Wu

The University of Wisconsin-Milwaukee, 2024
Under the Supervision of Professor Wei Wei and Professor Vytaras Brazauskas

The traditional pricing approach in the insurance industry assumes independence among insureds, yet overlooks the complexities of interdependent risk profiles. This dissertation addresses this limitation by proposing a premium pricing model tailored for managing dependent risks, drawing inspiration from conditional tail expectation (CTE) theory. In our model, each individual insured's premium is contingent upon the collective loss surpassing a predefined threshold.

To validate the efficacy of our model, we introduce several key properties to ensure fairness and stability in premium determination among insured individuals, including diversification and monotonicity. Diversification ensures that adding one policyholder to the insured group does not unjustly increase the premiums of others, while monotonicity ensures that others' premiums do not increase due to the increased riskiness of individual policyholders.

We analyze these properties under various distributional assumptions, such as normal, exponential, and Pareto distributions. By establishing the explicit CTE-induced premium and conducting comprehensive parameter analyses and simulations, we investigate the pricing dynamics under different scenarios, demonstrating the robustness and efficacy of our model.

In conclusion, this study emphasizes the importance of integrating nuanced risk dependencies into insurance pricing models. Our proposed model, rooted in conditional tail expectation theory, not only enhances risk management capabilities but also facilitates more equitable premium determination, thereby enhancing the resilience and stability of the insurance sector. This research lays the groundwork for broader adoption in various real-world applications.

To my parents

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

I extend my deepest appreciation to Professor Vytaras Brazauskas and Professor Wei Wei, my dissertation advisors, for their patient guidance, constant encouragement, and invaluable critiques of this research. Your mentorship in both academic pursuits and career development has been invaluable to me.

I am also grateful to Professor Chao Zhu, Professor Chudamani Poudyal, and Professor Lei Wang for their insightful comments and suggestions as members of my dissertation committee. Your contributions made my defense a truly rewarding experience.

Special thanks to Professor Yang Ho for introducing me to the captivating field of actuarial science. I am especially grateful to Mohammed Adjieteh and all my peers in the actuarial science program for sharing their experiences in both study and work.

My heartfelt gratitude goes to my family—my parents, grandparents, and my husband—for their unwavering support and love. Your faith in me has been a constant source of strength, even when I doubted myself.

Linjiao Wu
June 2024

Chapter 1

Introduction

1.1 Background and Motivation

In today's insurance industry, the occurrence of dependent risks is becoming increasingly widespread, underscoring the necessity of establishing a robust premium pricing framework to manage such risks. One prominent example of a dependent risk is the escalating threat of cyberattacks. According to a 2022 report on the website of the IT security software provider Check Point, the frequency of attacks peaked at 92 per week per organization globally by the end of 2021—a 50% increase compared to 2020. These attacks manifest in various forms such as ransomware attacks, identity theft, and data breaches. Importantly, victims of cyberattacks often share similar organizational traits or vulnerabilities, leading to positive dependency among the risks they face. According to a 2024 report on the website of the digital insurance brokerage Embroker, financial losses—including extortion demands and legal fees from ensuing lawsuits—are expected to grow at a rate of 15% yearly and reach an estimated 10.5 trillion dollars annually by 2025.

Another category of dependent risks is natural disasters. Events like tornadoes and floods typically result in collective losses for individuals residing in affected areas, leading to geographical perspective interdependence. Due to the difficulty of catastrophic forecasts and the complexity of disaster-related loss validation, not only the direct asset losses but also the subsequent reconstruction costs, decreased consumption, and the weakened capital markets together contribute to the underdevelopment of the insurance market according to Kusuma et al. (2019). The previous statement sheds light on the challenges of insuring dependent risks from sociological and economic perspectives. Acknowledging that diverse disciplines offer distinct lenses to analyze real-world phenomena, a mathematical expression of the difficulty will be the deficiency of the tools to effectively mitigate insolvency risk under dependent risk assumptions.

Before discussing the challenges of managing insolvency risk under dependent risk assumptions, it is crucial to clarify the definition of insolvency risk and why traditional insurance premium pricing frameworks have been able to address it under independent risk assumptions. Insolvency risk refers to the possibility that a company may be unable to meet its payment obligations. In the insurance industry, it signifies the possibility that an insurance company cannot cover losses with the collected premiums. According to Loss Data Analytics, an open text authored by the Actuarial Community, this concept can be formalized using the following accounting formula:

$$\text{premium} = \text{loss} + \text{expense} + \text{underwriting profit}.$$

Setting aside insurance company expenses, insolvency risk can be interpreted as the probability that the expected loss exceeds the expected premium. Assuming a loss X with distribution function $F(X)$, and given a premium on this loss, the insolvency risk is represented by $\mathbb{P}(X > \mathbb{E}(X))$, according to Young (2014). Insurers have always prioritized limiting insolvency risk when determining premium pricing.

When pricing a portfolio under independent risk assumptions, consisting of n losses $X_1, ..., X_n$, the total premium is the sum of premiums from all policyholders. The insolvency risk then becomes $\mathbb{P}(\sum_{i=1}^n X_i > \mathbb{E}[\sum_{i=1}^n X_i])$. According to the law of large numbers, as the portfolio size increases, this probability becomes less uncertain. Moreover, by the central limit theorem, $\mathbb{E}[\sum_{i=1}^n X_i]$ becomes more predictable as n increases. Overall, insolvency risk can be effectively managed under independent risk assumptions.

However, the scenario changes significantly under dependent risk assumptions. An extreme example is a portfolio consisting of identical losses, $n * X$. In this situation, insolvency risk retains its original definition, and managing it does not benefit from the law of large numbers and central limit theorem, potentially resulting in failure.

While individual perspective premium pricing may falter under dependent risk assumptions, it is important to note that insolvency can still be effectively managed by shifting our approach to the problem from pricing premiums individually to pricing premiums for groups. This pricing framework operates akin to the portfolio percentile premium principle, a concept originating from the field of life contingencies and long practiced (Dickson et al., 2019).

In actuarial sciences, the percentile is often referred to as the Value-at-Risk (VaR). For a portfolio comprising n individual losses $X_1, ..., X_n$, the aggregate loss is defined as:

$$S = \sum_{i=1}^{n} X_i$$

The formal definition of VaR is given by:

$$VaR_\alpha[S] = \inf\{x : \mathbb{P}\{S \leq x\} \geq \alpha\}$$

In this context, the symbol α represents a predetermined level of risk, guarantee that the insolvency probability is controlled under the level of $1-\alpha$. The value of α is always close to 1, such as 0.9 or 0.95. Throughout this dissertation, we will refer to the portfolio percentile premium as VaR premium to differentiate it from its application in the life contingency field. The calculation of the VaR premium hinges upon the assumption of independent risks and the prediction of aggregate loss through the application of the central limit theorem. The main distinction is that the VaR premium establishes a specified probability $1-\alpha$ of the aggregate loss and subsequently determines the premium using the inverse cumulative distribution function.

A more conservative alternative to the VaR premium is the Tail Value-at-Risk (TVaR) premium, defined as follows:

$$TVaR_\alpha[S] = \frac{1}{1-\alpha}\int_\alpha^1 VaR_\gamma(S)d\gamma = \frac{1}{1-\alpha}\int_{F^{-1}(\alpha)}^{\infty} sf(s)ds$$

Under the same risk level α, the TVaR premium demands a higher amount, making this pricing framework a safer option than VaR. When the risk variable is assumed to be continuous, TVaR is equivalent to Conditional Tail Expectation (CTE), defined as:

$$CTE_\alpha[S] = \mathbb{E}[S|S > VaR_\alpha[S]] \tag{1.1}$$

The focus of this dissertation is on CTE premium, which represents the variable's expectation given that it exceeds a certain percentile. When the variable represents losses, CTE signifies the average of the worst-case scenarios. This premium principle plays a crucial role in managing insolvency risk under the assumption of risk interdependence, primarily due to two advantages: the higher premium amount requirement and the coherent property. The latter provides the foundation of the CTE premium allocation and will be discussed in detail in the next section.

After determining the group premium, the subsequent consideration involves the fair and efficient allocation of premiums to each policyholder. Drawing inspiration from capital allocation principles (Dhaene et al., 2012), insurers aim to allocate premiums within a portfolio proportionally to each policyholder's risk level. In this context, the total risk capital (K) represents the sum of all individual risk capitals ($K_1, ..., K_n$) in the portfolio. The primary objective is to ensure that the total capital (K) is allocated among policyholders ($K_1, ..., K_n$) in a manner that satisfies the full allocation requirement:

$$\sum_{i=1}^{n} K_i = K$$

Various allocation principles exist, with the proportional allocation principle being the most general. It is defined as follows:

$$K_i = \frac{K}{\sum_{j=1}^{n} \rho[X_j]} \rho[X_i], \quad i = 1, ..., n.$$

Here, ρ represents the chosen risk measure, and K_i denotes the capital allocation to each unit i. Different allocation principles arise from selecting distinct risk measures. Some well-known examples include:

- The Conditional Tail Expectation (CTE) allocation principle, which closely relates to this paper. CTE, akin to TVaR premium, follows a similar structure:

$$K_i = \frac{K}{CTE_\alpha[S]} \mathbb{E}[X_i | S > F_S^{-1}(1 - \alpha)] \tag{1.2}$$

  This rule acknowledges risk interdependence, where units with higher expected losses are allocated greater premiums, particularly in situations where total loss is high.

- The covariance allocation principle also considers dependence, formulated as:

$$K_i = \frac{K}{Var[S]} Cov[X_i, S]$$

  Insurers with risk highly correlated with the aggregate risk will necessitate larger premiums.

Additionally, the haircut and quantile allocation principles utilize VaR and quantiles as risk measures, respectively. However, these methods may underestimate the effects of interdependence between risks, potentially leading to allocated capitals exceeding standalone capitals.

The premium pricing framework proposed in this paper combines the TVaR group premium and the CTE allocation principle.

1.2 Desirable Properties of Multivariate Premium Principle

As discussed in the previous section, premium allocation is closely tied to the underlying risk measure, with the risk measure often serving as the basis for the premium principle itself. In this section, we draw upon insights from the premium allocation literature (Dhaene et al., 2012 and Wei et.al., 2023) to explore the desirable properties of the premium principle.

Let $\rho(X)$ denote the capital amount required to adequately hedge against a loss X. We identify five desirable properties of the premium principle, and satisfying the last four leads to the establishment of a coherent risk measure, as outlined below:

1. Translational invariance. For any X_1, X_2 with $\mathbb{P}[X_1 \leq x] = \mathbb{P}[X_2 \leq x]$, $\rho[X_1] = \rho[X_2]$.
2. Monotonicity. For any X_1, X_2, $X_1 \leq X_2$ (almost surely) implies $\rho[X_1] \leq \rho[X_2]$.
3. Positive homogeneity. For $a > 0$, $\rho[aX] = a\rho[X]$.
4. Translation invariance. For $b \in \mathbb{R}$, $\rho[X + b] = \rho[X] + b$.
5. Subadditivity. For any X_1, X_2, $\rho[X_1 + X_2] \leq \rho[X_1] + \rho[X_2]$.

The axioms listed above provide a foundation for understanding the desirable properties of premium allocation principles. When applied to the multivariate premium principle, denoted by $\vec{X_n} = (X_1, X_2, ..., X_n)$ representing a portfolio of risks, and $\pi(X_i; \vec{X_n})$ representing the premium of individual risk X_i in the group, these axioms ideally guide the development of premium allocation principles. In this dissertation, we propose that insurers can ensure fair and efficient allocation of premiums across the portfolio by satisfying the following axioms:

1. Translational invariance. For any $\vec{c} \in \mathbb{R}^n$, $\pi(\vec{X_n} + \vec{c}) = \pi(\vec{X_n}) + \vec{c}$.

2. Positive homogeneity. For any $a > 0$, $\pi(a\vec{X_n}) = a\pi(\vec{X_n})$.

3. Diversification. Denote $\vec{X_n} = (X_1, ..., X_n)$ and $\vec{X}_{n+1} = (X_1, ..., X_n, X_{n+1})$,

$$\pi(X_i; \vec{X}_{n+1}) \leq \pi(X_i; \vec{X_n})$$

   for any $i = 1, ..., n$.

4. Monotonicity. Denote $\vec{X_a^i} = (X_1, ..., X_{i-1}, aX_i, X_{i+1}, ..., X_n)$ for some $a > 0$,

   (a) $\sum_{i \neq j} \pi(X_j; \vec{X_a^i}) + \pi(aX_i; \vec{X_a^i})$ increases when a increases;

   (b) $\pi(aX_i; \vec{X_a^i})$ increases when a increases;

   (c) $\pi(X_j; \vec{X_a^i})$ decreases when a increases for any $j \neq i$.

The first and second axioms express the transition of coherence from univariate risk measures to multivariate assumptions. These axioms assert that when risks are uniformly increased or scaled, the corresponding adjustments should be made to the premium amounts. This uniform increase demands consistent adjustments to premiums, ensuring coherence between risk and premium changes.

The third axiom, diversification, holds particular significance in group risk scenarios, as it states that the addition of one insured to the portfolio will decrease the premiums of the remaining risks. This phenomenon, defined as the diversification effect in this thesis, offers benefits to both insurers and insured parties. From the insurer's perspective, the diversification effect reduces the costs associated with business expansion and encourages participation in broader markets, including emerging ones such as the cyber insurance market, characterized by a high degree of uncertainty. The presence of the diversification effect in premium allocation principles can facilitate the development of the cyber insurance industry into a more mature and robust sector. From the insured's perspective, the diversification effect ensures that policyholders are not penalized for the risks posed by other insured parties, thereby safeguarding their interests against the operational risks of insurance companies. It is important to note the nuanced definition of the diversification effect in the context of premium allocation principles, which differs from its interpretation in traditional capital allocation frameworks. While in capital allocation, diversification entails allocating less capital to an asset within a portfolio compared to its standalone allocation, the divergence arises due to the distinct nature of insurance pricing and capital allocation problems.

The fourth axiom, monotonicity, indicates how individual policyholders should respond to changes in another policyholder's risk within the portfolio. Part (a) aligns with the positive homogeneity axiom, where an increase in one policyholder's risk leads to an overall increase in the total premium to reflect the heightened risk level. Similarly, part (b) adheres to positive homogeneity by reflecting the individual policyholder's increased risk in their premium amount. Part (c) ensures that no policyholder faces a higher premium due to another policyholder's risk escalation, maintaining fairness in premium distribution. This property encourages policyholders to manage their risk levels to secure lower premiums, as adjustments are made individually without benefiting other policyholders. Consequently, this mitigates moral hazard risks within insurance practices. Notably, the value of a encompasses all positive real numbers, naturally dividing into two parts: when $a \geq 1$, indicating that no one will suffer from others' high-risk behaviors, and when $0 < a \leq 1$, implying that no one will benefit from others' low-risk behaviors.

1.3 Technical Tools

In this section, we will introduce the technical tools necessary for understanding the CTE-induced premium pricing framework. Building upon the diversification and monotonicity properties discussed in the previous section, we will present preliminary analysis results and other tools to help us analyze these properties effectively. Specifically, we will demonstrate the subadditivity of CTE and its relationship to the diversification property. Additionally, we will explore how partial derivatives can be used to observe the monotonicity property in portfolios of size two. Finally, we will introduce the concept of conditionally pairwise negative correlation, which will be instrumental in analyzing diversification properties in future chapters.

1.3.1 CTE-induced Premium Formula and Preliminary Results

The CTE-induced premium pricing framework is a fusion of the CTE group premium, as described in (1.1), and the CTE allocation principle, as outlined in (1.2). By replacing the total capital K in (1.2) with the group premium in (1.1), we can easily derive the individual premium as the CTE of this single risk, conditioned on the group risk reaching a certain predetermined risk level. We will continue to use the same notation introduced in the previous section for consistency and clarity.

Definition 1.3.1. Denote the losses $\vec{X} = (X_1, X_2, ..., X_n)$ and the total loss $S = \sum_{i=1}^{n} X_i$. The premium for X_i, based on its contribution to $\text{CTE}_\alpha(S)$, is defined as

$$\pi(X_1; \vec{X}) = \mathbb{E}[X_i | S \geq VaR_\alpha(S)] \tag{1.3}$$

The CTE premium allocation principle satisfies the full allocation requirement and proportional allocation principle. Translational invariance and positive homogeneity are also demonstrated through straightforward manipulations applied to $\vec{X}$, thereby confirming its adherence to these properties.

We have obtained some preliminary results while investigating the conditions necessary to satisfy the diversification and monotonicity properties within the specific pricing framework outlined in (1.3). We initiate this exploration by analyzing a portfolio comprising only two insured parties.

Proposition 1.3.2. *Let X and Y be any two random variables. It holds that*

$$\pi(X; (X, Y)) \leq \pi(X; X)$$

Proof. By the subadditivity of CTE, we have

$$\pi((X + Y); (X, Y)) \leq \pi(X; X) + \pi(Y; Y) \tag{1.4}$$

Note that $\pi(Y;(X,Y)) \geq \pi(Y;Y)$. Rearranging (1.4), we have:

$$\begin{aligned}\pi((X+Y);(X,Y)) - \pi(Y;Y) &\leq \pi(X;X)\\ \pi(X;(X,Y)) + \pi(Y;(X,Y)) - \pi(Y;Y) &\leq \pi(X;X)\\ \pi(X;(X,Y)) &\leq \pi(X;X)\end{aligned}$$

□

The diversification property can be observed if expanding the assumption to a portfolio with size greater than two.

Corollary 1.3.3. *Let $\vec{X_n} = (X_1, ..., X_n)$ be a random vector. For any $i = 1, ..., n$, it holds that*

$$\pi(X_i; \vec{X_n}) \leq \pi(X_i; X_i)$$

Proof. By setting $X_i = X$ and $\sum_{j \neq i} X_j = Y$, it is a direct result from Proposition 1.3.2. □

The challenge in satisfying the diversification property arises when the group size increases from n to $n+1$. While numerical examples may provide insight into this phenomenon, constructing a rigorous algebraic proof remains an ongoing task.

One significant challenge in analyzing both diversification and monotonicity properties lies in the absence of explicit formulas, as defined in (1.3), which are essential for direct comparison. To address this challenge, we employ an approach that involves transforming the comparison between CTE premiums into a comparison between covariances by taking the derivative of the CTE with respect to the percentile.

Proposition 1.3.4. *Assume random vector (X, Y) has a joint density function, $\mathbb{E}[X^2] < \infty$, and $\mathbb{E}[Y^2] < \infty$. Denote $W_\alpha = X + aY$, $a > 0$. Then*

1. $\frac{d}{da} VaR_\alpha[W_\alpha] = \mathbb{E}[Y|W_\alpha = VaR_\alpha[W_\alpha]],$
2. $\frac{d}{da} CTE_\alpha[W_\alpha] = \mathbb{E}[Y|W_\alpha > VaR_\alpha[W_\alpha],$
3. $\frac{d}{da} \mathbb{E}[X|W_\alpha = VaR_\alpha[W_\alpha]] = \frac{f_{W_\alpha}(VaR_\alpha[W_\alpha])}{1-\alpha} Var[X|W_\alpha = VaR_\alpha[W_\alpha]],$
4. $\frac{d}{da} \mathbb{E}[Y|W_\alpha > VaR_\alpha[W_\alpha]] = \frac{f_{W_\alpha}(VaR_\alpha[W_\alpha])}{1-\alpha} Cov[X, Y|W_\alpha = VaR_\alpha[W_\alpha]].$

Proof. The first three outcomes can be established through direct differentiation. The comprehensive proof of the fourth outcome is provided below. For simplicity, let us denote $VaR_\alpha[W_\alpha] = q_\alpha^{W_\alpha}$, and the double integral inside CTE as a "joint-expectation" as follows:

$$\mathbb{E}[Y, \{W_\alpha > q_\alpha^{W_\alpha}\}] = \int_{-\infty}^{\infty} \int_{q_\alpha^{W_\alpha} - ay}^{\infty} y f(x, y)\, dx\, dy. \tag{1.5}$$

By using this notation, the CTE expression can be written as:

$$\mathbb{E}[Y \mid W_a > VaR_\alpha[W_a]] = \frac{1}{1-\alpha} \mathbb{E}[Y, \{W_a > q_\alpha^{W_a}\}] \tag{1.6}$$

Differentiating (1.5) with respect to a yields:

$$\frac{d}{da}\mathbb{E}[Y, \{W_\alpha > q_\alpha^{W_\alpha}\}] = \int_{-\infty}^{\infty} \left(-\frac{d}{da} q_\alpha^{W_\alpha} + y\right) \left(q_\alpha^{W_\alpha} - ay\right) f(q_\alpha^{W_a} - ay, y)\, dy.$$

Here, $\frac{d}{da} q_\alpha^{W_a} = \mathbb{E}[Y \mid W_a = q_\alpha^{W_a}]$. Also $\dfrac{f(q_\alpha^{W_a} - ay, y)}{f_{W_a}(q_\alpha^{W_a})} = f_{Y|W_a}(y)$, which is the conditional density function of Y given $W_a = q_\alpha^{W_a}$. Consequently, we have:

$$\begin{aligned}
\frac{d}{da}\mathbb{E}[Y, \{W_a > q_\alpha^{W_a}\}] &= -f_{W_a}(q_\alpha^{W_a}) \int_{-\infty}^{\infty} \left(\frac{d}{da} q_\alpha^{W_a} - y\right) \left(q_\alpha^{W_a} - ay\right) f_{Y|W_a}(y)\, dy \\
&= -f_{W_a}(q_\alpha^{W_a}) \mathbb{E}\left[(\mathbb{E}[Y \mid W_a = q_\alpha^{W_a}] - Y)(q_\alpha^{W_a} - aY) \mid W_a = q_\alpha^{W_a}\right] \\
&= f_{W_a}(q_\alpha^{W_a}) \mathbb{E}\left[(Y - \mathbb{E}[Y \mid W_a = q_\alpha^{W_a}])(X \mid W_a = q_\alpha^{W_a})\right] \\
&= f_{W_a}(q_\alpha^{W_a}) Cov[X, Y \mid W_a = VaR_\alpha[W_a]].
\end{aligned}$$

This completes the proof of the differentiation of (1.6) with respect to a, and thus, the proof of (4). □

Similar with Proposition 1.3.4, the correlation between monotonicty property and covariance can be derived when expanding the portfolio with size greater than two.

Proposition 1.3.5. *Assume* $(X_1, ..., X_n)$ *has a joint density function, with* $\mathbb{E}[X_i^2] < \infty$, $i = 1, 2, ..., n$. *Consider random vector* $(a_1 X_1, .., a_n X_n)$ *with* $a_1 \geq 0$ *for* $i = 1, ..., n$. *Denote* $S_a = \sum_{i=1}^n a_i X_1$, $q_\alpha^{S_a} = VaR_\alpha[S_a]$, *and denote by* $f_{S_a}(s)$ *the density function of* S_a. *It holds that, for each* $i = 1, ..., n$,

1. $\frac{\partial}{\partial a_i}(\sum_{k=1}^{n}\pi(a_k X_k)) = \mathbb{E}[X_i|S_a \geq q_\alpha^{S_a}]$,

2. $\frac{\partial}{\partial a_i}\pi(a_i X_i) = \mathbb{E}[X_i|S_a > q_\alpha^{S_a}] + \frac{a_i f_{S_a}(q_\alpha^{S_a})}{1-\alpha} Var[X_i|S_a = q_\alpha^{S_a}]$,

3. $\frac{\partial}{\partial a_i}\pi(a_j X_j) = \frac{a_j f_{S_a}(q_\alpha^{S_a})}{1-\alpha} Cov[X_i, X_j|S_a = q_\alpha^{S_a}]$, *for* $i \neq j$.

Proof. A direct derivation from Proposition 1.3.4, (2), (3), and (4). □

The third result from Proposition 1.3.5 provides us with the relationship between the value of covariance and the satisfaction of the monotonicity property. If the covariance between X_i and X_j is negative, then the monotonicity property is automatically satisfied. However, the question arises when the covariance is positive. In the future chapters, we will explore the dynamics between covariance and the satisfaction of the monotonicity property under specific loss distribution assumptions.

1.3.2 Conditionally Pairwise Negative Correlation

In this subsection, we introduce the concept of Conditionally Pairwise Negative Correlation (CPNC). Given the significance of covariance in our investigations, CPNC provides a specific definition to describe negative correlation under certain assumptions. This concept will play a crucial role in our subsequent discussions and analyses.

Definition 1.3.6. A random vector $(Y_1, ..., Y_n)$ is said to be conditionally pairwise negative correlated (CPNC) if

$$Cov[Y_i, Y_j|S = s] \leq 0$$

for any $i \neq j$ and any s.

The concept of CPNC can be demonstrated by considering a bivariate random vector such as (Y_1, Y_2):

$$\begin{aligned} Cov[Y_1, Y_2|Y_1 + Y_2 = s] &= Cov[Y_1, S - Y_1|Y_1 + Y_2 = s] \\ &= Cov[Y_1, S|Y_1 + Y_2 = s] - Cov[Y_1, Y_1|Y_1 + Y_2 = s] \\ &= Cov[Y_1, S|S = s] - Var[Y_1|Y_1 + Y_2 = s] \end{aligned}$$

$$= -Var[Y_1|Y_1 + Y_2 = s] \leq 0$$

Here, we use the fact that $Cov[Y_1, S|S = s] = 0$ because S is a constant given the condition. However, the relationship becomes uncertain in the trivariate case. For instance, consider the random vector (Y_1, Y_2, Y_3) with $Y_1 = Y_2$. In this case, (Y_1, Y_2, Y_3) is not CPNC because:

$$Cov[Y_1, Y_2|S = s] = Var[Y_1|S = s] \geq 0$$

The demonstration above provides valuable insights into analyzing the conditions required to achieve the monotonicity property and the role of covariance between policyholders. One straightforward scenario arises within the same bivariate random vector setting.

Lemma 1.3.7. *For a bivariate random vector* (Y_1, Y_2)*, if the conditional expectation* $\mathbb{E}[Y_1|Y_2 = y_2]$ *is decreasing as* y_2 *increases, then* $Cov[Y_1, Y_2] \leq 0$ *for any constant* s*.*

Proof. For simplicity, define $g(y_2) = \mathbb{E}[Y_1|Y_2 = y_2]$ for any value y_2, by the law of total expectation,

$$\mathbb{E}[Y_1Y_2] = \mathbb{E}[\mathbb{E}[Y_1Y_2|Y_2 = y_2]] = \mathbb{E}[y_2 \cdot \mathbb{E}[Y_1|Y_2 = y_2]] = \mathbb{E}[y_2 \cdot g(y_2)]$$

To compute the covariance,

$$\begin{aligned} Cov[Y_1, Y_2] &= \mathbb{E}[Y_1Y_2] - \mathbb{E}[Y_1]\mathbb{E}[Y_2] \\ &= \mathbb{E}[y_2 \cdot g(y_2)] - \mathbb{E}[\mathbb{E}[Y_1] \cdot Y_2] \\ &= \mathbb{E}[y_2 \cdot (g(y_2) - \mathbb{E}[Y_1])] \end{aligned}$$

Note that $\mathbb{E}[g(Y_2)] = \mathbb{E}[\mathbb{E}[Y_1|Y_2]] = \mathbb{E}[Y_1]$, implying that there exists y_0 such that $g(y_0) = \mathbb{E}[Y_1]$. If we assume that $g(y_2)$ is decreasing as y_2 increases, we have $y_2 - y_0 \geq 0$ and $g(y_2) - g(y_0) \leq 0$. After multiplication:

$$\begin{aligned} (y_2 - y_0)(g(y_2) - g(y_0)) &\leq 0 \\ y_2(g(y_2) - g(y_0)) - y_0(g(y_2) - g(y_0)) &\leq 0 \end{aligned}$$

$$y_2(g(y_2) - g(y_0)) \leq y_0(g(y_2) - g(y_0))$$

$$\mathbb{E}[y_2(g(y_2) - g(y_0))] \leq \mathbb{E}[y_0(g(y_2) - g(y_0))]$$

$$\mathbb{E}[y_2(g(y_2) - \mathbb{E}[Y_1]] \leq y_0\mathbb{E}[g(y_2) - g(y_0)]$$

$$Cov[Y_1, Y_2] \leq y_0\mathbb{E}[\mathbb{E}[Y_1|Y_2 = y_2] - \mathbb{E}[Y_1]]$$

$$Cov[Y_1, Y_2] \leq y_0[\mathbb{E}[Y_1] - \mathbb{E}[Y_1]]$$

$$Cov[Y_1, Y_2] \leq 0$$

Thus, if $\mathbb{E}[Y_1|Y_2 = y_2]$ decreases as y_2 increases, then $Cov(Y_1, Y_2)$ is negative. □

Assuming $Y_1 + Y_2 = S$, it is evident that $\mathbb{E}[Y_1|Y_2 = y_2, S = s]$ decreases as y_2 increases for any value of s. By incorporating this observation into the lemma, we deduce that $Cov(Y_1, Y_2|S = s) \leq 0$ always holds under the bivariate risk assumption. This statement can be extended to a multivariate case, such as a vector $(Y_1, Y_2, ..., Y_n)$ where $\sum_{i=1}^{n} Y_i = S$, with n being any positive integer greater than 2. In this case, we have:

$$\mathbb{E}[Y_1|Y_2 = y_2, S = s] = \mathbb{E}[Y_1|S - Y_2 = s - y_2]$$

Lemma 1.3.7 can be expanded to if $\mathbb{E}[Y_1|S - Y_2 = s - y_2]$ increases as y_2 decreases, then $Cov(Y_1, Y_2|S = s) \leq 0$. Therefore, according to Proposition 1.3.5 (3), the monotonicity property has been satisfied.

In summary, in this section, we introduced the key CTE-induced premium pricing expression (Equation 1.3). We then delved into the analysis of two desirable yet challenging properties: diversification and monotonicity. By proposing technical tools to quantify the triggering conditions for these properties, we laid the groundwork for their exploration. The diversification property will primarily be explored through numerical examples in subsequent chapters, while for the monotonicity property, our focus will be on identifying conditions that render the conditional expectation dependent on the covariance between the risks, with the total risk predetermined.

1.4 Thesis Organization

The remainder of the dissertation is structured as follows:

In Chapter 2, we delve into the CTE-induced formula under the assumption that the risk vector follows a multivariate normal distribution. Leveraging formulas derived by Landsman and Valdez (2003), we explore the conditions required to satisfy both the diversification and monotonicity properties. Through specific examples, we demonstrate how these properties manifest in practice.

Chapter 3 focuses on the CTE-induced formula when the risk vector follows a multivariate exponential distribution. By incorporating covariance considerations, we introduce a novel approach to pricing dependent risks.

In Chapter 4, we examine the CTE-induced formula for the case where the risk vector follows a multivariate Pareto distribution. Despite challenges posed by the heavy-tailed nature of the Pareto distribution, we present a stochastic representation and observe a fraction decomposition formula. Our analysis prioritizes $\mathbb{E}[X_i|S = VaR_\alpha(S)]$ over $\mathbb{E}[X_i|S > VaR_\alpha(S)]$ due to the distribution's characteristics.

Finally, in Chapter 5, we summarize the key findings of this dissertation and outline future research.

Chapter 2

Normal Distribution

In this chapter, we will assume that the loss variable follows a normal distribution. This assumption is primarily made because the normal distribution is well-studied in the literature, allowing us to leverage prior research results, such as those by Landsman and Valdez (2003). We will begin with a concise overview of the multivariate normal distribution setup, quote the CTE-induced formula from the earlier study by Landsman and Valdez (2003), and investigate the conditions necessary to satisfy both the diversification and monotonicity properties. This will involve employing partial derivatives and subsequent algebraic manipulations to clarify these relationships.

Despite the advantage of extensive research and analytical tractability, this assumption has a notable drawback: a normally distributed random variable can take on negative values, which is unrealistic for representing losses. However, we proceed with this assumption to explore any potential insights that may arise from this theoretical framework.

2.1 Introduction

The expressions for Value-at-Risk and Tail Value-at-Risk for a general univariate normal distribution, represented as $X \sim N(\mu, \sigma^2)$, were established by Landsman and Valdez (2003). The notation "CTE" is used in place of "TCE" when referencing their work. For any random variable X following a normal distribution with mean μ and variance σ^2, $\phi(\cdot)$ denotes the probability density function of the standard normal distribution. Also,

$$\begin{aligned} VaR_\alpha[X] &= \mu + \sigma VaR_\alpha[Z] \\ CTE_\alpha[X] &= \mu + \sigma CTE_\alpha[Z] = \mu + \sigma \frac{\phi(VaR_\alpha[Z]))}{1-\alpha} \end{aligned} \tag{2.1}$$

The expressions mentioned above can be derived through direct standardization. It's noteworthy that the CTE is obtained by adding the mean to the standard deviation evaluated at the specified risk level. However, extending this to the multivariate normal distribution poses additional challenges due to the presence of covariance. To address this, we will begin with a review of the fundamental setup.

Definition 2.1.1. For multivariate normal distribution, $\vec{X} = (X_1, ..., X_n) \sim MVN(\vec{\mu}, \Sigma)$, where $\mu = (E[X_1], ..., E[X_n])$ and

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 & \rho_{1n}\sigma_1\sigma_n \\ \rho_{12}\sigma_1\sigma_2 & \sigma_2^2 & \rho_{2n}\sigma_2\sigma_n \\ & ... & \\ \rho_1 n\sigma_1\sigma_n & \rho_{2n}\sigma_2\sigma_n & \sigma_n^2 \end{bmatrix}$$

is the covariance matrix.

Considering the distribution of the summation of n normally distributed random variables, $S_n = \sum_{i=1}^n X_i$, will follow the distribution such that

$$S_n \sim N\left(\sum_{i=1}^n \mu_i, \sum_{i,j} \rho_{ij}\sigma_i\sigma_j\right)$$

where $\rho_{ii} = 1$, for $i = 1, 2, ..., n$.

When analyzing the joint distribution of a single loss random variable X_i and the total portfolio loss S, represented as (X_i, S), we observe that it conforms to a bivariate normal distribution. This distribution is characterized by several parameters, including the means of the individual loss variable and the total portfolio loss, the correlation between X_i and S, the variance of X_i, and the covariance matrix of S. The notation for this distribution is expressed as follows:

$$(X_i, S) \sim BVN\left(\mu_i, \sum_{i=1}^{n} \mu_i; Corr(X_i, S), \sigma_i^2, \sum_{i,j} \rho_{ij}\sigma_i\sigma_j\right)$$

Here, $\sum_{i=1}^{n} \mu_i = \mu_s$, $\sigma_s^2 = \sum_{i,j} \rho_{ij}\sigma_i\sigma_j$.

Utilizing the formula established by Landsman and Valdez (2003), we can determine the conditional expectation of risk X_i given that the total risk S is greater than a predetermind value s. This expression reveals

$$\mathbb{E}[X_i|S > s] = \mu_i + \frac{Cov(X_i, S)}{\sigma_s} \cdot \frac{\phi(\frac{s-\mu_s}{\sigma_s})}{1 - \Phi(\frac{s-\mu_s}{\sigma_s})} \tag{2.2}$$

In the following section, we will integrate this formula into the premium pricing framework proposed in Equation (1.3) and investigate the conditions under which desired properties can be observed.

2.2 CTE Premiums

By refining the predetermined total loss value, denoted as s, to the $(1-\alpha)100\%$ percentile of the total loss within the conditional expectation outlined in (2.2), we unveil the premium pricing formula proposed in (1.3). Notably, the definition of Conditional Tail Expectation (CTE), as elucidated for a standard normal random variable in (2.1), enables the derivation of $CTE_\alpha[Z] = \frac{\phi(VaR_\alpha[Z])}{1-\alpha}$. Substituting these expressions into the conditional expectation yields the CTE-induced premium pricing formula defined in (1.3), under the assumption of a normal distribution for all losses.

Proposition 2.2.1. *When* $(X_1, X_2, ..., X_n) \sim MVN(\vec{0}; \Sigma)$*, denote* $Var[X_k] = \sigma_k^2$*,* $Cov[X_i, X_j] = \rho_{ij}\sigma_i\sigma_j$*,* $S_n = \Sigma_{i=1}^{n} X_i$ *and* $S_{n-j} = S_n - X_j$ *for any k,i,j. Then:*

$$\pi(X_i) = \mathbb{E}[X_i|S_n > VaR_\alpha(S_n)] = \mu_i + \frac{Cov(X_i, S_n)}{\sqrt{Var(S_n)}} \cdot CTE_\alpha[Z] \tag{2.3}$$

Proof. Substitute s with $VaR_\alpha(S_n)$ in (2.2). □

A key distinction between (2.3) and (2.1) lies in the direct proportionality of the CTE-induced premium to the covariance between X_i and S_n. A higher covariance leads to a higher CTE premium, which aligns with the rationale behind premium allocation. When the loss variable of a policyholder exhibits stronger correlation with the total loss, they bear greater responsibility in the event of a loss and, therefore, are expected to pay a higher premium. When comparing premium pricing frameworks across different policyholders, the primary differentiating factors are the individual mean and the covariance between the individual loss and the total loss.

2.3 Examples

After deriving the explicit CTE-induced premium formula, our subsequent aim is to explore the conditions necessary to observe specific desired properties. As noted in Proposition 1.3.5(3), the monotonicity property hinges on the covariance between the altered and unaltered loss random variables, under the condition that the total loss equals a predetermined value. This particular covariance can be derived from (2.3) through algebraic transformations.

Proposition 2.3.1. *Under the same setting as in Proposition 2.2.1, the following properties hold:*

1. $\frac{\partial}{\partial \sigma_i}\pi(X_i) \geq 0$

2. $\frac{\partial}{\partial \sigma_j}\pi(X_i) \leq 0$

3. *In addition, it follows from (2) that* $\frac{Cov[X_i, X_j]}{Cov[X_i, S_n]} \leq \frac{Cov[X_j, S_n]}{Var[S_n]}$.

Proof. By the definition of multivariate normal distribution, we have:

$$Cov[X_i, S_n] = \sum_{k=1}^{n} \rho_{ik}\sigma_i\sigma_k = \rho_{ii}\sigma_i^2 + \sum_{k\neq i} \rho_{ik}\sigma_i\sigma_k$$

$$Var[S_n] = Var[X_j + \sum_{k\neq j} X_k] = \sigma_j^2 + Var\left[\sum_{k\neq j} X_k\right] + 2\sum_{k\neq j} \rho_{jk}\sigma_j\sigma_k$$

Taking derivatives with respect to σ_j for the above terms, we find:

$$\frac{\partial}{\partial \sigma_j} Cov[X_i, S_n] = \rho_{ij}\sigma_i = \frac{Cov[X_i, X_j]}{\sigma_j} \tag{2.4}$$

$$\frac{\partial}{\partial \sigma_j} Var[S_n] = 2\sigma_j + 2\sum_{k\neq j} \rho_{jk}\sigma_k = \frac{2}{\sigma_j}\left(\sigma_j^2 + \sum_{k\neq j}\rho_{jk}\sigma_j\sigma_k\right) = \frac{2}{\sigma_j}\cdot Cov[X_j, S_n] \tag{2.5}$$

If the derivative of $\pi(X_i)$ with respect to σ_j is positive, then the derivative of $\ln(\pi(X_i))$ is also positive. Omitting μ_i and $CTE_\alpha[Z]$ in differentiation due to their independence from σ_j, we have:

$$\frac{\partial}{\partial \sigma_j} \ln \pi(X_i) = \frac{\frac{\partial}{\partial \sigma_j} Cov[X_i, S_n]}{Cov[X_i, S]} - \frac{1}{2}\cdot \frac{\frac{\partial}{\partial \sigma_j} Var[S_n]}{Var[S_n]}$$

Substituting the earlier derivatives from (2.4) and (2.5), the desirable observation $\frac{\partial}{\partial \sigma_j}\pi(X_i)$ can also be expressed as:

$$\frac{Cov[X_i, X_j]}{Cov[X_i, S_n]} \leq \frac{Cov[X_j, S_n]}{Var[S_n]} \tag{2.6}$$

□

We can now investigate whether the monotonicity property based on the precondition outlined in Proposition 2.3.1(2) is satisfied or not. This precondition depends on covariance and variance. To clarify our conclusion, let's consider a simple example involving three policyholders in the portfolio, each with identical standard deviations. In this case, the only variable to consider is the correlation coefficient.

Proposition 2.3.2. *If $S_3 = X_1 + X_2 + X_3$ and the standard deviations are the same, and the correlation coefficient $\rho_{ij} \leq \frac{\sqrt{5}-1}{2}$ for any $i, j \in \{1, 2, 3\}$, then the inequality $\frac{Cov[X_i, X_j]}{Cov[X_i, S_n]} \leq \frac{Cov[X_j, S_n]}{Var[S_n]}$ holds.*

Proof. Without loss of generality, assume $i = 1$ and $j = 2$, and $S_3 = X_1 + X_2 + X_3$. Then the inequality defined in (2.6) becomes

$$\frac{Cov[X_1, X_2]}{Cov[X_1, S_3]} \leq \frac{Cov[X_2, S_3]}{Var[S_3]}$$

and therefore

$$Cov[X_1, X_2]Var[S_3] \leq Cov[X_1, S]Cov[X_2, S_3]$$

Because we assume all the standard deviations of X_1, X_2 and X_3 are the same, we can assume it to be 1. Then:

$$\rho_{12}(2\rho_{12} + 2\rho_{13} + 2\rho_{23} + 3) \leq (1 + \rho_{12} + \rho_{13})(1 + \rho_{12} + \rho_{23})$$

$$\rho_{12}^2 + \rho_{12}\rho_{13} + \rho_{12}\rho_{23} + \rho_{12} \leq 1 + \rho_{23} + \rho_{13} + \rho_{13}\rho_{23}$$

$$\rho_{12}^2 + \rho_{12} + (\rho_{12} - 1)\rho_{13} + (\rho_{12} - 1)\rho_{23} \leq 1 + \rho_{13}\rho_{23}$$

Define $m = \min\{\rho_{13}, \rho_{23}\}$. Because $\rho_{12} - 1 < 0$, replacement of all such terms by m is an amplication of the left hand side of the inequality. The minification of the right hand side can also be achieved by replacing the $\rho_{13}\rho_{23}$ by m^2. Thus, the following steps can be justified:

$$\rho_{12}^2 + \rho_{12} + m(\rho_{12} - 1) + m(\rho_{12} - 1) \leq 1 + m^2$$

$$\rho_{12}^2 + (2m + 1)\rho_{12} \leq (m + 1)^2$$

$$\rho_{12} \leq \sqrt{(m + 1)^2 + (m + \frac{1}{2})^2} - (m + \frac{1}{2})$$

Define the function f such that $f(m) = \sqrt{(m+1)^2 + (m + \frac{1}{2})^2} - (m + \frac{1}{2})$, the function reaches its minimum when $f(0) = \sqrt{1 + \frac{1}{4}} - \frac{1}{2} = \frac{\sqrt{5} - 1}{2}$. □

Thus, if we can control the correlation coefficient such that $\rho_{12} \leq \frac{\sqrt{5}-1}{2}$, the premium of policyholder X_i would decrease as the standard deviation of the policyholder X_j increases, satisfying the monotonicity property.

To expand the research to the case when $n \geq 3$, we define $m = \min\{\rho_{ij}\}$, $M = \max\{\rho_{i,j}\}$ for all $i \neq j$, and $r = \dfrac{m}{M}$. The inequality defined in (2.6) can be derived under the same logic as the previous three random variable case:

$$1 + [(n-2)(2r + \frac{1}{M}) + (n-2)(n-3)] \leq (\frac{1}{M} + (n-2)r)^2$$

$$1 + 2(n-2)r + \frac{n-2}{M} + (n-2)(n-3) \leq \frac{1}{M^2} + \frac{2(n-2)}{M}r + (n-2)^2 r^2$$

$$\frac{1}{(n-2)^2} + \frac{2r}{n-2} + \frac{1}{M(n-2)} + \frac{n-3}{n-2} \leq \frac{1}{M^2(n-2)^2} + \frac{2}{M(n-2)}r + r^2$$
$$\frac{1}{(n-2)^2} + \frac{1}{M(n-2)} + \frac{n-3}{n-2} - \frac{1}{M^2(n-2)^2} \leq r^2 + 2\frac{1-M}{M(n-2)}r$$
$$\frac{1}{(n-2)^2} + \frac{1}{M(n-2)} + \frac{n-3}{n-2} - \frac{1}{M^2(n-2)^2} + (\frac{1-M}{M(n-2)})^2 \leq (r + \frac{1-M}{M(n-2)})^2$$
$$\sqrt{\frac{2M+n-4}{M(n-2)^2} + \frac{n-3}{n-2}} - \frac{1-M}{M(n-2)} \leq r$$

Now, let

$$g(n, M) = \sqrt{\frac{2M+n-4}{M(n-2)^2} + \frac{n-3}{n-2}} - \frac{1-M}{M(n-2)} \tag{2.7}$$

Then

$$\frac{\partial g(n, M)}{\partial M} = \frac{1}{2}(\frac{2M+n-4}{M(n-2)^2} + \frac{n-3}{n-2})^{-\frac{1}{2}}(\frac{4-n}{M^2(n-2)^2}) + \frac{1}{M^2(n-2)}$$
$$> 0$$

indicating that $g(n, M)$ is increasing with M. Similarly:

$$\frac{\partial g(n, M)}{\partial n} = \frac{1}{2}(\frac{2M+n-4}{M(n-2)^2} + \frac{n-3}{n-2})^{-\frac{1}{2}}(\frac{6-4M-n}{M(n-2)^3}) + \frac{1}{(n-2)^2} + \frac{1-M}{M(n-2)^2}$$
$$> 0$$

indicating that $g(n, M)$ is increasing with n. Combining these two results, we can conclude that the monotonicity property would be more difficult to observe when the maximum correlation coefficient is large and the portfolio size is large.

The lower bound of r can always be calculated when the values of M and n are assumed. Specifically, when $M = 1$, $g(n, M)$ will always be 1, requiring the lower bound of r to be 1. However, this can only occur when all the correlation coefficients are 1, indicating that when all the policyholders in the portfolio are perfectly correlated, it is almost impossible to satisfy the monotonicity property.

Alternatively, setting $M = \frac{1}{2}$ updates the function $g(n, M)$ as $g(n, M) = \sqrt{\frac{n(n-3)}{(n-2)^2}} - \frac{1}{n-2}$.

For instance, when $n = 4$, the minimum value of r is $r = \frac{1}{2}$; when $n = 5$, the lower bound of r would be $r = \frac{\sqrt{10}}{3} - \frac{1}{3}$. This indicates that when the ratio between the highest and lowest correlation coefficients is controlled, as the portfolio size increases, it becomes more difficult to observe the monotonicity property.

The value of $g(n, M)$ will determine the feasibility of the diversification property, and the calculation is based on two variables. To examine some calculation results, we chose thresholds of 0.25, 0.5, and 0.75 for M, where 0.25 represents weak correlation, 0.5 represents medium correlation, and 0.75 represents strong correlation. For n, we used values ranging from 4 to 50 to observe how the conditions respond to changes in portfolio size.

n\M	0.25	0.5	0.75
3	N/A	-1	0.4832
4	-0.5000	0.5000	0.8333
5	0.1547	0.7208	0.9072
10	0.7569	0.9208	0.9738
15	0.8625	0.9551	0.9851
20	0.9049	0.9688	0.9897
30	0.9414	0.9807	0.9936
50	0.9670	0.9891	0.9964

Table 2.1: Values of $g(n, M)$ for various n and M.

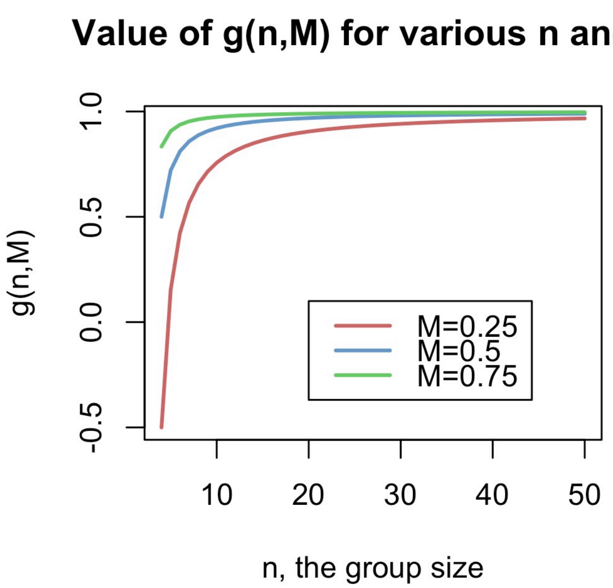

Figure 2.1: Values of $g(n, M)$ for various n and M.

The overall observation would be that, when M is fixed, the lower bound of the function increasingly approaches 1 as n increases. This implies that the difference between m and M needs to decrease, requiring a more homogeneous distribution among the variables to observe the diversification effect. This observation makes sense because it should be more challenging to satisfy the diversification property as the portfolio size increases. Conversely, when n is fixed, the value of r increases as M increases. This indicates that with higher correlations among variables, it becomes harder to achieve diversification. Overall, these results align with theoretical expectations. Notably, the lower bound of the function approaches 1 rather quickly when M exceeds 0.5. This suggests that with relatively strong correlations, it is difficult to satisfy the diversification effect unless all covariances between pairs of variables are highly similar.

Chapter 3

Exponential Distribution

In this chapter, we assume that the loss variables follow a multivariate exponential distribution. This distribution is commonly used in actuarial science due to its simplicity and analytical tractability, and it also addresses the limitation of the previous chapter, specifically the possibility of negative values. We will start with straightforward calculations of probabilities for the sum of independently exponentially distributed variables. Subsequently, we will explore the joint distribution of a specific loss variable X_i and the sum of all the remaining variables in the portfolio.

We will derive the expectation of this loss variable given that the sum equals a certain level, followed by the expectation given that the sum is greater than a certain level. Once these foundational tools are in place, we will introduce dependence into the model by assuming that all the loss variables in the portfolio share a common component. We will then analyze the preconditions necessary for the desired properties to hold.

3.1 Survival Function of Sum of Variables

In this section, we provide an explicit probability density function (PDF) of the summation of independent exponentially distributed random variables using a moment generating function (MGF) approach. First, we introduce an algebraic expressions, $A_{i,n}$ and $A^{\bar{i}}_{j,n}$, which serve as coefficients for distinct exponentially distributed random variables involved in the summation process.

Definition 3.1.1. Define $A_{i,n}$ and $A_{i,n}^{\bar{j}}$ such that $i, j, n \in \mathbb{Z}^+$, $i, j \leq n$, and $i \neq j$.

$$A_{i,n} = \prod_{j=1, j\neq i}^{n} \frac{\lambda_j}{\lambda_j - \lambda_i} \qquad\qquad A_{j,n}^{\bar{i}} = \prod_{j=1,k=1,j\neq i,k,k\neq i,j}^{n} \frac{\lambda_k}{\lambda_k - \lambda_j}$$

To illustrate, let's consider the case when $n = 3$. Then we have $A_{1,3}$, $A_{2,3}^{\bar{1}}$ and $A_{3,3}^{\bar{1}}$ such that:

$$A_{1,3} = \frac{\lambda_2}{\lambda_2 - \lambda_1}\frac{\lambda_3}{\lambda_3 - \lambda_1}, \quad A_{2,3}^{\bar{1}} = \frac{\lambda_3}{\lambda_3 - \lambda_2}, \quad A_{3,3}^{\bar{1}} = \frac{\lambda_2}{\lambda_2 - \lambda_3}.$$

Proposition 3.1.2. *Properties of $A_{i,n}$:*

1. $\sum_{i=1}^{n} \lambda_i A_{i,n} = 0$, *which can also be derived as* $\lambda_j A_{j,n} = -\sum_{i=1, i\neq j}^{n} \lambda_i A_{i,n}$
2. $\sum_{i=1}^{n} A_{i,n} = 1$
3. $\sum_{i=1}^{n} \frac{A_{i,n}}{\lambda_i} = \sum_{i=1}^{n} \frac{1}{\lambda_i}$

Proof. 1.

$$\begin{aligned}
\sum_{i=1}^{n} \lambda_i A_{i,n} &= \sum_{i=1}^{n} \lambda_i \prod_{j=1,j\neq i}^{n} \frac{\lambda_j}{\lambda_j - \lambda_i} \\
&= \sum_{i=1}^{n} \frac{\prod_{i=1}^{n} \lambda_i}{\prod_{j=1,j\neq i}^{n} \lambda_j - \lambda_i} \\
&= \prod_{i=1}^{n} \lambda_i \sum_{i=1}^{n} \frac{1}{\prod_{j=1,j\neq i}^{n} (\lambda_j - \lambda_i)} \\
&= \prod_{i=1}^{n} \lambda_i \sum_{i=1}^{n} \frac{1}{C} \sum_{j=1,j\neq i}^{n} \frac{1}{(\lambda_j - \lambda_i)} \\
&= \prod_{i=1}^{n} \frac{\lambda_i}{C} \sum_{i=1}^{n} \sum_{j=1,j\neq i}^{n} \frac{1}{(\lambda_j - \lambda_i)} \\
&= \prod_{i=1}^{n} \frac{\lambda_i}{C} \cdot 0 \\
&= 0
\end{aligned}$$

Notice that the term $\frac{1}{\prod_{j=1,j\neq i}^{n}(\lambda_j - \lambda_i)}$ can be partially decomposed. This decomposition involves expressing it as the product of a constant term, $\frac{1}{C}$, and the summation of fractions, $\sum_{j=1,j\neq i}^{n} \frac{1}{\lambda_j - \lambda_i}$. When combined with the outer summation, for every term $\frac{1}{\lambda_j - \lambda_i}$, there exists a corresponding term $\frac{1}{\lambda_i - \lambda_j}$. Since $\frac{1}{\lambda_j - \lambda_i} = -\frac{1}{\lambda_i - \lambda_j}$, the overall summation equals zero.

2. Proof by Induction. Base Case: when $n = 2$,

$$\sum_{i=1}^{2} A_{i,2} = \frac{\lambda_2}{\lambda_2 - \lambda_1} + \frac{\lambda_1}{\lambda_1 - \lambda_2} = 1$$

Inductive Step: Assume that for $n = k$, the statement holds:

$$\sum_{i=1}^{k} A_{i,k} = 1$$

We need to prove that for $n = k + 1$,

$$\sum_{i=1}^{k+1} A_{i,k+1} = 1$$

Starting from the inductive hypothesis, we have:

$$\begin{aligned}
\sum_{i=1}^{k+1} A_{i,k+1} &= \sum_{i=1}^{k} A_{i,k} \cdot \frac{\lambda_{k+1}}{\lambda_{k+1} - \lambda_i} + A_{k+1,k+1} \\
&= \sum_{i=1}^{k} A_{i,k} \frac{\lambda_i}{\lambda_i} \frac{\lambda_{k+1}}{\lambda_{k+1} - \lambda_i} + A_{k+1,k+1} \\
&= \sum_{i=1}^{k} A_{i,k} \lambda_i \left(\frac{1}{\lambda_i} + \frac{1}{\lambda_{k+1} - \lambda_i} \right) + A_{k+1,k+1} \\
&= \sum_{i=1}^{k} A_{i,k} \frac{\lambda_i}{\lambda_i} + \sum_{i=1}^{k} A_{i,k} \frac{\lambda_i}{\lambda_{k+1} - \lambda_i} + A_{k+1,k+1} \\
&= 1 + \sum_{i=1}^{k} A_{i,k} \frac{\lambda_{k+1}}{\lambda_{k+1}} \frac{\lambda_i}{\lambda_{k+1} - \lambda_i} + A_{k+1,k+1}
\end{aligned}$$

$$
\begin{aligned}
&= 1 + \frac{1}{\lambda_{k+1}} \sum_{i=1}^{k} A_{i,k} \lambda_i \frac{\lambda_{k+1}}{\lambda_{k+1} - \lambda_i} + A_{k+1,k+1} \\
&= 1 + \frac{1}{\lambda_{k+1}} \sum_{i=1}^{k} A_{i,k+1} \lambda_i + A_{k+1,k+1} \\
&= 1 - \frac{1}{\lambda_{k+1}} \lambda_{k+1} A_{k+1,k+1} + A_{k+1,k+1} \\
&= 1
\end{aligned}
$$

In this proof, we have applied the result from the first part that $\sum_{i=1}^{k} \lambda_i A_{i,k+1} = \lambda_{k+1} A_{k+1,k+1}$.

3.

$$
\begin{aligned}
\sum_{i=1}^{n} \frac{A_{i,n}}{\lambda_i} &= \sum_{i=1}^{n} \prod_{j=1,j\neq i}^{n} \frac{\lambda_j}{(\lambda_j - \lambda_i)\lambda_i} \\
&= \sum_{i=1}^{n} \sum_{j=1,j\neq i}^{n} (\frac{1}{\lambda_i} + \frac{1}{\lambda_j - \lambda_i}) \\
&= \sum_{i=1}^{n} \frac{1}{\lambda_i} + \sum_{i=1}^{n} \sum_{j=1,j\neq i}^{n} \frac{1}{\lambda_j - \lambda_i} \\
&= \sum_{i=1}^{n} \frac{1}{\lambda_i}
\end{aligned}
$$

Again, we use the result that $\displaystyle\sum_{i=1}^{n} \sum_{j=1,j\neq i}^{n} \frac{1}{\lambda_j - \lambda_i} = 0.$

□

From the earlier proof, we see that we utilized partial fraction decomposition techniques extensively. The properties of $A_{i,n}$ are a direct result of these techniques. In the following demonstration, we will observe that the $A_{i,n}$'s are derived as derivatives from the product of moment generating functions (MGFs) of exponential distributions. While these derivatives look like weights in the summation expression, it is important to note that they are not true weights, as not every term is nonnegative. A similar proof, derived from the convolution of probability density functions (pdf), can be found in Ross (2019).

Proposition 3.1.3. *Let $\{X_1, X_2, ..., X_n\}$ be mutually independent Exponential random vari-*

ables such that $X_i \sim Exp(\frac{1}{\lambda_i})$*,* $E[X_i] = \frac{1}{\lambda_i}$*, all* $\lambda_i' s$ *are distinct, and the probability density function of* X_i *is* $\lambda_i e^{-\lambda_i x_i}$*. Denote* $S = \sum_{i=1}^{n} X_i$*. Then the survival function of* S *is given by*

$$P(S > s) = \sum_{i=1}^{n} A_{i,n} e^{-\lambda_i s} \tag{3.1}$$

where $A_{i,n}$ *has been defined in Definition 3.1.1.*

Proof. Proving by induction, we begin with the base case when $n = 2$. The proof is based on the basic property of MGF and the repeated application of partial fraction decomposition.

$$\begin{aligned} m_{X_1+X_2}(t) &= m_{X_1}(t) \cdot m_{X_2}(t) \\ &= \frac{\lambda_1}{\lambda_1 - t} \cdot \frac{\lambda_2}{\lambda_2 - t} \\ &= \frac{\lambda_1 \lambda_2}{\lambda_2 - \lambda_1} \cdot (\frac{1}{\lambda_1 - t} - \frac{1}{\lambda_2 - t}) \\ &= \frac{\lambda_2}{\lambda_2 - \lambda_1} \frac{\lambda_1}{\lambda_1 - t} + \frac{\lambda_1}{\lambda_1 - \lambda_2} \frac{\lambda_2}{\lambda_2 - t} \\ &= A_{1,2} \frac{\lambda_1}{\lambda_1 - t} + A_{2,2} \frac{\lambda_2}{\lambda_2 - t} \\ &= \sum_{i=1}^{2} A_{i,2} \frac{\lambda_i}{\lambda_i - t} \end{aligned}$$

Inductive step: when $n = k$ for $k \in \mathbb{Z}^+$, we proceed with the inductive step.

$$\begin{aligned} m_{\sum_{i=1}^{k} X_i}(t) &= \prod_{i=1}^{k} m_{X_i}(t) \\ &= \sum_{i=1}^{k} A_{i,k} \frac{\lambda_i}{\lambda_i - t} \end{aligned}$$

To complete the proof, we establish the result for the case when $n = k + 1$. Proposition 3.1.2 (3) is applied at the last step.

$$m_{\sum_{i=1}^{k+1} X_i}(t) = \prod_{i=1}^{k+1} m_{X_i}(t)$$

$$
\begin{aligned}
&= \prod_{i=1}^{k} m_{X_i}(t) \cdot m_{X_{k+1}}(t) \\
&= \sum_{i=1}^{k} A_{i,k} \frac{\lambda_i}{\lambda_i - t} \cdot \frac{\lambda_{k+1}}{\lambda_{k+1} - t} \\
&= \sum_{i=1}^{k} A_{i,k} \frac{\lambda_i \lambda_{k+1}}{\lambda_{k+1} - \lambda_i} \left(\frac{1}{\lambda_i - t} - \frac{1}{\lambda_{k+1} - t} \right) \\
&= \sum_{i=1}^{k} A_{i,k+1} \frac{\lambda_i}{\lambda_i - t} + \sum_{i=1}^{k} A_{i,k} \frac{\lambda_{k+1}}{\lambda_{k+1} - \lambda_i} \cdot \frac{\lambda_i}{\lambda_{k+1}} \cdot \frac{\lambda_{k+1}}{\lambda_{k+1} - t} \\
&= \sum_{i=1}^{k} A_{i,k+1} \frac{\lambda_i}{\lambda_i - t} - \sum_{i=1}^{k} A_{i,k+1} \lambda_i \cdot \frac{1}{\lambda_{k+1}} \cdot \frac{\lambda_{k+1}}{\lambda_{k+1} - t} \\
&= \sum_{i=1}^{k} A_{i,k+1} \frac{\lambda_i}{\lambda_i - t} + A_{i,k+1} \lambda_{k+1} \cdot \frac{1}{\lambda_{k+1}} \cdot \frac{\lambda_{k+1}}{\lambda_{k+1} - t} \\
&= \sum_{i=1}^{k+1} A_{i,k+1} \frac{\lambda_i}{\lambda_i - t}
\end{aligned}
$$

Now, we recognize that this is a weighted average of exponential moment generating functions. Hence, we have successfully derived the survival function $P(S > s)$ and the probability density function $f_S(s)$ for the sum S of n independent exponential random variables. These expressions, given by $\sum_{i=1}^{n} A_{i,n} e^{-\lambda_i s}$ and $\sum_{i=1}^{n} \lambda_i A_{i,n} e^{-\lambda_i s}$ respectively, indicate that the distribution of S is a linear combination of $X_i's$. □

After the derivation process, we can verify the properties of $A_{i,n}$ both algebraically and in the context of probability theory. We observe that $\sum_{i=1}^{n} A_{i,n} = 1$ because the weights across all terms in the summation must sum to one. This is easily verified by the fact that $m_S(0) = 1$. Additionally, the third property is evident when we find the first moment of S using the MGF. Specifically,

$$
m_S^{(1)}(0) = \sum_{i=1}^{n} A_{i,n} \frac{\lambda_i}{(\lambda_i - 0)^2} = \sum_{i=1}^{n} \frac{A_{i,n}}{\lambda_i}
$$

is equivalent to

$$
\mathbb{E}[S] = \sum_{i=1}^{n} \mathbb{E}[X_i] = \sum_{i=1}^{n} \frac{1}{\lambda_i}.
$$

This equivalence further demonstrates the consistency of the properties of $A_{i,n}$ within the framework of probability theory.

3.2 CTE Premiums

After deriving the explicit formula for the sum of independent exponentially distributed random variables, we are equipped to derive the explicit CTE-induced conditional expectation formula, alongside the variance and covariance formulas. To facilitate these calculations, we establish several frequently used results.

Definition 3.2.1. In this chapter, we use the following notations:

$$\begin{aligned} g_1(\lambda_i, \lambda_j, s) &= \int_0^s x_i e^{-(\lambda_i - \lambda_j)x_i} dx_i \\ &= \frac{1 - e^{-(\lambda_i - \lambda_j)s} - (\lambda_i - \lambda_j)se^{-(\lambda_i - \lambda_j)s}}{(\lambda_i - \lambda_j)^2} \end{aligned} \tag{3.2}$$

$$\begin{aligned} g_2(\lambda_i, \lambda_j, s) &= \int_s^\infty e^{-\lambda_j t} g_1(\lambda_i, \lambda_j, t) dt \\ &= \frac{1}{(\lambda_i - \lambda_j)^2} \Big(\frac{e^{-\lambda_j s}}{\lambda_j} - \frac{((\lambda_i - \lambda_j)\lambda_i s + 2\lambda_i - \lambda_j)e^{-\lambda_i s}}{\lambda_i^2} \Big) \end{aligned} \tag{3.3}$$

$$\begin{aligned} h_1(\lambda_i, \lambda_j, s) &= \int_0^s x_i^2 e^{-(\lambda_i - \lambda_j)x_i} dx_i \\ &= \frac{2 - 2e^{-(\lambda_i - \lambda_j)s} - 2(\lambda_i - \lambda_j)se^{-(\lambda_i - \lambda_j)s} - (\lambda_i - \lambda_j)^2 s^2 e^{-(\lambda_i - \lambda_j)s}}{(\lambda_i - \lambda_j)^3} \end{aligned} \tag{3.4}$$

$$\begin{aligned} h_2(\lambda_i, \lambda_j, s) &= \int_s^\infty e^{-\lambda_j t} h_1(\lambda_i, \lambda_j, t) dt \\ &= \frac{1}{(\lambda_i - \lambda_j)^3} \Big[\frac{2e^{-\lambda_j s}}{\lambda_j} - \left(\frac{2}{\lambda_i} + 2(\lambda_i - \lambda_j)\Big(\frac{s}{\lambda_i} + \frac{1}{\lambda_i^2}\Big) + (\lambda_i - \lambda_j)^2 \Big(\frac{s^2}{\lambda_i} + \frac{2s}{\lambda_i^2} + \frac{2}{\lambda_i^3} \Big) \right) e^{-\lambda_i s} \Big] \end{aligned} \tag{3.5}$$

$$\begin{aligned} g_3(\lambda_i, \lambda_j, \lambda_k, s) &= \int_0^s x_j e^{-(\lambda_j - \lambda_k)x_j} g_1(\lambda_i, \lambda_k, s - x_j) dx_j \\ &= \frac{1}{(\lambda_i - \lambda_k)^2} \Big[g_1(\lambda_j, \lambda_k, s) - \left(e^{-(\lambda_i - \lambda_k)s} + (\lambda_i - \lambda_k)e^{-(\lambda_i - \lambda_k)s} s \right) g_1(\lambda_j, \lambda_i, s) \\ &\quad + (\lambda_i - \lambda_k)e^{-(\lambda_i - \lambda_k)s} h_1(\lambda_j, \lambda_i, s) \Big] \end{aligned} \tag{3.6}$$

The expressions g_1 and g_2 are utilized in the formulas about the first moment, while the functions h_1 and h_2 are employed in the formulas about the second moments. The expression

g_3 is incorporated in the covariance formula.

When deriving the formulas, we will consider three distinct scenarios: (1) when all the λ_i's are distinct, (2) when all the λ_i's are the same, and (3) when some λ_i's are identical while others differ. The results of the first scenario will be presented in detail in the following subsection. However, due to space constraints and the supplementary nature of the information, results from scenarios (2) and (3) will be provided in the appendix for cross-verification purposes.

Given the expression for the summation of random variables, it is necessary to assume that all λ_i's are distinct. Otherwise, we encounter division by zero in the denominator, rendering the calculation invalid.

When deriving the expectation of X_i given that the sum equals a predetermined value, due to the independence of all the random variables, we can represent the joint distribution of X_i and the sum of X_j for all $j \neq i$ as the product of their individual distributions. With this joint probability density function (pdf) established, we can then calculate the expectation using the standard definition.

Proposition 3.2.2. *Under the same assumptions as in Proposition 3.1.3, for the simplicity, let $A_{i,n} = A_i$ and $A_{i,n}^{\bar{j}} = A_i^{\bar{j}}$ for any $i, j, n \in \mathbb{Z}^+$ and $i, j < n$, $i \neq j$, the conditional expectation of X_i given that $S_n = s$ exists and is equal to*

$$\mathbb{E}[X_i|S_n = s] = \frac{\lambda_i}{\sum_{k=1}^n \lambda_k A_{k,n} e^{-\lambda_k s}} \sum_{j\neq i, j=1}^{n} \lambda_j A_{j,n}^{\bar{i}} e^{-\lambda_j s} g_1(\lambda_i, \lambda_j, s) \tag{3.7}$$

Proof.

$$\begin{aligned}
\mathbb{E}[X_i|S_n = s] &= \int_0^s x_i f_{X_i|S}(x_i|s) dx_i \\
&= \int_0^s x_i \frac{f_{X_i}(x_i) f_{\sum_{j\neq i,j=1}^n X_j}(s - x_i)}{f_{S_n}(s)} dx_i \\
&= \int_0^s x_i \frac{\lambda_i e^{-\lambda_i x_i} \sum_{j\neq i,j=1}^n \lambda_j A_j^{\bar{i}} e^{-\lambda_j (s-x_i)}}{\sum_{k=1}^n \lambda_k A_k e^{-\lambda_k s}} dx_i \\
&= \frac{\lambda_i \sum_{j\neq i,j=1}^n A_j^{\bar{i}} \lambda_j e^{-\lambda_j s}}{\sum_{k=1}^n \lambda_k A_k e^{-\lambda_k s}} \int_0^s x_i e^{-(\lambda_i - \lambda_j) x_i} dx_i \\
&= \frac{\lambda_i \sum_{j\neq i,j=1}^n A_j^{\bar{i}} \lambda_j e^{-\lambda_j s}}{\sum_{k=1}^n \lambda_k A_k e^{-\lambda_k s}} g_1(\lambda_i, \lambda_j, s)
\end{aligned} \tag{3.8}$$

□

Building upon the above result, we proceed to derive the CTE-induced premium pricing formula proposed at the beginning of this paper, the formula for the conditional expectation of X_i given that $S_n > s$. Following the definition, this can be accomplished by integrating the product of the conditional expectation $\mathbb{E}[X_i|S_n = t]$ and the probability density function $f_{S_n}(t)$ over the interval from s to ∞.

Proposition 3.2.3. *Under the same assumptions as in Proposition 3.2.2, the conditional expectation of X_i given that $S_n > s$ exists and is equal to*

$$\mathbb{E}[X_i|S_n > s] = \lambda_i \sum_{j\neq i, j=1}^{n} \lambda_j A_j g_2(\lambda_i, \lambda_j, s)$$

Proof.

$$\begin{aligned}
\mathbb{E}[X_i|S_n > s] &= \int_s^\infty f_S(t)\mathbb{E}[X_i|S = t]dt \\
&= \int_s^\infty \sum_{k=1}^{n} \lambda_k A_k e^{-\lambda_k t} \left(\frac{\lambda_i \sum_{j\neq i, j=1}^{n} \lambda_j A_j e^{-\lambda_j t}}{\sum_{k=1}^{n} \lambda_k A_k e^{-\lambda_k t}} g_1(\lambda_i, \lambda_j, t) \right) dt \\
&= \int_s^\infty \lambda_i \sum_{j\neq i, j=1}^{n} \lambda_j A_j e^{-\lambda_j t} g_1(\lambda_i, \lambda_j, t) dt \\
&= \lambda_i \sum_{j\neq i, j=1}^{n} \lambda_j A_j \int_s^\infty e^{-\lambda_j t} g_1(\lambda_i, \lambda_j, t) dt \\
&= \lambda_i \sum_{j\neq i, j=1}^{n} \lambda_j A_j g_2(\lambda_i, \lambda_j, s) \qquad (3.9)
\end{aligned}$$

□

The formula for the CTE-induced premium pricing may appear complex and challenging to interpret due to its direct integral nature. However, this structure offers an advantage that every component is integrable, which enables us to proceed with deriving variance and covariance expressions.

Proposition 3.2.4. *Under the same assumptions as in Proposition 3.2.2, we have*

$$Var[X_i|S_n = s] = \frac{\lambda_i \sum_{j\neq i,j=1}^{n} \lambda_j A_j e^{-\lambda_j s}}{\sum_{k=1}^{n} \lambda_k A_k e^{-\lambda_k s}} h_1(\lambda_i, \lambda_j, s) - \left[\frac{\lambda_i \sum_{j\neq i,j=1}^{n} \lambda_j A_j e^{-\lambda_j s}}{\sum_{k=1}^{n} \lambda_k A_k e^{-\lambda_k s}} g_1(\lambda_i, \lambda_j, s)\right]^2$$

Proof.

$$\begin{aligned}
\mathbb{E}[X_i^2|S = s] &= \int_0^s x_i^2 f_{X_i|S}(x_i|s) dx_i \\
&= \frac{\lambda_i \sum_{j\neq i,j=1}^{n} \lambda_j A_j e^{-\lambda_j s}}{f_S(s)} \int_0^s x_i^2 e^{-(\lambda_i - \lambda_j)x_i} dx_i \\
&= \frac{\lambda_i \sum_{j\neq i,j=1}^{n} \lambda_j A_j e^{-\lambda_j s}}{\sum_{k=1}^{n} \lambda_k A_k e^{-\lambda_k s_0}} h_1(\lambda_i, \lambda_j, s)
\end{aligned}$$

So the variance $V[X_i|S_n = s] = \mathbb{E}[X_i^2|S_n = s] - \Big(\mathbb{E}[X_i|S = s]\Big)^2$. □

Proposition 3.2.5. *Under the same assumptions as in Proposition 3.2.2, we have:*

$$Var[X_i|S_n > s] = \lambda_i \sum_{j\neq i,j=1}^{n} \lambda_j A_j h_2(\lambda_i, \lambda_j, s) - \left[\lambda_i \sum_{j\neq i,j=1}^{n} \lambda_j A_j g_2(\lambda_i, \lambda_j, s)\right]^2$$

Proof.

$$\begin{aligned}
\mathbb{E}[X_i^2|S_n > s] &= \int_s^\infty f_{S_n}(t) \mathbb{E}[X_i^2|S_n = t] dt \\
&= \int_s^\infty f_{S_n}(t) \frac{\lambda_i \sum_{j\neq i,j=1}^{n} \lambda_j A_j e^{-\lambda_j t}}{f_S(t)} h_1(\lambda_i, \lambda_j, t) dt \\
&= \int_s^\infty \lambda_i \sum_{j\neq i,j=1}^{n} \lambda_j A_j e^{-\lambda_j t} h_1(\lambda_i, \lambda_j, t) dt \\
&= \lambda_i \sum_{j\neq i,j=1}^{n} \lambda_j A_j \int_s^\infty e^{-\lambda_j t} h_1(\lambda_i, \lambda_j, t) dt \\
&= \lambda_i \sum_{j\neq i,j=1}^{n} \lambda_j A_j h_2(\lambda_i, \lambda_j, s)
\end{aligned}$$

So the variance $Var[X_i|S_n > s] = \mathbb{E}[X_i^2|S_n > s] - \Big(\mathbb{E}[X_i|S_n > s]\Big)^2$. □

The covariance between any pair of X_i and X_j conditional on $S_n = s$ plays a pivotal role in analyzing the preconditions necessary for the desired properties to hold.

Following the previous methodology, we segment the portfolio into 5 distinct subsets: X_i, X_j, $\sum_{p=1,p\neq i}^{n} X_p$, $\sum_{q=1,q\neq j}^{n} X_q$, and $\sum_{k=1,k\neq i,j}^{n} X_k$. Despite the complexity of the resulting covariance formula, each component is precisely defined.

Proposition 3.2.6. *Under the same assumptions of Proposition 3.2.2, we have*

$$\begin{aligned}
Cov(X_i, X_j|S_n = s) &= \frac{\lambda_i\lambda_j}{f_{S_n}(s)} \sum_{k=1,k\neq i,j}^{n} \lambda_k e^{-\lambda_k s} g_3(\lambda_i, \lambda_j, \lambda_k, s) \\
&\quad - \Big(\frac{\lambda_i \sum_{p\neq i,p=1}^{n} \lambda_p A_p e^{-\lambda_p s}}{f_{S_n}(s)} g_1(\lambda_i, \lambda_p, s)\Big)\Big(\frac{\lambda_j \sum_{q\neq j,q=1}^{n} \lambda_q A_q e^{-\lambda_q s}}{f_{S-n}(s)} g_1(\lambda_j, \lambda_q, s)\Big)
\end{aligned}$$

Proof. The joint distribution of X_i, X_j and $S_n - X_i - X_j$ is the product of marginal distributions due to the mutual independence of the variables:

$$\begin{aligned}
&f_{X_i,X_j,S_n-X_i-X_j}(x_i, x_j, s - x_i - x_j) \\
&= f_{X_i}(x_i) f_{X_j}(x_j) f_{S-X_i-X_j}(x_i, x_j, s - x_i - x_j) \\
&= \lambda_i\lambda_j e^{-\lambda_i x_i} e^{-\lambda_j x_j} \sum_{k=1,k\neq i,j}^{n} \lambda_k A_k e^{-\lambda_k(s-x_i-x_j)}
\end{aligned}$$

By following the previous derivations, we have:

$$\begin{aligned}
&\mathbb{E}[X_iX_j|S_n = s] \\
&= \int_0^s \int_0^{s-x_j} x_i x_j \frac{f_{X_i,X_j,S-X_i-X_j}(x_i, x_j, s - x_i - x_j)}{f_{S_n}(s)} dx_i dx_j \\
&= \frac{1}{f_S(s)} \int_0^s \int_0^{s-x_j} x_i x_j \lambda_i\lambda_j e^{-\lambda_i x_i} e^{-\lambda_j x_j} \sum_{k=1,k\neq i,j}^{n} \lambda_k A_k e^{-\lambda_k(s-x_i-x_j)} dx_i dx_j \\
&= \frac{\lambda_i\lambda_j}{f_S(s)} \int_0^s x_j e^{-\lambda_j x_j} \int_0^{s-x_j} x_i e^{-\lambda_i x_i} \sum_{k=1,k\neq i,j}^{n} \lambda_k A_k e^{-\lambda_k(s-x_i-x_j)} dx_i dx_j \\
&= \frac{\lambda_i\lambda_j}{f_S(s)} \int_0^s x_j e^{-\lambda_j x_j} \sum_{k=1,k\neq i,j}^{n} \lambda_k e^{-\lambda_k(s-x_j)} \int_0^{s-x_j} x_i e^{-(\lambda_i-\lambda_k)x_i} dx_i dx_j
\end{aligned}$$

$$= \frac{\lambda_i\lambda_j}{f_S(s)} \int_0^s x_j e^{-\lambda_j x_j} \sum_{k=1,k\neq i,j}^{n} \lambda_k e^{-\lambda_k(s-x_j)} g_1(\lambda_i, \lambda_k, s - x_j) dx_j$$
$$= \frac{\lambda_i\lambda_j}{f_S(s)} \sum_{k=1,k\neq i,j}^{n} \lambda_k e^{-\lambda_k s} \int_0^s x_j e^{-(\lambda_j-\lambda_k)x_j} g_1(\lambda_i, \lambda_k, s - x_j) dx_j$$
$$= \frac{\lambda_i\lambda_j}{f_S(s)} \sum_{k=1,k\neq i,j}^{n} \lambda_k e^{-\lambda_k s} g_3(\lambda_i, \lambda_j, \lambda_k, s) \tag{3.10}$$

So the covariance $Cov(X_i, X_j|S_n = s) = \mathbb{E}[X_iX_j|S_n = s] - \mathbb{E}[X_i|S_n = s]\mathbb{E}[X_j|S_n = s]$. □

In this section, we have derived essential formulas including the CTE-induced premium pricing formula, the covariance formula, and other important derivative expressions, focusing specifically on scenarios where all λ_i's are distinct. These formulas constitute the foundational framework for the expessions under dependence assumption.

3.3 Examples

In this section, we introduce a model aimed at understanding dependencies among variables. Our approach builds upon the independence assumption, laying the groundwork for exploring dependent assumptions later on.

To create these dependencies, we construct a set of variables $\{X_1 + Z, X_2 + Z, X_3 + Z\}$, where $X_i \sim Exp(\lambda_i)$ for $i = 1, 2, 3$, and $Z \sim Exp(\lambda_0)$, with all X_i and Z being mutually independent. By introducing the common factor Z, we control interdependence, facilitating our observation of the trigger conditions for desired properties.

We define the total sum $S = \sum_{i=1}^{3} X_i + 3Z = S_X + 3Z$, where $3Z \sim Exp(\frac{3}{\lambda_0})$. To simplify notation, we denote $3Z$ as X_4 and $\frac{\lambda_0}{3}$ as λ_4.

Given the relatively small size of the portfolio, we opt to present all formulas in detail. To illustrate, we begin by defining the probability density function (pdf) for summation as follows:

$$f_S(s) = \sum_{i=4}^{4} A_{i,4}\lambda_i e^{-\lambda_i s}$$
$$= \frac{\lambda_1\lambda_2\lambda_3\lambda_4}{(\lambda_2 - \lambda_1)(\lambda_3 - \lambda_1)(\lambda_4 - \lambda_1)} e^{-\lambda_1 s} + \frac{\lambda_1\lambda_2\lambda_3\lambda_4}{(\lambda_1 - \lambda_2)(\lambda_3 - \lambda_2)(\lambda_4 - \lambda_2)} e^{-\lambda_2 s}$$

$$+\frac{\lambda_1\lambda_2\lambda_3\lambda_4}{(\lambda_1-\lambda_3)(\lambda_2-\lambda_3)(\lambda_4-\lambda_3)}e^{-\lambda_3 s}+\frac{\lambda_1\lambda_2\lambda_3\lambda_4}{(\lambda_1-\lambda_4)(\lambda_2-\lambda_4)(\lambda_3-\lambda_4)}e^{-\lambda_4 s} \tag{3.11}$$

In the next step, without loss of generality, we find the expectation of X_1 given that the portfolio's sum is fixed. Specifically, we aim to compute $\mathbb{E}[X_1|S=s]$ rather than $\mathbb{E}[X_1+Z|S=s]$ for ease of computation with the joint probability density function (pdf), to work with the joint pdf of X_1 and X_2+X_3+3Z is more straightforward than that of X_1+Z and X_2+X_3+2Z. Thus, we have:

$$\begin{aligned} f_{X_1,X_2+X_3+3Z}(x_1,x_2+x_3+3z) &= f_{X_1}(x_1)\cdot f_{X_2+X_3+X_4}(x_2+x_3+X_4) \\ &= \lambda_1 e^{-\lambda_1 x_1}\cdot\sum_{j=2}^{4} A_{j,4}^{\bar{1}} e^{-\lambda_j X_j} \end{aligned}$$

We will use the derived joint pdf to determine $\mathbb{E}[X_1|S=s]$ and $\mathbb{E}[X_1|S>s]$. The same methodology is used to determine $\mathbb{E}[Z|S=s]$ and $\mathbb{E}[Z|S>s]$.

$$\begin{aligned} \mathbb{E}[X_1|S=s] &= \frac{1}{f_S(s)}\int_0^s x_1\lambda_1 e^{-\lambda_1 x_1}\sum_{j=2}^{4} A_j^{\bar{1}}\lambda_j e^{-\lambda_j(s-x_1)}dx_1 \\ &= \frac{\lambda_1}{f_S(s)}\sum_{j=2}^{4} A_j^{\bar{1}}\lambda_j e^{-\lambda_j s} g_1(\lambda_1,\lambda_j,s) \end{aligned} \tag{3.12}$$

$$\mathbb{E}[X_1|S>s] = \lambda_1\sum_{j=2}^{4} A_j^{\bar{1}}\lambda_j g_2(\lambda_1,\lambda_j,s) \tag{3.13}$$

$$\begin{aligned} \mathbb{E}[Z|S=s] &= \frac{1}{3}\mathbb{E}[3Z|S=s] \\ &= \frac{1}{3}\mathbb{E}[X_4|S=s] \\ &= \frac{1}{3f_S(s)}\int_0^s x_4\lambda_4 e^{-\lambda_4 x_4}\sum_{j=1}^{3} A_j^{\bar{4}}\lambda_j e^{-\lambda_j(s-x_4)}dx_4 \\ &= \frac{\lambda_4}{3f_S(s)}\sum_{j=1}^{3} A_j^{\bar{4}}\lambda_j e^{-\lambda_j s} g_1(\lambda_4,\lambda_j,s) \end{aligned} \tag{3.14}$$

$$\mathbb{E}[Z|S>s] = \lambda_4\sum_{j=1}^{3} A_j^{\bar{4}}\lambda_j g_2(\lambda_4,\lambda_j,s) \tag{3.15}$$

The CTE-induced premium pricing formula for the first policyholder is defined as: $\mathbb{E}[X_1+Z|S>s]=\mathbb{E}[X_1|S>s]+\mathbb{E}[Z|S>s]$. It is derived from the summation of (3.13) and (3.15).

When we switch our attention to covariance, we can still group the loss variables due to the independence assumption. To define the expectation of the product of X_1 and X_2 , we will split the portfolio into three parts: X_1, X_2 and $\sum_{j=3}^4 X_4$.

$$
\begin{aligned}
&f_{X_1}(x_1)f_{X_2}(x_2)f_{S-X_1-X_2}(s-x_1-x_2)\\
=&\lambda_1 e^{-\lambda_1 x_1}\lambda_2 e^{-\lambda_2 x_2}\sum_{j=3}^{4} A_j^{\bar{1},\bar{2}}\lambda_j e^{-\lambda_j(s-x_1-x_2)}\\
=&\lambda_1\lambda_2\sum_{j=3}^{4} A_j^{\bar{1},\bar{2}}\lambda_j e^{-\lambda_j s}e^{-(\lambda_1-\lambda_j)x_1}e^{-(\lambda_2-\lambda_j)x_2}
\end{aligned}
\tag{3.16}
$$

Using the joint pdf to find the expectation, we have:

$$
\begin{aligned}
\mathbb{E}[X_1X_2|S=s] &= \frac{\lambda_1\lambda_2}{f_S(s)}\int_0^s\int_0^{s-x_1} x_1x_2\sum_{j=3}^{4} A_j^{\bar{1},\bar{2}}\lambda_j e^{-\lambda_j s}e^{-(\lambda_1-\lambda_j)x_1}e^{-(\lambda_2-\lambda_j)x_2}dx_2dx_1\\
&= \frac{\lambda_1\lambda_2}{f_S(s)}\int_0^s x_1\sum_{j=3}^{4} A_j^{\bar{1},\bar{2}}\lambda_j e^{-\lambda_j s}e^{-(\lambda_1-\lambda_j)x_1}\int_0^{s-x_1} x_2 e^{-(\lambda_2-\lambda_j)x_2}dx_2dx_1\\
&= \frac{\lambda_1\lambda_2}{f_S(s)}\int_0^s x_1\sum_{j=3}^{4} A_j^{\bar{1},\bar{2}}\lambda_j e^{-\lambda_j s}e^{-(\lambda_1-\lambda_j)x_1}g_1(\lambda_2,\lambda_j,s-x_1)dx_1\\
&= \frac{\lambda_1\lambda_2}{f_S(s)}\sum_{j=3}^{4} A_j^{\bar{1},\bar{2}}\lambda_j e^{-\lambda_j s}\int_0^s x_1 e^{-(\lambda_1-\lambda_j)x_1}g_1(\lambda_2,\lambda_j,s-x_1)dx_1\\
&= \frac{\lambda_1\lambda_2}{f_S(s)}\sum_{j=3}^{4} A_j^{\bar{1},\bar{2}}\lambda_j e^{-\lambda_j s}g_3(\lambda_2,\lambda_1,\lambda_j,s)
\end{aligned}
\tag{3.17}
$$

$$
\begin{aligned}
\mathbb{E}[X_1Z|S=s] &= \frac{1}{3}\mathbb{E}[X_1X_4|S=s]\\
&= \frac{1}{3f_S(s)}\int_0^s\int_0^{s-x_1} x_1x_4 f_{X_1}(x_1)f_{X_4}(x_4)f_{X_2+X_3}(s-x_1-x_4)dx_4dx_1\\
&= \frac{\lambda_1\lambda_4}{3f_S(s)}\int_0^s\int_0^{s-x_1} x_1x_4\sum_{j=2}^{3} A_j^{\bar{1},\bar{4}}\lambda_j e^{-\lambda_j s}e^{-(\lambda_1-\lambda_j)x_1}e^{-(\lambda_4-\lambda_j)x_4}dx_4dx_1\\
&= \frac{\lambda_1\lambda_4}{3f_S(s)}\int_0^s x_1\sum_{j=2}^{3} A_j^{\bar{1},\bar{4}}\lambda_j e^{-\lambda_j s}e^{-(\lambda_1-\lambda_j)x_1}\int_0^{s-x_1} x_4 e^{-(\lambda_4-\lambda_j)x_4}dx_4dx_1\\
&= \frac{\lambda_1\lambda_4}{3f_S(s)}\sum_{j=2}^{3} A_j^{\bar{1},\bar{4}}\lambda_j e^{-\lambda_j s}g_3(\lambda_4,\lambda_1,\lambda_j,s)
\end{aligned}
\tag{3.18}
$$

Combining all the previous results, we can derive the covariance between X_1 and X_2 by the definition, $Cov[X_1, X_2|S = s] = \mathbb{E}[X_1X_2|S = s] - \mathbb{E}[X_1|S = s]\mathbb{E}[X_2|S = s]$. And the covariance between X_1 and Z is $Cov(X_1, Z|S = s) = \mathbb{E}[X_1Z|S = s] - \mathbb{E}[X_1|S = s]\mathbb{E}[Z|S = s]$. The last needed piece is the variance of Z.

$$
\begin{aligned}
Var(Z|S = s) &= \mathbb{E}[Z^2|S = s] - \Big(\mathbb{E}[Z|S = s]\Big)^2 \\
&= \frac{1}{9}\mathbb{E}[X_4^2|S = s] - \Big(\frac{1}{3}\mathbb{E}[X_4|S = s]\Big)^2 \\
&= \frac{\lambda_4 \sum_{j\neq 4, j=1}^{3} \lambda_j A_j^{\bar{4}} e^{-\lambda_j s}}{9 f_S(s)} h_1(\lambda_4, \lambda_j, s) - \Big(\frac{\lambda_4}{3 f_S(s)} \sum_{j=1}^{3} A_j^{\bar{4}} \lambda_j e^{-\lambda_j s} g_1(\lambda_4, \lambda_j, s)\Big)^2
\end{aligned}
\tag{3.19}
$$

Hence, we have all the pieces needed to construct the covariance between $X_1 + Z$ and $X_2 + Z$. The covariance formula is given by:

$$
\begin{aligned}
&Cov[X_1 + Z, X_2 + Z|S = s] \\
=&Cov[X_1, X_2|S = s] + Cov[X_1, Z|S = s] + Cov[X_2, Z|S = s] + Var[Z|S = s]
\end{aligned}
$$

To gain insights into whether the diversification property is satisfied, we need to determine if the covariance is negative. We will examine the calculation results by using specific values of λ_i for $i = 0, 1, 2, 3$. The parameter values 0.1, 0.2, 0.3, and 0.4 will be rotated and observed. Specifically, when $\lambda_0 = 0.3$, it will result in a situation where the denominator becomes zero. To avoid this issue, we will use $\lambda_0 = 0.31$ instead.

λ_1	λ_2	λ_3	λ_0	$VaR_{95\%}(S)$	ρ_{X_1+Z, X_2+Z}
0.1	0.2	0.3	0.4	77.3898	-0.6039
0.1	0.3	0.2	0.4	77.3898	-0.5237
0.2	0.3	0.4	0.1	122.3257	-0.3498
0.3	0.4	0.1	0.2	92.3684	0.6635
0.4	0.1	0.2	0.31	81.4163	-0.5369
0.4	0.3	0.2	0.31	61.4447	0.0629
0.4	0.3	0.2	0.1	122.3257	-0.0463
0.4	0.2	0.3	0.31	61.4447	-0.3882

Table 3.1: Values of $VaR_{95\%}(S)$ and $Cov(X_1 + Z, X_2 + Z)$ for various λ_i.

In the model, we are using Z as the common factor, and the positivity of covariance between variables is mainly determined by which parameter is dominant in the model. Several observations can be made based on the parameter values: (1) When the parameter of Z is relatively large, meaning that λ_0 is greater than or not significantly less than λ_1 and λ_2, we are more likely to observe a negative correlation coefficient between the first two loss variables. This is reasonable because the covariance is derived under the assumption that the loss summation equals the 95% VaR. (2) When λ_0 is relatively small compared to λ_1 and λ_2, the influence of Z is minimized, and we are more likely to observe a positive covariance. Under these conditions, the common factor Z has less impact on the overall correlation structure.

Chapter 4

Pareto Distributions

Since it was first proposed by the Italian economist Vilfredo Pareto, the Pareto distribution has been an appropriate description for the distribution of wealth, natural phenomena, and other social and scientific phenomena that exhibit the long-tail characteristic. This characteristic implies that while a large number of small outcomes are common, a small number of extreme outcomes are rare. Because of this, the Pareto distribution is particularly significant in actuarial science. In the insurance industry, for instance, a small number of claims can often account for a large portion of the total payout, and the heavy-tail nature of the Pareto distribution effectively captures this phenomenon. Modeling the tail of the loss distribution also aids in risk assessment, reserve allocation, and, as discussed in this dissertation, setting premiums.

Compared to the exponential distribution assumption we made in the previous chapter, the Pareto distribution captures the heavy-tail characteristics that the exponential distribution lacks, providing a more suitable framework for modeling rare and extreme events in insurance and actuarial science.

4.1 Introduction and Stochastic Representation

As time has progressed, the ability to describe the Pareto distribution in greater detail has improved by introducing additional parameters into the model. For the completeness, we will begin with the definition of four types of Pareto distributions using the notations provided by Arnold (2014), increasing in complexity. We will briefly discuss the significance of the parameters and how they evolve from one type to another.

To begin with, the Type I Pareto distribution, denoted as $P(\mathrm{I})(\sigma, \alpha)$, has a survival function defined by

$$S(x) = \left(\frac{x}{\sigma}\right)^{-\alpha}$$

for $x \geq \sigma$. In this definition, σ is the scale parameter, which determines the minimum possible value, and α is the shape parameter, which determines the steepness of the distribution.

If the variables need to be shifted by a level, the location parameter is introduced in the distribution, resulting in the Type II Pareto distribution. The survival function is given by

$$S(x) = \left(1 + \frac{x-\mu}{\sigma}\right)^{-\alpha}$$

for $x \geq \mu$, denoted as $P(\mathrm{II})(\mu, \sigma, \alpha)$. Most of the time, μ is positive because the Pareto distribution commonly describes wealth or income, and it is reasonable to assume these quantities are positive. However, in some special situations, negative income is also possible in the real world. Specifically, if we take the location parameter μ to be the same as the scale parameter σ, the two types of Pareto distributions would coincide.

Another way to capture the similar tail characteristics of the Type II Pareto distribution is provided by the Type III Pareto distribution. The survival function is given by

$$S(x) = \left(1 + \left(\frac{x-\mu}{\sigma}\right)^{\frac{1}{\gamma}}\right)^{-1}$$

for $x \geq \mu$, denoted as $P(\mathrm{III})(\mu, \sigma, \gamma)$. In addition to the location and scale parameters, this distribution introduces γ, called the inequality parameter. Similar to the shape parameter α, the inequality parameter γ also defines the tail shape of the distribution. As γ increases, more extreme values are observed.

To incorporate both the shape parameter and the inequality parameter into a single expression, we use the Type IV Pareto distribution. The survival function is given by

$$S(x) = \left(1 + \left(\frac{x-\mu}{\sigma}\right)^{\frac{1}{\gamma}}\right)^{-\alpha}$$

denoted as $P(\text{IV})(\mu, \sigma, \gamma, \alpha)$. Here, $\mu \in \mathbb{R}$, $\sigma > 0$, $\gamma > 0$, $\alpha > 0$, and $x \geq \mu$. The Type IV Pareto distribution is the most comprehensive definition, as the Type I, II, and III Pareto distributions can all be expressed as special cases of the Type IV Pareto distribution:

$$\begin{aligned} P(\text{I})(\sigma, \alpha) &= P(\text{IV})(\sigma, \sigma, 1, \alpha) \\ P(\text{II})(\mu, \sigma, \alpha) &= P(\text{IV})(\mu, \sigma, 1, \alpha) \\ P(\text{III})(\mu, \sigma, \gamma) &= P(\text{IV})(\mu, \sigma, \gamma, 1) \end{aligned}$$

We will utilize the Pareto Type II distribution in this chapter due to its stochastic representation derived from the exponential distribution.

Proposition 4.1.1 (Pareto Type II as a Log-Exponential Distribution)**.** *Let $V \sim Exp(1)$. Then, $X = \mu + \sigma(e^{\frac{V}{\alpha}} - 1) \sim P(\text{II})(\mu, \sigma, \alpha)$. In particular, $Y = \sigma e^{\frac{V}{\alpha}} \sim P(\text{I})(\sigma, \alpha)$.*

Proof.

$$\begin{aligned} \mathbb{P}(X > x) &= \mathbb{P}\left(e^{\frac{V}{\alpha}} - 1 > \frac{x - \mu}{\sigma}\right) \\ &= \mathbb{P}\left(e^{\frac{V}{\alpha}} > \frac{x - \mu}{\sigma} + 1\right) \\ &= \mathbb{P}\left(V > \alpha \ln\left(1 + \frac{x - \mu}{\sigma}\right)\right) \\ &= \exp\left(-\alpha \ln\left(1 + \frac{x - \mu}{\sigma}\right)\right) \\ &= \left(1 + \frac{x - \mu}{\sigma}\right)^{-\alpha} \end{aligned}$$

which is the survival function of $P(\text{II})(\mu, \sigma, \alpha)$.

In particular, define a random variable $Y = \sigma e^{\frac{V}{\alpha}}$, we have

$$\begin{aligned} \mathbb{P}(Y > y) &= \mathbb{P}\left(\sigma e^{\frac{V}{\alpha}} > y\right) \\ &= \mathbb{P}\left(V > \alpha \ln\left(\frac{y}{\sigma}\right)\right) \\ &= \exp\left(-\alpha \ln\left(\frac{y}{\sigma}\right)\right) \end{aligned}$$

$$= \left(\frac{y}{\sigma}\right)^{-\alpha}$$

which is the survival function of $P(\mathrm{I})(\sigma, \alpha)$. □

The advantage of using the stochastic representation related to the exponential distribution is that when deriving the distribution of the summation, it is natural to construct a similar structure using the gamma distribution.

Proposition 4.1.2. *Assume* $X_i \overset{i.i.d}{\sim} Exp(\lambda_i)$ *and* $V \sim Gamma(\alpha, \lambda_0)$*,* V *is independent with any* X_i*, then* $\frac{X_1}{V} \sim P(\mathrm{II})(0, \frac{\lambda_0}{\lambda_i}, \alpha)$ *and* $S = \frac{\sum_{i=1}^n X_i}{V}$ *follows a mixture Pareto distribution which is the summation of* $P(\mathrm{II})(0, \frac{\lambda_0}{\lambda_i}, \alpha)$ *with the weight of* A_i*.*

Proof. Let $\frac{X_i}{V} = Y_i$, $\sum_{i=1}^n X_i = S_X$ and the probability density function of S_X is given by $f_{S_X}(s_X) = \sum_{i=1}^n A_i \lambda_i e^{-\lambda_i s_X}$. Using the results of Chapter 3, the following steps can be justified:

$$\begin{aligned}
\mathbb{P}(S > s) &= \mathbb{P}\left(\frac{S_X}{V} > s\right) \\
&= \mathbb{P}(S_X > sV) \\
&= \int_0^\infty \int_{sv}^\infty f_{S_X}(s_X) dS_X f_V(v) dv \\
&= \int_0^\infty \int_{sv}^\infty \sum_{i=1}^n A_i \lambda_i e^{-\lambda_i S_X} dS_X \frac{(\lambda_0 v)^\alpha e^{-\lambda_0 v}}{v\Gamma(\alpha)} dv \\
&= \int_0^\infty \left[\sum_{i=1}^n A_i \lambda_i \frac{e^{-\lambda_i S_X}}{-\lambda_i}\Bigg|_{sv}^\infty\right] \frac{(\lambda_0 v)^\alpha e^{-\lambda_0 v}}{v\Gamma(\alpha)} dv \\
&= \int_0^\infty \sum_{i=1}^n A_i e^{-\lambda_i sv} \frac{(\lambda_0 v)^\alpha e^{-\lambda_0 v}}{v\Gamma(\alpha)} dv \\
&= \sum_{i=1}^n \frac{A_i \lambda_0^\alpha}{\Gamma(\alpha)} \int_0^\infty v^{\alpha-1} e^{-(\lambda_0 + \lambda_i s)v} dv \\
&= \sum_{i=1}^n \frac{A_i \lambda_0^\alpha}{\Gamma(\alpha)} \frac{\Gamma(\alpha)}{(\lambda_0 + \lambda_i s)^\alpha} \\
&= \sum_{i=1}^n A_i \frac{\lambda_0^\alpha}{(\lambda_0 + \lambda_i s)^\alpha} \\
&= \sum_{i=1}^n A_i \left(1 + \frac{s}{\frac{\lambda_0}{\lambda_i}}\right)^{-\alpha}
\end{aligned}$$

Then the pdf of S will be

$$f_S(s) = \sum_{i=1}^{n} A_i \frac{\alpha \lambda_0^\alpha \lambda_i}{(\lambda_0 + \lambda_i s)^{\alpha+1}}$$

□

4.2 CTE Premiums

Consider a portfolio of independent loss random variables $\{Y_1, Y_2, ..., Y_n\}$ and denote the total loss as S. Then, the conditional distribution of any variable Y_i given that S equals to some predetermined value s can be derived as:

$$\begin{aligned}
f_{Y_i|S}(y_i|s) &= \frac{f_{Y_i}(y_i) f_{\sum_{j\neq i, j=1}^{n} Y_j}(s-y_i)}{f_S(s)} \\
&= \frac{\frac{\alpha\lambda_0\lambda_i}{(\lambda_0+\lambda_i y_i)^{\alpha+1}} \sum_{j=1, j\neq i}^{n} \frac{\alpha A_j^{\bar{i}} \lambda_0 \lambda_j}{(\lambda_0+\lambda_j(s-y_i))^{\alpha+1}}}{\sum_{k=1}^{n} \frac{\alpha A_i \lambda_0 \lambda_k}{(\lambda_0+\lambda_k s)^{\alpha+1}}} \\
&= \frac{\alpha^2 \lambda_0 \lambda_i \sum_{j=1, j\neq i}^{n} \frac{A_j^{\bar{i}} \lambda_j}{(\lambda_0+\lambda_i y_i)^{\alpha+1}(\lambda_0+\lambda_j(s-y_i))^{\alpha+1}}}{\sum_{k=1}^{n} \frac{\alpha A_i \lambda_k}{(\lambda_0+\lambda_k s)^{\alpha+1}}}
\end{aligned}$$

After this, the conditional expectation can be derived by straightforward integration.

$$\begin{aligned}
\mathbb{E}[Y_i|S=s] &= \int_0^s y_i f_{Y_i|S}(y_i|s) dy_i \\
&= \int_0^s y_i \frac{\alpha^2 \lambda_0 \lambda_i \sum_{j=1, j\neq i}^{n} \frac{A_j^{\bar{i}} \lambda_j}{(\lambda_0+\lambda_i y_i)^{\alpha+1}(\lambda_0+\lambda_j(s-y_i))^{\alpha+1}}}{\sum_{k=1}^{n} \frac{\alpha A_i \lambda_k}{(\lambda_0+\lambda_k s)^{\alpha+1}}} dy_i \\
&= \frac{\alpha^2 \lambda_0 \lambda_i}{\sum_{k=1}^{n} \frac{\alpha A_i \lambda_k}{(\lambda_0+\lambda_k s)^{\alpha+1}}} \sum_{j=1, j\neq i}^{n} A_j^{\bar{i}} \lambda_j \int_0^s \frac{y_i}{(\lambda_0+\lambda_i y_i)^{\alpha+1}(\lambda_0+\lambda_j(s-y_i))^{\alpha+1}} dy_i \quad (4.1)
\end{aligned}$$

When examining the integrand in the final step, $\frac{y_i}{(\lambda_0+\lambda_i y_i)^{\alpha+1}(\lambda_0+\lambda_j(s-y_i))^{\alpha+1}}$, if the exponent term α is an integer, then the term can be decomposed using partial fraction decomposition into a summation of simpler fraction terms such as $\frac{1}{(a+b)^n}$ for some a, b, and integer n. We explore this process and derive a formula accordingly.

Definition 4.2.1. For $p, q \geq 1$, $p, q \in \mathbb{Z}$, define

$$m_x(p,q) = \frac{1}{(a+x)^p(b-x)^q},$$

where a, b are any real constants.

Proposition 4.2.2. *Assume* $\binom{0}{0} = 1$ *and* $\binom{s}{t} = 0$ *for any* $s < t$*,* $m_x(p,q)$ *can be decomposed by an iteration form such as*

$$\begin{aligned} m_x(p,q) = &\sum_{i=1}^{p-1} \frac{1}{(a+b)^{q+i}} \sum_{j=0}^{i-1} \binom{q}{j+1}\binom{i-1}{j} m_x(p-i,0) + \frac{1}{(a+b)^q} m_x(p,0) \\ &+ \sum_{i=1}^{q-1} \frac{1}{(a+b)^{p+i}} \sum_{j=0}^{i-1} \binom{p}{j+1}\binom{i-1}{j} m_x(0,q-i) + \frac{1}{(a+b)^p} m_x(0,q) \end{aligned} \tag{4.2}$$

The result can be proven by mathematical induction, and the proof will be detailed in the appendix. This iterative method not only applies to fractions with a constant numerator but can also be extended to fractions with a variable numerator after some algebraic transformations. For example, consider x and x^2, then the breakdown process is as follows:

$$\begin{aligned} \frac{x}{(a+x)^p(b-x)^q} &= \frac{a+x-a}{(a+x)^p(b-x)^q} \\ &= \frac{1}{(a+x)^{p-1}(b-x)^q} - \frac{a}{(a+x)^p(b-x)^q} \\ &= m_x(p-1,q) - a m_x(p,q) \\ \frac{x^2}{(a+x)^p(b-x)^q} &= \frac{(a+x)^2 - 2a(a+x) + a^2}{(a+x)^p(b-x)^q} \\ &= \frac{1}{(a+x)^{p-2}(b-x)^q} - \frac{2a}{(a+x)^{p-1}(b-x)^q} + \frac{a^2}{(a+x)^p(b-x)^q} \\ &= m_x(p-2,q) - 2a m_x(p-1,q) + a^2 m_x(p,q) \end{aligned}$$

The previous expression can be further decomposed into a linear combination of $m_x(i,0)$ and $m_x(0,j)$ for $i = 1, 2, ..., p$ and $j = 1, 2, ..., q$. It is evident that every term comprising the final decomposition result is integral, as follows:

$$\int_0^s m(p,0)dx = \begin{cases} \ln|a+x| - \ln|a| & \text{if } p=1 \\ \frac{1}{1-p}\left[(a+s)^{1-p} - a^{1-p}\right] & \text{if } p \geq 2 \end{cases}$$

$$\int_0^s m(0,q)dx = \begin{cases} \ln|b| - \ln|b-s| & \text{if } p=1 \\ \frac{1}{q-1}\left[(b-s)^{1-q} - b^{1-q}\right] & \text{if } p \geq 2 \end{cases}$$

So we can have the explicit expression of $\mathbb{E}[Y_i|S=s]$ by substituting $a = \frac{\lambda_0}{\lambda_i}$, $b = \frac{\lambda_0}{\lambda_j} + s$, $\alpha > 2$:

$$\mathbb{E}[Y_i|S=s] = \frac{\alpha^2\lambda_0\lambda_i}{\sum_{k=1}^n \frac{\alpha A_i\lambda_k}{(\lambda_0+\lambda_k s)^{\alpha+1}}} \sum_{j=1,j\neq i}^n A_j^{\bar{i}}\lambda_i^{\alpha+1}\lambda_j^{\alpha+2}\theta_1(a,b,s) \tag{4.3}$$

Where $\theta_1(a,b,s)$ is the full explicit integral result, and due to its length, it will be provided in the appendix. The second moment expectation can be derived after one step breakdown.

$$\begin{aligned} \mathbb{E}[Y_i^2|S=s] &= \int_0^s y_i^2 f_{Y_i|S}(y_i|s)dy_i \\ &= \frac{\alpha^2\lambda_0\lambda_i}{\sum_{k=1}^n \frac{\alpha A_i\lambda_k}{(\lambda_0+\lambda_k s)^{\alpha+1}}} \sum_{j=1,j\neq i}^n A_j^{\bar{i}}\lambda_j \int_0^s \frac{y_i^2}{(\lambda_0+\lambda_i y_i)^{\alpha+1}(\lambda_0+\lambda_j(s-y_i))^{\alpha+1}}dy_i \\ &= \frac{\alpha^2\lambda_0\lambda_i}{\sum_{k=1}^n \frac{\alpha A_i\lambda_k}{(\lambda_0+\lambda_k s)^{\alpha+1}}} \sum_{j=1,j\neq i}^n A_j^{\bar{i}}\lambda_j\theta_2(a,b,s) \end{aligned} \tag{4.4}$$

Where $\theta_2(a,b,s)$ is the full explicit integral result, and will be provided in the appendix.

By the variance definition, it's straightforward to show that

$$Var[Y_i|S=s] = \frac{\alpha^2\lambda_0\lambda_i}{f_S(s)} \sum_{j=1,j\neq i}^n A_j^{\bar{i}}\lambda_j\theta_2(a,b,s) - \left[\frac{\alpha^2\lambda_0\lambda_i}{f_S(s)} \sum_{j=1,j\neq i}^n A_j^{\bar{i}}\lambda_j\theta_1(a,b,s)\right]^2,$$

where $f_S(s) = \sum_{k=1}^n \frac{\alpha A_i\lambda_k}{(\lambda_0+\lambda_k s)^{\alpha+1}}$.

An issue arises when attempting to evaluate $\mathbb{E}[Y_i|S>s]$, as the integral $\int_s^\infty \ln|a+t|dt$

diverges. Given this limitation, this chapter will focus solely on the condition where $S = s$.

In order to introduce these dependencies, we intend to replicate a similar approach to that outlined in Chapter 3. This involves constructing a portfolio comprising Y_1, Y_2, and Y_3, where each Y_i is defined as $Y_i = \frac{X_i+Z}{V}$. Here, $X_i \sim Exp(\lambda_i)$, $Z \sim Exp(\lambda_0)$, and $V \sim Gamma(\alpha, \lambda_0)$ for $i = 1, 2, 3$, with all X_i and Z being mutually independent. The total sum is then given by $S = \sum_{i=1}^{3} \frac{X_i+Z}{V} = \frac{S_X+3Z}{V}$, where $S_x = \sum_{i=1}^{3} X_i$. This allows us to establish the joint pdf as follows:

$$\begin{aligned}
&f_{Y_i,Y_j,S-Y_i-Y_j}(y_i, y_j, s - y_i - y_j) \\
=&f_{Y_i}(y_i)f_{y_j}(y_j)f_{S-Y_i-Y_j}(s - y_i - y_j) \\
=&\frac{\alpha\lambda_0\lambda_i}{(\lambda_0 + \lambda_i y_i)^{\alpha+1}} \frac{\alpha\lambda_0\lambda_j}{(\lambda_0 + \lambda_j y_j)^{\alpha+1}} \sum_{k=1,k\neq i,j}^{n} A_k^{\overline{i,j}} \frac{\alpha\lambda_0^{\alpha}\lambda_k}{(\lambda_0 + \lambda_k(s - y_i - y_j))^{\alpha+1}}
\end{aligned} \tag{4.5}$$

Subsequently, our analysis will involve calculating $\mathbb{E}[Y_iY_j|S = s]$,

$$\begin{aligned}
&\mathbb{E}[Y_iY_j|S = s] \\
=&\frac{\alpha^3\lambda_0^{2+\alpha}\lambda_i\lambda_j\lambda_k}{f_S(s)} \sum_{k=1,k\neq i,j}^{n} A_k^{\overline{i,j}} \int_0^s \frac{y_i}{(\lambda_0 + \lambda_j y_j)^{\alpha+1}} \lambda_i^{\alpha+1}\lambda_k^{\alpha+1}\theta_1(\frac{\lambda_0}{\lambda_i}, \frac{\lambda_0}{\lambda_k} + s, s - y_j)dy_j
\end{aligned} \tag{4.6}$$

However, in our attempt to establish the formula, we encountered a similar challenge when computing $\mathbb{E}[Y_i|S > s]$. Specifically, we encountered divergence in the integral of θ_1. This divergence poses a significant obstacle in our analysis, requiring further investigation and potentially alternative approaches to address it effectively.

Chapter 5

Conclusions and Future Research

5.1 Summary

In this dissertation, we have developed a premium pricing framework aimed at controlling insolvency risk when loss variables are interdependent. Traditionally, insurance pricing assumes all loss variables are independent, allowing the application of the Central Limit Theorem and the Law of Large Numbers to ensure the total loss of a group of insureds does not exceed a predetermined level. However, this independence assumption is increasingly challenged by the rising occurrence of cyber-attacks and natural disasters, introducing significant uncertainty into the premium pricing model due to the dependence between variables.

To address this issue, we adopted the concept of the Conditional Tail Expectation (CTE) allocation principle from capital allocation principles. Individual premiums are made proportional to their covariance with the total loss. From this idea, we proposed a premium pricing formula based on the expectation of a specific loss variable, given that the total loss exceeds a predetermined level. To ensure the fairness of premium allocation, we posited that the framework should satisfy two desirable properties: diversification and monotonicity.

The diversification property ensures that adding one more policyholder to the portfolio will not increase the premiums of all existing policyholders, implying that the framework is not adversely affected by the expansion of the plan. The monotonicity property ensures that if one policyholder becomes riskier, other policyholders' premiums will not increase, protecting low-risk insureds from the risks posed by higher-risk insureds. We derived the corresponding mathematical inequalities to reflect these properties.

We assumed the risk variables follow normal, exponential, and Pareto distributions. Under the normal distribution assumption, we utilized the formula established by Landsman and Valdez (2003), expanding our observations by applying different numerical settings. Under the exponential distribution assumption, we first established the distribution of the summation of independent exponential random variables, derived the CTE-induced premium pricing formula, and finally calculated the covariance between two random variables given that the summation exceeds a certain level. To incorporate dependence, we designed an example where each policyholder's loss random variable is the sum of distinct exponential random variables and a common exponential random variable. The covariance was determined using previous discoveries and numerical explorations. Under the Pareto distribution assumption, we used the stochastic representation of the Type II Pareto distribution to derive the summation of independent variables and the expectation of a specific variable given that the summation exceeds a predetermined level. The exact CTE-induced premium pricing formula was not achieved due to the divergence of the resulting integral.

When considering the feasibility of a concept, it is essential to evaluate both its theoretical and practical applications. This dissertation has focused primarily on the theoretical feasibility of the CTE-induced premium pricing framework.

5.2 Future work

The first potential enhancement for this dissertation lies in the continued exploration of the premium pricing formula under the Pareto distribution assumption. As discussed in Chapter 4, the proposed CTE-induced formula has not been explicitly established due to the divergence of the integral of the natural logarithm terms. This necessitates further algebraic efforts or possibly a solution outside the realm of analytical approaches.

Additionally, conducting a simulation study to cross-check the derived explicit formula could yield valuable insights. Such a study might uncover new patterns or relationships that are not apparent from the theoretical analysis alone.

Another significant potential for this dissertation is its application to real-world data. Real-data applications introduce numerous factors, such as the mixture of distributions, deductibles, and policy limits, which must be considered. Testing the proposed formula with real-world data would provide a more comprehensive understanding of its practical utility and limitations.

Overall, while this dissertation has focused on identifying the preconditions for the desired properties—monotonicity and diversification—it has primarily emphasized trends in the data rather than providing extensive numerical conclusions. The complexity of the formulas and the construction of sample portfolios present challenges for deeper analytical analysis. Future research could further refine the premium pricing formula by continuing to explore the Pareto distribution, conducting simulation studies, and applying the model to real-world data.

Bibliography

[1] Arnold, B.C. (2014). *Pareto Distributions*. 2nd edition. Chapman & Hall.

[2] Check Point, *Check Point Research: Cyber Attacks Increased 50% Year over Year*
`https://blog.checkpoint.com/2022/01/10/check-point-research-cyber-attacks-increased-50-year-over-year/`
(Last accessed: June 3, 2024)

[3] Denault, M. (2001). Coherent allocation of risk capital. *Journal of Risk*, **4**(1), 1-34.

[4] Denuit, M., Dhaene, Goovaerts, M. and Kaas, R. (2006). *Actuarial Theory for Dependent Risks: Measures, Orders and Models*. Wiley.

[5] Dhaene, J., Henrard, L., Landsman, Z., Vandendorpe, A. and Vanduffel, S. (2008) Some results on the CTE-based capital allocation rule. *Insurance: Mathematics and Economics*, **42**(2), 855-863.

[6] Dhaene, J., Tsanakas, A., Valdez, E.A., and Vanduffel, S. (2012). Optimal capital allocation principles. *Journal of Risk and Insurance*, **79**(1), 1-28.

[7] Dickson, D.C.M., Hardy, M.R., Waters, H. (2020). *Actuarial Mathematics for Life Contingent Risks*. 3rd edition. Cambridge University Press.

[8] Embroker, *2024 Must-Know Cyber Attack Statistics and Trends.*
`https://www.embroker.com/blog/cyber-attack-statistics/`
(Last accessed: June 3, 2024)

[9] *Loss Data Analytics* (2022). Frees, E.W. (editor). *https://openacttexts.github.io/Loss-Data-Analytics/*. (Last accessed: June 5, 2024)

[10] Kusuma, A., Nguyen, C., and Noy, I. (2019). Insurance for Catastrophes: Why Are Natural Hazards Underinsured, and Does It Matter?. In: Okuyama, Y., Rose, A. (eds). *Advances in Spatial and Economic Modeling of Disaster Impacts*. Springer.

[11] Landsman, Z.M., and Valdez, E.A. (2003). Tail conditional expectations for elliptical distributions. *North American Actuarial Journal*, **7**(4), 55-71.

[12] Liu, Z., Wei, W., Wang, L., Ten, C.W., and Rho, Y. (2020). An actuarial framework for power system reliability considering cybersecurity threats. *IEEE Transactions on Power Systems*, **36**(2), 851-864.

[13] Ross, S.M. (2019). *Introduction to Probability Models*. 12th edition. Elsevier Science.

[14] Wei, W., and Wang, L. (2023). Pricing interdependent insurance risks: TVaR premium allocation principle and its properties, *SSRN Electronic Journal*, **01**.

Appendices

Appendix A: Proposition 3.1.3 Formula Verification

To verify Proposition 3.1.3, assume that $\lambda_i = \lambda_j$ for all $i \neq j$. Let $\{X_1, X_2, \ldots, X_n\}$ be identically and independently exponentially distributed random variables such that $X_i \sim \text{Exp}\left(\frac{1}{\lambda}\right)$. Denote $S = \sum_{i=1}^{n} X_i$. Then $S \sim \text{Gamma}(\alpha = n, \beta = \frac{1}{\lambda})$. The survival function is given by:

$$\mathbb{P}(S > s) = e^{-\lambda s} \sum_{i=0}^{n-1} \frac{(\lambda s)^i}{i!}$$

For the case when $n = 2$, consider two portfolios. Portfolio 1 consists of $\{Y_1, Y_2\}$, where $Y_1 \sim \text{Exp}(\frac{1}{\lambda_1}$ and $Y_2 \sim \text{Exp}(\frac{1}{\lambda_2})$ with $\lambda_1 \neq \lambda_2$. Portfolio 2 consists of $\{Y_1', Y_2'\}$ where both variables follow $\text{Exp}(\frac{1}{\lambda_1})$. Define $\epsilon = \lambda_2 - \lambda_1$.

We analyze the limit as $\epsilon \to 0$:

$$\begin{aligned}
\lim_{\epsilon \to 0} \left(\frac{\lambda_1 + \epsilon}{\epsilon} e^{-\lambda_1 s} + \frac{\lambda_1}{-\epsilon} e^{-(\lambda_1 + \epsilon)s} \right) &= \lim_{\epsilon \to 0} \left(\frac{(\lambda_1 + \epsilon) e^{-\lambda_1 s} - \lambda_1 e^{-\lambda_1 s} e^{-\epsilon s}}{\epsilon} \right) \\
&= \lim_{\epsilon \to 0} \left(\frac{e^{-\lambda_1 s}(1 + s\lambda_1 e^{-\epsilon s})}{1} \right) \\
&= e^{-\lambda_1 s} + s\lambda_1 e^{-\lambda_1 s} \\
&= e^{-\lambda_1 s} \sum_{i=0}^{1} \frac{(\lambda_1 s)^i}{i!}
\end{aligned}$$

This limit holds for $n \geq 2$ and shows that when λ_i and λ_j are the same for all $i \neq j$, the distribution of the sum $S = \sum_{i=1}^{n} Y_i$ becomes a gamma distribution with parameter n. Therefore, Proposition 3.1.3 is verified.

Appendix B: Full Expression in (4.3) and (4.4)

The full expression of $\theta_1(a,b,s)$ in (4.3).

$$
\begin{aligned}
&\theta_1(a,b,s)\\
=&\int_0^s \frac{y_i}{(a+y_i)^{\alpha+1}(b-y_i)^{\alpha+1}}dy_i\\
=&\int_0^s \Big[m_{y_i}(\alpha,\alpha+1)-am_{y_i}(\alpha+1,\alpha+1)\Big]dy_i\\
=&\int_0^s \Bigg[\sum_{i=1}^{\alpha-1}\frac{1}{(a+b)^{\alpha+1+i}}\sum_{j=0}^{i-1}\binom{\alpha+1}{j+1}\binom{i-1}{j}m_{y_i}(\alpha-i,0)+\frac{1}{(a+b)^{\alpha+1}}m_{y_i}(\alpha,0)\\
&+\sum_{i=1}^{\alpha}\frac{1}{(a+b)^{\alpha+i}}\sum_{j=0}^{i-1}\binom{\alpha}{j+1}\binom{i-1}{j}m_{y_i}(0,\alpha+1-i)+\frac{1}{(a+b)^{\alpha}}m_{y_i}(0,\alpha+1)\\
&-\sum_{i=1}^{\alpha}\frac{a}{(a+b)^{\alpha+1+i}}\sum_{j=0}^{i-1}\binom{\alpha+1}{j+1}\binom{i-1}{j}[m_{y_i}(\alpha+1-i,0)+m_{y_i}(0,\alpha+1-i)]\\
&-\frac{a}{(a+b)^{\alpha+1}}[m_{y_i}(\alpha+1,0)+m_{y_i}(0,\alpha+1)]\Bigg]dy_i\\
=&\sum_{i=1}^{\alpha-2}\frac{1}{(a+b)^{\alpha+1+i}}\sum_{j=0}^{i-1}\binom{\alpha+1}{j+1}\binom{i-1}{j}\frac{1}{1-\alpha+i}[(a+s)^{1-\alpha+i}-a^{1-\alpha+i}]\\
&+\frac{1}{(a+b)^{2\alpha}}\sum_{j=0}^{\alpha-2}\binom{\alpha+1}{j+1}\binom{\alpha-2}{j}[ln|a+s|-ln|a|]+\frac{1}{(a+b)^{\alpha+1}}\frac{1}{1-\alpha}[(a+s)^{1-\alpha}-a^{1-\alpha}]\\
&+\sum_{i=1}^{\alpha-1}\frac{1}{(a+b)^{\alpha+i}}\sum_{j=0}^{i-1}\binom{\alpha}{j+1}\binom{i-1}{j}\frac{1}{\alpha-i}[(b-s)^{i-\alpha}-b^{i-\alpha}]\\
&+\frac{1}{(a+b)^{2\alpha}}\sum_{j=0}^{\alpha-1}\binom{\alpha}{j+1}\binom{\alpha-1}{j}[ln|b|-ln|b-s|]+\frac{1}{(a+b)^{\alpha}}\frac{1}{\alpha}[(b-s)^{-\alpha}-b^{-\alpha}]\\
&-\sum_{i=1}^{\alpha-1}\frac{a}{(a+b)^{\alpha+1+i}}\sum_{j=0}^{i-1}\binom{\alpha+1}{j+1}\binom{i-1}{j}\frac{1}{\alpha-i}[(b-s)^{i-\alpha}-b^{i-\alpha}-(a+s)^{i-\alpha}+a^{i-\alpha}]\\
&-\frac{a}{(a+b)^{2\alpha+1}}\sum_{j=0}^{\alpha-1}\binom{\alpha+1}{j+1}\binom{\alpha-1}{j}\frac{1}{\alpha}[ln|a+s|-ln|a|+ln|b|-ln|b-s|]\\
&-\frac{a}{(a+b)^{\alpha+1}}\frac{1}{\alpha}[(b-s)^{-\alpha}-b^{-\alpha}-(a+s)^{-\alpha}+a^{-\alpha}]
\end{aligned}
$$

The full expression of $\theta_2(a,b,s)$ in (4.4).

$$
\begin{aligned}
&\theta_2(a,b,s)\\
=&\int_0^s \left[m_{y_i}(\alpha-1,\alpha+1)-2am_{y_i}(\alpha,\alpha+1)+a^2m_{y_i}(\alpha+1,\alpha+1)\right]dy_i\\
=&\sum_{i=1}^{\alpha-3}\frac{1}{(a+b)^{\alpha+1+i}}\sum_{j=0}^{i-1}\binom{\alpha+1}{j+1}\binom{i-1}{j}\frac{1}{2-\alpha+i}[(a+s)^{2-\alpha+i}-a^{2-\alpha+i}]\\
&+\frac{1}{(a+b)^{2\alpha}-1}\sum_{j=0}^{\alpha-3}\binom{\alpha+1}{j+1}\binom{\alpha-3}{j}[ln|a+s|-ln|a|]+\frac{1}{(a+b)^{\alpha+1}}\frac{1}{2-\alpha}[(a+s)^{2-\alpha}-a^{2-\alpha}]\\
&+\sum_{i=1}^{\alpha-1}\frac{1}{(a+b)^{\alpha-1+i}}\sum_{j=0}^{i-1}\binom{\alpha-1}{j+1}\binom{i-1}{j}\frac{1}{\alpha-i}[(b-s)^{i-\alpha}-b^{i-\alpha}]\\
&+\frac{1}{(a+b)^{2\alpha}-1}\sum_{j=0}^{\alpha-1}\binom{\alpha-1}{j+1}\binom{\alpha-1}{j}[ln|b|-ln|b-s|]+\frac{1}{(a+b)^{\alpha-1}}\frac{1}{\alpha}[(b-s)^{-\alpha}-b^{-\alpha}]\\
&-\sum_{i=1}^{\alpha-2}\frac{2a}{(a+b)^{\alpha+1+i}}\sum_{j=0}^{i-1}\binom{\alpha+1}{j+1}\binom{i-1}{j}\frac{1}{1-\alpha+i}[(a+s)^{1-\alpha+i}-a^{1-\alpha+i}]\\
&-\frac{2a}{(a+b)^{2\alpha}}\sum_{j=0}^{\alpha-2}\binom{\alpha+1}{j+1}\binom{\alpha-2}{j}[ln|a+s|-ln|a|]-\frac{2a}{(a+b)^{\alpha+1}}\frac{1}{1-\alpha}[(a+s)^{1-\alpha}-a^{1-\alpha}]\\
&-\sum_{i=1}^{\alpha-1}\frac{2a}{(a+b)^{\alpha+i}}\sum_{j=0}^{i-1}\binom{\alpha}{j+1}\binom{i-1}{j}\frac{1}{\alpha-i}[(b-s)^{i-\alpha}-b^{i-\alpha}]\\
&-\frac{2a}{(a+b)^{2\alpha}}\sum_{j=0}^{\alpha-1}\binom{\alpha}{j+1}\binom{\alpha-1}{j}[ln|b|-ln|b-s|]-\frac{2a}{(a+b)^{\alpha}}\frac{1}{\alpha}[(b-s)^{-\alpha}-b^{-\alpha}]\\
&+\sum_{i=1}^{\alpha-1}\frac{a^2}{(a+b)^{\alpha+1+i}}\sum_{j=0}^{i-1}\binom{\alpha+1}{j+1}\binom{i-1}{j}\frac{1}{\alpha-i}[(b-s)^{i-\alpha}-b^{i-\alpha}-(a+s)^{i-\alpha}+a^{i-\alpha}]\\
&+\frac{a^2}{(a+b)^{2\alpha+1}}\sum_{j=0}^{\alpha-1}\binom{\alpha+1}{j+1}\binom{\alpha-1}{j}\frac{1}{\alpha}[ln|a+s|-ln|a|+ln|b|-ln|b-s|]\\
&+\frac{a^2}{(a+b)^{\alpha+1}}\frac{1}{\alpha}[(b-s)^{-\alpha}-b^{-\alpha}-(a+s)^{-\alpha}+a^{-\alpha}]
\end{aligned}
$$

Appendix C: Proof of Proposition 4.2.2

Proof. Prove by induction. Base Case: when $p = 1$, $q = 1$

$$\begin{aligned}
m_x(1,1) &= \frac{1}{(a+x)(b-x)} \\
&= \frac{1}{(a+b)}\left(\frac{1}{(a+x)} + \frac{1}{(b-x)}\right) \\
&= \frac{1}{(a+b)}\frac{1}{(a+x)} + \frac{1}{(a+b)}\frac{1}{(b-x)} \\
&= \frac{1}{(a+b)}m_x(1,0) + \frac{1}{(a+b)}m_x(0,1)
\end{aligned}$$

Inductive Step: If the proposition holds,

$$\begin{aligned}
m_x(p,q-1) = &\sum_{i=1}^{p-1}\frac{1}{(a+b)^{q-1+i}}\sum_{j=0}^{i-1}\binom{q-1}{j+1}\binom{i-1}{j}m_x(p-i,0) + \frac{1}{(a+b)^{q-1}}m_x(p,0) \\
&+ \sum_{i=1}^{q-2}\frac{1}{(a+b)^{p+i}}\sum_{j=0}^{i-1}\binom{p}{j+1}\binom{i-1}{j}m_x(0,q-1-i) + \frac{1}{(a+b)^{p}}m_x(0,q-1) \\
m_x(p-1,q) = &\sum_{i=1}^{p-2}\frac{1}{(a+b)^{q+i}}\sum_{j=0}^{i-1}\binom{q}{j+1}\binom{i-1}{j}m_x(p-1-i,0) + \frac{1}{(a+b)^{q}}m_x(p-1,0) \\
&+ \sum_{i=1}^{q-1}\frac{1}{(a+b)^{p-1+i}}\sum_{j=0}^{i-1}\binom{p-1}{j+1}\binom{i-1}{j}m_x(0,q-i) + \frac{1}{(a+b)^{p-1}}m_x(0,q)
\end{aligned}$$

Starting from the inductive hypothesis, we have

$$\begin{aligned}
m_x(p,q) =&\frac{1}{(a+b)}[m_x(p-1,q) + m_x(p-1,q)] \\
=&\frac{1}{(a+b)}\Bigg[\sum_{i=1}^{p-1}\frac{1}{(a+b)^{q-1+i}}\sum_{j=0}^{i-1}\binom{q-1}{j+1}\binom{i-1}{j}m_x(p-i,0) + \frac{1}{(a+b)^{q-1}}m_x(p,0) \\
&+ \sum_{i=1}^{q-2}\frac{1}{(a+b)^{p+i}}\sum_{j=0}^{i-1}\binom{p}{j+1}\binom{i-1}{j}m_x(0,q-1-i) + \frac{1}{(a+b)^{p}}m_x(0,q-1) \\
&+ \sum_{i=1}^{p-2}\frac{1}{(a+b)^{q+i}}\sum_{j=0}^{i-1}\binom{q}{j+1}\binom{i-1}{j}m_x(p-1-i,0) + \frac{1}{(a+b)^{q}}m_x(p-1,0) \\
&+ \sum_{i=1}^{q-1}\frac{1}{(a+b)^{p-1+i}}\sum_{j=0}^{i-1}\binom{p-1}{j+1}\binom{i-1}{j}m_x(0,q-i) + \frac{1}{(a+b)^{p-1}}m_x(0,q)\Bigg] \\
&= \sum_{i=1}^{p-1}\frac{1}{(a+b)^{q+i}}\sum_{j=0}^{i-1}\binom{q-1}{j+1}\binom{i-1}{j}m_x(p-i,0) + \frac{1}{(a+b)^{q}}m_x(p,0)
\end{aligned}$$

$$+\sum_{i=2}^{q-1}\frac{1}{(a+b)^{p+i}}\sum_{j=0}^{i-1}\binom{p}{j+1}\binom{i-1}{j}m_x(0,q-1-i)+\frac{1}{(a+b)^{p+1}}m_x(0,q-1)$$
$$+\sum_{i=2}^{p-1}\frac{1}{(a+b)^{q+i}}\sum_{j=0}^{i-1}\binom{q}{j+1}\binom{i-1}{j}m_x(p-1-i,0)+\frac{1}{(a+b)^{q+1}}m_x(p-1,0)$$
$$+\sum_{i=1}^{q-1}\frac{1}{(a+b)^{p+i}}\sum_{j=0}^{i-1}\binom{p-1}{j+1}\binom{i-1}{j}m_x(0,q-i)+\frac{1}{(a+b)^{p}}m_x(0,q)$$

For $k \in \mathbb{Z}^+$, because the structure of terms $m_x(k,0)$ and $m_x(0,k)$ are the same, so we will focus on the terms $m_x(k,0)$. We want to check that if the following equation holds:

$$\sum_{i=1}^{p-1}\frac{1}{(a+b)^{q+i}}\sum_{j=0}^{i-1}\binom{q-1}{j+1}\binom{i-1}{j}m_x(p-i,0)+\frac{1}{(a+b)^{q}}m_x(p,0)$$
$$+\sum_{i=2}^{p-1}\frac{1}{(a+b)^{q+i}}\sum_{j=0}^{i-1}\binom{q}{j+1}\binom{i-1}{j}m_x(p-1-i,0)+\frac{1}{(a+b)^{q+1}}m_x(p-1,0)$$
$$=\sum_{i=1}^{p-1}\frac{1}{(a+b)^{q+i}}\sum_{j=0}^{i-1}\binom{q}{j+1}\binom{i-1}{j}m_x(p-i,0)+\frac{1}{(a+b)^{q}}m_x(p,0)$$

The above equation reduces to:

$$\sum_{j=0}^{i-1}\binom{q-1}{j+1}\binom{i-1}{j}+\sum_{j=0}^{i-2}\binom{q}{j+1}\binom{i-2}{j}=\sum_{j=0}^{i-1}\binom{q}{j+1}\binom{i-1}{j} \tag{5.1}$$

To prove (5.1) by calculation, we can define a function in R as follows:

```
iteration <- function(q, i) {
  # The first summation on the left-hand side
  s1 <- sum(choose(q-1, 1:i) * choose(i-1, 0:(i-1)))
  # The second summation on the left-hand side
  s2 <- sum(choose(q, 1:(i-1)) * choose(i-2, 0:(i-2)))
  # The summation on the right-hand side
  s3 <- sum(choose(q, 1:i) * choose(i-1, 0:(i-1)))
    return(s1 + s2 - s3)
}
```

This function will always return 0, indicating that Proposition 4.2.2 is correct. □