การคิดเบี้ยประกันภัยแบบเบย์สำหรับแบบจำลองการเรียกค่าสินไหมทดแทน ที่ไม่เป็นอิสระต่อกันโดยใช้คอมมอนเอฟเฟค

นายธิปไตย พงษ์ศาสตร์

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรดุษฎีบัณฑิต
สาขาวิชาคณิตศาสตร์ประยุกต์
มหาวิทยาลัยเทคโนโลยีสุรนารี
ปีการศึกษา 2558

BAYES PREMIUM FOR A CLAIM DEPENDENCE MODEL WITH COMMON EFFECT

Tippatai Pongsart

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in Applied Mathematics

Suranaree University of Technology

Academic Year 2015

BAYES PREMIUM FOR A CLAIM DEPENDENCE MODEL WITH COMMON EFFECT

Suranaree University of Technology has approved this thesis submitted in partial fulfillment of the requirements for the Degree of Doctor of Philosophy.

Thesis Examining Committee

(Asst. Prof. Dr. Eckart Schulz)
Chairperson

(Prof. Dr. Pairote Sattayatham)
Member (Thesis Advisor)

(Prof. Dr. Samruam Chongcharoen)
Member

(Asst. Prof. Dr. Arjuna Chaiyasena)
Member

(Asst. Prof. Dr. Benjawan Rodjanadid)
Member

(Prof. Dr. Sukit Limpijumnong)
Vice Rector for Academic Affairs and Innovation

(Prof. Dr. Santi Maensiri)
Dean of Institute of Science

ธิปไตย พงษ์ศาสตร์ : การคิดเบี้ยประกันภัยแบบเบย์สำหรับแบบจำลองการเรียกค่าสินไหมทดแทนที่ไม่เป็นอิสระต่อกันโดยใช้คอมมอนเอฟเฟค (BAYES PREMIUM FOR A CLAIM DEPENDENCE MODEL WITH COMMON EFFECT) อาจารย์ที่ปรึกษา : ศาสตราจารย์ ดร.ไพโรจน์ สัตยธรรม, 85 หน้า.

ในการศึกษาเกี่ยวกับทฤษฎีความเสี่ยง การคิดเบี้ยประกันภัยจะมาจากแบบจำลองซึ่งใช้ข้อมูลการเรียกค่าสินไหมทดแทนจากผู้เอาประกัน(เคลม)ประกอบการพิจารณา ข้อมูลนี้เป็นข้อมูลเชิง 2 มิติ กล่าวคือในส่วนแรกจะเป็นข้อมูลประวัติการเคลมสำหรับผู้เอาประกันของบุคคลหนึ่ง(ในหลายๆ ช่วงเวลา) อีกส่วนหนึ่งเป็นการเคลมของผู้เอาประกันรายอื่นในแต่ละช่วงเวลาเดียวกัน

ในการศึกษาแบบจำลองการเรียกค่าสินไหมทดแทนพบว่า มีหลายแบบจำลองที่ยอมรับสมมติฐานที่ว่า ค่าสินไหมทดแทนที่ถูกเรียกจากผู้เอาประกันในแต่ละรายเป็นอิสระต่อกัน ซึ่งอาจจะมีส่วนทำให้การคำนวณเบี้ยประกันภัยสะดวกขึ้น ทว่าในความเป็นจริงนั้น สมมติฐานนี้อาจจะไม่สอดคล้องในบางสถานการณ์ วิทยานิพนธ์นี้ได้ศึกษาในกรอบของสมมติฐานที่ว่า ค่าสินไหมทดแทนที่ถูกเรียกจากผู้เอาประกันในแต่ละรายไม่เป็นอิสระต่อกันโดยใช้คอมมอนเอฟเฟค ประกอบการพิจารณา โดยมีวัตถุประสงค์เพื่อที่จะคำนวณหาเบี้ยประกันภัยแบบเบย์ ในการศึกษาได้แยกพิจารณาออกเป็นสองส่วน ดังนี้

ส่วนที่หนึ่ง ได้มีการนำเสนอแบบจำลองที่การเรียกค่าสินไหมทดแทนจากผู้เอาประกันแต่ละรายไม่เป็นอิสระต่อกันภายใต้กรอบที่ว่า รู้ฟังก์ชันการแจกแจงของเคลมและคอมมอนเอฟเฟค และทำการศึกษาคุณสมบัติเบื้องต้นของเบี้ยประกันแบบต่างๆ ภายใต้เงื่อนไขของแบบจำลองนี้

ส่วนที่สอง กระบวนการคำนวณเบี้ยประกันภัยแบบเบย์ถูกพิจารณา โดยในขั้นต้นได้ทำการศึกษาสูตรของเบี้ยประกันภัยแบบเบย์ที่ทำให้ง่ายต่อการคิดคำนวณ จากนั้นทำการศึกษาการเคลมที่มีการแจกแจงแบบลอกนอร์มอลและการแจกแจงแบบปกติ โดยที่คอมมอนเอฟเฟคมีการแจกแจงแบบปกติ และได้ทำการสร้างสูตรของเบี้ยประกันภัยแบบเบย์สำหรับการเคลมที่มีการแจกแจงแบบลอกนอร์มอล และการแจกแจงปกติ

ในส่วนของบทประยุกต์ ได้ทำการนำแบบจำลองไปประยุกต์ใช้กับข้อมูลการจ่ายค่าสินไหมทดแทนของการประกันภัยรถยนต์ข้อมูลรายเดี่ยวในปี 2552 ซึ่งการประกันภัยเป็นประเภทความคุ้มครองที่ 5 จำนวน 1,296 ข้อมูล ซึ่งได้รับความอนุเคราะห์จากบริษัทประกันวินาศภัยรายหนึ่งในประเทศไทย เพื่อที่จะศึกษาผลกระทบของคอมมอนเอฟเฟคที่มีต่อเบี้ยประกันภัยแบบเบย์

สาขาวิชาคณิตศาสตร์ ลายมือชื่อนักศึกษา T. Pongsart

ปีการศึกษา 2558 ลายมือชื่ออาจารย์ที่ปรึกษา ________

TIPPATAI PONGSART : BAYES PREMIUM FOR A CLAIM DEPENDENCE MODEL WITH COMMON EFFECT. THESIS ADVISOR : PROF. PAIROTE SATTAYATHAM, Ph.D. 85 PP.

CLAIM DEPENDENCE MODEL/ COMMON EFFECT/ BAYESIAN PREMIUM/ LOGNORMAL DISTRIBUTION/ MOTOR INSURANCE

In risk theory, insurance premiums are calculated from a model using claim data which can be constructed in two dimensions with one dimension representing time and the other representing distinct insured individuals. Several models found in the literature allow for independence assumptions across different risks for the sake of convenience and mathematical tractability. However, these assumptions may be violated in some practical situations. In this thesis, modelling claim dependence is built under the common effect in the framework for investigating the Bayesian Premium. The study is separated into two parts.

In the first part, model descriptions and preliminaries are introduced. We study the basic properties of some types of premiums corresponding to the model.

In the second part, we derive some results in order to find the Bayesian premium under square-error loss function for arbitrary distributions of both claim amounts and common effect. We also establish the Bayesian premiums for lognormal and normal claim amount distributions while the common effect of both are normally distributed.

As an application of this part, we illustrate how the common effect influences the Bayesian premiums by using an actual motor insurance positive claim data set of

1,296 observations for the year 2009. These data were supplied by a non-life insurance company in Thailand.

School of Mathematics Student's Signature T. Pongsart

Academic Year 2015 Advisor's Signature P. Sattayatham

ACKNOWLEDGEMENTS

I am grateful to my advisor, Prof. Dr. Pairote Sattayatham for the many encouraging conversations, for suggestions which improved the work, and for helping me throughout the process. I would like to thank Dr. Tosaporn Talangtam for her guidance and useful discussions. Thanks to the anonymous Thai actuary for sample data sets.

I very much appreciate the assistance of Asst. Prof. Dr. Eckart Schulz for a very careful reading of my manuscript, including my thesis proposal. I would also like to thank Prof. Dr. Samruam Chongcharoen for his many useful comments and guidance, as well as Asst. Prof Dr. Arjuna Chaiyasena and Asst. Prof Dr. Benjawan Rodjanadid for their advice and helpful comments.

Thanks to my friend for the great experience and very interesting time during my life as a Ph.D. student at SUT. Special thanks are due to my family for their encouragement and support.

This work is partially supported by the Centre of Excellence in Mathematics in Thailand.

Tippatai Pongsart

CONTENTS

CONTENTS (Continued)

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS AND SYMBOLS

CDF	Cumulative distribution function
PDF	Probability density function
$(\Omega, \mathcal{F}, \mathrm{P})$	Probability space
Ω	Outcome space
$\mathcal{F}$	σ-field
P	Probability measure
X	Random variable
$\mathbb{N}$	Set of positive integers
$\mathbb{R}$	Real line
$\mathbb{R}^n$	n-dimensional Euclidean space
inf	Infimum (greatest lower bound)
sup	Supremum (least upper bound)
$L^2(\mathcal{F})$	Hilbert space of all random variables $X : \Omega \to \mathbb{R}$ having finite second moment
$\sigma(X)$	σ-field generated by random variable X
$(X \in B)$	$\{\omega \in \Omega : X(\omega) \in B\}$
$(X \leq x)$	$\{\omega \in \Omega : X(\omega) \leq x\}$
$F_X(x)$	CDF of X
$f_X(x)$	PDF of X
$\mathrm{E}[X]$	Expectation of the random variable X
$E[X\|\mathcal{F}]$	Conditional expectation of the random variable X given the σ-field $\mathcal{F}$
$Var(X)$	Variance of X

ABBREVIATIONS AND SYMBOLS (Continued)

$\vec{X}_i = (X_{i,1}, \ldots, X_{i,n})'$	a point of $\mathbb{R}^n$ as an n dimensional vector
	The symbol $'$ is used to indicate transposition
$\vec{X} = (\vec{X}_1, \ldots, \vec{X}_I)$	A row vector with components $\vec{X}_1, \ldots, \vec{X}_I$
$N(\mu, \sigma^2)$	Normal distribution with mean μ and variance σ^2
$LN(\mu, \sigma^2)$	Lognormal distribution with parameters μ and σ^2
$\varphi(x)$	Standard normal density

CHAPTER I

INTRODUCTION

One crucial task both in the practical management of an insurance company and in theoretical considerations is to determine premiums adequate to cover all risks. These premiums are calculated based on the chosen model, information in the insurance contract (e.g. claim experience), and a loss function which in mathematical terms belongs to the area of Bayesian statistics.

It is a common practice to group individual risks, so that the risks within each group are as homogeneous as possible, in order to reach a fair and equitable premium across all individuals. A collective premium, also know as the manual premium, is then calculated and charged for this group. But in reality, not all risks in any general class are truly homogeneous. No matter how detailed the underwriting procedure, there still remains some heterogeneity with respect to risk characteristics within a rating class.

In risk theory, each risk X for an individual is characterized by a risk parameter θ (possibly vector valued) due to the heterogeneity over policies in the concerned portfolio being examined. All values θ associated with each risk are modeled by the random variable Θ. Let $\Pi(\theta)$ be the cumulative distribution function of Θ and assume that the density of the random variable Θ exists and is denoted by $\pi(\theta)$. The function $\pi(\theta)$ is referred to as a structure function in actuarial studies and prior distribution in statistical theory. In order to predict a possible future loss for the risk X, we require a sequence of historical claims including accurately summarized information from the observed data to estimate the distribution $\pi(\theta)$.

1.1 Claim Dependence Modelling

There have been many studies of modelling the structure of claims which is an essential part of insurance pricing. In this section, an overview of the literature on claim dependence modelling is presented which leads to the purpose of the study.

In the process of modelling claims structure, many studies assume independence of claims which may be appropriate in some practical situations, including mathematical tractability. However, in real applications, there are some situations in which these assumptions may be violated; for example, in house insurance where geographic proximity of the insured may result in exposure to a common catastrophe, and in motor insurance where one collision may involve several insured parties.

The concept of modelling dependence began with a consideration of time dependence, but not of dependence across individuals. An early paper by Gerber and Jones (1975) and a paper by Frees et al. (1999) are examples of credibility models with time dependence for claims. Heliman (1986) and Hürliman (1993) have investigated the effect of dependencies of risks on stop-loss premiums. Wang (1998) proposed a set of statistic tools for modelling dependencies of risks in an insurance portfolio. Purcaru and Denuit (2002) provided a kind of dependence for claim frequency induced by introducing common latent variables in the annual numbers of claims reported by several policyholders. Several generalizations and alternative models of dependence have been suggested; however, in the context of credibility pricing, dependence over individuals has not received adequate attention from researchers and practitioners so far. However, in 2006, Yeo and Valdez investigated dependence over individuals by using common effects in a framework for developing credibility premiums. They used two-level common effects to describe the dependence structure of claims across individuals for a fixed time period and across time periods for a fixed individual. They further investigated credibility premiums when the claims are assumed to be normally distributed. There have been many remarkable efforts in the literature to study the dependence structure across individuals induced by common effects, e.g., Went et al. (2009) and H. Weizhong et al. (2012). Both of these groups pursued a variation from the work of Yeo and Valdez (2006). They extended the Bühlmann and Bühlmann-Straub credibility models with a dependence structure induced by common effects and derived corresponding credibility estimators.

In this thesis, we study the problem of dependence over individuals by using the one-level common effect in the framework for calculating the Bayesian premium. In addition, we are interested in the situation where claim amounts are lognormally distributed. This distribution is often used to describe the features of heavy-tailed claim amounts.

1.2 Concept of Insurance Premiums

An insurance premium is composed of a pure premium and the necessary loading. The pure premium of the insured loss is defined as the expected value of the claim amounts to be paid by the insurer. In practice, the insurer will add a risk (loss) loading to the pure premium. The sum of the pure premium and the loss loading is called the *net premium.* Adding the acquisition, expenses, and administration costs to this net premium, one obtains the *gross premium* that will be charged to the insured or policyholder. In this thesis, I shall consider only the pure premium. Next, we briefly review some concepts of premium calculation principles which relate to insurance premiums (more details in section (2.4)).

A premium calculation principle is a functional that assigns a real number, the premium, to any risk X (with probability function $f_{X|\Theta}(x|\theta)$, where x takes value in the sample space $\mathcal{X}$ and θ is considered to be a realization of a parameter space Θ). Let $L : \mathbb{R}^2 \to \mathbb{R}$ be a loss function that assigns to any $(x, P) \in \mathbb{R}^2$ the loss sustained by a decision maker who takes the action P, the premium charged, and is faced with the outcome x of a random experiment. $P(X)$ must be determined such that the expected loss is minimized, i.e., the minimum point of the mapping $P \longrightarrow E[L(X, P)]$ must be found. In risk theory many loss functions are used: the quadratic loss function gives the net premium principle, the exponential loss function results in the exponential principle, etc.

In this study, we are interested in the Bayesian premiums used to predict expected claims given the history of all observable claims. It is well-known that this conditional expectation actually gives the best predictor of the next period claim for a single individual under the mean-squared error loss function (see Klugman, 1992). More specifically, let $\vec{X}_i = (X_{i,1}, X_{i,2}, \ldots, X_{i,T})'$ be the random vector representing the vector of claims for a particular individual $i = 1, 2, \ldots, I$ within T time periods. Define a subspace H of $L^2(\mathcal{F})$ (the Hilbert space for all random variables having finite second moment) by

$$H \ := \ L^2(\sigma(\vec{X}_1, \vec{X}_2, \ldots, \vec{X}_I)).$$

Since H is clearly a closed subspace of $L^2(\mathcal{F})$, then for a fixed $i = 1, 2, \ldots, I$, the projection theorem in Hilbert space yields the unique existence of $p_i^* \in H$ satisfying

$$E[(X_{i,T+1} - p_i^*)^2] \ = \ \inf_{p \in H} E[(X_{i,T+1} - p)^2].$$

The solution p_i^* satisfies $p_i^* = E[X_{i,T+1}|\vec{X}]$ and is known as the Bayesian premium for risk $\vec{X}_i$.

1.3 Scope of Research

This thesis studies how to derive an explicit formula of the Bayesian premium under mean-squared error loss function in a claims dependence model. For this purpose, I begin with proposing the claims dependence model and then investigate the Bayesian premium with the assumption that claim amounts have known probability function.

There are many tools in mathematics that could be used to model claim dependencies; for example, the concept of copulas. Although copulas offer the flexibility of dependent random variables, they offer very limited mathematical tractability. In this thesis, I use only the common effect (in the terminology of Yeo and Valdez, 2006) to describe the dependency of claims. This concept can allow, besides the intuitive appeal, mathematical tractability in modelling claim dependencies. However, copula models are still appropriate tools and may be explored in future work.

When considerating of claim amounts distributions, we restrict our discussion to lognormal and normal claim amount distributions. As for the distribution of the common effect, we study only normal common effect distributions.

1.4 Objective and Overview of the Thesis

The aim of this thesis is to study a model of claim dependence which includes the concept of common effect across individual risks in its framework, in order to investigate Bayesian premiums with some well-know claim amounts distributions. The thesis is organized as follows.

In Chapter II, we introduce some notations, terminologies, some mathematical tools and statistical background which are used in Chapters III and IV.

In Chapter III, model descriptions and background are introduced. We also derive some results in order to find the Bayesian premium under the square-error loss function.

Chapter IV establishes the Bayesian premiums for lognornal and normal claim amounts distributions when the common effects of both are normally distributed. In addition, we find the Bayesian premium under normal claim amounts distribution is in credibility formula.

In Chapter V, we apply our model to an actual data set of claims which has been supplied by a non-life insurance public company in Thailand, in order to demonstrate the process of calculating the Bayesian premium and illustrate how the common effect influences the premium.

Conclusions, discussion and further research are shown in Chapter VI.

CHAPTER II

PRELIMINARIES

This chapter introduces the concepts and theories of some mathematical and statistic materials which are useful for the claim dependence modelling and insurance pricing, and provides some terminologies including the background information on insurance premiums.

2.1 Random Variables

By definition, a random variable X is a function whose domain is a sample space and whose range is a subset of the real numbers. In actuarial science, the actuary deals with objects such as random variables. For example, taking samples of insurance portfolio X might represent the number of annual claims, or the amount of annual claim for a policyholder associated with the occurrence of an automobile accident.

The notation $X(s) = x$ means that x is the value associated with the outcome s by the random variable X.

There are three types of random variables: discrete random variables, continuous random variables, and mixed random variables.

A discrete random variable is usually the result of a count and therefore the range consists of integers. A continuous random variable is usually the result of a measurement. As a result the range is any subset of the set of all real numbers. A mixed random variable is partially discrete and partially continuous.

2.2 Distribution Function

In probability theory and statistics, the cumulative distribution function (CDF), or just distribution function, describes the probability that a real-valued random variable X will have a value less than or equal to x. For example, let X be a random variable representing the total claim amount generated by some policyholder. Then the cumulative distribution function of X is the probability that this policyholder produces a total claim amount of at most x. The distribution function is important because it makes sense for any type of random variable, regardless of whether the distribution is discrete, continuous, or even mixed.

Definition 2.1. Let $(\Omega, \mathcal{F}, \mathrm{P})$ be a probability space. The (cumulative) distribution function of a random variable X is defined by

$$F_X(x) = P\{\omega \in \Omega \ ; \ X(\omega) \leq x\}$$

Using the abbreviated notation, we shall typical write the less explicit expression

$$F_X(x) = P\{X \leq x\}$$

for the distribution function.

Properties of distribution function

(i) F_X is non-decreasing ($x_1 \leq x_2$ implies $F_X(x_1) \leq F_X(x_2)$);

(ii) $\lim\limits_{x\to\infty} F_X(x) = 1, \ \lim\limits_{x\to-\infty} F_X(x) = 0$;

(iii) F_X is right continuous ($\lim_{h\to 0^+} F(x+h) = F(x)$ for all $x \in \mathbb{R}$).

Definition 2.2. A random variable X is called *discrete* if it takes values in some countable subset $\{x_1, x_2, \ldots\}$ of $\mathbb{R}$. The discrete random variable X has probability mass function (PMF), $f : \mathbb{R} \longrightarrow [0, 1]$ given by

$$f(x) = P\{X = x\}.$$

Definition 2.3. A random variable X is called *continuous* if its distribution function can be expressed as

$$F(x) = \int_{-\infty}^{x} f(u)du \quad ; x \in \mathbb{R},$$

for some nonnegative integrable function $f : \mathbb{R} \longrightarrow [0, \infty)$. The function f is called the probability density function (PDF) of X.

Notice that there exist distributions that are neither continuous nor discrete.

2.2.1 Normal Distribution

The normal distribution is the most widely known and used of all distributions. Because the normal distribution approximates many natural phenomena very well, it has developed into a standard of reference for many probability problems.

Assume that a random variable X has the normal distribution with parameters μ and σ^2, abbreviated $X \sim N(\mu, \sigma^2)$. Then we have

(1) Cumulative distribution function (CDF) : $F(x) = \Phi(\frac{x-\mu}{\sigma})$ where

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{\frac{-t^2}{2}} dt,$$

is the area under the "Bell curve function" $(\frac{1}{\sqrt{2\pi}})e^{\frac{-t^2}{2}}$ between $-\infty$ to z.

(2) Probability density function (PDF) :

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2}, \quad -\infty < x < \infty.$$

(3) Expectation : $E[X] = \mu$.

(4) Variance : $var(X) = \sigma^2$.

(5) Moment-generating function : $M_X(t) = E[e^{tX}] = e^{\mu t + \frac{1}{2}\sigma^2 t^2}$.

2.2.2 Lognormal Distribution

The lognormal distribution is useful as a model for the claim size distribution. A random variable X is said to have the lognormal distribution with parameters μ and σ^2 if $Y = lnX$ has the normal distribution with mean μ and standard deviation σ. A random variable which is lognormally distributed takes only positive real values.

Assume that a random variable X has the lognormal distribution with parameters μ and σ^2, abbreviated $X \sim LN(\mu, \sigma^2)$. Then we have

(1) Cumulative distribution function (CDF) : $F(x) = \Phi(\frac{lnx-\mu}{\sigma})$; $\mu \in \mathbb{R}$, $\sigma > 0$, $x > 0$ where

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{\frac{-t^2}{2}} dt.$$

(2) Probability density function (PDF) :

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{lnx-\mu}{\sigma})^2}.$$

(3) Expectation : $E[X] = e^{\mu+\frac{\sigma^2}{2}}$.

(4) Variance : $var(X) = (e^{\sigma^2} - 1)e^{2\mu+\sigma^2}$.

(5) Moment-generating function : $M_X(t) = E[e^{Xt}] = \infty$ for any $t > 0$.

Summary of the Propositions

Proposition 2.1. *If X is $N(\mu, \sigma^2)$ then $aX + b$ is $N(a\mu + b, a^2\sigma^2)$.*

Proposition 2.2. *If X is $N(\mu_1, \sigma_1^2)$ and Y is $N(\mu_2, \sigma_2^2)$, and X and Y are independent, then $X + Y$ is $N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$.*

Corollary 2.3. *If X_i are independent $N(\mu, \sigma^2)$ for $i = 1, 2, \ldots, n$ then $\sum_{i=1}^{n} X_i$ is $N(n\mu, n\sigma^2)$.*

Corollary 2.4. *If Y_i are independent $LN(\mu, \sigma^2)$ for $i = 1, 2, \ldots, n$ then $\Pi_{i=1}^{n} Y_i$ is $LN(n\mu, n\sigma^2)$.*

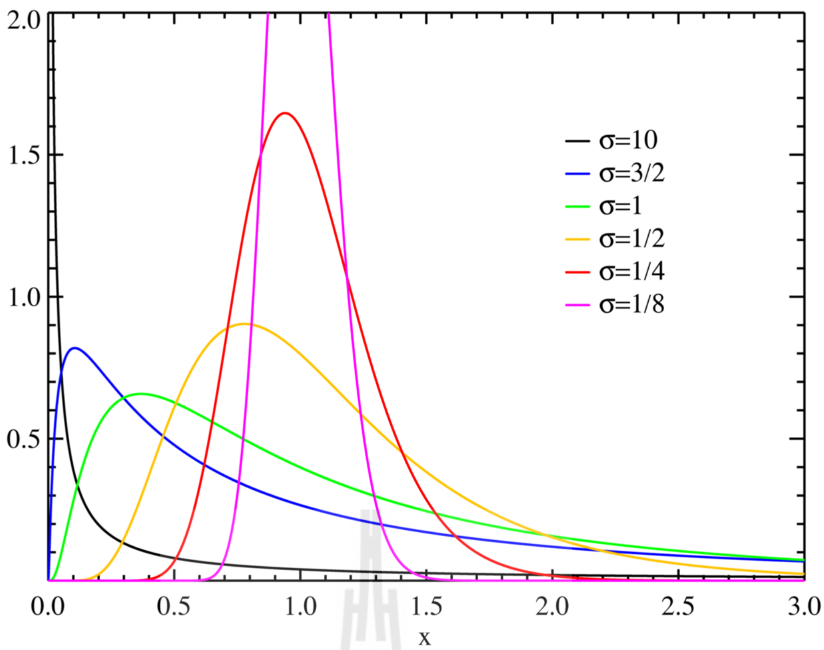

Figure 2.1 The PDF of lognormal distribution with parameters (μ, σ^2) in case of $\mu = 0$.

2.3 Properties of Expectation

The expected value of a random variable is the weighted average of the possible values of random variable X and also is the center of the distribution of the variable.

Definition 2.4. Let X be a discrete random variable with probability mass function $p(x)$. The expected value of X is given by

$$E(X) = \sum_x xp(x)$$

provided that the sum is finite.

For a continuous random variable X with probability density function $f(x)$, the expected value is given by

$$E[X] = \int_{-\infty}^{\infty} xf(x)dx$$

provided that the improper integral is convergent.

2.3.1 Expected value of a Function of two Random Variables

In this subsection, we introduce some equalities and inequalities about the expectation of random variables. First, we introduce the definition of expectation of a function of two random variables.

Definition 2.5. Suppose that X and Y are two random variables taking values in S_X and S_Y respectively. For a function $g : S_X \times S_Y \to \mathbb{R}$ the expected value of $g(X, Y)$ is

$$E[g(X,Y)] = \sum_{x \in S_X} \sum_{y \in S_Y} g(x,y) p_{XY}(x,y)$$

if X and Y are discrete with joint probability mass function $p_{X,Y}(x, y)$ and

$$E[g(X,Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x,y) f_{X,Y}(x,y) dx dy$$

if X and Y are continuous with joint probability density function $f_{X,Y}(x, y)$.

An important application of the above definition is the following result.

Proposition 2.5. *The expected value of the sum/difference of two random variables is equal to the sum/difference of their expectations. That is,*

$$E[X+Y] = E[X] + E[Y]$$

and

$$E[X-Y] = E[X] - E[Y].$$

Proof. We proof the result for discrete random variables X and Y with joint probability mass function $p_{X,Y}(xy)$. Letting $g(X, Y) = X \pm Y$ we have

$$\begin{aligned} E[X \pm Y] &= \sum_x \sum_y (x \pm y) p_{XY}(x,y) \\ &= \sum_x \sum_y x p_{XY}(x,y) \pm \sum_x \sum_y y p_{XY}(x,y) \\ &= \sum_x x \sum_y p_{XY}(x,y) \pm \sum_x y \sum_y p_{XY}(x,y) \\ &= \sum_x x p(x) \pm \sum_y y p(y) \\ &= E[X] \pm E[Y]. \end{aligned}$$

A similar proof holds for the continuous case where we just need to replace the sums by improper integrals and the joint probability mass function by the joint probability density function. □

Proposition 2.6. *Let X be nonnegative random variable, then $E[X] \geq 0$. Suppose X and Y are two random variables such that $X > Y$. Then $E[X] \geq E[Y]$.*

Proof. We proof the result for the continuous case. We have

$$\begin{aligned} E[X] &= \int_{-\infty}^{\infty} xf(x)dx \\ &= \int_{0}^{\infty} xf(x)dx \geq 0 \end{aligned}$$

since $f(x) \geq 0$ so the integrand is nonnegative. Now, if $X \geq Y$ then $X - Y \geq 0$ so that by the previous preposition we can write

$$E[X] - E[Y] \;=\; E[X - Y] \;\geq\; 0.$$

This concludes the proof. □

Proposition 2.7. *If X and Y are independent random variables then for any function h and g we have*

$$E[g(X)h(Y)] \;=\; E[g(X)]E[h(Y)].$$

In particular, $E[XY] = E[X]E[Y]$.

Proof. We proof the result for the continuous case. The proof of the discrete case is similar. Let X and Y be two independent random variables with joint density function $f_{X,Y}(x, y)$. Then

$$\begin{aligned} E[g(X)h(X)] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x)h(y)f_{X,Y}(x,y)dxdy \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x)h(y)f_X(x)f_Y(y)dxdy \\ &= \Big(\int_{-\infty}^{\infty} g(x)f_X(x)dx\Big)\Big(\int_{-\infty}^{\infty} h(y)f_Y(y)dy\Big) \\ &= E[g(X)]E[h(Y)]. \end{aligned}$$

□

We note that the expected values need not multiply if the random variables are not independent.

2.3.2 Conditional Expectation

Since conditional probability measures are probability measures (that is, they possess all of the properties of unconditional probability measures), conditional expectations inherit all of the properties of regular expectations.

Definition 2.6. Let X and Y be random variables. In the discrete case, the conditional expectation of X given that $Y = y$ is defined by

$$\begin{aligned} E[X|Y=y] &= \sum_x xP(X=x|Y=y) \\ &= \sum_x xp_{X|Y}(x|y) \end{aligned}$$

where $p_{X|Y}$ is the conditional probability mass function of X, given that $Y = y$ which is given by

$$p_{X|Y}(x|y) = P(X=x|Y=y) = \frac{p_{X,Y}(x,y)}{p_Y(y)}.$$

This is defined for non-zero $p_Y(y)$.

In the continuous case we have

$$E[X|Y=y] = \int_{-\infty}^{\infty} xf_{X|Y}(x|y)dx$$

where

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}.$$

Also in this case it necessary that $f_Y(y) > 0$.

For any function $g(x)$, the conditional expected value of g given $Y = y$ is, in the continuous case,

$$E[g(X)|Y=y] = \int_{-\infty}^{\infty} g(x)f_{X|Y}(x|y)dx$$

if the integral exists. For the discrete case, we have a sum instead of an integral. That is, the conditional expectation of g given $Y = y$ is

$$E[g(X)|Y=y] = \sum_x g(x)p_{X|Y}(x|y).$$

The proof of this result is identical to the unconditional case.

Theorem 2.8. *(Double Expectation Property)*

Let X and Y be random variables. Then

$$E[X] = E[E[X|Y]].$$

Proof. We proof the result for the case X and Y are continuous random variables.

$$\begin{aligned} E[E[X|Y]] &= \int_{-\infty}^{\infty} E[X|Y=y]f_Y(y)dy \\ &= \int_{-\infty}^{\infty} \Big(\int_{-\infty}^{\infty} xf_{X|Y}(x|y)dx\Big) f_Y(y)dy \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xf_{X|Y}(x|y)dxf_Y(y)dy \\ &= \int_{-\infty}^{\infty} x\int_{-\infty}^{\infty} f_{X|Y}(x|y)f_Y(y)dydx \\ &= \int_{-\infty}^{\infty} xf_X(x)dx \\ &= E[X]. \end{aligned}$$

□

Definition 2.7. (The Conditional Variance)

Let X and Y be random variables. The conditional variance of X given Y is defined by

$$Var(X|Y=y) = E[(X-E[X|Y])^2|Y=y].$$

Note that the conditional variance is a random variable since it is a function of Y.

Proposition 2.9. *Let X and Y be random variables. Then*

(*a*) $Var(X|Y) = E[X^2|Y] - (E[X|Y])^2$

(*b*) $E[var(X|Y)] = E[E[X^2|Y] - (E[X|Y])^2] = E[X^2] - E[(E[X|Y])^2]$

(*c*) $Var(E[X|Y]) = E[(E[X|Y])^2] - (E[X])^2$

(*d*) *Law of Total Variance :* $Var(X) = E[Var(X|Y)] + Var(E[X|Y])$.

Proof. (a) we have

$$\begin{aligned} Var(X|Y) &= E[(X - E[X|Y])^2|Y] \\ &= E[(X^2 - 2XE[X|Y]) + (E[X|Y])^2|Y] \\ &= E[X^2|Y] - 2E[X|Y]E[X|Y] + (E[X|Y])^2 \\ &= E[X^2|Y] - (E[X|Y])^2. \end{aligned}$$

(b) Taking E of both sides of the result in (a) we find

$$\begin{aligned} E[Var(X|Y)] &= E[E[X^2|Y] - (E[X|Y])^2] \\ &= E[X^2] - E[(E[X|Y])^2]. \end{aligned}$$

(c) Since $E[E[X|Y]] = E[X]$ we have

$$Var(E[X|Y]) = E[(E[X|Y])^2] - (E[X|Y])^2.$$

(d) The result follows by adding the two equations in (b) and (c). □

2.4 Premium Calculation and Insurance Pricing

The price of insurance is the monetary value for which two parties agree to exchange risk and certainty. There are two commonly encountered situations in which the price of insurance is the subject of consideration: when an individual agent (for example, a household), bearing an insurance risk, buys insurance from an insurer at an agreed periodic premium; and when insurance portfolios (that is, a collection of insurance contracts) are traded in the financial industry (e.g., being transferred from an insurer to another insurer or from insurer to the financial market (securitization).

Pricing in the former situation is usually referred to as *premium calculation* while pricing in the latter situation is usually referred to as *insurance pricing*, although such a distinction is not unambiguous.

We first mention the concepts of risk and refer to premium calculation principle later. These definitions are stated as follows.

Definition 2.8 (**Fundamentals).** We fix a measurable space $(\Omega, \mathcal{F})$ where Ω is the outcome space and $\mathcal{F}$ is a (σ-)algebra defined on it. A risk is a random variable defined on $(\Omega, \mathcal{F})$; that is, $X : \Omega \to \mathbb{R}$ is a risk if $X^{-1}((-\infty, x]) \in \mathcal{F}$ for all $x \in \mathbb{R}$. A risk represents the final net loss of a position (contingency) currently held. When $X > 0$, we call it a loss, whereas when $X \leq 0$, we call it a gain. The class of all random variables on $(\Omega, \mathcal{F})$ is denoted by $\mathcal{X}$.

In the insurance industry, the main types of risk are classified as follows:

(1) The market risk, the credit risk, the operational risk, the model risk and the liquidity risk. These are the main types of risks encountered in the financial industry.

(2) The underwriting risk: the risks inherent in insurance policies that have been sold:

- The risk that premiums will not be sufficient to cover future incurred losses and that losses and loss adjustment expenses' current reserves are not sufficient although the distributions of losses have been well assessed.

- The risk that may arise from an inaccurate assessment of the risks entailed in writing an insurance policy or from factors that are not under the insurer's control (changes in patterns of natural catastrophes, changes in demographic tables underlying long-date life products, changes in customer behaviour, so on)

The families of risk measures: for measurement of both financial and insurance risks, is computed by P-quantile risk measure, risk measures based on expected utility theory, risk measures based on distorted expectation theory and calculation principle.

Definition 2.9 (**Premium calculation principle (pricing principle)).** A premium (calculation) principle or pricing principle h is a functional assigning a real number to any random variable defined on $(\Omega, \mathcal{F})$; that is, h is a mapping from $\mathcal{X}$ to $\mathbb{R}$.

Remark 2.1. In general, no integrability conditions need to be imposed on the elements of $\mathcal{X}$. In the absence of integrability conditions, some of the premium principles studied below will be on infinite subclasses of $\mathcal{X}$. Instead of imposing integrability conditions, we may extend the range in the definition of h to $\mathbb{R} \cup \{-\infty, \infty\}$. In case $h(X) = +\infty$, we say that the risk is unacceptable or non-insurable.

Classical actuarial pricing of insurance risks mainly relies on the economic theories of decision under uncertainty, in particular on Von Neumann & Morgenstern's (1947) expected utility theory and Savage's (1954) subjective expected utility theory. Using the principle of equivalent utility and specifying a utility function, various well-known premium principles can be derived. An important example is the exponential premium principle, obtained by using a (negative) exponential utility function, having a constant rate of risk aversion.

2.4.1 Premium Principles and Their Properties

Many of the (families of) premium principles can be (directly) characterized axiomatically. The general purpose of an axiomatic characterization is to demonstrate what are the essential assumptions to be imposed and what are relevant parameters or concepts to be determined. A premium principle is appropriate if and only if its characterizing axioms are. Axiomatizations can be used to justify a premium principle, but also criticize it.

A systematic study of the properties of premium calculation principles and their axiomatic characterizations was pioneered by Goovaerts, De Vylder & Haezendonck (1984). Here we give some details of properties that premiums principles may (or may not) satisfy.

Properties of Premium Principles

Definition 2.10 (Law invariance (independent, objectivity)). h is (P-) law invariant if $h(X) = h(Y)$ when $P(X \le x) = P(Y \le y)$ for all real x.

Definition 2.11 **(Monotonicity).** h is monotonic if $h(X) \leq h(Y)$ when $X(\omega) \leq Y(\omega)$ for all $\omega \in \Omega$. h is P-monotonic if $h(X) \leq h(Y)$ when $X \leq Y$ $P-$almost surely.

Definition 2.12 **(Preserving first-order stochastic dominance (FOSD)).** h preserves first-order stochastic dominance if $h(X) \leq h(Y)$ when $P(X \leq x) \geq P(Y \leq y)$ for all $x \in \mathbb{R}$.

Law invariance together with monotonicity implies preserving FOSD.

Definition 2.13 **(Preserving stop-loss order (SL)).** h preserves stop-loss order if $h(X) \leq h(Y)$ when $E[(X-d)_+]) \leq E[(Y-d)_+]$ for all $d \in \mathbb{R}$.

Definition 2.14 **(Risk loading).** h induces a risk loading if $h(X) \geq E[X]$.

Definition 2.15 **(Not unjustified).** h is not unjustified if $h(c) = c$ for all real c.

Definition 2.16 **(Additivity).** h is additive if $h(X+Y) = h(X) + h(Y)$.

Definition 2.17 **(Translation invariance).** h is translation invariant if $h(X+c) = h(X) + c$ for all real c.

Definition 2.18 **(Positive homogeneity (scale invariance)).** h is positively homogeneous if $h(aX) = ah(X)$ for all $a \geq 0$.

Definition 2.19 **(Subadditivity resp. Superaddivity).** h is subadditive (resp. super-additive) if $h(X+Y) \leq h(X) + h(Y)$.

Definition 2.20 **(Convexity).** h is convex if $h(\alpha X + (1-\alpha)Y) \leq \alpha h(X) + (1-\alpha)h(Y)$ for all $\alpha \in (0,1)$.

Definition 2.21 **(Independent additivity).** h is independent additive if $h(X+Y) = h(X) + h(Y)$ when X and Y are independent.

Definition 2.22 **(Comonotonic additivity).** h is comonotonic additive if $h(X+Y) = h(X) + h(Y)$ when X and Y are comonotonic (see Appendix C).

Definition 2.23 **(Iterativity).** h is iterative if $h(X) = h(h(X|Y))$.

Premium Principles

In this subsection, we list many well-known premium principles. Some premium principles (Esscher) arise from more than one theory. Other premium principles (Dutch) instead are not directly based on any theories nor on an axiomatic characterization, but rather on the nice properties that they exhibit.

Definition 2.24 (**Net premium principle).** The net premium principle is given by

$$h(X) = E[X].$$

Definition 2.25 (**Expected value principle).** The expected value principle is given by

$$h(X) = (1+\lambda)E[X], \qquad \lambda \geq 0.$$

Notice that if $\lambda = 0$, the net premium is obtained.

Definition 2.26 (**Mean value principle).** For a given non-decreasing and non-negative function f on $\mathbb{R}$ the mean value principle is the root of

$$f(h) = E[f(X)].$$

Definition 2.27 (**Variance principle).** The variance principle is given by

$$h(X) = E[X] + \lambda Var(X), \qquad \lambda > 0.$$

Definition 2.28 (**Standard deviation principle).** The standard deviation principle is given by

$$h(X) = E[X] + \lambda\sqrt{Var(X)}, \qquad \lambda > 0.$$

Definition 2.29 (**Exponential principle).** The exponential principle is given by

$$h(X) = \frac{1}{\alpha} log E[e^{\alpha X}], \qquad \alpha > 0.$$

Definition 2.30 (**Esscher principle).** The Esscher principle is given by

$$h(X) = \frac{E[Xe^{\alpha X}]}{E[e^{\alpha X}]}, \qquad \alpha > 0.$$

Definition 2.31 (**Swiss principle**). For a given non-negative and non-decreasing function w on $\mathbb{R}$ and a given parameter $0 \leq p \leq 1$ the Swiss premium is the root of

$$E[w(X - ph)] = w((1-p)h).$$

Definition 2.32 (**Dutch principle**). The Dutch principle is given by

$$h(X) = E[X] + \theta E[(X - \alpha E[X])_+], \qquad \alpha \geq 1, \quad 0 < \theta \leq 1.$$

Definition 2.33 (**Orlicz principle**). Let $X \geq 0$. For a given normalized Young function φ on $\mathbb{R}_+ \cup \{0\}$ the Orlicz premium is the root of

$$E[\varphi(\frac{X}{h})] = 1.$$

CHAPTER III

CLAIM DEPENDENCE MODELLING

In this chapter, we present a model of claim dependence induced by common effect (in the terminology of Yeo and Valdez (2006)) and preliminaries. For a loss function consideration, we are mainly interested in the expected squared error loss. We derive some basic results concerning the identification of premiums and of the losses attached to them.

3.1 Model Formulation and Preliminaries

3.1.1 Model Formulation

Let $(\Omega, \mathcal{F}, \mathrm{P})$ be a probability space, let $L^2(\mathcal{F})$ denoted the Hilbert space of all random variables $X : \Omega \longrightarrow \mathbb{R}$ having a finite second moment. All random variables that we shall work with will be in this space.

Let I and T be positive integers. Consider a portfolio of insurance contracts consisting of I insured individuals and each individual has available a history of T time periods. Denote by $X_{i,t}; 1 \leq i \leq I, 1 \leq t \leq T$, the claim amount for individual i during period t. Therefore, the random vector

$$\vec{X}_i \;=\; (X_{i,1}, X_{i,2}, \ldots, X_{i,t})'$$

represents the vector of claims for a particular individual $i = 1, 2, \ldots, I$. Our primary interest is to predict the next claim for each individual i based on all the observed claims

$$\vec{X} \;=\; (\vec{X}_1, \vec{X}_2, \ldots, \vec{X}_I). \tag{3.1}$$

This will be denoted by the random variable $X_{i,T+1}$.

The one-level common effect model of claim dependence.

As already mentioned in the introductory section, the model of dependence being proposed in this section below, will allow the dependence among the individual risks. The dependence among the individual risks will be described by a common effect random variable Λ whose density function will be assumed to be known and denoted by $f_\Lambda(\lambda)$. Realization of this common effect is denoted by λ. Conditional on this common effect, the random vectors $\vec{X}_i$ are independent. As Λ is a common effect among all risks, it will define the dependence structure between risks, and it can either be a discrete, continuous, or a mixture of discrete and continuous random variables. More precisely, we shall summarize these setting into the following assumptions.

A1. The common effect random variable Λ has known probability density function $f_\Lambda(\lambda)$ provided that $f_\Lambda(\lambda) > 0$ for all λ.

A2. For a fixed $i = 1, 2, \ldots, I,$ the random variables $X_{i,t}, t = 1, 2, \ldots, T$ are mutually independent and identically distributed.

A3. The random vectors $\vec{X}_i|\Lambda = \lambda,\ i = 1, 2, \ldots, I$ where $\vec{X}_i = (X_{i,1}, X_{i,2}, \ldots, X_{i,T})'$ are conditionally independent.

A4. For a fixed $i = 1, 2, \ldots, I$ and a fixed $t = 1, 2, \ldots, T,$ the conditional random variable $X_{i,t}$ given that $\Lambda = \lambda$ has known probability function denoted by

$$f_{X_{i,t}|\Lambda}(x_{i,t}|\lambda) \ =: \ \frac{f_{X_{i,t},\Lambda}(x_{i,t}, \lambda)}{f_\Lambda(\lambda)}.$$

One can think of Λ as the variable inducing dependence of claim among individuals, such as in the case of an epidemic in life insurance, a catastrophe in general insurance, or simply bad weather conditions on a day when automobile accidents are frequent.

3.2 Loss function Minimization

The word *risk* in section (3.1) usually refers to any general risk, while on individual's risk in the latter situation is referred to as the claim amount for an individual, unless otherwise stated.

Consider the proposed model in section (3.1.1), the nature of dependence which influences the claim amounts is represented by the common effect Λ. Conditional on this common effect, the claim amounts are independent. In this subsection, we show that the Bayesian premium is asymptotically optimal in terms of losses. Before we minimize loss by using a expected square error loss function, we first mention the mathematical tools and introduce some types of premiums which relate to the expected square error loss.

3.2.1 Background

We are interested in $X_{i,T+1}, i = 1, 2, \ldots, I,$ the claim amount of individual $i = 1, 2, \ldots, I$ for the next time period which is to be predicted by a premium $p \in \triangle$ minimizing the loss $E[(X_{i,T+1} - p)^2]$ over $\triangle$, where $\triangle \subset L^2(\mathcal{F})$ is a prescribed class of premiums to be specified below.

We assume that $varE[X_{i,T+1}|\Lambda] > 0$. Here $E[X_{i,T+1}|\Lambda]$ denotes the conditional expectation of $X_{i,T+1}$ with respect to the σ-algebra $\sigma(\Lambda)$ generated by Λ, and we have $E[X_{i,T+1}|\Lambda] \in L^2(\sigma(\Lambda))$; correspondingly, $var(X_{i,T+1}|\Lambda)$ denotes the conditional variance $E[(X_{i,T+1} - E[X_{i,T+1}|\Lambda])^2|\Lambda]$ of $X_{i,T+1}$ with respect to $\sigma(\Lambda)$. Let $\bar{X}$ denote the sample mean $\frac{1}{IT}(\sum_{i=1}^{I}\sum_{t=1}^{T} X_{i,t})$, and define

$$\begin{aligned} \mu &:= E[X_{i,T+1}] = E[E[X_{i,T+1}|\Lambda]], \\ v &:= E[var(X_{i,T+1}|\Lambda)], \\ a &:= var(E[X_{i,T+1}|\Lambda]). \end{aligned} \tag{3.2}$$

With these definitions, we have $var(X) = E[(X_{i,T+1} - \mu)^2] = v + a$; see also proposition (3.3) below.

We consider three classes of premiums:

$$\begin{aligned}\triangle_0 &:= \mathbb{R}\\ \triangle_1 &:= L^2(\sigma(\Lambda))\\ \triangle_I &:= L^2(\sigma(X_{1,1}, X_{1,2}, \ldots, X_{1,T}, \ldots, X_{I,1}, \ldots, X_{I,T}).\end{aligned} \tag{3.3}$$

Since each $\triangle_0, \triangle_1, \triangle_I$ is a closed subspace of $L^2(\mathcal{F})$, the projection theorem in Hilbert spaces yields the existence of unique $p_0 \in \triangle_0, p_1 \in \triangle_1, p_2 \in \triangle_I$ satisfying

$$\begin{aligned}E[(X_{i,T+1} - p_0)^2] &:= \inf_{\triangle_0} E[(X_{i,T+1} - p)^2]\\ E[(X_{i,T+1} - p_1)^2] &:= \inf_{\triangle_1} E[(X_{i,T+1} - p)^2]\\ E[(X_{i,T+1} - p_2)^2] &:= \inf_{\triangle_I} E[(X_{i,T+1} - p)^2],\end{aligned} \tag{3.4}$$

for a fixed individual $i = 1, 2, \ldots, I$. In what follows in this thesis, we shall call

p_0 the *collective premium*,

p_1 the *individual premium*, and

p_2 the *Bayesian premium*. (3.5)

We are mainly interested in the Bayesian premium, which is the best prediction of $X_{i,T+1}$ by an *arbitrary* function of $X_{1,1}, X_{1,2}, \ldots, X_{1,T}, \ldots, X_{I,1}, \ldots, X_{I,T}$. The collective premium and the individual premium may be interpreted as the Bayesian in the no-data case and serves mainly as a reference for comparisons. However, we note that there is no obvious relation between $\triangle_1$ and $\triangle_I$; nothing more than a suggestive notation which still has to be justified.

In subsection 3.2.2 we state some results concerning the identification of these premiums and of the losses attached to them.

3.2.2 Basic Results

Since $X_{i,T+1}$ and $X_{1,1}, X_{1,2}, \ldots, X_{1,T}, \ldots, X_{I,1}, \ldots, X_{I,T}$ are conditionally independent with respect to Λ for a fixed $i = 1, 2, \ldots, I$, the same is true for $X_{i,T+1}$

and each $p \in \triangle_I$. This yields the following useful result:

Lemma 3.1. *For a fixed* $i = 1, 2, \ldots, I$, *The identity*

$$E[(X_{i,T+1} - p)^2] \ = \ E[(X_{i,T+1} - E[X_{i,T+1}|\Lambda])^2] + E[(E[X_{i,T+1}|\Lambda] - p)^2]$$

holds for all $p \in \triangle_I$.

The proof of the lemma is shown in Appendix B.

Proposition 3.2. *For the optimum premiums* p_0, p_1, p_2 *and a fixed individual* $i = 1, 2, \ldots, I$ *we have:*

(a) $p_0 \ = \ \mu$

(b) $p_1 \ = \ E[X_{i,T+1}|\Lambda]$

(c) $p_2 \ = \ E[X_{i,T+1}|\vec{X}] \ = \ E[E[X_{i,T+1}|\Lambda)|\vec{X}]]$.

In particular, $p_2 \ = \ E[p_1|\vec{X}]$.

Proof. Firstly, we fix $i; i = 1, 2, \ldots, I$. From equation (3.3) and equation (3.4) then we have:

(a) The projection theorem in Hilbert spaces yields the existence of unique $p_0 \in \triangle_0$, satisfying

$$E[(X_{i,T+1} - p_0)^2] \ = \ \inf_{\triangle_0} E[(X_{i,T+1} - p)^2].$$

It follows immediately that $p_0 \ = \ E[X_{i,T+1}] \ = \ E[E[X_{i,T+1}\Lambda]]$.

(b) The projection theorem in Hilbert spaces yields the existence of unique $p_1 \in \triangle_1$, satisfying

$$E[(X_{i,T+1} - p_1)^2] \ = \ \inf_{\triangle_1} E[(X_{i,T+1} - p)^2].$$

It follows from the fact that the projection of $X_{i,T+1}$ onto $\triangle_1 = L^2(\sigma(\Lambda))$ is precisely the conditional expectation of $X_{i,T+1}$ with respect to $\sigma(\Lambda)$. Then $p_1 \ = \ E[X_{i,T+1}|\Lambda]$.

(c) Similarly to *(b)*, the projection of $X_{i,T+1}$ onto

$$\triangle_2 = L^2(\sigma(X_{1,1}, X_{1,2}, \ldots, X_{1,T}, \ldots, X_{I,1}, \ldots, X_{I,T}))$$

is precisely the conditional expectation of $X_{i,T+1}$ with respect to

$$\sigma(X_{1,1}, X_{1,2}, \ldots, X_{1,T}, \ldots, X_{I,1}, \ldots, X_{I,T}).$$

Then $p_2 = E[X_{i,T+1}|X_{1,1}, X_{1,2}, \ldots, X_{1,T}, \ldots, X_{I,1}, \ldots, X_{I,T}] = E[X_{i,T+1}|\vec{X}]$.

The final equality in proposition 3.2 follows from

$$p_2 = E[E[X_{i,T+1}|\Lambda)|\vec{X}]] = E[p_1|\vec{X}].$$

This concludes the proof. □

Proposition 3.3. *For the loss attached to these premiums and a fixed individual* $i = 1, 2, \ldots, I$, *we have:*

(*a*) $E[(X_{i,T+1} - p_0)^2] = v + a$

(*b*) $E[(X_{i,T+1} - p_1)^2] = v$

(*c*) $E[(X_{i,T+1} - p_2)^2] = v + E[var(E[X_{i,T+1}|\Lambda]|\vec{X})]$

In particular, $E[(X_{i,T+1} - p_1)^2] \leq E[(X_{i,T+1} - p_2)^2] \leq E[(X_{i,T+1} - p_0)^2]$.

Proof. Firstly, we fix $i; i = 1, 2, \ldots, I$. The proof is straightforward.

(*a*)

$$\begin{aligned} E[(X_{i,T+1} - p_0)^2] &= E[(X_{i,T+1} - E[X_{i,T+1}])^2] \\ &= var(X_{i,T+1}) \\ &= E[var(X_{i,T+1}|\Lambda)] + var(E[X_{i,T+1}|\Lambda]) \\ &= v + a. \end{aligned}$$

(*b*)

$$\begin{aligned} E[(X_{i,T+1} - p_1)^2] &= E[(X_{i,T+1} - E[X_{i,T+1}|\Lambda])^2] \\ &= E\big[E[(X_{i,T+1} - E[X_{i,T+1}|\Lambda)^2]|\Lambda\big] \\ &= E[var(X_{i,T+1}|\Lambda)] \\ &= v. \end{aligned}$$

(c)

$$\begin{aligned}
E[(X_{i,T+1} - p_2)^2] &= E[(X_{i,T+1} - E[X_{i,T+1}|X_{1,1}, \ldots, X_{1,T}, \ldots, X_{I,1}, \ldots, X_{I,T}])^2] \\
&= E[(X_{i,T+1} - E[X_{i,T+1}|\vec{X}])^2].
\end{aligned}$$

Now, using Lemma 3.1 yields

$$\begin{aligned}
E[(X_{i,T+1} - E[X_{i,T+1}|\vec{X}])^2] &= E[(X_{i,T+1} - E[X_{i,T+1}|\Lambda])^2] \\
&\quad + E[(E[X_{i,T+1}|\Lambda] - E[X_{i,T+1}|\vec{X}])^2] \\
&= E[var(X_{i,T+1}|\Lambda))] + E[(E[X_{i,T+1}|\Lambda] - E[X_{i,T+1}|\vec{X}])^2] \\
&= v + E\Big[E\big[(E[X_{i,T+1}|\Lambda] - E[X_{i,T+1}|\vec{X}])^2|\vec{X}\Big] \\
&= v + E\Big[E[(E[X_{i,T+1}|\Lambda] - E\big[E[X_{i,T+1}|\Lambda]|\vec{X}\big])^2|\vec{X}\Big] \\
&= v + E\Big[var(E[X_{i,T+1}|\Lambda]|\vec{X})\Big].
\end{aligned}$$

The final inequality in proposition 3.3 follows from

$$v \leq v + E[var(E[X_{i,T+1}|\Lambda]|\vec{X})]$$

and $\triangle_0 \subset \triangle_1$. This concludes the proof. $\square$

CHAPTER IV

MAIN RESULTS

From section (3.2), we know that for a fixed individual $j = 1, 2, \ldots, I$, the conditional expectation $E[X_{j,T+1}|\vec{X}]$ gives our best estimate of next period claim in the sense of the mean squared prediction error and also gives our desired premium. For convenience, denote the random vector

$$\vec{X} = (\vec{X_1}, \vec{X_2}, \ldots, \vec{X_I}) \tag{4.1}$$

where $\vec{X_j} = (X_{j,1}, X_{j,2}, \ldots, X_{j,T})'$ for $j = 1, 2, \ldots, I$, which gives all the observable claims from all individuals and across T time period. The conditional expectation which is the so-called "Bayesian premium" can then be conveniently expressed as

$$E[X_{j,T+1}|\vec{X}] = \int x_{j,T+1} \cdot f_{X_{j,T+1}|\vec{X}}(x_{j,T+1}|\vec{x})dx_{j,T+1} \tag{4.2}$$

and the integral is the Riemann-Stieljes integral.

4.1 Method to Find Bayesian Premium

The Bayesian premium $E[X_{j,T+1}|\vec{X}]$ for a fixed individual $j = 1, 2, \ldots, I$ requires an explicit formula of conditional density $f_{X_{j,T+1}|\vec{X}}(x_{j,T+1}|\vec{x})$. To achieve this, we need a following lemma.

Lemma 4.1. *Let Λ be a random variable satisfying the assumption **A1** to **A4** and $\vec{X}$ be the vector of all observable claims which is defined in (4.1). The joint density of $\vec{X}$ and the overall risk parameter Λ can be expressed as*

$$f_{\vec{X},\Lambda}(\vec{x}, \lambda) = \prod_{i=1}^{I} f_{\vec{X_i}|\Lambda}(\vec{x_i}|\lambda) \times f_\Lambda(\lambda). \tag{4.3}$$

Proof. By definition of conditional density and assumption **A3**, we have

$$\begin{aligned} f_{\vec{X},\Lambda}(\vec{x},\lambda) &= f_{\vec{X}|\Lambda}(\vec{x},|\lambda) \times f_\Lambda(\lambda) \\ &= \prod_{i=1}^{I} f_{\vec{X}_i|\Lambda}(\vec{x_i}|\lambda) \times f_\Lambda(\lambda). \end{aligned} \tag{4.4}$$

□

Next, we compute $f_{\Lambda|\vec{X}}(\lambda|\vec{x})$. Using the definition of conditional density and equation (4.4), we have

$$\begin{aligned} f_{\Lambda|\vec{X}}(\lambda|\vec{x}) &= f_{\vec{X},\Lambda}(\vec{x},\lambda) \times \frac{1}{f_{\vec{X}}(\vec{x})} \\ &= C \times \prod_{i=1}^{I} f_{\vec{X}_i|\Lambda}(\vec{x_i}|\lambda) \times f_\Lambda(\lambda) \end{aligned} \tag{4.5}$$

where C is a normalizing constant and can be expressed as

$$C = \frac{1}{f_{\vec{X}}(\vec{x})} = (\int f_{\vec{X},\Lambda}(\vec{x},\lambda) d\lambda)^{-1}.$$

We now state the result for desired conditional density of $X_{j,T+1}|\vec{X}$.

Theorem 4.2. *Suppose the random variable* Λ *and the random vector* $\vec{X}$ *satisfy all assumptions as in Lemma 4.1. The conditional density of* $X_{j,T+1}|\vec{X}$ *can be expressed as*

$$f_{X_{j,T+1}|\vec{X}}(x_{j,T+1}|\vec{x}) = \int f_{X_{j,T+1}|\Lambda}(x_{j,T+1}|\lambda) \times f_{\Lambda|\vec{X}}(\lambda|\vec{x})\ d\lambda. \tag{4.6}$$

Proof. By definition of conditional density, we have

$$f_{X_{j,T+1}|\vec{X}}(x_{j,T+1}|\vec{x}) = f_{X_{j,T+1},\vec{X}}(x_{j,T+1},\vec{x}) \times \frac{1}{f_{\vec{X}}(\vec{x})} \tag{4.7}$$

Considering the right-hand side of equation (4.7), we note that the term $f_{X_{j,T+1},\vec{X}}(x_{j,T+1},\vec{x})$ can be calculated by integrating $f_{X_{j,T+1},\vec{X},\Lambda}(x_{j,T+1},\vec{x},\lambda)$ with respect to λ. Hence, firstly we shall compute $f_{X_{j,T+1},\vec{X},\Lambda}(x_{j,T+1},\vec{x},\lambda)$. By definition of conditional density, assumptions **A3** and **A4** then we have

$$\begin{aligned} f_{X_{j,T+1},\vec{X},\Lambda}(x_{j,T+1},\vec{x},\lambda) &= f_{X_{j,T+1},\vec{X}|\Lambda}(x_{j,T+1},\vec{x}|\lambda) \times f_\Lambda(\lambda) \\ &= f_{X_{j,T+1}|\Lambda}(x_{j,T+1}|\lambda) \times f_{\vec{X}|\Lambda}(\vec{x}|\lambda) \times f_\Lambda(\lambda) \\ &= f_{X_{j,T+1}|\Lambda}(x_{j,T+1}|\lambda) \times \prod_{i=1}^{I} f_{X_i|\Lambda}(x_i,|\lambda) \times f_\Lambda(\lambda) \end{aligned} \tag{4.8}$$

By equation (4.5)

$$f_{\Lambda|\vec{X}}(\lambda|\vec{x}) \;=\; \prod_{i=1}^{I} f_{\vec{X}_i|\Lambda}(\vec{x_i}|\lambda) \times f_{\Lambda}(\lambda) \times \frac{1}{f_{\vec{X}}(\vec{x})},$$

so we have

$$\prod_{i=1}^{I} f_{\vec{X}_i|\Lambda}(\vec{x_i}|\lambda) \times f_{\Lambda}(\lambda) \;=\; f_{\Lambda|\vec{X}}(\lambda|\vec{x}) \times f_{\vec{X}}(\vec{x}). \tag{4.9}$$

Substituting (4.9) into (4.8) yields

$$f_{X_{j,T+1},\vec{X},\Lambda}(x_{j,T+1},\vec{x},\lambda) \;=\; f_{X_{j,T+1}|\Lambda}(x_{j,T+1}|\lambda) \times f_{\Lambda|\vec{X}}(\lambda|\vec{x}) \times f_{\vec{X}}(\vec{x}). \tag{4.10}$$

Next, integrating (4.10) with respect to λ, we have

$$\begin{aligned} f_{X_{j,T+1},\vec{X}}(x_{j,T+1},\vec{x}) &= \int f_{X_{j,T+1}|\Lambda}(x_{j,T+1}|\lambda) \times f_{\Lambda|\vec{X}}(\lambda|\vec{x}) \times f_{\vec{X}}(\vec{x})\; d\lambda \\ &= f_{\vec{X}}(\vec{x}) \int f_{X_{j,T+1}|\Lambda}(x_{i,T+1}|\lambda) \times f_{\Lambda|\vec{X}}(\lambda|\vec{x})\; d\lambda. \end{aligned} \tag{4.11}$$

Substituting (4.11) into (4.7), we obtain

$$f_{X_{j,T+1}|\vec{X}}(x_{j,T+1}|\vec{x}) \;=\; \int f_{X_{j,T+1}|\Lambda}(x_{j,T+1}|\lambda) \times f_{\Lambda|\vec{X}}(\lambda|\vec{x})\; d\lambda.$$

The proof is now complete. □

The purpose of the theorem above is to derive an explicit expression for the conditional density in terms of all available or given information. First, notice from this theorem that this conditional density involves the product of the conditional density which according to assumption **A4** is known and given, and that of

$$\begin{aligned} f_{\Lambda|\vec{X}}(\lambda|\vec{x}) &= f_{\Lambda,\vec{X}}(\lambda,\vec{x}) \times \frac{1}{f_{\vec{X}}(\vec{x})} \\ &= \frac{f_{\vec{X}|\Lambda}(\vec{x}|\lambda) \times f_{\Lambda}(\lambda)}{f_{\vec{X}}(\vec{x})}, \end{aligned}$$

for which the numerator can be evaluated using Lemma (4.1) together with the independence of common effect.

Although the setting of the proposed model renders it non-Bayesian superficially, the nature of it is still very much Bayesian. As such, it will inherit all the benefits from a Bayesian solution, most notably, the smallest mean squared prediction error. See Klugman (1992) for a discussion of this.

4.2 Bayesian Premium with Normal Common Effect

In this section, we shall use equation (4.2) by applying Theorem 4.2 to find the Bayesian premium when the common effect Λ is normally distributed and the claim amounts are lognormally or normally distributed. Moreover, we can use this result to derive an explicit expression for the predicted claim amount.

4.2.1 Bayesian Premium with Lognormal Claim Amounts

In this subsection, we assume that the claim amounts of each policyholder in the rating class are lognormally distributed and characterized by the common effect Λ introduced in previous chapter. For convenience, we write $X|\lambda =: X|\Lambda = \lambda$. Before we consider special cases of the normal common effect assumption, let us consider the more general case where we have I insured individuals and where the common effects have variances which are not necessarily unit. To carry out the derivation, we make the following assumptions:

L1. the random variables $X_{j,t}|\lambda$ are lognormally distributed, i.e.,

$$X_{j,t}|\lambda \sim LN(\mu_j + \lambda, \sigma_x^2) \quad \text{for } j = 1, 2, \ldots, I, \quad \text{and } t = 1, 2, \ldots, T,$$

where μ_j is a constant depending on individual j.
Then $X_{j,t}|\lambda$ has a mean of $e^{(\mu_j+\lambda)+\frac{\sigma_x^2}{2}}$ and a variance of $(e^{\sigma_x^2} - 1)(e^{2(\mu_j+\lambda)+\sigma_x^2})$,

L2. the *over all* common effect λ is normally distributed with mean μ_λ and variance σ_λ^2.

It follows therefore that we have

$$f_{X_{j,t}|\Lambda}(x_{j,t}|\lambda) = \frac{1}{x_{j,t}\sigma_x\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{ln x_{j,t} - (\mu_j+\lambda)}{\sigma_x}\right)^2}, \quad \text{and}$$

$$f_\Lambda(\lambda) = \frac{1}{\sigma_\lambda\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\lambda-\mu_\lambda}{\sigma_\lambda}\right)^2}.$$

A useful application of Theorem (4.2) appears in the following theorem.

Theorem 4.3. *Suppose the random variable* Λ *and the random vector* $\vec{X}$ *satisfy all assumptions as in Lemma 4.1. Assume further that* $X_{j,t}|\lambda$ *and common effect* Λ *satisfy* ***L1*** *and* ***L2****, respectively. Then the Bayesian premium can be written as*

$$E[X_{j,T+1}|\vec{X}] = e^{\left[\frac{\sigma_\lambda^2\left(\sum_{i=1}^{I}\sum_{t=1}^{T} lnX_{i,t} - T\sum_{i=1}^{I}\mu_i + \mu_j IT\right) + \sigma_x^2\left(\mu_\lambda + \mu_j\right)}{\sigma_\lambda^2 IT + \sigma_x^2}\right]} e^{\left[\frac{\left(\sigma_\lambda^2(IT+1)+\sigma_x^2\right)\sigma_x^2}{2(\sigma_\lambda^2 IT+\sigma_x^2)}\right]}. \quad (4.12)$$

for $j = 1, 2, \ldots, I$.

Proof. Assume assumptions **L1** and **L2** hold, i.e.,

$$f_{X_{j,t}|\Lambda}(x_{j,t}|\lambda) = \frac{1}{x_j\sigma_x\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(\frac{lnx_j-(\mu_j+\lambda)}{\sigma_x}\right)^2}, \quad \text{and}$$

$$f_\Lambda(\lambda) = \frac{1}{\sigma_\lambda\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(\frac{\lambda-\mu_\lambda}{\sigma_\lambda}\right)^2}.$$

Recall that

$$E[X_{j,T+1}|\vec{X}] = \int x_{j,T+1}\cdot f_{X_{j,T+1}|\vec{X}}(x_{j,T+1}|\vec{x})dx_{j,T+1}$$

The main purpose is to derive the density of $X_{j,T+1}|\vec{X}$ where without loss of generality, we fix $j = 1$. Applying Theorem (4.2), we have

$$\begin{aligned} f_{X_{1,T+1}|\vec{X}}(x_{1,T+1}|\vec{x}) &= \int f_{X_{1,T+1}|\Lambda}(x_{1,T+1}|\lambda)\times f_{\Lambda|\vec{X}}(\lambda|\vec{x})\, d\lambda \\ &= \int f_{X_{1,T+1}|\Lambda}(x_{1,T+1}|\lambda)\times f_{\Lambda,\vec{X}}(\lambda,\vec{x})\times\frac{1}{f_{\vec{X}}(\vec{x})}\, d\lambda \\ &= C_1\int f_{X_{1,T+1}|\Lambda}(x_{1,T+1}|\lambda)\times f_{\Lambda,\vec{X}}(\lambda,\vec{x})\, d\lambda. \quad (4.13) \end{aligned}$$

where $C_1 = \frac{1}{f_{\vec{X}}(\vec{x})}$ is just a normalizing constant and does not have to be solved for explicitly. Here, and in the subsequent development, the limits of the integrals are the entire real line. The conditional density $f_{X_{1,T+1}|\Lambda}(x_{1,T+1}|\lambda)$ is already known to be

$$f_{X_{1,T+1}|\Lambda}(x_{1,T+1}|\lambda) = \frac{1}{x_{1,T+1}\sigma_x\sqrt{2\pi}}\, e^{-\frac{1}{T+1}\left(\frac{lnx_{1,T+1}-(\mu_1+\lambda)}{\sigma_x}\right)^2}. \quad (4.14)$$

The joint density $f_{\Lambda,\vec{X}}(\lambda,\vec{x})$ can be derived by utilizing lemma (4.1), giving

$$\begin{aligned} f_{\vec{X},\Lambda}(\vec{x},\lambda) &= f_{\vec{X}|\Lambda}(\vec{x},|\lambda) \times f_{\Lambda}(\lambda) \\ &= \prod_{i=1}^{I} f_{\vec{X}_i|\Lambda}(\vec{x_i},|\lambda) \times f_{\Lambda}(\lambda) \\ &= \prod_{i=1}^{I}\Big(\prod_{t=1}^{T}\frac{1}{x_{i,t}\sigma_x\sqrt{2\pi}}\ e^{-\frac{1}{2}\left(\frac{lnx_{i,t}-(\mu_i+\lambda)}{\sigma_x}\right)^2}\Big)\frac{1}{\sigma_\lambda\sqrt{2\pi}}\ e^{-\frac{1}{2}\left(\frac{\lambda-\mu_\lambda}{\sigma_\lambda}\right)^2}. \end{aligned} \tag{4.15}$$

Substituting (4.14) and (4.15) into (4.13), we have

$$\begin{aligned} f_{X_{1,T+1}|\vec{X}}(x_{1,T+1}|\vec{x}) &= C_1\int f_{X_{1,T+1}|\Lambda}(x_{1,T+1}|\lambda) \times f_{\Lambda,\vec{X}}(\lambda,\vec{x})\ d\lambda \\ &= C_1\int \frac{1}{x_{1,T+1}\sigma_x\sqrt{2\pi}}\ e^{-\frac{1}{2}\left(\frac{lnx_{1,T+1}-(\mu_1+\lambda)}{\sigma_x}\right)^2} \\ &\quad \times\prod_{i=1}^{I}\Big(\prod_{t=1}^{T}\frac{1}{x_{i,t}\sigma_x\sqrt{2\pi}}\ e^{-\frac{1}{2}\left(\frac{lnx_{i,t}-(\mu_i+\lambda)}{\sigma_x}\right)^2}\Big)\frac{1}{\sigma_\lambda\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{\lambda-\mu_\lambda}{\sigma_\lambda}\right)^2} d\lambda \\ &= \int\frac{C_1}{x_{1,T+1}\prod_{i=1}^{I}\big(\prod_{t=1}^{T}x_{i,t}\big)(\sigma_x)^{IT+1}(2\pi)^{\frac{IT+2}{2}}\sigma_\lambda} \\ &\quad \times e^{-\frac{1}{2}\left[\sum_{t=1}^{T+1}\left(\frac{lnx_{1,t}-(\mu_1+\lambda)}{\sigma_x}\right)^2\right]} \\ &\quad \times e^{-\frac{1}{2}\left[\sum_{i=2}^{I}\sum_{t=1}^{T}\left(\frac{lnx_{i,t}-(\mu_i+\lambda)}{\sigma_x}\right)^2\right]} \times\ e^{-\frac{1}{2}\left(\frac{\lambda-\mu_\lambda}{\sigma_\lambda}\right)^2}\ d\lambda. \end{aligned} \tag{4.16}$$

Taking the terms containing λ from (4.16), after simplifying, we have

$$\begin{aligned} &\int\frac{1}{2\pi}\ e^{-\frac{1}{2}\left[\sum_{t=1}^{T+1}\left(\frac{lnx_{1,t}-(\mu_1+\lambda)}{\sigma_x}\right)^2+\sum_{i=2}^{I}\sum_{t=1}^{T}\left(\frac{lnx_{i,t}-(\mu_i+\lambda)}{\sigma_x}\right)^2\right]}\times e^{-\frac{1}{2}\left(\frac{\lambda-\mu_\lambda}{\sigma_\lambda}\right)^2}\ d\lambda \\ &= \int\frac{1}{2\pi}\ e^{-\frac{IT+1}{2\sigma_x^2}\left[\left(\lambda-\frac{\left(\sum_{i=1}^{I}\sum_{t=1}^{T}lnx_{i,t}+lnx_{1,T+1}\right)-\left(T\sum_{i=1}^{I}\mu_i+\mu_1\right)}{IT+1}\right)^2\right]}\ e^{-\frac{1}{2}\left(\frac{\lambda-\mu_\lambda}{\sigma_\lambda}\right)^2} \\ &\quad\times e^{-\frac{1}{2\sigma_x^2}\left[\frac{-\left(\left[\sum_{i=1}^{I}\sum_{t=1}^{T}lnx_{i,t}+lnx_{1,T+1}\right]-\left[T\sum_{i=1}^{I}\mu_i+\mu_1\right]\right)^2}{IT+1}\right]} \\ &\quad\times e^{-\frac{1}{2\sigma_x^2}\left[\ \left(\sum_{i=1}^{I}\sum_{t=1}^{T}(lnx_{i,t})^2+(lnx_{1,T+1})^2\right)-2\left(\sum_{i=1}^{I}(\mu_i\sum_{t=1}^{T}lnx_{i,t})+(lnx_{1,T+1}\cdot\mu_1)\right)\right]}d\lambda. \\ &\quad\times e^{-\frac{1}{2\sigma_x^2}\left[\left(T\sum_{i=1}^{I}\mu_i^2+\mu_1^2\right)\right]}d\lambda. \end{aligned} \tag{4.17}$$

(see proof of (4.17) in Appendix A1).

Notice the part of (4.17) can be simplified as follows:

$$
\begin{aligned}
&\frac{1}{2\pi}\ e^{-\frac{IT+1}{2\sigma_x^2}\left[\left(\lambda-\frac{\left(\sum_{i=1}^{I}\sum_{t=1}^{T}lnx_{i,t}+lnx_{1,T+1}\right)-\left(T\sum_{i=1}^{I}\mu_i+\mu_1\right)}{IT+1}\right)^2\right]}\ e^{-\frac{1}{2}\left(\frac{\lambda-\mu_\lambda}{\sigma_\lambda}\right)^2} \\
&= \frac{\sigma_x}{\sqrt{IT+1}}\Big(\frac{\sqrt{IT+1}}{\sigma_x}\times\frac{1}{\sqrt{2\pi}}\ e^{-\frac{IT+1}{2\sigma_x^2}\left[\left(\lambda-\frac{\left(\sum_{i=1}^{I}\sum_{t=1}^{T}lnx_{i,t}+lnx_{1,T+1}\right)-\left(T\sum_{i=1}^{I}\mu_i+\mu_1\right)}{IT+1}\right)^2\right]}\Big) \\
&\quad\times\sigma_\lambda\Big(\frac{1}{\sigma_\lambda}\times\frac{1}{\sqrt{2\pi}}\ e^{-\frac{1}{2}\left(\frac{\lambda-\mu_\lambda}{\sigma_\lambda}\right)^2}\Big) \\
&= \frac{\sigma_x\sigma_\lambda}{\sqrt{IT+1}}\times\varphi(\frac{\sqrt{IT+1}}{\sigma_x}\Big[\lambda-\frac{\big(\sum_{i=1}^{I}\sum_{t=1}^{T}lnx_{i,t}+lnx_{1,T+1}\big)-\big(T\sum_{i=1}^{I}\mu_i+\mu_1\big)}{IT+1}\Big]) \\
&\quad\times\varphi(\frac{\lambda-\mu_\lambda}{\sigma_\lambda}),
\end{aligned} \tag{4.18}
$$

where $\varphi(z)$ is the standard normal density.

At this point, we use a result from Valdez (2004) to simplify (4.18). This result states that for $\varphi(z)$ and any constants a and b, the following is true:

$$\int_{-\infty}^{\infty}\varphi(z)\varphi(a-bz)\ =\ \frac{1}{\sqrt{b^2+1}}\ \varphi(\sqrt{\frac{a^2}{b^2+1}}). \tag{4.19}$$

Thus, by letting $z=\frac{\lambda-\mu_\lambda}{\sigma_\lambda}$ so that $dz=\frac{1}{\sigma_\lambda}d\lambda$, then applying (4.19) to (4.18) and after simplifying, we have

$$\int \frac{1}{2\pi}\ e^{-\frac{1}{2}\left[\sum_{t=1}^{T+1}\left(\frac{lnx_{1,t}-(\mu_1+\lambda)}{\sigma_x}\right)^2+\sum_{i=2}^{I}\sum_{t=1}^{T}\left(\frac{lnx_{i,t}-(\mu_i+\lambda)}{\sigma_x}\right)^2\right]} \times\ e^{-\frac{1}{2}\left(\frac{\lambda-\mu_\lambda}{\sigma_\lambda}\right)^2}\ d\lambda$$

$$= \int \varphi(\frac{\sqrt{IT+1}}{\sigma_x}\ \left[\lambda - \frac{\left(\sum_{i=1}^{I}\sum_{t=1}^{T} lnx_{i,t}+lnx_{1,T+1}\right)-\left(T\sum_{i=1}^{I}\mu_i+\mu_1\right)}{IT+1}\right])$$

$$\times\ \frac{\sigma_x\sigma_\lambda}{\sqrt{IT+1}} \times\ \varphi(\frac{\lambda-\mu_\lambda}{\sigma_\lambda}) \times\ e^{-\frac{1}{2\sigma_x^2}\left[\frac{-\left(\left[\sum_{i=1}^{I}\sum_{t=1}^{T} lnx_{i,t}+lnx_{1,T+1}\right]-\left[T\sum_{i=1}^{I}\mu_i+\mu_1\right]\right)^2}{IT+1}\right]}$$

$$\times\ e^{-\frac{1}{2\sigma_x^2}\left[\ \left(\sum_{i=1}^{I}\sum_{t=1}^{T}(lnx_{i,t})^2+(lnx_{1,T+1})^2\right)-2\left(\sum_{i=1}^{I}(\mu_i\sum_{t=1}^{T} lnx_{i,t})+(lnx_{1,T+1}\cdot\mu_1)\right)\right]}$$

$$\times\ e^{-\frac{1}{2\sigma_x^2}\left[T\sum_{i=1}^{I}\mu_i^2+\mu_1^2\right]}d\lambda$$

$$= \varphi(\sqrt{\frac{IT+1}{\sigma_\lambda^2(IT+1)+\sigma_x^2}}\ \left[\frac{\left(\sum_{i=1}^{I}\sum_{t=1}^{T} lnx_{i,t}+lnx_{1,T+1}\right)-\left(T\sum_{i=1}^{I}\mu_i+\mu_1\right)}{IT+1}-\mu_\lambda\right])$$

$$\times\ \frac{\sigma_x\sigma_\lambda}{\sqrt{IT+1}} \times\ \frac{\sigma_\lambda}{\sqrt{\frac{\sigma_\lambda^2(IT+1)}{\sigma_x^2}+1}} \times\ e^{-\frac{1}{2\sigma_x^2}\left[\frac{-\left(\left[\sum_{i=1}^{I}\sum_{t=1}^{T} lnx_{i,t}+lnx_{1,T+1}\right]-\left[T\sum_{i=1}^{I}\mu_i+\mu_1\right]\right)^2}{IT+1}\right]}$$

$$\times\ e^{-\frac{1}{2\sigma_x^2}\left[\ \left(\sum_{i=1}^{I}\sum_{t=1}^{T}(lnx_{i,t})^2+(lnx_{1,T+1})^2\right)-2\left(\sum_{i=1}^{I}(\mu_i\sum_{t=1}^{T} lnx_{i,t})+(lnx_{1,T+1}\cdot\mu_1)\right)\right]}$$

$$\times\ e^{-\frac{1}{2\sigma_x^2}\left[T\sum_{i=1}^{I}\mu_i^2+\mu_1^2\right]}$$

$$= \varphi(\sqrt{\frac{IT+1}{\sigma_\lambda^2(IT+1)+\sigma_x^2}}\ \left[\frac{\left(\sum_{i=1}^{I}\sum_{t=1}^{T} lnx_{i,t}+lnx_{1,T+1}\right)-\left(T\sum_{i=1}^{I}\mu_i+\mu_1\right)}{IT+1}-\mu_\lambda\right])$$

$$\times\ \frac{\sigma_x\sigma_\lambda^2}{\sqrt{IT+1}\sqrt{\frac{\sigma_\lambda^2(IT+1)}{\sigma_x^2}+1}} \times e^{-\frac{1}{2\sigma_x^2}\left[\frac{-\left(\left[\sum_{i=1}^{I}\sum_{t=1}^{T} lnx_{i,t}+lnx_{1,T+1}\right]-\left[T\sum_{i=1}^{I}\mu_i+\mu_1\right]\right)^2}{IT+1}\right]}$$

$$\times\ e^{-\frac{1}{2\sigma_x^2}\left[\ \left(\sum_{i=1}^{I}\sum_{t=1}^{T}(lnx_{i,t})^2+(lnx_{1,T+1})^2\right)-2\left(\sum_{i=1}^{I}(\mu_i\sum_{t=1}^{T} lnx_{i,t})+(lnx_{1,T+1}\cdot\mu_1)\right)\right]}$$

$$\times\ e^{-\frac{1}{2\sigma_x^2}\left[T\sum_{i=1}^{I}\mu_i^2+\mu_1^2\right]}. \qquad (4.20)$$

(see proof of (4.20) in Appendix A2)

We now continue from (4.16) by substituting (4.20) back into the equation to obtain

$$
\begin{aligned}
& f_{X_{1,T+1}|\vec{X}}(x_{1,T+1}|\vec{x}) \\
= & \int \frac{C_1}{x_{1,T+1}(\prod_{i=1}^{I}\prod_{t=1}^{T} x_{i,t})(\sigma_x)^{IT+1}(2\pi)^{\frac{IT+2}{2}}\sigma_\lambda} \; e^{-\frac{1}{2}\left[\sum_{t=1}^{T+1}\left(\frac{lnx_{1,t}-(\mu_1+\lambda)}{\sigma_x}\right)^2\right]} \\
& \times e^{-\frac{1}{2}\left[\sum_{i=2}^{I}\sum_{t=1}^{T}\left(\frac{lnx_{i,t}-(\mu_i+\lambda)}{\sigma_x}\right)^2\right]} \times \; e^{-\frac{1}{2}\left(\frac{\lambda-\mu_\lambda}{\sigma_\lambda}\right)^2} \; d\lambda \\
= & \varphi(\sqrt{\frac{IT+1}{\sigma_\lambda^2(IT+1)+\sigma_x^2}} \; \Big[\frac{\left(\sum_{i=1}^{I}\sum_{t=1}^{T} lnx_{i,t} + lnx_{1,T+1}\right) - \left(T\sum_{i=1}^{I}\mu_i + \mu_1\right)}{IT+1} - \mu_\lambda\Big]) \\
& \times \; \frac{\sigma_x \sigma_\lambda^2}{\sqrt{IT+1}\sqrt{\frac{\sigma_\lambda^2(IT+1)}{\sigma_x^2}+1}} \times \; e^{-\frac{1}{2\sigma_x^2}\left[\frac{-\left(\left[\sum_{i=1}^{I}\sum_{t=1}^{T} lnx_{i,t}+lnx_{1,T+1}\right]-\left[T\sum_{i=1}^{I}\mu_i+\mu_1\right]\right)^2}{IT+1}\right]} \\
& \times \; e^{-\frac{1}{2\sigma_x^2}\left[\;\left(\sum_{i=1}^{I}\sum_{t=1}^{T}(lnx_{i,t})^2+(lnx_{1,T+1})^2\right)-2\left(\sum_{i=1}^{I}(\mu_i\sum_{t=1}^{T} lnx_{i,t})+(lnx_{1,T+1}\cdot\mu_1)\right)\right]} \\
& \times \; e^{-\frac{1}{2\sigma_x^2}\left[T\sum_{i=1}^{I}\mu_i^2+\mu_1^2\right]} \times \frac{C_1}{x_{1,T+1}(\prod_{i=1}^{I}\prod_{t=1}^{T} x_{i,t})(\sigma_x)^{IT+1}(2\pi)^{\frac{IT+1}{2}}\sigma_\lambda} \\
= & \frac{1}{\sqrt{2\pi}} \times e^{-\frac{1}{2}\left(\frac{IT+1}{\sigma_\lambda^2(IT+1)+\sigma_x^2}\;\left[\frac{\left(\sum_{i=1}^{I}\sum_{t=1}^{T} lnx_{i,t}+lnx_{1,T+1}\right)-\left(T\sum_{i=1}^{I}\mu_i+\mu_1\right)}{IT+1}-\mu_\lambda\right]^2\right)} \\
& \times \; e^{-\frac{1}{2\sigma_x^2}\left[\frac{-\left(\left[\sum_{i=1}^{I}\sum_{t=1}^{T} lnx_{i,t}+lnx_{1,T+1}\right]-\left[T\sum_{i=1}^{I}\mu_i+\mu_1\right]\right)^2}{IT+1}\right]} \\
& \times \; e^{-\frac{1}{2\sigma_x^2}\left[\;\left(\sum_{i=1}^{I}\sum_{t=1}^{T}(lnx_{i,t})^2+(lnx_{1,T+1})^2\right)-2\left(\sum_{i=1}^{I}(\mu_i\sum_{t=1}^{T} lnx_{i,t})+(lnx_{1,T+1}\cdot\mu_1)\right)\right]} \\
& \times \; e^{-\frac{1}{2\sigma_x^2}\left[T\sum_{i=1}^{I}\mu_i^2+\mu_1^2\right]} \\
& \times \frac{C_1\sigma_\lambda}{x_{1,T+1}(\prod_{i=1}^{I}\prod_{t=1}^{T} x_{i,t})(\sigma_x)^{IT}(2\pi)^{\frac{IT+1}{2}}\sqrt{IT+1}\sqrt{\frac{\sigma_\lambda^2(IT+1)}{\sigma_x^2}+1}}.
\end{aligned}
\tag{4.21}
$$

By setting a constant

$$
C_2 \;=\; \frac{C_1\sigma_\lambda}{(\prod_{i=1}^{I}\prod_{t=1}^{T} x_{i,t})(\sigma_x)^{IT}(2\pi)^{\frac{IT+1}{2}}\sqrt{IT+1}\sqrt{\frac{\sigma_\lambda^2(IT+1)}{\sigma_x^2}+1}},
$$

then we have

$$
\begin{aligned}
&f_{X_{1,T+1}|\vec{X}}(x_{1,T+1}|\vec{x}) \\
&= \frac{C_2}{X_{1,T+1}}\ e^{-\frac{1}{2}\left(\frac{IT+1}{\sigma_\lambda^2(IT+1)+\sigma_x^2}\ \left[\frac{\left(\sum_{i=1}^{I}\sum_{t=1}^{T} lnx_{i,t}+lnx_{1,T+1}\right)-\left(T\sum_{i=1}^{I}\mu_i+\mu_1\right)}{IT+1}-\mu_\lambda\right]^2\right)} \\
&\quad\times\ e^{-\frac{1}{2\sigma_x^2}\left[\frac{-\left(\left[\sum_{i=1}^{I}\sum_{t=1}^{T} lnx_{i,t}+lnx_{1,T+1}\right]-\left[T\sum_{i=1}^{I}\mu_i+\mu_1\right]\right)^2}{IT+1}\right]} \\
&\quad\times\ e^{-\frac{1}{2\sigma_x^2}\left[\left(\sum_{i=1}^{I}\sum_{t=1}^{T}(lnx_{i,t})^2+(lnx_{1,T+1})^2\right)-2\left(\sum_{i=1}^{I}(\mu_i\sum_{t=1}^{t} lnx_{i,t})+(lnx_{1,T+1}\cdot\mu_1)\right)\right]} \\
&\quad\times\ e^{-\frac{1}{2\sigma_x^2}\left[T\sum_{i=1}^{I}\mu_i^2+\mu_1^2\right]} \\
&= \frac{C_2}{X_{1,T+1}}\ e^{-\frac{1}{2}\left(\frac{IT+1}{\sigma_\lambda^2(IT+1)+\sigma_x^2}\ \left[\frac{\left(\sum_{i=1}^{I}\sum_{t=1}^{T} lnx_{i,t}+lnx_{1,T+1}\right)-\left(T\sum_{i=1}^{I}\mu_i+\mu_1\right)}{IT+1}\right]^2\right)} \\
&\quad\times\ e^{-\frac{1}{2}\left(\frac{IT+1}{\sigma_\lambda^2(IT+1)+\sigma_x^2}\ \left[-2\mu_\lambda\frac{\left(\sum_{i=1}^{I}\sum_{t=1}^{T} lnx_{i,t}+lnx_{1,T+1}\right)-\left(T\sum_{i=1}^{I}\mu_i+\mu_1\right)}{IT+1}+(\mu_\lambda)^2\right]\right)} \\
&\quad\times\ e^{-\frac{1}{2\sigma_x^2}\left[\frac{-\left(\left[\sum_{i=1}^{I}\sum_{t=1}^{T} lnx_{i,t}+lnx_{1,T+1}\right]-\left[T\sum_{i=1}^{I}\mu_i+\mu_1\right]\right)^2}{IT+1}+\frac{(IT+1)\left(\sum_{i=1}^{I}\sum_{t=1}^{T}(lnx_{i,t})^2+(lnx_{1,T+1})^2\right)}{IT+1}\right]} \\
&\quad\times\ e^{-\frac{1}{2\sigma_x^2}\left[\frac{-2(IT+1)\left(\sum_{i=1}^{I}(\mu_i\sum_{t=1}^{T} lnx_{i,t})+(lnx_{1,T+1}\cdot\mu_1)\right)+(IT+1)\left(T\sum_{i=1}^{I}\mu_i^2+\mu_1^2\right)}{IT+1}\right]} \\
&= \frac{C_2}{X_{1,T+1}}\ e^{-\frac{1}{2}\left(\left(\frac{\sigma_x^2}{\sigma_x^2}\right)\times\frac{\left[\left(\sum_{i=1}^{I}\sum_{t=1}^{T} lnx_{i,t}+lnx_{1,T+1}\right)-\left(T\sum_{i=1}^{I}\mu_i+\mu_1\right)\right]^2}{(\sigma_\lambda^2(IT+1)+\sigma_x^2)(IT+1)}\right)} \\
&\quad\times\ e^{-\frac{1}{2}\left(\left(\frac{\sigma_x^2}{\sigma_x^2}\right)\times\left[\frac{-2\mu_\lambda(IT+1)\left[\left(\sum_{i=1}^{I}\sum_{t=1}^{T} lnx_{i,t}+lnx_{1,T+1}\right)-\left(T\sum_{i=1}^{I}\mu_i+\mu_1\right)\right]+(\mu_\lambda)^2(IT+1)^2}{(\sigma_\lambda^2(IT+1)+\sigma_x^2)(IT+1)}\right]\right)} \\
&\quad\times\ e^{-\frac{1}{2}\left[\left(\frac{\sigma_\lambda^2(IT+1)+\sigma_x^2}{\sigma_\lambda^2(IT+1)+\sigma_x^2}\right)\times\frac{-\left(\left[\sum_{i=1}^{I}\sum_{t=1}^{T} lnx_{i,t}+lnx_{1,T+1}\right]-\left[T\sum_{i=1}^{I}\mu_i+\mu_1\right]\right)^2}{\sigma_x^2(IT+1)}\right]} \\
&\quad\times\ e^{-\frac{1}{2}\left[\left(\frac{\sigma_\lambda^2(IT+1)+\sigma_x^2}{\sigma_\lambda^2(IT+1)+\sigma_x^2}\right)\times\frac{(IT+1)\left(\sum_{i=1}^{I}\sum_{t=1}^{T}(lnx_{i,t})^2+(lnx_{1,T+1})^2\right)}{\sigma_x^2(IT+1)}\right]} \\
&\quad\times\ e^{-\frac{1}{2}\left[\left(\frac{\sigma_\lambda^2(IT+1)+\sigma_x^2}{\sigma_\lambda^2(IT+1)+\sigma_x^2}\right)\times\frac{-2(IT+1)\left(\sum_{i=1}^{I}(\mu_i\sum_{t=1}^{T} lnx_{i,t})+(lnx_{1,T+1}\cdot\mu_1)\right)+(IT+1)\left(T\sum_{i=1}^{I}\mu_i^2+\mu_1^2\right)}{\sigma_x^2(IT+1)}\right]}.
\end{aligned}
\tag{4.22}
$$

By extracting and regrouping (4.22), we have

$$
\begin{aligned}
& f_{X_{1,T+1}|\vec{X}}(x_{1,T+1}|\vec{x}) \\
&= \frac{C_2}{X_{1,T+1}}\; e^{-\frac{1}{2}\left(\frac{\sigma_x^2\left[\left(\sum_{i=1}^{I}\sum_{t=1}^{T} lnx_{i,t}+lnx_{1,T+1}\right)-\left(T\sum_{i=1}^{I}\mu_i+\mu_1\right)\right]^2}{(\sigma_\lambda^2(IT+1)+\sigma_x^2)\sigma_x^2(IT+1)}\right)} \\
&\quad\times\; e^{-\frac{1}{2}\left(\left[\frac{-2\mu_\lambda\sigma_x^2(IT+1)\left[\left(\sum_{i=1}^{I}\sum_{t=1}^{T} lnx_{i,t}+lnx_{1,T+1}\right)-\left(T\sum_{i=1}^{I}\mu_i+\mu_1\right)\right]+(\mu_\lambda)^2\sigma_x^2(IT+1)^2}{(\sigma_\lambda^2(IT+1)+\sigma_x^2)\sigma_x^2(IT+1)}\right]\right)} \\
&\quad\times\; e^{-\frac{1}{2}\left[\frac{-(\sigma_\lambda^2(IT+1))\left(\left[\sum_{i=1}^{I}\sum_{t=1}^{T} lnx_{i,t}+lnx_{1,T+1}\right]-\left[T\sum_{i=1}^{I}\mu_i+\mu_1\right]\right)^2}{(\sigma_\lambda^2(IT+1)+\sigma_x^2)\sigma_x^2(IT+1)}\right]} \\
&\quad\times\; e^{-\frac{1}{2}\left[\frac{\sigma_\lambda^2(IT+1)^2\left(\sum_{i=1}^{I}\sum_{t=1}^{T}(lnx_{i,t})^2+(lnx_{1,T+1})^2\right)}{(\sigma_\lambda^2(IT+1)+\sigma_x^2)\sigma_x^2(IT+1)}\right]} \\
&\quad\times\; e^{-\frac{1}{2}\left[\frac{-\sigma_x^2\left(\left[\sum_{i=1}^{I}\sum_{t=1}^{T} lnx_{i,t}+lnx_{1,T+1}\right]-\left[T\sum_{i=1}^{I}\mu_i+\mu_1\right]\right)^2}{(\sigma_\lambda^2(IT+1)+\sigma_x^2)\sigma_x^2(IT+1)}\right]} \\
&\quad\times\; e^{-\frac{1}{2}\left[\frac{\sigma_x^2(IT+1)\left(\sum_{i=1}^{I}\sum_{t=1}^{T}(lnx_{i,t})^2+(lnx_{1,T+1})^2\right)}{(\sigma_\lambda^2(IT+1)+\sigma_x^2)\sigma_x^2(IT+1)}\right]} \\
&\quad\times\; e^{-\frac{1}{2}\left[\frac{-2\sigma_\lambda^2(IT+1)^2\left(\sum_{i=1}^{I}(\mu_i\sum_{t=1}^{T} lnx_{i,t})+(lnx_{1,T+1}\cdot\mu_1)\right)+\sigma_\lambda^2(IT+1)^2\left(T\sum_{i=1}^{I}\mu_i^2+\mu_1^2\right)}{(\sigma_\lambda^2(IT+1)+\sigma_x^2)\sigma_x^2(IT+1)}\right]} \\
&\quad\times\; e^{-\frac{1}{2}\left[\frac{-2\sigma_x^2(IT+1)\left(\sum_{i=1}^{I}(\mu_i\sum_{t=1}^{T} lnx_{i,t})+(lnx_{1,T+1}\cdot\mu_1)\right)+\sigma_x^2(IT+1)\left(T\sum_{i=1}^{I}\mu_i^2+\mu_1^2\right)}{(\sigma_\lambda^2(IT+1)+\sigma_x^2)\sigma_x^2(IT+1)}\right]} \\
&= \frac{C_2}{X_{1,T+1}}\; e^{-\frac{1}{2}\left[\frac{-(\sigma_\lambda^2(IT+1))\left(\left[\sum_{i=1}^{I}\sum_{t=1}^{T} lnx_{i,t}+lnx_{1,T+1}\right]-\left[T\sum_{i=1}^{I}\mu_i+\mu_1\right]\right)^2}{(\sigma_\lambda^2(IT+1)+\sigma_x^2)\sigma_x^2(IT+1)}\right]} \\
&\quad\times\; e^{-\frac{1}{2}\left[\frac{\left(\sum_{i=1}^{I}\sum_{t=1}^{T}(lnx_{i,t})^2+(lnx_{1,T+1})^2\right)\left(\sigma_\lambda^2(IT+1)^2+\sigma_x^2(IT+1)\right)}{(\sigma_\lambda^2(IT+1)+\sigma_x^2)\sigma_x^2(IT+1)}\right]} \\
&\quad\times\; e^{-\frac{1}{2}\left[\frac{\left(\sum_{i=1}^{I}(\mu_i\sum_{t=1}^{T} lnx_{i,t})+(lnx_{1,T+1}\cdot\mu_1)\right)\left(-2\sigma_\lambda^2(IT+1)^2-2\sigma_x^2(IT+1)\right)}{(\sigma_\lambda^2(IT+1)+\sigma_x^2)\sigma_x^2(IT+1)}\right]} \\
&\quad\times\; e^{-\frac{1}{2}\left[\frac{\left(T\sum_{i=1}^{I}\mu_i^2+\mu_1^2\right)\left(\sigma_\lambda^2(IT+1)^2+\sigma_x^2(IT+1)\right)}{(\sigma_\lambda^2(IT+1)+\sigma_x^2)\sigma_x^2(IT+1)}\right]} \times\; e^{-\frac{1}{2}\left[\frac{-2\sigma_x^2\mu_\lambda(IT+1)\left(\sum_{i=1}^{I}\sum_{t=1}^{T} lnx_{i,t}+lnx_{1,T+1}\right)}{(\sigma_\lambda^2(IT+1)+\sigma_x^2)\sigma_x^2(IT+1)}\right]} \\
&\quad\times\; e^{-\frac{1}{2}\left[\frac{2\sigma_x^2\mu_\lambda(IT+1)\left(T\sum_{i=1}^{I}\mu_i+\mu_1\right)+\sigma_x^2(IT+1)^2\mu_\lambda^2}{(\sigma_\lambda^2(IT+1)+\sigma_x^2)\sigma_x^2(IT+1)}\right]}.
\end{aligned}
\tag{4.23}
$$

Extracting and grouping the terms containing $(ln\ x_{1,T+1})^2$ and $ln\ x_{1,T+1}$, to obtain

$$
\begin{aligned}
& f_{X_{1,T+1}|\vec{X}}(x_{1,T+1}|\vec{x}) \\
= \;& \frac{C_2}{X_{1,T+1}}\; e^{-\frac{1}{2}\left[\frac{-(\sigma_\lambda^2(IT+1))\left[\sum_{i=1}^{I}\sum_{t=1}^{T} lnx_{i,t}+lnx_{1,T+1}\right]^2}{(\sigma_\lambda^2(IT+1)+\sigma_x^2)\sigma_x^2(IT+1)}\right]} \\
& \times\; e^{-\frac{1}{2}\left[\frac{-(\sigma_\lambda^2(IT+1))\left(-2\left[\sum_{i=1}^{I}\sum_{t=1}^{T} lnx_{i,t}+lnx_{1,T+1}\right]\left[T\sum_{i=1}^{I}\mu_i+\mu_1\right]\right)-(\sigma_\lambda^2(IT+1))\left[T\sum_{i=1}^{I}\mu_i+\mu_1\right]^2}{(\sigma_\lambda^2(IT+1)+\sigma_x^2)\sigma_x^2(IT+1)}\right]} \\
& \times\; e^{-\frac{1}{2}\left[\frac{\left((lnx_{1,T+1})^2\right)\left(\sigma_\lambda^2(IT+1)^2+\sigma_x^2(IT+1)\right)+\left((lnx_{1,T+1})\right)\left(-2\mu_1\sigma_\lambda^2(I+1)^2-2\mu_1\sigma_x^2(IT+1)-2\sigma_x^2\mu_\lambda(IT+1)\right)}{(\sigma_\lambda^2(IT+1)+\sigma_x^2)\sigma_x^2(IT+1)}\right]} \\
& \times\; e^{-\frac{1}{2}\left[\frac{\left(\sum_{i=1}^{I}\sum_{t=1}^{T}(lnx_{i,t})^2\right)\left(\sigma_\lambda^2(IT+1)^2+\sigma_x^2(IT+1)\right)+\left(\sum_{i=1}^{I}(\mu_i\sum_{t=1}^{T} lnx_{i,t})\right)\left(-2\sigma_\lambda^2(IT+1)^2-2\sigma_x^2(IT+1)\right)}{(\sigma_\lambda^2(IT+1)+\sigma_x^2)\sigma_x^2(IT+1)}\right]} \\
& \times\; e^{-\frac{1}{2}\left[\frac{\left(T\sum_{i=1}^{I}\mu_i^2+\mu_1^2\right)\left(\sigma_\lambda^2(IT+1)^2+\sigma_x^2(IT+1)\right)-2\sigma_x^2\mu_\lambda(IT+1)\left(\sum_{i=1}^{I}\sum_{t=1}^{T} lnx_{i,t}\right)}{(\sigma_\lambda^2(IT+1)+\sigma_x^2)\sigma_x^2(IT+1)}\right]} \\
& \times\; e^{-\frac{1}{2}\left[\frac{2\sigma_x^2\mu_\lambda(IT+1)\left(T\sum_{i=1}^{I}\mu_i+\mu_1\right)+\sigma_x^2(IT+1)^2\mu_\lambda^2}{(\sigma_\lambda^2(IT+1)+\sigma_x^2)\sigma_x^2(IT+1)}\right]} \\
= \;& \frac{C_2}{X_{1,T+1}}\; e^{-\frac{1}{2}\left[\frac{-(\sigma_\lambda^2(IT+1))\left((\sum_{i=1}^{I}\sum_{t=1}^{T} lnx_{i,t})^2+2(\sum_{i=1}^{I}\sum_{t=1}^{T} lnx_{i,t})(lnx_{1,T+1})+(lnx_{1,T+1})^2\right)}{(\sigma_\lambda^2(IT+1)+\sigma_x^2)\sigma_x^2(IT+1)}\right]} \\
& \times\; e^{-\frac{1}{2}\left[\frac{2(\sigma_\lambda^2(IT+1))\left[(\sum_{i=1}^{I}\sum_{t=1}^{T} lnx_{i,t})(T\sum_{i=1}^{I}\mu_i+\mu_1)+(lnx_{1,T+1})(T\sum_{i=1}^{I}\mu_i+\mu_1)\right]}{(\sigma_\lambda^2(IT+1)+\sigma_x^2)\sigma_x^2(IT+1)}\right]} \\
& \times\; e^{-\frac{1}{2}\left[\frac{-(\sigma_\lambda^2(IT+1))\left[T\sum_{i=1}^{I}\mu_i+\mu_1\right]^2}{(\sigma_\lambda^2(IT+1)+\sigma_x^2)\sigma_x^2(IT+1)}\right]} \\
& \times\; e^{-\frac{1}{2}\left[\frac{\left((lnx_{1,T+1})^2\right)\left(\sigma_\lambda^2(IT+1)^2+\sigma_x^2(IT+1)\right)+\left((lnx_{1,T+1})\right)\left(-2\mu_1\sigma_\lambda^2(IT+1)^2-2\mu_1\sigma_x^2(IT+1)-2\sigma_x^2\mu_\lambda(IT+1)\right)}{(\sigma_\lambda^2(IT+1)+\sigma_x^2)\sigma_x^2(IT+1)}\right]} \\
& \times\; e^{-\frac{1}{2}\left[\frac{\left(\sum_{i=1}^{I}\sum_{t=1}^{T}(lnx_{i,t})^2\right)\left(\sigma_\lambda^2(IT+1)^2+\sigma_x^2(IT+1)\right)+\left(\sum_{i=1}^{I}(\mu_i\sum_{t=1}^{T} lnx_{i,t})\right)\left(-2\sigma_\lambda^2(IT+1)^2-2\sigma_x^2(IT+1)\right)}{(\sigma_\lambda^2(IT+1)+\sigma_x^2)\sigma_x^2(IT+1)}\right]} \\
& \times\; e^{-\frac{1}{2}\left[\frac{\left(T\sum_{i=1}^{I}\mu_i^2+\mu_1^2\right)\left(\sigma_\lambda^2(IT+1)^2+\sigma_x^2(IT+1)\right)-2\sigma_x^2\mu_\lambda(IT+1)\left(\sum_{i=1}^{I}\sum_{t=1}^{T} lnx_{i,t}\right)}{(\sigma_\lambda^2(IT+1)+\sigma_x^2)\sigma_x^2(IT+1)}\right]} \\
& \times\; e^{-\frac{1}{2}\left[\frac{2\sigma_x^2\mu_\lambda(IT+1)\left(T\sum_{i=1}^{I}\mu_i+\mu_1\right)+\sigma_x^2(IT+1)^2\mu_\lambda^2}{(\sigma_\lambda^2(IT+1)+\sigma_x^2)\sigma_x^2(IT+1)}\right]} \qquad (4.24)
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{C_2}{X_{1,T+1}}\ e^{-\frac{1}{2}\left[\frac{\left((lnx_{1,T+1})^2\right)\left(\sigma_\lambda^2(IT+1)^2+\sigma_x^2(IT+1)-\sigma_\lambda^2(IT+1)\right)}{(\sigma_\lambda^2(IT+1)+\sigma_x^2)\sigma_x^2(IT+1)}\right]} \\
&\times\ e^{-\frac{1}{2}\left[\frac{\left((lnx_{1,T+1})\right)\left(-2\mu_1\sigma_\lambda^2(IT+1)^2-2\mu_1\sigma_x^2(IT+1)-2\sigma_x^2\mu_\lambda(IT+1)-2\sigma_\lambda^2(IT+1)(\sum_{i=1}^{I}\sum_{t=1}^{T}lnx_{i,t})\right)}{(\sigma_\lambda^2(IT+1)+\sigma_x^2)\sigma_x^2(IT+1)}\right]} \\
&\times\ e^{-\frac{1}{2}\left[\frac{\left((lnx_{1,T+1})\right)\left(2\sigma_\lambda^2(IT+1)(T\sum_{i=1}^{I}\mu_i+\mu_1)\right)}{(\sigma_\lambda^2(IT+1)+\sigma_x^2)\sigma_x^2(IT+1)}\right]} \\
&\times\ e^{-\frac{1}{2}\left[\frac{-(\sigma_\lambda^2(IT+1))(\sum_{i=1}^{I}\sum_{t=1}^{T}lnx_{i,t})^2+2(\sigma_\lambda^2(IT+1))(\sum_{i=1}^{I}\sum_{t=1}^{T}lnx_{i,t})(T\sum_{i=1}^{I}\mu_i+\mu_1)}{(\sigma_\lambda^2(IT+1)+\sigma_x^2)\sigma_x^2(IT+1)}\right]} \\
&\times\ e^{-\frac{1}{2}\left[\frac{-(\sigma_\lambda^2(IT+1))\left[T\sum_{i=1}^{I}\mu_i+\mu_1\right]^2}{(\sigma_\lambda^2(IT+1)+\sigma_x^2)\sigma_x^2(IT+1)}\right]} \\
&\times\ e^{-\frac{1}{2}\left[\frac{\left(\sum_{i=1}^{I}\sum_{t=1}^{T}(lnx_{i,t})^2\right)\left(\sigma_\lambda^2(IT+1)^2+\sigma_x^2(IT+1)\right)+\sum_{i=1}^{I}\left(\mu_i\sum_{t=1}^{T}lnx_{i,t}\right)\left(-2\sigma_\lambda^2(IT+1)^2-2\sigma_x^2(IT+1)\right)}{(\sigma_\lambda^2(IT+1)+\sigma_x^2)\sigma_x^2(IT+1)}\right]} \\
&\times\ e^{-\frac{1}{2}\left[\frac{\left(T\sum_{i=1}^{I}\mu_i^2+\mu_1^2\right)\left(\sigma_\lambda^2(IT+1)^2+\sigma_x^2(IT+1)\right)-2\sigma_x^2\mu_\lambda(IT+1)\left(\sum_{i=1}^{I}\sum_{t=1}^{T}lnx_{i,t}\right)}{(\sigma_\lambda^2(IT+1)+\sigma_x^2)\sigma_x^2(IT+1)}\right]} \\
&\times\ e^{-\frac{1}{2}\left[\frac{2\sigma_x^2\mu_\lambda(IT+1)\left(T\sum_{i=1}^{I}\mu_i+\mu_1\right)+\sigma_x^2(IT+1)^2\mu_\lambda^2}{(\sigma_\lambda^2(IT+1)+\sigma_x^2)\sigma_x^2(IT+1)}\right]}. \qquad (4.25)
\end{aligned}
$$

Now, we can rewritten (4.25) in the form

$$
f_{X_{1,T+1}|\vec{X}}(x_{1,T+1}|\vec{x}) = \frac{C_2}{x_{1,T+1}} \times e^{-\frac{1}{2}\left[A(ln\ x_{1,T+1})^2-2B(ln\ x_{1,T+1})+K\right]}, \qquad (4.26)
$$

where

$$
\begin{aligned}
A &= \frac{\sigma_\lambda^2(IT+1)^2+\sigma_x^2(IT+1)-\sigma_\lambda^2(IT+1))}{(\sigma_\lambda^2(IT+1)+\sigma_x^2)\sigma_x^2(IT+1)} \\
&= \frac{\sigma_x^2+\sigma_\lambda^2 IT}{(\sigma_\lambda^2(IT+1)+\sigma_x^2)\sigma_x^2}, \\
B &= \frac{\mu_1\sigma_\lambda^2(IT+1)^2+\mu_1\sigma_x^2(IT+1)+\sigma_x^2\mu_\lambda(IT+1)}{(\sigma_\lambda^2(IT+1)+\sigma_x^2)\sigma_x^2(IT+1)} \\
&\quad + \frac{\sigma_\lambda^2(IT+1)(\sum_{i=1}^{I}\sum_{t=1}^{T}lnx_{i,t})-\sigma_\lambda^2(IT+1)(T\sum_{i=1}^{I}\mu_i+\mu_1)}{(\sigma_\lambda^2(IT+1)+\sigma_x^2)\sigma_x^2(IT+1)} \\
&= \frac{\mu_1\sigma_\lambda^2(IT+1)+\sigma_\lambda^2\left(\sum_{i=1}^{I}\sum_{t=1}^{T}lnx_{i,t}-(T\sum_{i=1}^{I}\mu_i+\mu_1)\right)+\sigma_x^2(\mu_\lambda+\mu_1)}{(\sigma_\lambda^2(IT+1)+\sigma_x^2)\sigma_x^2} \\
&= \frac{\sigma_\lambda^2\left(\sum_{i=1}^{I}\sum_{t=1}^{T}lnx_{i,t}-T\sum_{i=1}^{I}\mu_i+\mu_1 IT\right)+\sigma_x^2(\mu_\lambda+\mu_1)}{(\sigma_\lambda^2(IT+1)+\sigma_x^2)\sigma_x^2}, \quad \text{and}
\end{aligned}
\qquad (4.27)
$$

$$\begin{aligned} K &= \frac{-(\sigma_\lambda^2(IT+1))(\sum_{i=1}^{I}\sum_{t=1}^{T} lnx_{i,t})^2}{(\sigma_\lambda^2(IT+1)+\sigma_x^2)\sigma_x^2(IT+1)} \\ &+ \frac{2(\sigma_\lambda^2(IT+1))(\sum_{i=1}^{I}\sum_{t=1}^{T} lnx_{i,t})(T\sum_{i=1}^{I}\mu_i+\mu_1)}{(\sigma_\lambda^2(IT+1)+\sigma_x^2)\sigma_x^2(IT+1)} \\ &- \frac{(\sigma_\lambda^2(IT+1))\big[T\sum_{i=1}^{I}\mu_i+\mu_1\big]^2}{(\sigma_\lambda^2(IT+1)+\sigma_x^2)\sigma_x^2(IT+1)} \\ &+ \frac{\big(\sum_{i=1}^{I}\sum_{t=1}^{T}(lnx_{i,t})^2\big)\big(\sigma_\lambda^2(IT+1)^2+\sigma_x^2(IT+1)\big)}{(\sigma_\lambda^2(IT+1)+\sigma_x^2)\sigma_x^2(IT+1)} \\ &+ \frac{\big(\sum_{i=1}^{I}(\mu_i\sum_{t=1}^{T} lnx_{i,t})\big)\big(-2\sigma_\lambda^2(IT+1)^2-2\sigma_x^2(IT+1)\big)}{(\sigma_\lambda^2(IT+1)+\sigma_x^2)\sigma_x^2(IT+1)} \\ &+ \frac{\big(T\sum_{i=1}^{I}\mu_i^2+\mu_1^2\big)\big(\sigma_\lambda^2(IT+1)^2+\sigma_x^2(IT+1)\big)}{(\sigma_\lambda^2(IT+1)+\sigma_x^2)\sigma_x^2(IT+1)} \\ &- \frac{2\sigma_x^2\mu_\lambda(IT+1)\big(\sum_{i=1}^{I}\sum_{t=1}^{T} lnx_{i,t}\big)}{(\sigma_\lambda^2(IT+1)+\sigma_x^2)\sigma_x^2(IT+1)} \\ &+ \frac{2\sigma_x^2\mu_\lambda(IT+1)\big(T\sum_{i=1}^{I}\mu_i+\mu_1\big)+\sigma_x^2(IT+1)^2\mu_\lambda^2}{(\sigma_\lambda^2(IT+1)+\sigma_x^2)\sigma_x^2(IT+1)}. \end{aligned} \tag{4.28}$$

Using the square operation

$$e^{-\frac{1}{2}\left[Ax^2-2x+K\right]} = e^{-\frac{1}{2}\left[K-\frac{B^2}{A}\right]} \times e^{-\frac{1}{2}\left[\frac{\left(x-\frac{B}{A}\right)^2}{\frac{1}{A}}\right]},$$

then (4.28) becomes

$$f_{X_{1,T+1}|\vec{X}}(x_{1,T+1}|\vec{x}) = \frac{C_2}{x_{1,T+1}} e^{-\frac{1}{2}\left[\frac{\left(lnx_{1,T+1}-\frac{B}{A}\right)^2}{\frac{1}{A}}\right]} \times e^{-\frac{1}{2}\left[K-\frac{B^2}{A}\right]}. \tag{4.29}$$

We observe that $\frac{1}{x_{1,T+1}} e^{-\frac{1}{2}\left[\frac{\left(lnx_{1,T+1}-\frac{B}{A}\right)^2}{\frac{1}{A}}\right]}$ is the kernel of lognormal distribution. Therefore, it can be concluded that $X_{1,T+1}|\vec{X} \sim LN(\mu_{1,T+1}, \sigma_{1,T+1}^2)$ where $\mu_{1,T+1} = \frac{B}{A}$ and $\sigma_{1,T+1}^2 = \frac{1}{A}$. Thus

$$\begin{aligned} E[X_{1,T+1}|\vec{X}] &= e^{\mu_{1,T+1}+\frac{\sigma_{1,T+1}^2}{2}} \\ &= e^{\left[\frac{\sigma_\lambda^2\left(\sum_{i=1}^{I}\sum_{t=1}^{T} lnx_{i,t}-T\sum_{i=1}^{I}\mu_i+\mu_1 IT\right)+\sigma_x^2(\mu_\lambda+\mu_1)}{\sigma_x^2+\sigma_\lambda^2 IT}\right]} \times e^{\left[\frac{\left(\sigma_\lambda^2(IT+1)+\sigma_x^2\right)\sigma_x^2}{2(\sigma_x^2+\sigma_\lambda^2 IT)}\right]}. \end{aligned} \tag{4.30}$$

This concludes the proof. □

4.2.2 Bayesian Premiums with Normal Claim Amounts

In this subsection, we consider the case in which the risks in a portfolio are homogeneous, i.e., each individual's claim amounts have the same mean and variance, and the claims of each individual in the group policy are normally distributed with mean $(\mu + \lambda)$ and variance σ_x^2. More precisely, we make the assumption

N1. the random variables $X_{j,t}|\lambda$ are normally distributed, i.e.,

$$X_{j,t}|\lambda \sim N(\mu + \lambda, \sigma_x^2) \quad \text{for } j = 1, 2, \dots, I, \quad \text{and } t = 1, 2, \dots, T,$$

where μ is a constant which is used for all individuals and needs to be chosen. We assume the common effect Λ satisfies assumption **L2**. In this case, we can write the Bayesian premium in the more compact form of the credibility formula. That is we have the following theorem.

Theorem 4.4. *Suppose the random variable Λ and the random vector $\vec{X}$ satisfy all assumptions as in Lemma 4.1. Assume further that $X_{j,t}|\lambda$ and common effect Λ satisfy **N1** and **L2**, respectively. Then the Bayesian premium can be written as*

$$\begin{aligned} E[X_{j,T+1}|\vec{X}] &= \frac{\sigma_\lambda^2 IT \frac{\left(\sum_{i=1}^{I}\sum_{t=1}^{T} x_{i,t}\right)}{IT} + \sigma_x^2(\mu_\lambda + \mu)}{\sigma_\lambda^2 IT + \sigma_x^2} \\ &= w_1 \bar{\bar{X}} + w_2(\mu_\lambda + \mu) \end{aligned} \tag{4.31}$$

for $j = 1, 2, \dots, I$, where $w_1 = \frac{\sigma_\lambda^2 IT}{\sigma_\lambda^2 IT + \sigma_x^2}$, $w_2 = 1 - w_1 = \frac{\sigma_x^2}{\sigma_\lambda^2 IT + \sigma_x^2}$, and $\bar{\bar{X}} = \frac{\left(\sum_{i=1}^{I}\sum_{t=1}^{T} x_{i,t}\right)}{IT}$.

Proof. The proof for this theorem is similar that for Theorem (4.3). By just substituting normal density for lognormal density in equation (4.15) and then continuing the proof in the same manner as in Theorem (4.3) one gets the credibility formula (4.31) □

The form of expressions (4.31) is the same for all individuals in a group. The credibility premium to be charged to the group in the next period would thus be $I[w_1\bar{\bar{X}} + w_2(\mu_\lambda + \mu_1)]$ for I members in the next time period.

Asymtotic Properties

There are some interesting asymptotic properties that can be derived from the credibility premium formulae (4.31). These asymptotic properties are summarized below:

P1. Lack of past claims experience.

If we let $T \longrightarrow 0$, that is, past experience is lacking for all individuals, then we find that

$$w_1 = \frac{\sigma_\lambda^2 IT}{\sigma_\lambda^2 IT + \sigma_x^2} \longrightarrow 0$$

and

$$w_2 = \frac{\sigma_x^2}{\sigma_\lambda^2 IT + \sigma_x^2} \longrightarrow 1.$$

The less experience available to the insurer to asses future claims experience, the more weight it will attach to what it believes (that is, the prior) it should be.

P2. Abundant past claims experience.

If we let $T \longrightarrow \infty$, that is, there is abundance of past experience for all individuals, then we find that

$$w_1 = \frac{\sigma_\lambda^2 IT}{\sigma_\lambda^2 IT + \sigma_x^2} \longrightarrow 1$$

and

$$w_2 = \frac{\sigma_x^2}{\sigma_\lambda^2 IT + \sigma_x^2} \longrightarrow 0.$$

This is intuitively appealing as one would expect to attach more weight to individual's own experience.

P3. Abundant group experience.

If we let $I \longrightarrow \infty$, that is, there is abundance of group experience, then we find that

$$w_1 = \frac{\sigma_\lambda^2 IT}{\sigma_\lambda^2 IT + \sigma_x^2} \longrightarrow 1$$

and

$$w_2 = \frac{\sigma_x^2}{\sigma_\lambda^2 IT + \sigma_x^2} \longrightarrow 0.$$

Since each individual's own experience in a group has the same distribution and there are more experiences available, the weight will attach to individual's experience more than what it believes (that is, the prior).

P4. Large variation of individual claims.

We find that as the variability of individual claims increases, that is $\sigma_x^2 \longrightarrow \infty$, then we have

$$w_1 = \frac{\sigma_\lambda^2 IT}{\sigma_\lambda^2 IT + \sigma_x^2} \longrightarrow 0$$

and

$$w_2 = \frac{\sigma_x^2}{\sigma_\lambda^2 IT + \sigma_x^2} \longrightarrow 1.$$

Thus, when variability in recent experience is high and therefore comparatively unreliable for risk assessment, we would attach all the weights to our prior beliefs.

P5. Large variation in overall risk parameter.

If $\sigma_\lambda^2 \longrightarrow \infty$, that is, a large variability in overall risk parameter, then we have

$$w_1 = \frac{\sigma_\lambda^2 IT}{\sigma_\lambda^2 IT + \sigma_x^2} \longrightarrow 1$$

and

$$w_2 = \frac{\sigma_x^2}{\sigma_\lambda^2 IT + \sigma_x^2} \longrightarrow 0.$$

Thus, when variability in risk parameter is high, we would attach all the weights to recent experience.

CHAPTER V

APPLICATION TO MOTOR INSURANCE

The one-level common effect model of claim dependence described in Chapter III can be applied to many types of insurance problems, for example, private household, economic environment or motor insurance. The main purpose of this chapter is to illustrate the prediction of future claim amounts for each individual by applying the Bayesian formula (4.12) to an actual motor insurance positive claim data set of 1,296 observations of class 5 for the year 2009. These data were supplied by a non-life insurance company in Thailand. Since these claims were observed in one year only, then the observed time experience (or period) is $t = 1$. Let $x_{j,1},\ j = 1, 2, \ldots, 1296$ denote these observations and $X_{j,1},\ j = 1, 2, \ldots, 1296$ be the corresponding random variables.

Using the K-S test, these 1,296 observations can be matched to a lognormal distribution, i.e. $X \sim LN(\alpha, \beta^2)$, with a pdf

$$f_X(x) = \frac{1}{\beta\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{lnx-\alpha}{\beta}\right)^2},$$

at a significance level of 0.10 with the estimated parameters,

$$\alpha = 8.9662 \ \text{ and } \ \beta = 1.1804. \tag{5.1}$$

We note that the mean of this lognormal is

$$e^{8.9662+\frac{(1.1804)^2}{2}} = 15,738.60798. \tag{5.2}$$

The historical data of claim severity and claim frequency are illustrated in figure 5.1 and figure 5.2, respectively.

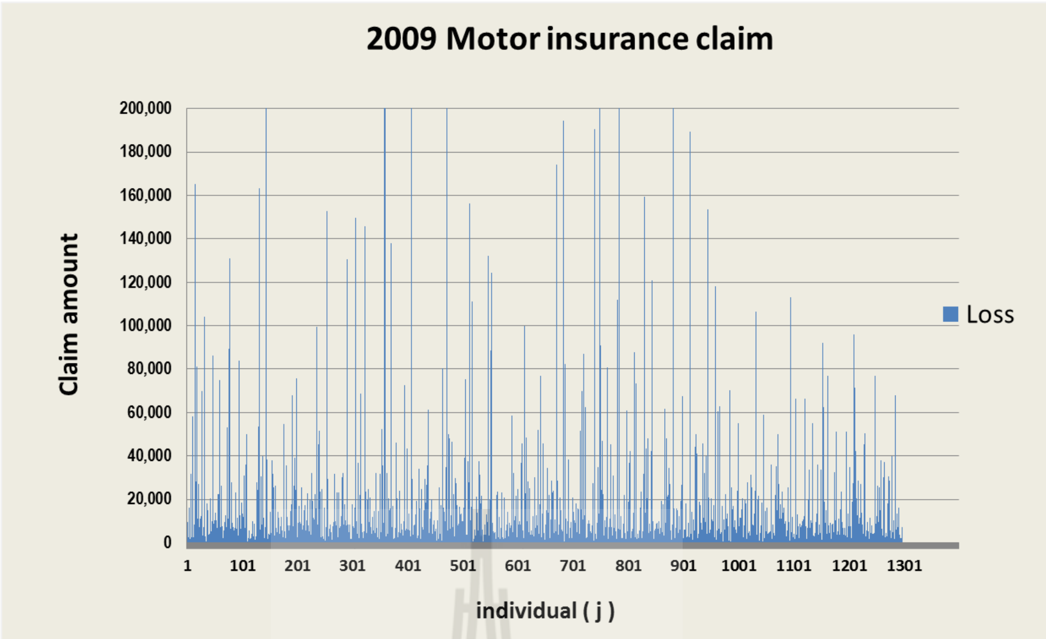

Figure 5.1 Actual claim amounts(Baht) from 1,296 observations($t = 1$).

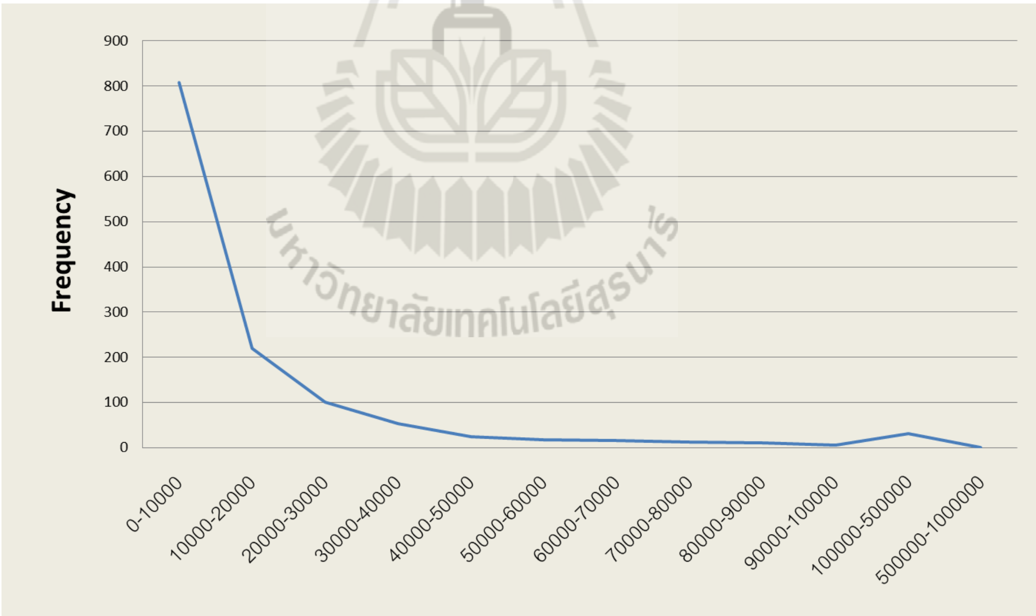

Figure 5.2 Frequency of claim amounts(Baht) from data.

5.1 Model Assumptions and Parameter used in Illustration

5.1.1 Claim amount assumptions

In order to predict future claim amounts, i.e. $E[X_{j,2}|\vec{X}],\ j = 1, 2, \ldots, 1296$, we divide the prediction into two cases.

Case 1. We consider the case in which individual risks $X_{j,1},\ j = 1, 2, \ldots, 1296$ are homogeneous, i.e., each individual's claim amounts has the same parameter $\mu + \lambda$ and σ_x, and each $X_{j,1}$ is lognormally distributed. So we have

$$X_{j,1}|\lambda \sim LN(\mu + \lambda, \sigma_x^2) \quad \text{for } j = 1, 2, \ldots, 1296,$$

where μ and σ_x are constants and use for all individuals and need to be chosen

Case 2. We consider the case in which individual risks $X_{j,1},\ j = 1, 2, \ldots, 1296$ are inhomogeneous, i.e., each individual's claim amounts has distinct parameters $\mu_j + \lambda$ where μ_j is a constant depending on individual j but has the same parameter σ_x. So we have

$$X_{j,1}|\lambda \sim LN(\mu_j + \lambda, \sigma_x^2) \quad \text{for } j = 1, 2, \ldots, 1296.$$

5.1.2 Process for Calculating Bayesian premium

As an illustration, we show how to compute the Bayesian premium $E[X_{j,2}|\vec{X}],\ j = 1, 2, \ldots, 1296$, for each individual j by using equation (4.12) when the individual risks are inhomogeneous. We use model descriptions as in MD2 (see Table 1).

To find the Bayesian premium, we firstly measure the dependence between individual risks, which will be denoted by λ. In order to illustrate how the common effect has an influence on the Bayesian premium, we let, for example,

$$\mu_\lambda = 5, \sigma_\lambda^2 = 100, \text{ and } \lambda = 6.$$

We assume further that

$$\sigma_x = \beta = 1.1804,$$

where β appears in (5.1). Now, only the value of μ_j needs to be chosen in a suitable manner. We propose one method for handling this problem by letting $\mu_j, j = 1, 2, , 1296$ be the convex summand of w_1 and m, i.e

$$\mu_j = w_1 \tilde{\mu_j} + (1 - w_1)m, \quad 0 \leq w_1 \leq 1 \tag{5.3}$$

where $\tilde{\mu_j}$ and m solve the following equations

$$x_{j,1} = e^{(\tilde{\mu_j}+\lambda)+\frac{\sigma_x^2}{2}} \quad \text{and} \quad 15,738.60798 = e^{(m+\lambda)+\frac{\sigma_x^2}{2}}, \tag{5.4}$$

respectively.

Specifically, let $j = 10$ so $x_{10,1} = 2,500$ (the tenth element in the set of actual data) and by putting $\lambda = 6, \sigma_x = 1.1804$ into equation (5.4), one obtains $m = 2.9672$ and $\tilde{\mu_j} = 1.12737$. Hence, (5.3) implies

$$\mu_j = w_1(1.12737) + (1 - w_1)(2.9672).$$

Finally, choose $w_1 = 0.5$, one obtains $\mu_1 = 2.04729$. Then equation (11) yields

$$E[X_{1,2}|\vec{X}] \; = \; 8,891.19699$$

as appears in the second row and sixth column of Table 5.2.

In the homogeneous cases, The Bayesian premiums are computed in a similar manner by setting $\mu_j = \mu = 2.9672$, $j = 1, 2, \ldots, 1296$ and using model descriptions as in MD1 (see Table 5.1). One gets

$$E[X_{j,2}|\vec{X}] \; = \; 15,746.94027, \quad j = 1, 2, \ldots, 1296$$

as appears in the third column of Table 5.2.

Table 5.1 Summary of given information from the actual data, assumptions and parameters used in models.

Specification	Model Descriptions
Case1: Homogeneous Class	
Conditional density	$X_{j,t}\|\lambda \sim LN(\mu+\lambda, \sigma_x^2)$,
	for $j = 1, 2, \ldots, I$ and $t = 1, 2, \ldots, T$
Overall common effect	$\lambda \sim N(\mu_\lambda, \sigma_\lambda^2)$
Given information	$I = 1,296,\ \sum_{i=1}^{1296} lnX_{i,1} = 11,621.48,\ \sigma_x = 1.1804$
Assumption	$\mu + \lambda = 8.9672$
Parameter values (**MD1**)	$\mu_\lambda = 5,\ \sigma_\lambda^2 = 100,\ \lambda = 6$
Case2: Inhomogeneous Class	(various w_1)
Conditional density	$X_{j,t}\|\lambda \sim LN(\mu_j+\lambda, \sigma_x^2)$,
	for $j = 1, 2, \ldots, I$ and $t = 1, 2, \ldots, T$
Overall common effect	$\lambda \sim N(\mu_\lambda, \sigma_\lambda^2)$
Given information	$I = 1,296,\ \sum_{i=1}^{1296} lnX_{i,1} = 11,621.48,\ \sigma_x = 1.1804$
Assumption	$\mu_i = w_1\tilde{\mu}_i + (1-w_1)m,\quad 0 < w_1 < 1$
	where $\tilde{\mu}_i = ln\left(\frac{X_i}{e^{\lambda+\frac{\sigma_x^2}{2}}}\right),\ m = 2.96720$
Parameter values (**MD2.1**)	$\mu_\lambda = 5,\ \sigma_\lambda^2 = 100,\ \lambda = 6,\ w_1 = 0.1$
Parameter values (**MD2.2**)	$\mu_\lambda = 5,\ \sigma_\lambda^2 = 100,\ \lambda = 6,\ w_1 = 0.3$
Parameter values (**MD2.3**)	$\mu_\lambda = 5,\ \sigma_\lambda^2 = 100,\ \lambda = 6,\ w_1 = 0.5$
Parameter values (**MD2.4**)	$\mu_\lambda = 5,\ \sigma_\lambda^2 = 100,\ \lambda = 6,\ w_1 = 0.7$
Parameter values (**MD2.5**)	$\mu_\lambda = 5,\ \sigma_\lambda^2 = 100,\ \lambda = 6,\ w_1 = 0.9$

5.2 Results and Discussion

When using the one-level common effect model with the information in Table 1, the Bayesian premiums corresponding to equation (4.12) are illustrated as follows. Figure 5.3 shows the resulting Bayesian premiums using model MD1 compared with observed claim of each individual in the year 2009. Figure 5.4 to Figure 5.8 show the resulting Bayesian premiums using model MD2.1 - MD2.5, respectively.

Some Bayesian premiums corresponding to observed claim $x_{j,1}$ and the result statistics for prediction are shown in Table 2 and Table 3, respectively.

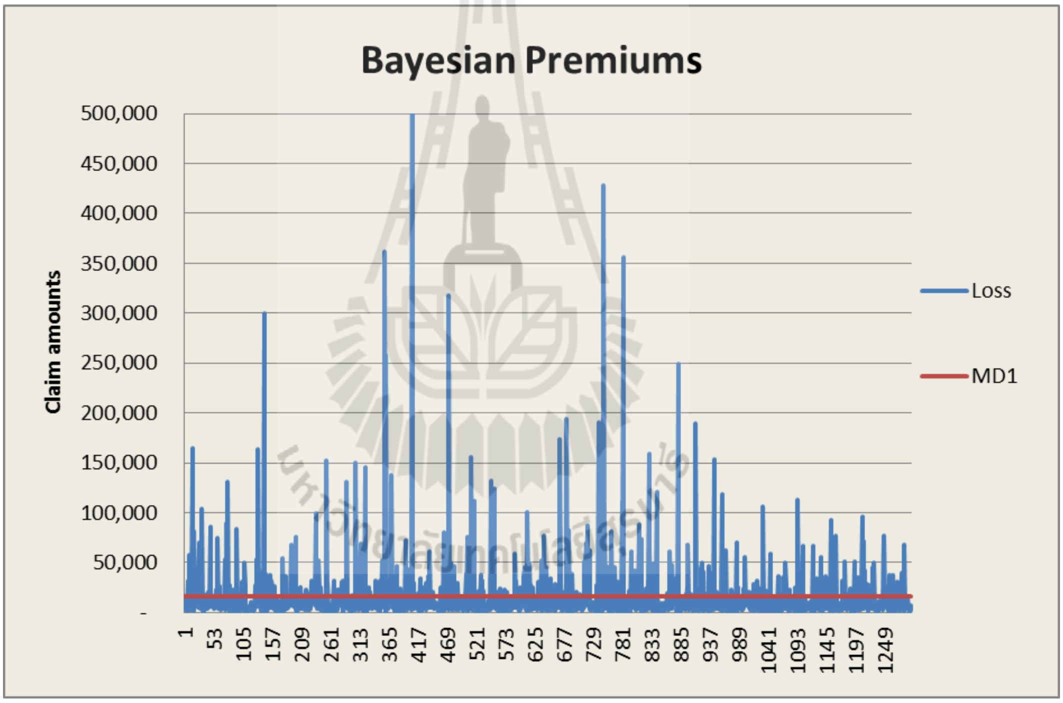

Figure 5.3 Bayesian premiums computed by models MD1.

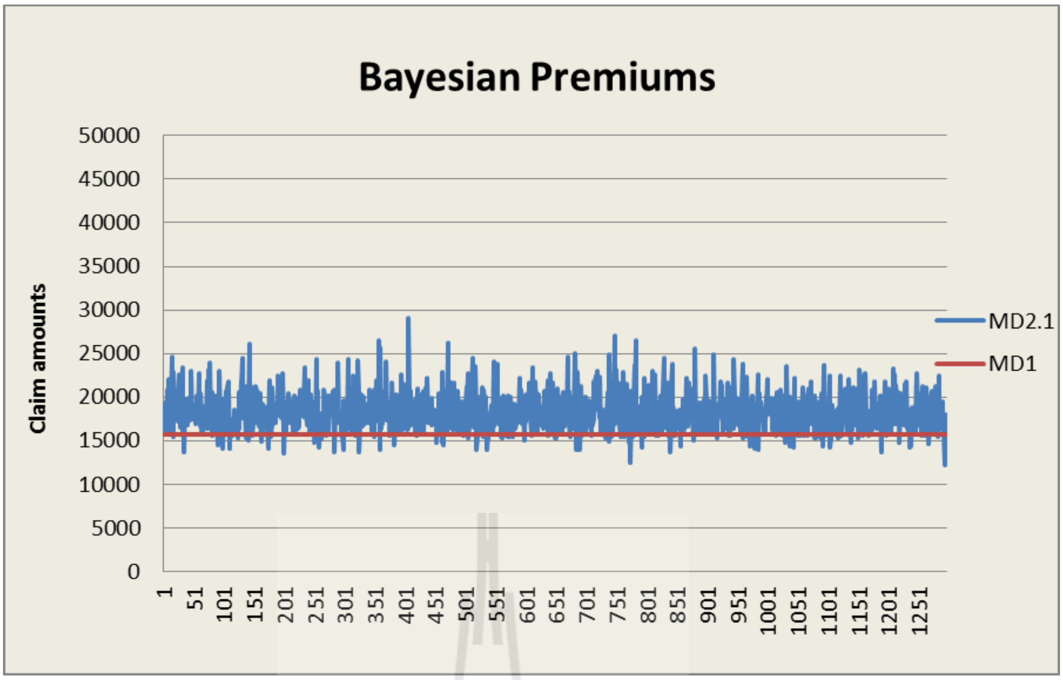

Figure 5.4 Bayesian premiums computed by models MD2.1.

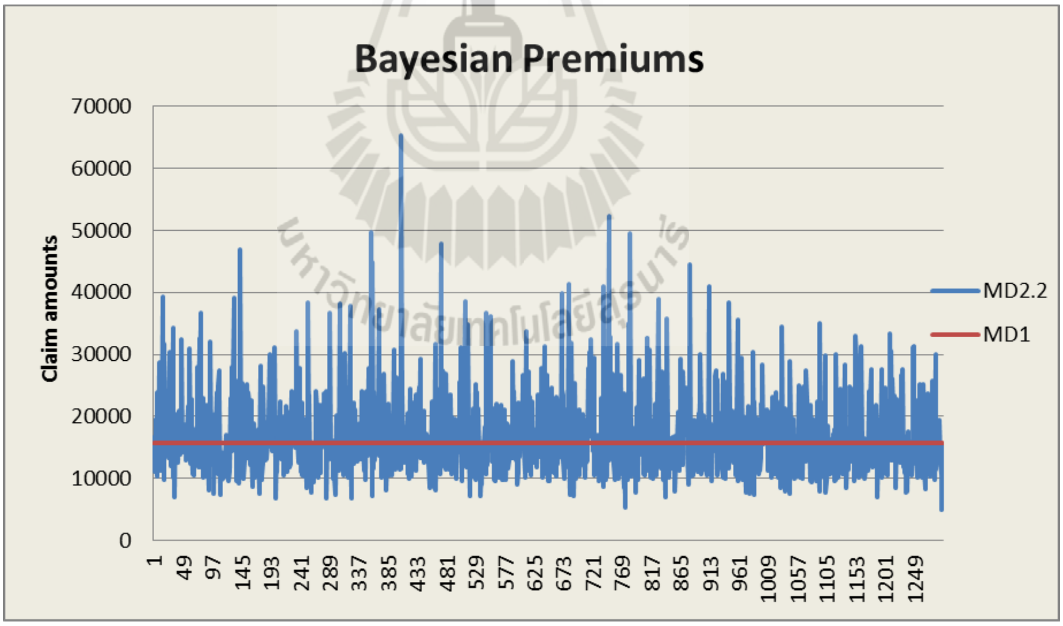

Figure 5.5 Bayesian premiums computed by models MD2.2.

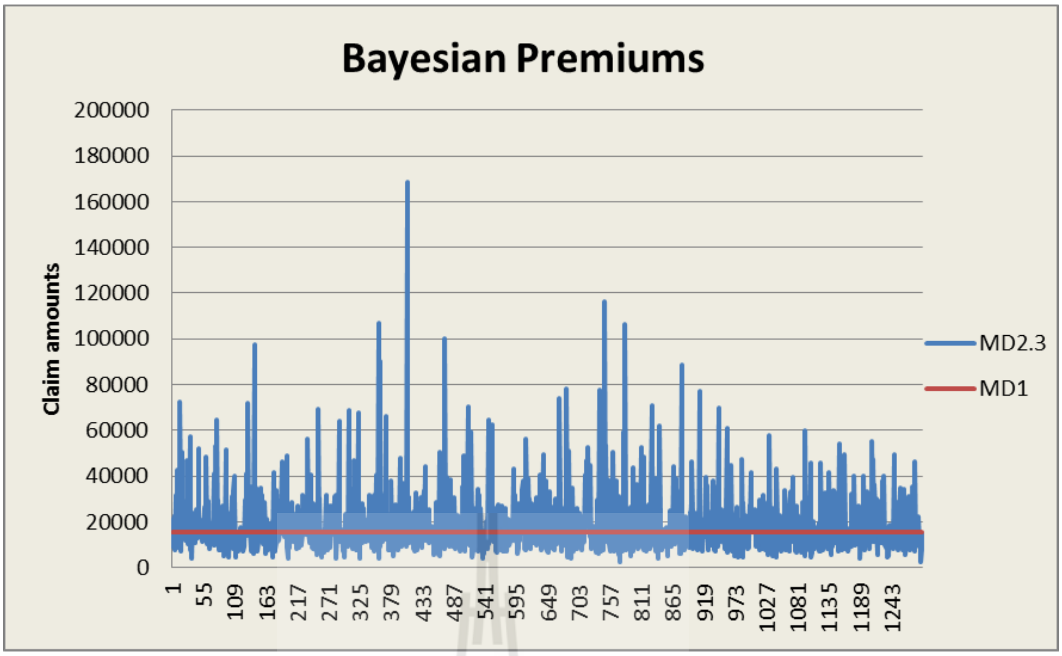

Figure 5.6 Bayesian premiums computed by models MD2.3.

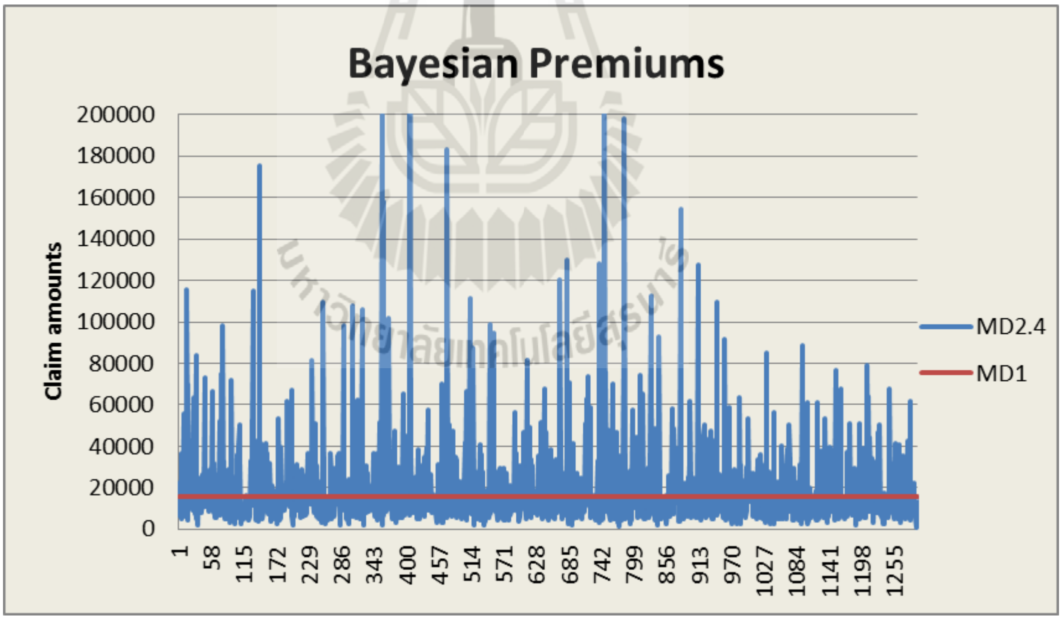

Figure 5.7 Bayesian premiums computed by models MD2.4.

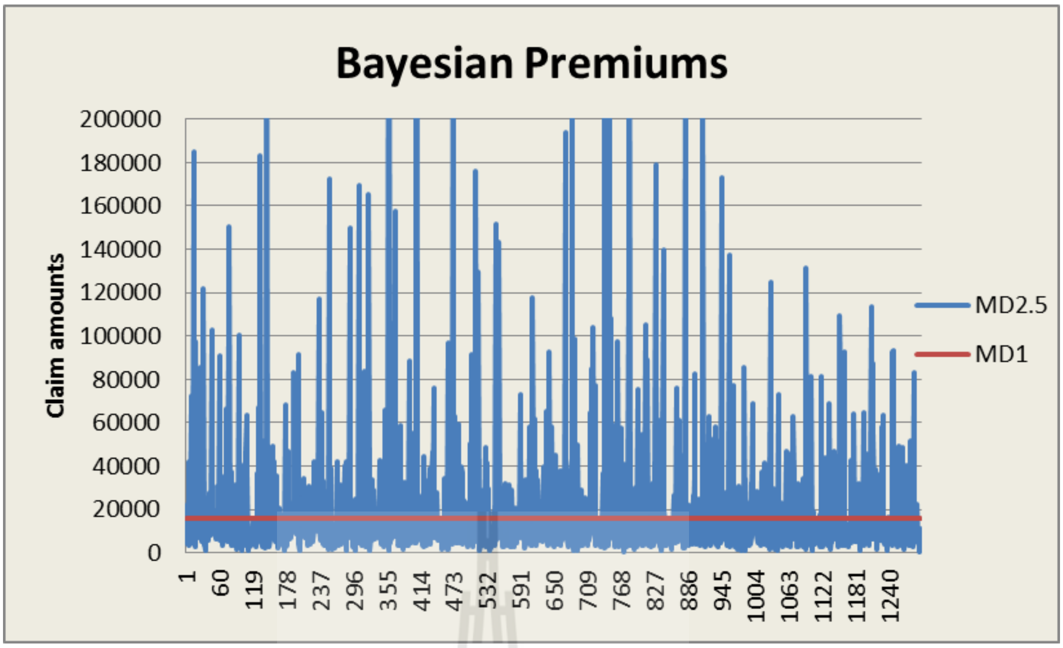

Figure 5.8 Bayesian premiums computed by models MD2.5.

Table 5.2 Some results for Bayesian premiums, according to MD1-MD2.5.

Individual no. (j)	Observed Claim($x_{j,1}$)	MD1	MD2.1	MD2.2	MD2.3	MD2.4	MD2.5
36	500	15,747	11,958	6,895	3,976	2,293	1,322
10	2,500	15,747	14,046	11,175	8,891	7,074	5,628
49	5,500	15,747	15,198	14,157	13,188	12,285	11,443
12	9,500	15,747	16,052	16,680	17,332	18,010	18,715
39	15,130	15,747	16,817	19,179	21,873	24,946	28,450
269	20,957	15,747	17,373	21,148	25,743	31,336	38,144
282	30,323	15,747	18,027	23,627	30,965	40,583	53,189
926	40,987	15,747	18,579	25,862	36,001	50,114	69,761
476	50,029	15,747	18,953	27,456	39,774	57,619	83,470
61	74,779	15,747	19,730	30,974	48,627	76,340	119,848
613	100,000	15,747	20,312	33,796	56,233	93,564	155,678
255	152,800	15,747	21,192	38,380	69,511	125,891	228,002
683	194,405	15,747	21,708	41,255	78,405	149,006	283,184
144	300,000	15,747	22,671	46,990	97,398	201,879	418,444
749	428,012	15,747	23,491	52,276	116,337	258,897	576,154
408	899,879	15,747	25,303	65,332	168,687	435,548	1,124,590

Table 5.3 Some descriptive statistic for Bayesian premiums, according to MD1-MD2.5.

Statistics	Observed Claim($x_{j,1}$)	MD1	MD2.1	MD2.2	MD2.3	MD2.4	MD2.5
Minimum	159	15,747	11,958	6,895	3,976	2,293	1,322
Maximum	899,879	15,747	14,046	11,175	8,891	7,074	5,628
Mean	17,662	15,747	15,198	14,157	13,188	12,285	11,443
Median	7,297	15,747	16,052	16,680	17,332	18,010	18,715
Mode	1,800	15,747	16,817	19,179	21,873	24,946	28,450
Variance ($\times 10^6$)	$\approx$1,708	$\approx$0	$\approx$3.69	$\approx$4.38	$\approx$197	$\approx$781	$\approx$3,252
Aggregate claim($\times 10^6$)	$\approx$22.89	$\approx$20.41	$\approx$20.55	$\approx$21.79	$\approx$24.63	$\approx$29.86	$\approx$39.03

From Table 5.3, we note that the aggregate claim amounts increase if the weight w_1 corresponding to equation (5.3) approaches 1.

Next, we illustrate how the common effect has affects the Bayesian premium by using the same parameters as in model MD1-MD2.5, i.e, set $\mu_\lambda = 5, \sigma_\lambda^2 = 100, w_1 = 0.5$ but vary in λ. The model descriptions are shown in Table 5.4

Table 5.4 Summary of given information from the actual data, assumptions and parameters used in models (with various λ).

Specification	Model Descriptions
Case2: Inhomogeneous Class	
Conditional density	$X_{j,t}\|\lambda \sim LN(\mu_j + \lambda, \sigma_x^2)$,
	for $j = 1, 2, \ldots, I$ and $t = 1, 2, \ldots, T$
Overall common effect	$\lambda \sim N(\mu_\lambda, \sigma_\lambda^2)$
Given information	$I = 1,296,\ \sum_{i=1}^{1296} lnX_{i,1} = 11,621.48,\ \sigma_x = 1.1804$
Assumption	$\mu_i = w_1\tilde{\mu}_i + (1 - w_1)m,\ \ 0 < w_1 < 1$
	where $\tilde{\mu}_i = ln\left(\frac{X_i}{e^{\lambda + \frac{\sigma_x^2}{2}}}\right),\ m = 2.96720$
Parameter values (**MD3.1**)	$\mu_\lambda = 5,\ \sigma_\lambda^2 = 100,\ \ w_1 = 0.5,\ \lambda = 2$
Parameter values (**MD3.2**)	$\mu_\lambda = 5,\ \sigma_\lambda^2 = 100,\ \ w_1 = 0.5,\ \lambda = 4$
Parameter values (**MD3.3**)	$\mu_\lambda = 5,\ \sigma_\lambda^2 = 100,\ \ w_1 = 0.5,\ \lambda = 8$
Parameter values (**MD3.4**)	$\mu_\lambda = 5,\ \sigma_\lambda^2 = 100,\ \ w_1 = 0.5,\ \lambda = 10$

Figure 5.9 to Figure 5.12 show the resulting Bayesian premiums using model MD3.1 - MD3.4, respectively. Some Bayesian premiums corresponding to observed claim $x_{j,1}$ and the results statistics for prediction are shown in Table 5.5 and Table 5.6, respectively.

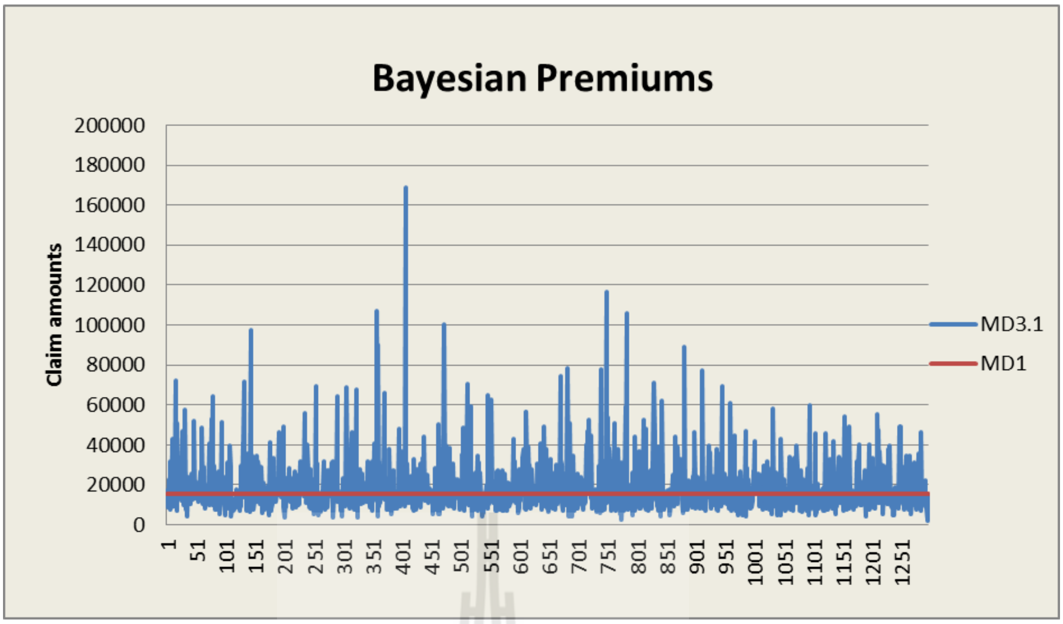

Figure 5.9 Bayesian premiums computed by models MD3.1.

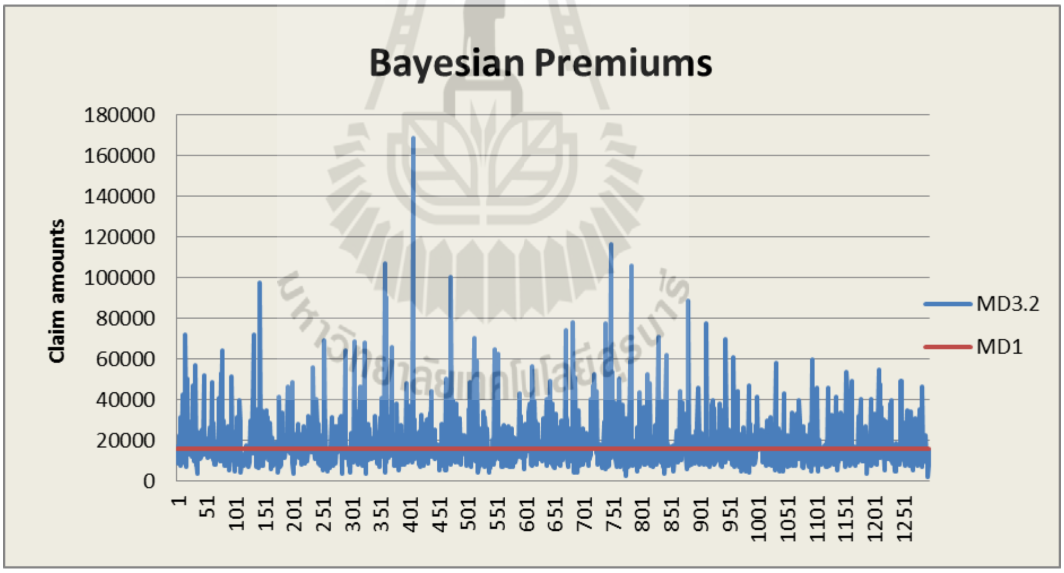

Figure 5.10 Bayesian premiums computed by models MD3.2.

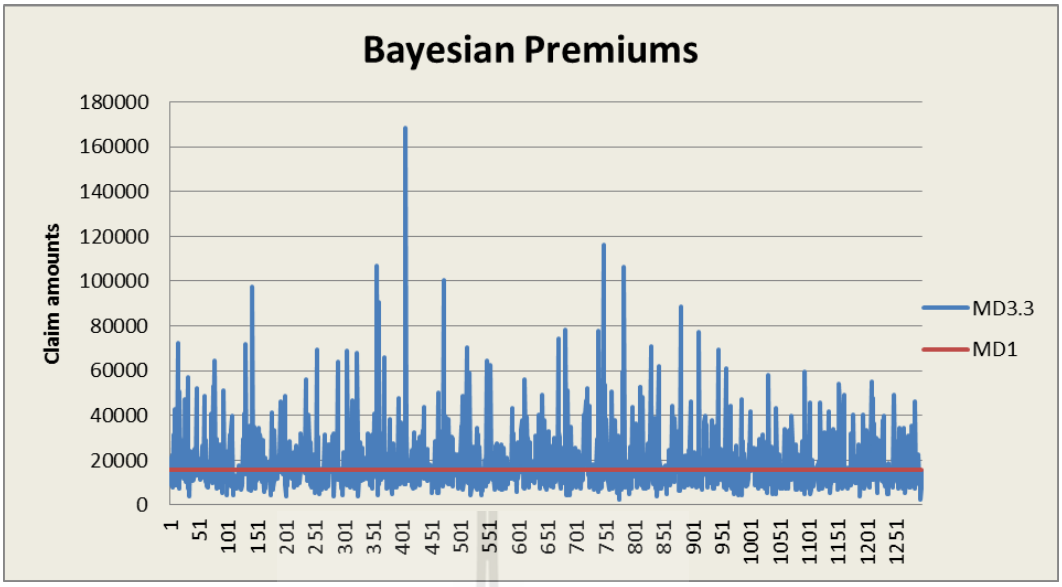

Figure 5.11 Bayesian premiums computed by models MD3.3.

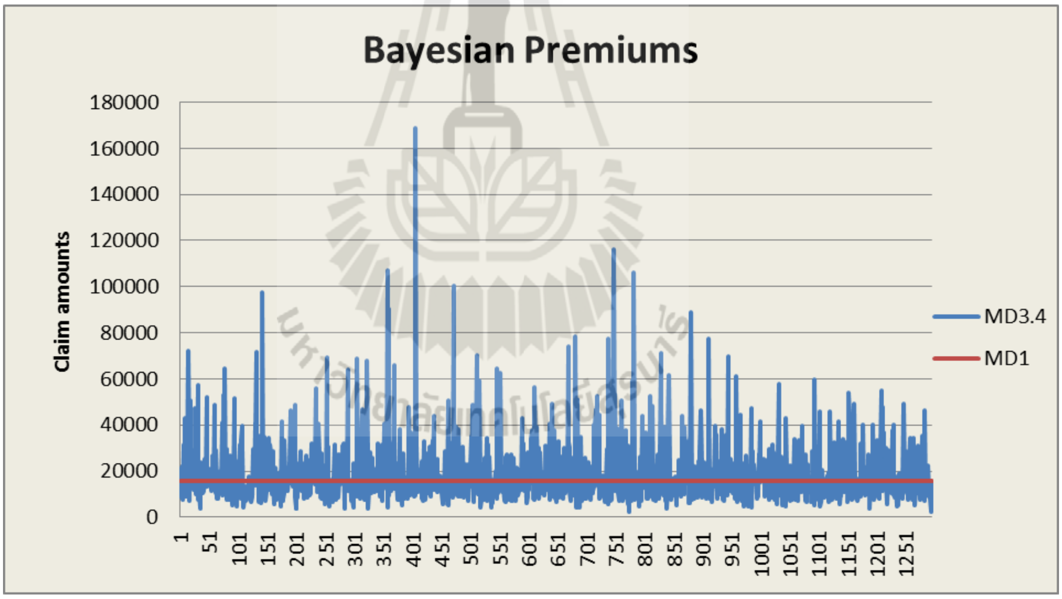

Figure 5.12 Bayesian premiums computed by models MD3.4.

Table 5.5 Some results for Bayesian premiums, according to MD3.1-MD3.4.

Individual no. (j)	Observed Claim($x_{j,1}$)	MD1	MD3.1	MD3.2	MD3.3	MD3.4
36	500	15,747	3,976.42	3,976.34	3,976.19	3,976.10
10	2,500	15,747	8,891.95	8,891.38	8,891.02	8,890.84
49	5,500	15,747	13,188.30	13,188.04	13,187.51	13,187.25
12	9,500	15,747	17,332.82	17,332.48	17,331.79	17,331.44
39	15,130	15,747	21,873.94	21,873.50	21,872.63	21,872.19
269	20,957	15,747	15,743.77	25,743.26	25,742.23	25,741.71
282	30,323	15,747	30,966.61	30,965.99	30,964.75	30,964.13
926	40,987	15,747	36,002.33	36,001.61	36,000.17	35,999.45
476	50,029	15,747	39,775.76	39,774.97	39,773.38	39,772.58
61	74,779	15,747	48,629.23	48,628.26	48,626.32	48,625.34
613	100,000	15,747	56,235.11	56,233.99	56,231.74	56,230.62
255	152,800	15,747	69,513.51	69,512.13	69,509.35	69,507.96
683	194,405	15,747	78,408.16	78,406.60	78,403.47	78,401.90
144	300,000	15,747	97,402.06	97,400.13	97,396.24	97,394.29
749	428,012	15,747	116,341.73	116,339.40	116,334.80	116,332.40
408	899,879	15,747	168,693.98	168,690.63	168,683.89	168,680.51

Table 5.6 Some descriptive statistic for Bayesian premiums, according to MD3.1-MD3.4.

Statistics	Observed Claim($x_{j,1}$)	MD1	MD3.1	MD3.2	MD3.3	MD3.4
Minimum	159	15,747	2,242.36	2,242.32	2,242.23	2,242.18
Maximum	899,879	15,747	168,693.98	168,690.63	168,683.89	168,680.51
Mean	17,662	15,747	19,007.41	19,007.03	19,006.27	19,005.89
Median	7,297	15,747	15,190.24	15,189.94	15,189.33	15,189.03
Mode	1,800	15,747	7,544.73	7,544.58	7,544.28	7,544.13
Variance ($\times 10^6$)	$\approx$1,708	$\approx$0	$\approx$197.425	$\approx$197.418	$\approx$197.402	$\approx$197.394
Aggregate claim($\times 10^6$)	$\approx$22.89	$\approx$20.41	$\approx$24.6336	$\approx$24.6331	$\approx$24.6321	$\approx$24.6316

From Table 5.5, we see that these values of the Bayesian premiums change only a little bit as the value of λ changes. However, note that these parameter values which are used in model ($\lambda, \mu_\lambda = 5, \sigma_\lambda = 100, w_1$), are chosen in order to illustrate the process for calculating Bayesian premiums, they are nothing more than suggestive values which still need to be justified.

In this study, we don't investigate the process of measuring the common effect between risks. To obtain this process is still an interesting problem needing further investigation, the same is true for the problem of justifying the value of w_1. However, one can include other factors from their portfolios (e.g. cause of claim) to justify the appropriacy of this value.

CHAPTER VI

CONCLUSIONS

This thesis is devoted to the study of claim dependence models. In this study, we propose a model for claim dependence between risks (claim dependence across insured individuals) by using common effect and also investigate the Bayesian premium for each individual according to the model. The results obtained are separated into two parts.

In the first part, we study the claim dependence model with the following assumptions:

A1. The common effect random variable Λ has known probability density function $f_\Lambda(\lambda)$ provided that $f_\Lambda(\lambda) > 0$ for all λ.

A2. For a fixed $i = 1, 2, \ldots, I,$ the random variables $X_{i,t}, t = 1, 2, \ldots, T$ are mutually independent and identically distributed.

A3. The random vectors $\vec{X}_i|\Lambda = \lambda,\ i = 1, 2, \ldots, I$ where $\vec{X}_i = (X_{i,1}, X_{i,2}, \ldots, X_{i,T})'$ are conditionally independent.

A4. For a fixed $i = 1, 2, \ldots, I$ and a fixed $t = 1, 2, \ldots, T,$ the conditional random variable $X_{i,t}$ given that $\Lambda = \lambda$ has known probability function denoted by

$$f_{X_{i,t}|\Lambda}(x_{i,t}|\lambda) =: \frac{f_{X_{i,t},\Lambda}(x_{i,t}, \lambda)}{f_\Lambda(\lambda)}.$$

We consider three classes of premiums:

$$\begin{aligned} \triangle_0 &:= \mathbb{R} \\ \triangle_1 &:= L^2(\sigma(\Lambda)) \\ \triangle_I &:= L^2(\sigma(X_{1,1}, X_{1,2}, \ldots, X_{1,T}, \ldots, X_{I,1}, \ldots, X_{I,T}) \end{aligned} \tag{6.1}$$

and study three premiums: $p_0,\ p_1$ and p_2 satisfying

$$\begin{aligned} E[(X_{i,T+1}-p_0)^2] &:= \inf_{\triangle_0} E[(X_{i,T+1}-p)^2] \\ E[(X_{i,T+1}-p_1)^2] &:= \inf_{\triangle_1} E[(X_{i,T+1}-p)^2] \\ E[(X_{i,T+1}-p_2)^2] &:= \inf_{\triangle_I} E[(X_{i,T+1}-p)^2]. \end{aligned} \tag{6.2}$$

Based on these assumptions, we define

$$\begin{aligned} \mu &:= E[X_{i,T+1}] = E[E[X_{i,T+1}|\Lambda]], \\ v &:= E[var(X_{i,T+1}|\Lambda)], \\ a &:= var(E[X_{i,T+1}|\Lambda]), \end{aligned} \tag{6.3}$$

The results are summarized as follows:

Proposition 6.1. *For the optimum premiums* p_0, p_1, p_2 *and a fixed individual* $i = 1, 2, \ldots, I$ *we have:*

$(a)\quad p_0 = \mu$

$(b)\quad p_1 = E[X_{i,T+1}|\Lambda]$

$(c)\quad p_2 = E[X_{i,T+1}|\vec{X}] = E[E[X_{i,T+1}|\Lambda)|\vec{X}]].$

In particular, $p_2 = E[p_1|\vec{X}]$.

Proposition 6.2. *For the loss attached to these premiums and a fixed individual* $i = 1, 2, \ldots, I$, *we have:*

$(a)\quad E[(X_{i,T+1}-p_0)^2] = v + a$

$(b)\quad E[(X_{i,T+1}-p_1)^2] = v$

$(c)\quad E[(X_{i,T+1}-p_2)^2] = v + E[var(E[X_{i,T+1}|\Lambda]|\vec{X})]$

In particular, $E[(X_{i,T+1}-p_1)^2] \le E[(X_{i,T+1}-p_2)^2] \le E[(X_{i,T+1}-p_0)^2]$.

Further, this thesis is mainly interested in the Bayesian premium which can be conveniently expressed as

$$E[X_{j,T+1}|\vec{X}] = \int x_{j,T+1} \cdot f_{X_{j,T+1}|\vec{X}}(x_{j,T+1}|\vec{x})dx_{j,T+1}. \tag{6.4}$$

To obtain an explicit form of the premium for arbitrary claim amounts distribution and common effect distribution, we need the following theorems which are summarized as follows:

Lemma 6.3. *Let* Λ *be a random variable satisfying the assumptions* ***A1*** *to* ***A4*** *and* $\vec{X}$ *be the vector of all observable claims which is defined in (4.1). The joint density of* $\vec{X}$ *and the overall risk parameter* Λ *can be expressed as*

$$f_{\vec{X},\Lambda}(\vec{x},\lambda) \;=\; \prod_{i=1}^{I} f_{\vec{X}_i|\Lambda}(\vec{x}_i|\lambda) \times f_{\Lambda}(\lambda). \tag{6.5}$$

Theorem 6.4. *Suppose the random variable* Λ *and the random vector* $\vec{X}$ *satisfy all assumptions as in Lemma 6.3. The conditional density of* $X_{j,T+1}|\vec{X}$ *can be expressed as*

$$f_{X_{j,T+1}|\vec{X}}(x_{j,T+1}|\vec{x}) \;=\; \int f_{X_{j,T+1}|\Lambda}(x_{j,T+1}|\lambda) \times f_{\Lambda|\vec{X}}(\lambda|\vec{x}) \; d\lambda. \tag{6.6}$$

In the second part, we use equation (6.4) to find the Bayesian premium by applying Theorem 6.4 when the common effect is normally distributed and the claim amounts follow lognormal or normal distribution. We divide our investigation into two cases.

Case 1. Bayesian premiums with lognormal claim amounts.

We make the following assumptions:

L1. the random variables $X_{j,t}|\lambda$ are lognormally distributed, i.e.,

$$X_{j,t}|\lambda \sim LN(\mu_j+\lambda,\sigma_x^2) \quad \text{for } j=1,2,\ldots,I, \quad \text{and } t=1,2,\ldots,T,$$

where μ_j is a constant depending on individual j.

L2. the *overall* common effect λ is normally distributed with mean μ_λ and variance σ_λ^2.

A useful application of theorem 6.4 appears in the following theorem.

Theorem 6.5. *Suppose the random variable* Λ *and the random vector* $\vec{X}$ *satisfy all assumptions as in lemma 6.3. Assume further that* $X_{j,t}|\lambda$ *and common effect* Λ *satisfy*

L1 *and* ***L2****, respectively. Then the Bayesian premium can be written as*

$$E[X_{j,T+1}|\vec{X}] \;=\; e^{\left[\frac{\sigma_\lambda^2\left(\sum_{i=1}^{I}\sum_{t=1}^{T} lnX_{i,t} - T\sum_{i=1}^{I}\mu_i + \mu_j IT\right) + \sigma_x^2\left(\mu_\lambda + \mu_j\right)}{\sigma_\lambda^2 IT + \sigma_x^2}\right]} e^{\left[\frac{\left(\sigma_\lambda^2(IT+1)+\sigma_x^2\right)\sigma_x^2}{2(\sigma_\lambda^2 IT + \sigma_x^2)}\right]}. \quad (6.7)$$

for $j = 1, 2, \ldots, I$.

Case 2. Bayesian premiums with normal claim amounts

We make the following assumptions:

N1. the random variables $X_{j,t}|\lambda$ are normally distributed, i.e.,

$$X_{j,t}|\lambda \sim N(\mu + \lambda, \sigma_x^2) \quad \text{for } j = 1, 2, \ldots, I, \quad \text{and } t = 1, 2, \ldots, T.$$

where μ is a constant which is used for all individuals and need to be chosen. We assume a common effect Λ satisfies assumption **L2**.

Theorem 6.6. *Suppose the random variable* Λ *and the random vector* $\vec{X}$ *satisfy all assumptions as in lemma 6.3. Assume further that* $X_{j,t}|\lambda$ *and common effect* Λ *satisfy* ***N1*** *and* ***L2****, respectively. Then the Bayesian premium can be written as*

$$\begin{aligned} E[X_{j,T+1}|\vec{X}] &= \frac{\sigma_\lambda^2 IT\frac{\left(\sum_{i=1}^{I}\sum_{t=1}^{T} x_{i,t}\right)}{IT} + \sigma_x^2(\mu_\lambda + \mu)}{\sigma_\lambda^2 IT + \sigma_x^2} \\ &= w_1\bar{\bar{X}} + w_2(\mu_\lambda + \mu) \end{aligned} \quad (6.8)$$

for $j = 1, 2, \ldots, I$, *where* $w_1 = \frac{\sigma_\lambda^2 IT}{\sigma_\lambda^2 IT + \sigma_x^2}$, $w_2 = 1 - w_1 = \frac{\sigma_x^2}{\sigma_\lambda^2 IT + \sigma_x^2}$, *and* $\bar{\bar{X}} = \frac{\left(\sum_{i=1}^{I}\sum_{t=1}^{T} x_{i,t}\right)}{IT}$.

An application part of lognormal claim amounts (chapter VI) shows the process for calculating the Bayesian premiums. The interpretation of a common effect in the claim dependence model is illustrated in the context of automobile insurance by using an actual claims data set.

Finally, we should observe that further problems can be considered. Firstly, for the Bayesian premium with lognormal claims (see, equation (5.3)) one needs to choose the weight attached to each individual. To obtain a suitable weight is still an interesting problem needing further investigation. However, one can include other factors from their portfolios to justify the appropriacy of this value. Secondly, the modeling claim dependence using common effect in the proposed model requires distribution formulas for both risks and common effect which could lead to a cumbersome process for obtaining the required premiums. One can conduct further investigations by omitting the form for distributions and using other methodologies such as the means of the projection method involving significant constraints (analogous to Wen et al., 2009).

REFERENCES

REFERENCES

Klugman, S. A., Panjer, H., and Willmot, G. E. (2008). **From Data to Decisions.** (3rd ed.). New Jersey: Wiley.

Roussas, G. G. (1997). **A Course in Mathematical Statistics.** (2nd ed.). New York: Academic Press.

Denuit, M., Dhaene, J., Goovaerts, M., and Kaas, R. (2005). **Actuarial Theory for Dependent Risks Measures, Orders and Models.** Chichester, West Sussex: John Wiley & Sons..

Apostol, T. M. (1974). **Mathematical Analysis**. (2nd ed.). Reading, MA: Addison-Wesley.

Capiński, M., and Kopp, E. (2004). **Measure, Integral and Probability**. London: Springer-Verlag.

Aggoun, L., and Elliod, R. J. (2004). **Measure Theory and Filtering: Introduction and Application.** New York: Cambridge University Press.

Kreyszig, E. (1978). **Introductory Functional Analysis with Applications**. USA: John Wiley & Sons.

Bühlmann, H. (1967). Experience rating and credibility. **ASTIN Bulletin**. 4(1): 199-207.

Bühlmann, H., and Straub, E. (1970). Credibility for loss ratio. **Bulletin of Swiss Ass. of Act.**. 70(1): 111-133.

Finan, M. B. (2013). **A Probability Course for the Actuaries, A Preparation for Exam P/1.** (Preliminary Draft).

Gerber, H., and Jones, D. (1975). Credibility formulas of the updating type. **Transactions of the Society of Actuaries**. 27(1): 31-52.

Heilmann, W.-R. (1986). On the impact of independence or risks on stop-loss transform. **Insurance: Mathematics and Economics**. 5(1): 197-199.

Heilmann, W.-R. (1989). Decision theoretic foundations of credibility. **Insurance: Mathematics and Economics**. 8(1): 77-95.

Schmidt, K. D. (1991). Convergence of Bayes and Credibility premiums. **Astin Bulletin**. 20(2): 167-172.

Hürlimann, W. (1993). Bivariate distributions with diatomic conditional and stop-loss transform of random sum. **Statistics and Probability Letters**. 17(1): 329-335.

Wang, S. (1998). Aggregation of correlated risk portfolios. **Proceedings of the Casualty Actuarial Society**. 85(1): 848-939.

Purcaru, O., and Denuit, M. (2002). On the dependence induced by frequency credibility models. **Belgian Actuarial Bulletin**. 2(1): 73-79.

Mashayekhi, M. (2002). On asymptotic optimality in empirical Bayes credibility. **Insurance: Mathematics and Economics**. 31(1): 285-295.

Valdez, E. A. (2004). Some less-known but useful results for normal distribution. **Working paper**, UNSW, Sydney, Australia.

Yeo, K. L., and Valdez, E. A. (2006). Claim dependence with common effects in credibility models. **Insurance: Mathematics and Economics**. 38(1): 609-629.

Wen, L., Xianyi, W., and Xian, Z. (2009). The credibility premiums for models with dependence induced by common effects. **Insurance: Mathematics and Economics**. 44(1): 19-25.

Weizhong, H., and Xianyi, W. (2012). The credibility premiums with common effects obtained under balanced loss functions. **Chinese Journal of Applied Probability and Statistics**. 28(2): 203-216.

Leaven, R. J. A., and Goovaerts, M. J. (2011). **Premium Calculation and Insurance Pricing.** (Unpublished manuscript).

Neumann, V., and John& Oscar, M. (1997). **Theory of Games and Economic Behavior.** (2nd ed.). Princeton: Princeton University Press.

Savage, L. J. (1954). **The Foundation of Statistics.** New York: Wiley.

Goovaerts, M. J., De Vylder F. C. E., and Haezendonck, J. (1984). **Insurance Premiums.** Amsterdam: North Holland Publishing.

Vyncke, D. (2004). **Comotonicity.** (Unpublished manuscript).

Klugmann, S. A. (1992). **Bayesian Statistics in Actuarial Science with Emphasis on Credibility.** Boston/Dordrecht/London: Kluwer Academic Publishing.

APPENDICES

APPENDIX A

THE PROOF OF SOME PARTS IN THEOREM 4.3

Appendix A1

We show that equation (4.17) holds,i.e.,

$$
\begin{aligned}
&\int \frac{1}{2\pi}\, e^{-\frac{1}{2}\left[\sum_{t=1}^{T+1}\left(\frac{lnx_{1,t}-(\mu_1+\lambda)}{\sigma_x}\right)^2+\sum_{i=2}^{I}\sum_{t=1}^{T}\left(\frac{lnx_{i,t}-(\mu_i+\lambda)}{\sigma_x}\right)^2\right]} \times\ e^{-\frac{1}{2}\left(\frac{\lambda-\mu_\lambda}{\sigma_\lambda}\right)^2}\, d\lambda \\
= &\int \frac{1}{2\pi}\, e^{-\frac{IT+1}{2\sigma_x^2}\left[\left(\lambda-\frac{\left(\sum_{i=1}^{I}\sum_{t=1}^{T} lnx_{i,t}+lnx_{1,T+1}\right)-\left(T\sum_{i=1}^{I}\mu_i+\mu_1\right)}{IT+1}\right)^2\right]}\, e^{-\frac{1}{2}\left(\frac{\lambda-\mu_\lambda}{\sigma_\lambda}\right)^2} \\
&\times\ e^{-\frac{1}{2\sigma_x^2}\left[\frac{-\left(\left[\sum_{i=1}^{I}\sum_{t=1}^{T} lnx_{i,t}+lnx_{1,T+1}\right]-\left[T\sum_{i=1}^{I}\mu_i+\mu_1\right]\right)^2}{IT+1}\right]} \\
&\times\ e^{-\frac{1}{2\sigma_x^2}\left[\left(\sum_{i=1}^{I}\sum_{t=1}^{T}(lnx_{i,t})^2+(lnx_{1,T+1})^2\right)-2\left(\sum_{i=1}^{I}(\mu_i\sum_{t=1}^{T} lnx_{i,t})+(lnx_{1,T+1}\cdot\mu_1)\right)\right]} \\
&\times\ e^{-\frac{1}{2\sigma_x^2}\left[\left(T\sum_{i=1}^{I}\mu_i^2+\mu_1^2\right)\right]}\, d\lambda.
\end{aligned}
$$

It suffices to show that

$$
\begin{aligned}
&e^{-\frac{1}{2}\left[\sum_{t=1}^{T+1}\left(\frac{lnx_{1,t}-(\mu_1+\lambda)}{\sigma_x}\right)^2+\sum_{i=2}^{I}\sum_{t=1}^{T}\left(\frac{lnx_{i,t}-(\mu_i+\lambda)}{\sigma_x}\right)^2\right]} \\
= \ &e^{-\frac{IT+1}{2\sigma_x^2}\left[\left(\lambda-\frac{\left(\sum_{i=1}^{I}\sum_{t=1}^{T} lnx_{i,t}+lnx_{1,T+1}\right)-\left(T\sum_{i=1}^{I}\mu_i+\mu_1\right)}{IT+1}\right)^2\right]} \\
&\times\ e^{-\frac{1}{2\sigma_x^2}\left[\frac{-\left(\left[\sum_{i=1}^{I}\sum_{t=1}^{T} lnx_{i,t}+lnx_{1,T+1}\right]-\left[T\sum_{i=1}^{I}\mu_i+\mu_1\right]\right)^2}{IT+1}\right]} \\
&\times\ e^{-\frac{1}{2\sigma_x^2}\left[\left(\sum_{i=1}^{I}\sum_{t=1}^{T}(lnx_{i,t})^2+(lnx_{1,T+1})^2\right)-2\left(\sum_{i=1}^{I}(\mu_i\sum_{t=1}^{T} lnx_{i,t})+(lnx_{1,T+1}\cdot\mu_1)\right)\right]} \\
&\times\ e^{-\frac{1}{2\sigma_x^2}\left[\left(T\sum_{i=1}^{I}\mu_i^2+\mu_1^2\right)\right]}.
\end{aligned}
$$

Proof. The proof is straightforward by applying the square operation to grouping the term of λ in the form of normally distributed.

$$e^{-\frac{1}{2}\left[\sum_{t=1}^{T+1}\left(\frac{lnx_{1,t}-(\mu_1+\lambda)}{\sigma_x}\right)^2+\sum_{i=2}^{I}\sum_{t=1}^{T}\left(\frac{lnx_{i,t}-(\mu_i+\lambda)}{\sigma_x}\right)^2\right]}$$

$$= e^{-\frac{1}{2}\left[\sum_{i=1}^{I}\sum_{t=1}^{T}\left(\frac{lnx_{i,t}-(\lambda+\mu_i)}{\sigma_x}\right)^2+\left(\frac{lnx_{1,T+1}-(\lambda+\mu_1)}{\sigma_x}\right)^2\right]}$$

$$= e^{-\frac{1}{2\sigma_x^2}\left[\sum_{i=1}^{I}\sum_{t=1}^{T}\left((lnx_{i,t})^2-2lnx_{i,t}\lambda-2lnx_{i,t}\mu_i+\lambda^2+2\lambda\mu_i+\mu_i^2\right)\right]}$$

$$\times e^{-\frac{1}{2\sigma_x^2}\left[\left((lnx_{1,T+1})^2-2lnx_{1,T+1}\lambda-2lnx_{1,T+1}\mu_1+\lambda^2+2\lambda\mu_1+\mu_1^2\right)\right]}$$

$$= e^{-\frac{1}{2\sigma_x^2}\left[\sum_{i=1}^{I}\sum_{t=1}^{T}(lnx_{i,t})^2-2\lambda\sum_{i=1}^{I}\sum_{t=1}^{T}lnx_{i,t}-2\sum_{i=1}^{I}\sum_{t=1}^{T}lnx_{i,t}\mu_i+IT\lambda^2+2\lambda T\sum_{i=1}^{I}\mu_i\right]}$$

$$\times e^{-\frac{1}{2\sigma_x^2}\left[T\sum_{i=1}^{I}\mu_i^2+(lnx_{1,T+1})^2-2lnx_{1,T+1}\lambda-2lnx_{1,T+1}\mu_1+\lambda^2+2\lambda\mu_1+\mu_1^2\right]}$$

$$= e^{-\frac{1}{2\sigma_x^2}\left[\lambda^2(IT+1)-2\lambda(\sum_{i=1}^{I}\sum_{t=1}^{T}lnx_{i,t}+lnx_{1,T+1})+2\lambda(T\sum_{i=1}^{I}\mu_i+\mu_1)\right]}$$

$$\times\ e^{-\frac{1}{2\sigma_x^2}\left[\left(\sum_{i=1}^{I}\sum_{t=1}^{T}(lnx_{i,t})^2+(lnx_{1,T+1})^2\right)-2\left(\sum_{i=1}^{I}\sum_{t=1}^{T}(lnx_{i,t}\cdot\mu_i)+(lnx_{1,T+1}\cdot\mu_1)\right)\right]}$$

$$\times\ e^{-\frac{1}{2\sigma_x^2}\left[T\sum_{i=1}^{I}\mu_i^2+\mu_1^2\right]}$$

$$= e^{-\frac{IT+1}{2\sigma_x^2}\left[\lambda^2-2\lambda\frac{\left((\sum_{i=1}^{I}\sum_{t=1}^{T}lnx_{i,t}+lnx_{1,T+1})-(T\sum_{i=1}^{I}\mu_i+\mu_1)\right)}{IT+1}\right]}$$

$$\times\ e^{-\frac{1}{2\sigma_x^2}\left[\left(\sum_{i=1}^{I}\sum_{t=1}^{T}(lnx_{i,t})^2+(lnx_{1,T+1})^2\right)-2\left(\sum_{i=1}^{I}\sum_{t=1}^{T}(lnx_{i,t}\cdot\mu_i)+(lnx_{1,T+1}\cdot\mu_1)\right)\right]}$$

$$\times\ e^{-\frac{1}{2\sigma_x^2}\left[T\sum_{i=1}^{I}\mu_i^2+\mu_1^2\right]}$$

$$= e^{-\frac{IT+1}{2\sigma_x^2}\left[\left(\lambda-\frac{\left(\sum_{i=1}^{I}\sum_{t=1}^{T}lnx_{i,t}+lnx_{1,T+1}\right)-\left(T\sum_{i=1}^{I}\mu_i+\mu_1\right)}{IT+1}\right)^2\right]}$$

$$\times\ e^{-\frac{1}{2\sigma_x^2}\left[\frac{-\left(\left[\sum_{i=1}^{I}\sum_{t=1}^{T}lnx_{i,t}+lnx_{1,T+1}\right]-\left[T\sum_{i=1}^{I}\mu_i+\mu_1\right]\right)^2}{IT+1}\right]}$$

$$\times\ e^{-\frac{1}{2\sigma_x^2}\left[\left(\sum_{i=1}^{I}\sum_{t=1}^{T}(lnx_{i,t})^2+(lnx_{1,T+1})^2\right)-2\left(\sum_{i=1}^{I}(\mu_i\sum_{t=1}^{T}lnx_{i,t})+(lnx_{1,T+1}\cdot\mu_1)\right)\right]}$$

$$\times\ e^{-\frac{1}{2\sigma_x^2}\left[T\sum_{i=1}^{I}\mu_i^2+\mu_1^2\right]}.$$

This concludes the proof. □

Appendix A2

We show that equation (4.20) is true. First, we recall the proposition of Valdez (2004):

Proposition 1.1. *Suppose $Z \sim N(0,1)$, the standard normal random variable. Then the following holds for any constant a and b:*

$$E[\varphi(a-bz)] \;=\; \frac{1}{\sqrt{1+b^2}} \cdot \varphi(\frac{a}{\sqrt{1+b^2}}). \tag{A.1}$$

Proof. The proof is straightforward integration problem. First, notice that we can re-write

$$\begin{aligned}
\varphi(z)\varphi(a-bz) &= \frac{1}{\sqrt{2\pi}} \cdot \frac{1}{\sqrt{2\pi}} \cdot e^{\left(-\frac{1}{2}[z^2+(a-bz)^2]\right)} \\
&= \frac{1}{\sqrt{2\pi}} \cdot \frac{1}{\sqrt{2\pi}} \cdot e^{\left(-\frac{1}{2}[z^2(1+b)^2-2abz+a^2]\right)} \\
&= \frac{1}{\sqrt{2\pi}} e^{\left(-\frac{1}{2}[a^2-\frac{a^2b^2}{1+b^2}]\right)} \cdot \frac{1}{\sqrt{2\pi}} e^{\left(-\frac{1}{2}\left[\frac{z-\frac{ab}{1+b^2}}{\frac{1}{\sqrt{1+b^2}}}\right]^2\right)}.
\end{aligned}$$

After completing squares. Further simplifying, we have

$$\begin{aligned}
\varphi(z)\varphi(a-bz) &= \frac{1}{\sqrt{2\pi}} e^{\left(-\frac{1}{2}[\frac{a}{\sqrt{1+b^2}}]\right)} \cdot \frac{1}{\sqrt{2\pi}} e^{\left(-\frac{1}{2}\left[\frac{z-\frac{ab}{1+b^2}}{\frac{1}{\sqrt{1+b^2}}}\right]^2\right)} \\
&= \varphi(\frac{a}{\sqrt{1+b^2}})\varphi(\frac{z-\frac{ab}{1+b^2}}{\frac{1}{\sqrt{1+b^2}}}).
\end{aligned}$$

It follows that

$$\begin{aligned}
E[\varphi(a-bz)] &= \int_{-\infty}^{\infty} \varphi(z) \cdot \varphi(a-bz)dz \\
&= \varphi(\frac{a}{\sqrt{1+b^2}}) \int_{-\infty}^{\infty} \varphi(\frac{z-\frac{ab}{1+b^2}}{\frac{1}{\sqrt{1+b^2}}})dz \\
&= \varphi(\frac{a}{\sqrt{1+b^2}}) \cdot \frac{1}{\sqrt{1+b^2}}.
\end{aligned}$$

In other words,

$$\int_{-\infty}^{\infty} \varphi(z) \cdot \varphi(a-bz)dz \;=\; \frac{1}{\sqrt{1+b^2}} \cdot \varphi(\frac{a}{\sqrt{1+b^2}}). \tag{A.2}$$

This concludes the proof. □

Next, we apply the useful result (A.2) to our work, i.e., to show

$$\int \varphi\Big(\frac{\sqrt{IT+1}}{\sigma_x}\,\Big[\lambda - \frac{\big(\sum_{i=1}^{I}\sum_{t=1}^{T} lnx_{i,t} + lnx_{1,T+1}\big) - \big(T\sum_{i=1}^{I}\mu_i + \mu_1\big)}{IT+1}\Big]\Big)$$
$$\times\ \varphi\Big(\frac{\lambda-\mu_\lambda}{\sigma_\lambda}\Big)\, d\lambda$$
$$=\ \varphi\Big(\sqrt{\frac{IT+1}{\sigma_\lambda^2(IT+1)+\sigma_x^2}}\,\Big[\frac{\big(\sum_{i=1}^{I}\sum_{t=1}^{T} lnx_{i,t} + lnx_{1,T+1}\big) - \big(T\sum_{i=1}^{I}\mu_i + \mu_1\big)}{IT+1} - \mu_\lambda\Big]\Big)$$
$$\times\ \frac{\sigma_\lambda}{\sqrt{\frac{\sigma_\lambda^2(IT+1)}{\sigma_x^2}+1}}.$$

Proof. Let $z = \frac{\lambda-\mu_\lambda}{\sigma_\lambda}$ so that $dz = \frac{1}{\sigma_\lambda}d\lambda$. Set

$$a\ =\ \frac{\sqrt{IT+1}}{\sigma_x}\Big[\frac{\big(\sum_{i=1}^{I}\sum_{t=1}^{T} lnx_{i,t} + lnx_{1,T+1}\big) - \big(T\sum_{i=1}^{I}\mu_i + \mu_1\big)}{IT+1} - \mu_\lambda\Big],$$
$$b\ =\ \frac{\sigma_\lambda\sqrt{IT+1}}{\sigma_x}$$

then we have

$$b\cdot z\ =\ \frac{\sqrt{IT+1}(\lambda-\mu_\lambda)}{\sigma_x}.$$

Consider the left-hand side of (A.2)

$$\int \varphi(z)\cdot\varphi(a-bz)dz$$
$$= \int \varphi\Big(\frac{\lambda-\mu_\lambda}{\sigma_\lambda}\Big)\varphi\Big(\frac{\sqrt{IT+1}}{\sigma_x}\Big[\frac{\big(\sum_{i=1}^{I}\sum_{t=1}^{T} lnx_{i,t} + lnx_{1,T+1}\big) - \big(T\sum_{i=1}^{I}\mu_i + \mu_1\big)}{IT+1} - \mu_\lambda\Big]$$
$$-\ \frac{\sqrt{IT+1}(\lambda-\mu_\lambda)}{\sigma_x}\Big)\cdot\frac{1}{\sigma_\lambda}d\lambda$$
$$= \int \varphi\Big(\frac{\lambda-\mu_\lambda}{\sigma_\lambda}\Big)\varphi\Big(\Big[\frac{\big(\sum_{i=1}^{I}\sum_{t=1}^{T} lnx_{i,t} + lnx_{1,T+1}\big) - \big(T\sum_{i=1}^{I}\mu_i + \mu_1\big)}{\sigma_x\sqrt{IT+1}} - \mu_\lambda\frac{\sqrt{IT+1}}{\sigma_x}\Big]$$
$$-\ \frac{\sqrt{IT+1}\lambda}{\sigma_x} + \mu_\lambda\frac{\sqrt{IT+1}}{\sigma_x}\Big)\cdot\frac{1}{\sigma_\lambda}d\lambda$$
$$= \int \varphi\Big(\frac{\sqrt{IT+1}}{\sigma_x}\Big[\frac{\big(\sum_{i=1}^{I}\sum_{t=1}^{T} lnx_{i,t} + lnx_{1,T+1}\big) - \big(T\sum_{i=1}^{I}\mu_i + \mu_1\big)}{IT+1} - \lambda\Big]$$
$$\times\,\varphi\Big(\frac{\lambda-\mu_\lambda}{\sigma_\lambda}\Big)\cdot\frac{1}{\sigma_\lambda}d\lambda.$$

And the right-hand side of (A.2) yields

$$
\begin{aligned}
& \frac{1}{\sqrt{1+b^2}} \cdot \varphi(\frac{a}{\sqrt{1+b^2}}) \\
= & \frac{1}{\sqrt{1+(\frac{\sigma_\lambda\sqrt{IT+1}}{\sigma_x})^2}} \cdot \varphi(\frac{\frac{\sqrt{IT+1}}{\sigma_x}\Big[\frac{\big(\sum_{i=1}^{I}\sum_{t=1}^{T} lnx_{i,t}+lnx_{1,T+1}\big)-\big(T\sum_{i=1}^{I}\mu_i+\mu_1\big)}{IT+1}-\mu_\lambda\Big]}{\sqrt{1+(\frac{\sigma_\lambda\sqrt{IT+1}}{\sigma_x})^2}}) \\
= & \frac{1}{\sqrt{1+(\frac{\sigma_\lambda\sqrt{IT+1}}{\sigma_x})^2}} \cdot \varphi(\sqrt{\frac{IT+1}{\sigma_\lambda^2(IT+1)+\sigma_x^2}}\Big[\frac{\big(\sum_{i=1}^{I}\sum_{t=1}^{T} lnx_{i,t}+lnx_{1,T+1}\big)}{IT+1} \\
& - \frac{\big(T\sum_{i=1}^{I}\mu_i+\mu_1\big)}{IT+1}-\mu_\lambda\Big]).
\end{aligned}
$$

From equation (A.2), it follows

$$
\begin{aligned}
& \int \varphi(\frac{\lambda-\mu_\lambda}{\sigma_\lambda})\varphi(\frac{\sqrt{IT+1}}{\sigma_x}\Big[\frac{\big(\sum_{i=1}^{I}\sum_{t=1}^{T} lnx_{i,t}+lnx_{1,T+1}\big)-\big(T\sum_{i=1}^{I}\mu_i+\mu_1\big)}{IT+1}-\lambda\Big]d\lambda \\
= & \frac{\sigma_\lambda}{\sqrt{1+(\frac{\sigma_\lambda\sqrt{IT+1}}{\sigma_x})^2}} \cdot \varphi(\sqrt{\frac{IT+1}{\sigma_\lambda^2(IT+1)+\sigma_x^2}}\Big[\frac{\big(\sum_{i=1}^{I}\sum_{t=1}^{T} lnx_{i,t}+lnx_{1,T+1}\big)}{IT+1} \\
& - \frac{\big(T\sum_{i=1}^{I}\mu_i+\mu_1\big)}{IT+1}-\mu_\lambda\Big]).
\end{aligned}
$$

The proof is now complete. □

APPENDIX B

THE PROOF OF LEMMA 3.1

Lemma 3.1. *For a fixed* $i = 1, 2, \ldots, I$, *The identity*

$$E[(X_{i,T+1} - p)^2] \;=\; E[(X_{i,T+1} - E[X_{i,T+1}|\Lambda])^2] + E[(E[X_{i,T+1}|\Lambda] - p)^2]$$

holds for all $p \in \triangle_I$.

Proof. Let $X_{1,1}, X_{1,2}, \ldots, X_{1,T}, \ldots, X_{I,1}, \ldots, X_{I,T}, X$ be conditionally independent random variables with respect to Λ and

$$\triangle_I \;:=\; L^2(\sigma(X_{1,1}, X_{1,2}, \ldots, X_{1,T}, \ldots, X_{I,1}, \ldots, X_{I,T}).$$

It suffices to show that

$$E[(X - p)^2] \;=\; E[(X - E[X|\Lambda])^2] + E[(E[X|\Lambda] - p)^2]$$

holds for all $p \in \triangle_I$. Consider

$$\begin{aligned}
& E[(X - E[X|\Lambda])^2] + E[(E[X|\Lambda] - p)^2] \\
&= E[(X^2 - 2XE[X|\Lambda] + (E[X|\Lambda])^2] + E[(E[X|\Lambda])^2 - 2E[X|\Lambda]p + p^2] \\
&= E[X^2] - 2E[XE[X|\Lambda]] + E[(E[X|\Lambda])^2] + E[(E[X|\Lambda])^2] - 2E[E[X|\Lambda]p] + E[p^2] \\
&= E[X^2] - 2E[X]E[E[X|\Lambda]] + 2(E[E[X|\Lambda]])^2 - 2E[E[X|\Lambda]]E[p] + E[p^2] \\
&= E[X^2] - 2(E[X])^2 + 2([E[X])^2 - 2E[E[X|\Lambda]]E[p] + E[p^2] \\
&= E[X^2] - 2E[X]E[p] + E[p^2] \\
&= E[(X - p)^2].
\end{aligned}$$

This concludes the proof. □

APPENDIX C

PROBABILITY THEORY

We recall some definitions and theorems in probability theory. Most of these results can be found in Brzeźniak and Zastawniak (1999), Capiński and Kopp (2004), and Aggoun and Elliott (2004).

Definition C.1. Let Ω be a non-empty set. A *σ-field* $\mathcal{F}$ on Ω is a family of a subsets of Ω such that

1. the empty set $\emptyset$ belongs to $\mathcal{F}$;
2. if A belongs to $\mathcal{F}$, then so does the complement $\Omega \setminus A$;
3. if $A_1, A_2, \cdots$ is a sequence of sets in $\mathcal{F}$, then their union $A_1 \cup A_2 \cup \cdots$ also belongs to $\mathcal{F}$.

Definition C.2. Let $\mathcal{F}$ be a σ-field on Ω. A *probability measure* P is a function

$$\mathrm{P} : \mathcal{F} \to [0, 1]$$

such that

1. $\mathrm{P}(\Omega) = 1$;
2. if $A_1, A_2, \cdots$ are pairwise disjoint set (that is, $A_i \cap A_j = \emptyset$ for $i \neq j$) belonging to $\mathcal{F}$, then
$$\mathrm{P}(A_1 \cup A_2 \cup \cdots) = \mathrm{P}(A_1) + \mathrm{P}(A_2) + \cdots .$$

The triple $(\Omega, \mathcal{F}, \mathrm{P})$ is called a *probability space*. The sets belonging to $\mathcal{F}$ is called *events*. An event A is said to occur *almost surely* (a.s.) whenever $\mathrm{P}(A) = 1$.

Definition C.3. If $\mathcal{F}$ is a σ-field on Ω, then a function $X : \Omega \to \mathbb{R}$ is said to be $\mathcal{F}$-*measurable* if

$$(X \in B) := \{\omega \in \Omega : X(\omega) \in B\} = X^{-1}(B) \in \mathcal{F}$$

for every Borel set $B \in \mathcal{B}(\mathbb{R})$. If $(\Omega, \mathcal{F}, \mathrm{P})$ is a probability space, then such a function X is called a *random variable.*

Definition C.4. The σ-field $\sigma(X)$ *generated* by a random variable $X : \Omega \to \mathbb{R}$ consists of all sets of the form $(X \in B)$, where B is a Borel set in $\mathbb{R}$.

Lemma C.1 (Doob-Dynkin)**.** Let X be a random variable. Then each $\sigma(X)$-measurable random variable Y can be written as

$$Y = f(X)$$

for some Borel function $f : \mathbb{R} \to \mathbb{R}$.

Definition C.5. Every random variable $X : \Omega \to \mathbb{R}$ gives rise to a probability measure

$$\mathrm{P}_X(B) = \mathrm{P}(X \in B)$$

on $\mathbb{R}$ defined on the σ-field of Berel sets $B \in \mathcal{B}(\mathbb{R})$. We call P_X the distribution of X. The function $F_X : \mathbb{R} \to [0, 1]$ defined by

$$F_X(x) = \mathrm{P}(X \leq x)$$

is called the *distribution function* of X.

Definition C.6. If there is a Borel function $f_X : \mathbb{R} \to \mathbb{R}$ such that for any Borel set $B \subset \mathbb{R}$

$$\mathrm{P}(X \in B) = \int_B f_X(x)dx$$

then X is said to be a random variable with *absolutely continuous distribution* and f_X is called *density* of X. If there is a (finite or infinite) sequence of pairwise distinct real numbers $x_1, x_2, \cdots$ such that for any Borel set $B \subset \mathbb{R}$

$$\mathrm{P}(X \in B) = \sum_{x_i \in B} \mathrm{P}(X = x_i),$$

then X is said to have *discrete distribution* with value $x_1, x_2, \cdots$ and *mass* $\mathrm{P}(X = x_i)$ at x_i.

Definition C.7. A random variable $X : \Omega \to \mathbb{R}$ is said to be *integrable* if

$$\int_\Omega |X| d\mathrm{P} < \infty.$$

Then

$$\mathrm{E}[X] := \int_\Omega X d\mathrm{P}$$

exist and is called the *expectation* of X.

Definition C.8. Two events $A, B \in \mathcal{F}$ are called *independent* if

$$\mathrm{P}(A \cap B) = \mathrm{P}(A)\mathrm{P}(B).$$

In general, we say that n events $A_1, A_2, \cdots, A_n \in \mathcal{F}$ are *independent* if

$$\mathrm{P}(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}) = \mathrm{P}(A_{i_1})\mathrm{P}(A_{i_2}) \cdots \mathrm{P}(A_{i_k})$$

for any indices $1 \leq i_1 < i_2 < \cdots < t_k \leq n$.

Definition C.9. Two random variable X and Y are called *independent* if for any Borel sets $A, B \in \mathcal{B}(\mathbb{R})$ the two events

$$(X \in A) \text{ and } (Y \in B)$$

are independent. We say that n random variable $X_1, X_2, \cdots, X_n$ are *independent* if for any Borel sets $B_1, B_2, \cdots, B_n \in \mathcal{B}(\mathbb{R})$ the events

$$(X_1 \in B_1),\ (X_2 \in B_2), \cdots, (X_n \in B_n)$$

are independent.

Definition C.10. Two σ-fields $\mathcal{G}$ and $\mathcal{H}$ contain in $\mathcal{F}$ are called *independent* if any two events $A \in \mathcal{G}$ and $B \in \mathcal{H}$ are independent. Similarly, any finite number of σ-fields $\mathcal{G}_1, \mathcal{G}_2, \cdots, \mathcal{G}_n$ contained in $\mathcal{F}$ are *independent* if any n events

$$A_1 \in \mathcal{G}_1, A_2 \in \mathcal{G}_2, \cdots, A_n \in \mathcal{G}_n$$

are independent.

Definition C.11. We say that a random variable X is *independent* of σ-field $\mathcal{G}$ if the σ-fields $\sigma(X)$ and $\mathcal{G}$ are independent.

Theorem C.2 (Lebesgue's Dominated Convergence Theorem)**.** Suppose $\{X_n, n \in \mathbb{N}\}$ is a sequence of random variables such that $|X_n| \leq Y$ a.s. where Y is an integrable random variable. If X_n converses to X a.s., then X_n and X are integrable,

$$\lim_{n\to\infty} \int_\Omega X_n d\mathrm{P} = \lim_{n\to\infty} \int_\Omega X d\mathrm{P}$$

and

$$\lim_{n\to\infty} \int_\Omega |X_n - X| d\mathrm{P} = 0.$$

Theorem C.3. Let $(\Omega, \mathcal{F}, \mathrm{P})$ be a probability space. Given a random variable $X : \Omega \to \mathbb{R}$,

$$\int_\Omega g(X(\omega)) d\mathrm{P}(\omega) = \int_\mathbb{R} g(x) d\mathrm{P}_X(x).$$

Theorem C.4. If P_X defined on $\mathbb{R}^n$ is absolutely continuous with density f_X, $g : \mathbb{R}^n \to \mathbb{R}$ is integrable with respect to P_X, then

$$\int_{\mathbb{R}^n} g(x) d\mathrm{P}_X(x) = \int_{\mathbb{R}^n} f_X(x) g(x) dx.$$

Corollary C.5. In the situation of the previous theorem we have

$$\int_\Omega g(X) d\mathrm{P} = \int_{\mathbb{R}^n} f_X(x) g(x) dx.$$

Theorem C.6. Let $(\Omega, \mathcal{F}, \mathrm{P})$ be a probability space. Let X be a real random variable and B a Borel set. Then

$$\int_B g(x) dF_X(x) = \int_{X^{-1}(B)} g(X(\omega)) d\mathrm{P}(\omega).$$

Here g is a Borel function and where $B = \mathbb{R}$

$$\int_\mathbb{R} g(x) dF_X(x) = \int_\Omega g(X(\omega)) d\mathrm{P}(\omega).$$

Comotonicity

Let $\vec{x} = (x_1, x_2, \ldots, x_n)$ be a set of n-vector in $\mathbb{R}^n$. For two n-vector $\vec{x}$ and $\vec{y}$, the notation $\vec{x} \leq \vec{y}$ will be used for the componentwise order which is defined by $x_i \leq y_i$ for all $i = 1, 2, \ldots, n$.

Definition C.12 (Comotonic set)**.** The set $A \subset \mathbb{R}^n$ is comotonic if for any $\vec{x}$ and $\vec{y}$ in A, either $\vec{x} \leq \vec{y}$ or $\vec{y} \leq \vec{x}$ holds.

So, a set $A \subset \mathbb{R}^n$ is comotonic if for any $\vec{x}$ and $\vec{y}$ in A, the inequality $x_i \leq y_i$ for some i, implies that $\vec{x} \leq \vec{y}$. As a comotonic set is simultaneously non-decreasing in each component, it is also called a nondecreasing set. Notice that any subset of comotonic set is also comotonic.

Next, we define a comotonic random vector $\vec{X} = (X_1, X_2, \ldots, X_n)$ through its support. A support of random vector $\vec{X}$ is a set $A \subset \mathbb{R}^n$ for which $P(\vec{X} \in A) = 1$.

Definition C.13 (Comotonic random vector)**.** A random vector $\vec{X} = (X_1, X_2, \ldots, X_n)$ is comotonic if it has a comotonic support.

From the definition, we can conclude that comotonicity is a very strong positive dependency structure. Indeed, if $\vec{x}$ and $\vec{y}$ are elements of the (comotonic) support of $\vec{X}$, i.e. $\vec{x}$ and $\vec{y}$ are possible outcomes of $\vec{X}$, then they must be ordered componentwise. This explains why the term comotomic (common monotonic) is used.

APPENDIX D

FUNCTIONAL ANALYSIS

We recall some definition and theorem from functional analysis. Most of these results can be found in Kreyszig (1998) and Apostol (1974).

Theorem D.1 (Continuous Mapping)**.** A mapping $T : X \to Y$ of a metric space (X, d) into a metric space $(Y, \tilde{d})$ is continuous at a point $x_0 \in X$ if and only if

$$x_n \to x_0 \text{ implies } Tx_n \to Tx_0.$$

Definition D.1. A metric space X is said to be *compact* if every sequence in X has a convergent subsequence. A subset M of X is said to be *compact* if M is compact considered as a subspace of X, that is, if every sequence in M has a convergent subsequence whose limit is an element in M.

Theorem D.2 (Compactness)**.** In a finite dimensional normed space X, any subset $M \subset X$ is compact if and only if M is closed and bounded.

Theorem D.3 (Continuous Mapping)**.** Let X and Y be metric spaces and $T : X \to Y$. Then the image of a compact subset M of X under T is compact.

Corollary D.4 (Maximum and Minimum)**.** A continuous mapping T of a compact subset M of a metric space X into $\mathbb{R}$ assumes a maximum and a minimum at some points of M.

Theorem D.5 (Canter Intersection Theorem)**.**

Let $\{F_1, F_2, F_3, \cdots\}$ be a countable collection of nonempty sets in $\mathbb{R}^m$ such that:

(i) $F_{n+1} \subset F_n,\ n \in \mathbb{N}$;

(ii) each set F_n is closed and F_1 is bounded. Then the intersection $\cap_{n=1}^{\infty} F_n$ is closed and nonempty.

Definition D.2 (Orthogonality)**.** An element x of an inner product space X is said to be *orthogonal* to an element $y \in X$ if

$$\langle x, y \rangle = 0.$$

Theorem D.6 (Schwarz Inequality)**.** An inner product and the corresponding norm satisfy the Schwarz inequality, i.e.,

$$|\langle x, y \rangle| \leq \|x\| \|y\|.$$

Theorem D.7 (Minimizing Vector)**.** Let X be an inner product space and $M \neq \emptyset$ a convex subset which is complete (in the metric induced by the inner product). Then for every given $x \in X$ there exists a unique $y \in M$ such that

$$\delta = \inf_{\tilde{y} \in M} \|x - \tilde{y}\| = \|x - y\|.$$

Theorem D.8 (Orthogonality)**.** In theorem D.7, let M be a complete subspace Y and $x \in X$ fixed. Then $z = x - y$ is orthogonal to Y.

Definition D.3 (Direct Sum)**.** A vector space X is said to be the *direct sum* of two subspaces of Y and Z , written

$$X = Y \oplus Z,$$

if each $x \in X$ has a unique representation

$$x = y + z$$

for some $y \in Y$ and $z \in Z$. Then Z is called *algebraic complement* of Y in X and vice versa, and Y, Z is called a *complement pair* of subspaces in X.

In the case of a general Hilbert space H, the main interest concerns representations of H as a direct sum of a closed subspace Y and its orthogonal complement

$$Y^{\perp} := \{z \in H : z \perp Y\},$$

which is the set of all vectors orthogonal to Y.

Theorem D.9 (Direct Sum)**.** Let Y be any closed subset of a Hilbert space H. Then $H = Y \oplus Z$ and $Y \cap Z = \{0\}$ when $Z = Y^{\perp}$.

In theorem D.9, we found that for every $x \in H$ there exists and unique a $y \in Y$ and $z \in Y^{\perp}$ such that $x = y + z$, y is called the *orthogonal projection* of x on Y. Define a mapping

$$\begin{array}{rccl} \rho : & H & \rightarrow & Y \\ & x & \mapsto & y = \rho(x). \end{array}$$

ρ is called the (*orthogonal*) *projection* of H onto Y.

CURRICULUM VITAE

NAME: Tippatai Pongsart.

GENDER: Male.

NATIONALITY: Thai.

DATE OF BIRTH: June 14, 1985.

EDUCATIONAL BACKGROUND:

- B.Sc. in Mathematics, Khonkaen University, Khonkaen, Thailand, 2007.
- M.Sc. in Mathematics, Khonkaen University, Khonkaen, Thailand, 2010.
  (Academic Excellence Award from the Professor Dr. Tab Nilanidhi Foundation: year 2010)

SCHOLARSHIPS:

- Centre of Excellence in Mathematics of Thailand (CEM).