



Some Bayesian Premiums Obtained by Using the Common Effect in Claim Dependence Model

T. Pongsart[†], K. Sangaroon[‡] and P. Sattayatham^{†,1}

[†]School of Mathematics, Institute of Science, Suranaree University of
Technology, Nakhon Ratchasima, Thailand
e-mail : fanpongsart@yahoo.com (T. Pongsart)
pairote@sut.ac.th (P. Sattayatham)

[‡]Department of Mathematics, Faculty of Science, Khon Kaen University
Khon Kaen, Thailand
e-mail : skiat@kku.ac.th (K. Sangaroon)

Abstract : 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 independent assumptions across different risks for the sake of convenience and mathematical tractability. However, these assumptions may be violated in some practical situations. In this paper, modeling claim dependence is built under the common effect. According to the model, we express the expected claims given the history of all observable claims (Bayesian premiums) in explicit form with lognormal claim amounts distribution.

Keywords : Bayesian premiums; claim dependence; common effect; lognormal claims.

2010 Mathematics Subject Classification : 62P05; 97M30.

This work was supported by the Centre of Excellence in Mathematics in Thailand
¹ Corresponding author.

1 Introduction

One crucial task both in the practical management of an insurance company and in theoretical consideration is to determine premiums adequate to cover any risks. These premiums are calculated based on the chosen model, information in the insurance contract (e.g. claims 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 known 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. However, no matter how detailed the underwriting procedure, there still remains some heterogeneity with respect to risk characteristics within a rating class. In this paper, we note that an individual's risk refers to the claim amounts for an individual.

In risk theory, each risk X for an individual is characterized by a risk parameter θ (possibly vector valued) due to heterogeneity over policies in the 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 will be 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 $\pi(\theta)$.

In the modeling of claims, several studies have assumed independence of claims which may be appropriate in some practical situations, including mathematical tractability. In real applications, we agree there are some situations in which these assumptions may be violated; for example, in house insurance in which geographic proximity of the insured may result in exposure to a common catastrophe, and in motor insurance in which one collision may involve several insured parties.

The concept of modeling dependence began with a consideration of time dependence but not of dependence across individuals. An early paper by Gerber and Jones (1975)[1] and a paper by Frees et al. (1999)[2] are examples of credibility models with time dependence for claims. Works by Wang (1998)[3] proposed a set of statistical tools for modeling dependencies of risks in an insurance portfolio. A paper by Purcaru and Denuit (2002)[4] provided a kind of dependence for claim frequency induced by introducing common latent variables. 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 [5] proposed a claims dependent model under the assumptions as follows: 1. claims in a portfolio are dependent across time periods for a fixed individual and 2. claims in a portfolio are dependent across individuals for a fixed time period.

In this paper, we introduce a different claims dependent model under the

assumption that claims in a portfolio are only dependence across individuals for a fixed time periods.

In the present model, the common effect is used to describe the common dependence among individuals. This model is suitable for some situations such as the data of dependence between time periods for each individual in a portfolio may not appropriate or unavailable. The main purpose is to derive the Bayesian premium corresponding to the model.

Furthermore, the explicit form of the Bayesian premium requires distributions of both claim amounts and common effect. A paper by Yeo and Valdez (2006)[5] considered both of these two distributions are only normally distributed. In our work, we are especially interested in lognormal distribution. This distribution is a long-tailed distribution that is widely used to describe a feature of claim amounts in non-life insurance; for example, motor insurance, fire insurance or allied perils insurance. We derive explicit form of the Bayesian premium under normal common effect and lognormal claim amounts. We also obtain the explicit form of the Bayesian premium under normal claim amounts which can be further rewritten in a credibility formula.

The structure for the rest of the paper has been made as follows. In section 2, model descriptions and preliminaries are introduced. We also derive some results in order to find the Bayesian premium under the square-error loss function. Section 3 establishes the Bayesian premiums for lognormal and normal claim amounts distributions while the common effects of both are normally distributed. Finally, we provide some concluding remarks in section 4.

2 Model Formulation and Preliminaries

Let (Ω, \mathcal{F}, P) be a probability space, let $L^2(\mathcal{F})$ denoted the Hilbert space of all random variables $X : \Omega \rightarrow \mathbb{R}$ 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

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

represents the vector of claims for a particular individual $i = 1, 2, \dots, I$. We are interested in the prediction of a future claim for each individual i based on all the observed claims

$$\vec{X} = (\vec{X}_1, \vec{X}_2, \dots, \vec{X}_I). \quad (2.1)$$

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

Define a subspace H of $L^2(\mathcal{F})$ by

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

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

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

It is well known from statistical theory that this solution p_i^* satisfies $p_i^* = E[X_{i,T+1} | \vec{X} = \vec{x}]$ and it is known as the Bayesian premium for risk \vec{X}_i .

The One-Level Common Effect Model

As already mentioned in the introductory section, our main purpose is to study the Bayesian premium under a type of dependence structure among individuals which will be described by a common effect random variable Λ and its realization λ . Conditional on this common effect, the random vectors \vec{X}_i are independent. This common effect will define the dependence structure between individual risks and it can either be a discrete, continuous, or a mixture of discrete and continuous random variables. We assume the density function $f_\Lambda(\lambda)$ is provided. More precisely, we shall summarize these with the following assumptions.

A1. The common effect random variable Λ has known probability density function $f_\Lambda(\lambda)$.

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

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

A4. For a fixed $i = 1, 2, \dots, I$ and a fixed $t = 1, 2, \dots, T+1$, 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 into the claim among individuals, such as a catastrophe in general insurance, or simply bad weather conditions on a day when automobile accidents are frequent. In order to estimate parameters for the model, better parameter estimates can be acquired if the historical claims data can be separated into categories of *catastrophe* and *non-catastrophe*. This categorization information, keeping track of claims, especially the cause of claims, should be readily available for most insurance companies as they come in reported.

Method to find Bayesian premiums: For a fixed individual $j = 1, 2, \dots, I$, the Bayesian premiums which can be conveniently expressed as

$$E[X_{j,T+1} | \vec{X} = \vec{x}] = \int x_{j,T+1} \cdot f_{X_{j,T+1}|\vec{X}}(x_{j,T+1}|\vec{x}) dx_{j,T+1} \quad (2.2)$$

requires an explicit formula for conditional density $f_{X_{j,T+1}|\vec{X}}(x_{j,T+1}|\vec{x})$. To achieve this, we need a lemma.

Lemma 2.1. *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 (2.1). The joint density of \vec{X} and the overall risk parameter Λ can be expressed as*

$$f_{\vec{X},\Lambda}(\vec{x}, \lambda) = \prod_{i=1}^I f_{\vec{X}_i|\Lambda}(\vec{x}_i|\lambda) \times f_{\Lambda}(\lambda). \quad (2.3)$$

Proof. From the definition for conditional density and assumption **A3**, we have

$$\begin{aligned} f_{\vec{X},\Lambda}(\vec{x}, \lambda) &= f_{\vec{X}|\Lambda}(\vec{x}, |\lambda) \times f_{\Lambda}(\lambda) \\ &= \prod_{i=1}^I f_{\vec{X}_i|\Lambda}(\vec{x}_i|\lambda) \times f_{\Lambda}(\lambda). \end{aligned} \quad (2.4)$$

The proof is now complete. \square

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

$$f_{\Lambda|\vec{X}}(\lambda|\vec{x}) = f_{\vec{X},\Lambda}(\vec{x}, \lambda) \times \frac{1}{f_{\vec{X}}(\vec{x})}. \quad (2.5)$$

Substituting (2.4) into the right-hand side of (2.5), one gets

$$f_{\Lambda|\vec{X}}(\lambda|\vec{x}) = C \times \prod_{i=1}^I f_{\vec{X}_i|\Lambda}(\vec{x}_i|\lambda) \times f_{\Lambda}(\lambda), \quad (2.6)$$

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

Theorem 2.1. *Suppose the random variable Λ and the random vector \vec{X} satisfy all assumptions as in Lemma 2.1. The conditional density of $X_{j,T+1}|\vec{X}$ can be expressed as*

$$f_{X_{j,T+1}|\vec{X}}(x_{j,T+1}|\vec{x}) = \int f_{X_{j,T+1}|\Lambda}(x_{j,T+1}|\lambda) \times f_{\Lambda|\vec{X}}(\lambda|\vec{x}) d\lambda. \quad (2.7)$$

Proof. In the definition of conditional density, we have

$$f_{X_{j,T+1}|\vec{X}}(x_{j,T+1}|\vec{x}) = f_{X_{j,T+1},\vec{X}}(x_{j,T+1}, \vec{x}) \times \frac{1}{f_{\vec{X}}(\vec{x})}. \quad (2.8)$$

We note that the density $f_{X_{j,T+1},\vec{X}}(x_{j,T+1}, \vec{x})$ 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 $f_{X_{j,T+1},\vec{X},\Lambda}(x_{j,T+1}, \vec{x}, \lambda)$. Using the definition of conditional density, assumptions **A3** and **A4** then we have

$$f_{X_{j,T+1},\vec{X},\Lambda}(x_{j,T+1}, \vec{x}, \lambda) = f_{X_{j,T+1}|\Lambda}(x_{j,T+1}|\lambda) \times \prod_{i=1}^I f_{X_i|\Lambda}(x_i, |\lambda) \times f_{\Lambda}(\lambda). \quad (2.9)$$

By equation (2.6), one gets

$$\prod_{i=1}^I f_{\vec{x}_i|\Lambda}(\vec{x}_i|\lambda) \times f_{\Lambda}(\lambda) = f_{\Lambda|\vec{X}}(\lambda|\vec{x}) \times f_{\vec{X}}(\vec{x}). \quad (2.10)$$

Substituting (2.10) into the right-hand side of (2.9) yields

$$f_{X_{j,T+1},\vec{X},\Lambda}(x_{j,T+1},\vec{x},\lambda) = f_{X_{j,T+1}|\Lambda}(x_{j,T+1}|\lambda) \times f_{\Lambda|\vec{X}}(\lambda|\vec{x}) \times f_{\vec{X}}(\vec{x}). \quad (2.11)$$

Next, integrating (2.11) with respect to λ , we have

$$f_{X_{j,T+1},\vec{X}}(x_{j,T+1},\vec{x}) = f_{\vec{X}}(\vec{x}) \int f_{X_{j,T+1}|\Lambda}(x_{j,T+1}|\lambda) \times f_{\Lambda|\vec{X}}(\lambda|\vec{x}) d\lambda. \quad (2.12)$$

Substituting (2.12) into the right-hand side of (2.8), 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. \square

The objective of the theorem above is to derive an explicit expression for the conditional density in terms of all available or given information. Notice that the conditional density $f_{X_{j,T+1}|\Lambda}(x_{j,T+1}|\lambda)$ which according to assumption A4 is known and given, and that of

$$f_{\Lambda|\vec{X}}(\lambda|\vec{x}) = f_{\Lambda,\vec{X}}(\lambda,\vec{x}) \times \frac{1}{f_{\vec{X}}(\vec{x})}, \quad (2.13)$$

for which the numerator can be evaluated using Lemma (2.1) together with the independence of the common effect Λ .

3 Bayesian Premiums with Normal Common Effect

The normal distribution is the most widely known and used of all distributions because the normal distribution approximates many natural phenomena. In this section, we shall use equation (2.2) to find the Bayesian premium when the common effect λ is assumed to be normally distributed. We consider the claim amounts following lognormal or normal distribution. We divide our investigation into two cases.

3.1 Bayesian Premiums with Lognormal Claim Amounts

In this case, we make the following assumptions: for convenience, we write $X|\lambda := X|\Lambda = \lambda$.

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

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

where μ_j is a constant depending on individual j and needs to be chosen.

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 common effect λ is normally distributed with mean μ_λ and variance σ_λ^2 .

A useful application of Theorem 2.1 appears in the following theorem.

Theorem 3.1. *Suppose the random variable Λ and the random vector \vec{X} satisfy all assumptions as in Lemma 2.1. Assume further that $X_{j,t}|\lambda$ and common effect λ satisfy **L1** and **L2**, respectively. Then the Bayesian premium can be written as*

$$E[X_{j,T+1}|\vec{X} = \vec{x}] = e^{\left[\frac{\sigma_\lambda^2 (\sum_{i=1}^I \sum_{t=1}^T \ln X_{i,t} - T \sum_{i=1}^I \mu_i + \mu_j IT) + \sigma_x^2 (\mu_\lambda + \mu_j)}{\sigma_\lambda^2 IT + \sigma_x^2} \right]} \times e^{\left[\frac{(\sigma_\lambda^2 (IT+1) + \sigma_x^2) \sigma_x^2}{2(\sigma_\lambda^2 IT + \sigma_x^2)} \right]} \quad (3.1)$$

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

Proof. We proceed the proof by following steps:

STEP 1. We recall equation (2.2):

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

Our main interest is to derive the density of $X_{j,T+1}|\vec{X} = \vec{x}$ where without loss of generality, we fix $j = 1$.

From assumptions **L1** and **L2**, we have

$$f_{X_{j,t}|\Lambda}(x_{j,t}|\lambda) = \frac{1}{x_{j,t}\sigma_x\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\ln x_{j,t} - (\mu_j + \lambda)}{\sigma_x}\right)^2}, \quad \text{and} \quad (3.2)$$

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

And the conditional density $f_{X_{1,T+1}|\Lambda}(x_{1,T+1}|\lambda)$ is already known to be

$$f_{X_{1,T+1}|\Lambda}(x_{1,T+1}|\lambda) = \frac{1}{x_{1,T+1}\sigma_x\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\ln x_{1,T+1} - (\mu_1 + \lambda)}{\sigma_x}\right)^2}. \quad (3.3)$$

Applying Theorem (2.1), we have

$$f_{X_{1,T+1}|\vec{X}}(x_{1,T+1}|\vec{x}) = C_1 \int f_{X_{1,T+1}|\Lambda}(x_{1,T+1}|\lambda) \times f_{\Lambda,\vec{X}}(\lambda, \vec{x}) d\lambda. \quad (3.4)$$

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 joint density $f_{\Lambda, \vec{X}}(\lambda, \vec{x})$ can be derived by using lemma (2.1), giving

$$f_{\vec{X}, \Lambda}(\vec{x}, \lambda) = \prod_{i=1}^I \left(\prod_{t=1}^T \frac{1}{x_{i,t} \sigma_x \sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{\ln x_{i,t} - (\mu_i + \lambda)}{\sigma_x} \right)^2} \right) \frac{1}{\sigma_\lambda \sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{\lambda - \mu_\lambda}{\sigma_\lambda} \right)^2}. \quad (3.5)$$

Substituting (3.3) and (3.5) into (3.4), we have the density of $X_{1,T+1} | \vec{X} = \vec{x}$ is in the form

$$f_{X_{1,T+1} | \vec{X}}(x_{1,T+1} | \vec{x}) = \int \frac{C_1}{x_{1,T+1} \prod_{i=1}^I \left(\prod_{t=1}^T x_{i,t} \right) (\sigma_x)^{IT+1} (2\pi)^{\frac{IT+2}{2}} \sigma_\lambda} \times e^{-\frac{1}{2} \left[\sum_{t=1}^{T+1} \left(\frac{\ln x_{1,t} - (\mu_1 + \lambda)}{\sigma_x} \right)^2 \right]} \times e^{-\frac{1}{2} \left(\frac{\lambda - \mu_\lambda}{\sigma_\lambda} \right)^2} \times e^{-\frac{1}{2} \left[\sum_{i=2}^I \sum_{t=1}^T \left(\frac{\ln x_{i,t} - (\mu_i + \lambda)}{\sigma_x} \right)^2 \right]} d\lambda. \quad (3.6)$$

STEP 2. Consider the right-hand side of (3.6), we show that

$$\begin{aligned} & \int \frac{1}{2\pi} e^{-\frac{1}{2} \left[\sum_{t=1}^{T+1} \left(\frac{\ln x_{1,t} - (\mu_1 + \lambda)}{\sigma_x} \right)^2 + \sum_{i=2}^I \sum_{t=1}^T \left(\frac{\ln x_{i,t} - (\mu_i + \lambda)}{\sigma_x} \right)^2 + \left(\frac{\lambda - \mu_\lambda}{\sigma_\lambda} \right)^2 \right]} d\lambda \\ &= \varphi \left(\sqrt{\frac{IT+1}{\sigma_\lambda^2 (IT+1) + \sigma_x^2}} \left[\frac{(\sum_{i=1}^I \sum_{t=1}^T \ln x_{i,t} + \ln x_{1,T+1}) - (T \sum_{i=1}^I \mu_i + \mu_1)}{IT+1} - \mu_\lambda \right] \right) \\ & \times \frac{\sigma_x \sigma_\lambda^2}{\sqrt{IT+1} \sqrt{\frac{\sigma_\lambda^2 (IT+1)}{\sigma_x^2} + 1}} \times e^{-\frac{1}{2\sigma_x^2} \left[- \left(\frac{[\sum_{i=1}^I \sum_{t=1}^T \ln x_{i,t} + \ln x_{1,T+1}] - [T \sum_{i=1}^I \mu_i + \mu_1]}{IT+1} \right)^2 \right]} \\ & \times e^{-\frac{1}{2\sigma_x^2} \left[(\sum_{i=1}^I \sum_{t=1}^T (\ln x_{i,t})^2 + (\ln x_{1,T+1})^2) - 2(\sum_{i=1}^I (\mu_i \sum_{t=1}^T \ln x_{i,t}) + (\ln x_{1,T+1} \cdot \mu_1)) \right]} \\ & \times e^{-\frac{1}{2\sigma_x^2} [T \sum_{i=1}^I \mu_i^2 + \mu_1^2]}. \quad (3.7) \end{aligned}$$

From equation (3.6), the term containing λ can be written in the form

$$\int \frac{1}{2\pi} e^{-\frac{1}{2} \left[\sum_{t=1}^{T+1} \left(\frac{\ln x_{1,t} - (\mu_1 + \lambda)}{\sigma_x} \right)^2 + \sum_{i=2}^I \sum_{t=1}^T \left(\frac{\ln x_{i,t} - (\mu_i + \lambda)}{\sigma_x} \right)^2 + \left(\frac{\lambda - \mu_\lambda}{\sigma_\lambda} \right)^2 \right]} d\lambda. \quad (3.8)$$

By extracting and regrouping (3.8), we get a term which can be simplified as follow:

$$\begin{aligned} & \frac{1}{2\pi} e^{-\frac{IT+1}{2\sigma_x^2} \left[\left(\lambda - \frac{(\sum_{i=1}^I \sum_{t=1}^T \ln x_{i,t} + \ln x_{1,T+1}) - (T \sum_{i=1}^I \mu_i + \mu_1)}{IT+1} \right)^2 \right]} e^{-\frac{1}{2} \left(\frac{\lambda - \mu_\lambda}{\sigma_\lambda} \right)^2} \\ &= \frac{\sigma_x \sigma_\lambda}{\sqrt{IT+1}} \times \varphi \left(\frac{\sqrt{IT+1}}{\sigma_x} \left[\lambda - \frac{(\sum_{i=1}^I \sum_{t=1}^T \ln x_{i,t} + \ln x_{1,T+1}) - (T \sum_{i=1}^I \mu_i + \mu_1)}{IT+1} \right] \right) \\ & \times \varphi \left(\frac{\lambda - \mu_\lambda}{\sigma_\lambda} \right), \quad (3.9) \end{aligned}$$

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

We use a result from Valdez (2004)[6] to simplify (3.9). This result states that for $\varphi(z)$ and any constants a and b , the following is true:

$$\int_{-\infty}^{\infty} \varphi(z)\varphi(a - bz)dz = \frac{1}{\sqrt{b^2 + 1}} \varphi\left(\sqrt{\frac{a^2}{b^2 + 1}}\right). \quad (3.10)$$

Thus, by letting $z = \frac{\lambda - \mu_\lambda}{\sigma_\lambda}$ so that $dz = \frac{1}{\sigma_\lambda}d\lambda$, then applying (3.10) to (3.9) and after simplifying, one gets equation (3.7).

STEP 3. To verify the equation

$$f_{X_{1,T+1}|\bar{X}}(x_{1,T+1}|\bar{x}) = \frac{C_2}{x_{1,T+1}} \times e^{-\frac{1}{2}[A(\ln x_{1,T+1})^2 - 2B(\ln x_{1,T+1}) + K]}, \quad (3.11)$$

where

$$\begin{aligned} A &= \frac{\sigma_x^2 + \sigma_\lambda^2 IT}{(\sigma_\lambda^2(IT + 1) + \sigma_x^2)\sigma_x^2}, \\ B &= \frac{\sigma_\lambda^2(\sum_{i=1}^I \sum_{t=1}^T \ln x_{i,t} - T \sum_{i=1}^I \mu_i + \mu_1 IT) + \sigma_x^2(\mu_\lambda + \mu_1)}{(\sigma_\lambda^2(IT + 1) + \sigma_x^2)\sigma_x^2}, \quad \text{and} \\ K &= \frac{-(\sigma_\lambda^2(IT + 1))(\sum_{i=1}^I \sum_{t=1}^T \ln x_{i,t})^2 - (\sigma_\lambda^2(IT + 1))[T \sum_{i=1}^I \mu_i + \mu_1]^2}{(\sigma_\lambda^2(IT + 1) + \sigma_x^2)\sigma_x^2(IT + 1)} \\ &+ \frac{2(\sigma_\lambda^2(IT + 1))(\sum_{i=1}^I \sum_{t=1}^T \ln x_{i,t})(T \sum_{i=1}^I \mu_i + \mu_1)}{(\sigma_\lambda^2(IT + 1) + \sigma_x^2)\sigma_x^2(IT + 1)} \\ &+ \frac{(\sum_{i=1}^I \sum_{t=1}^T (\ln x_{i,t})^2)(\sigma_\lambda^2(IT + 1)^2 + \sigma_x^2(IT + 1))}{(\sigma_\lambda^2(IT + 1) + \sigma_x^2)\sigma_x^2(IT + 1)} \\ &+ \frac{(\sum_{i=1}^I (\mu_i \sum_{t=1}^T \ln x_{i,t}))(-2\sigma_\lambda^2(IT + 1)^2 - 2\sigma_x^2(IT + 1))}{(\sigma_\lambda^2(IT + 1) + \sigma_x^2)\sigma_x^2(IT + 1)} \\ &+ \frac{(T \sum_{i=1}^I \mu_i^2 + \mu_1^2)(\sigma_\lambda^2(IT + 1)^2 + \sigma_x^2(IT + 1))}{(\sigma_\lambda^2(IT + 1) + \sigma_x^2)\sigma_x^2(IT + 1)} \\ &- \frac{2\sigma_x^2 \mu_\lambda (IT + 1)(\sum_{i=1}^I \sum_{t=1}^T \ln x_{i,t})}{(\sigma_\lambda^2(IT + 1) + \sigma_x^2)\sigma_x^2(IT + 1)} \\ &+ \frac{2\sigma_x^2 \mu_\lambda (IT + 1)(T \sum_{i=1}^I \mu_i + \mu_1) + \sigma_x^2(IT + 1)^2 \mu_\lambda^2}{(\sigma_\lambda^2(IT + 1) + \sigma_x^2)\sigma_x^2(IT + 1)}. \end{aligned} \quad (3.12)$$

We continue from (3.6) by substituting (3.7) back into the equation to obtain

$$\begin{aligned}
 & f_{X_{1,T+1}|\vec{X}}(x_{1,T+1}|\vec{x}) \\
 &= \frac{1}{\sqrt{2\pi}} \times e^{-\frac{1}{2} \left(\frac{IT+1}{\sigma_\lambda^2(IT+1)+\sigma_x^2} \left[\frac{(\sum_{i=1}^I \sum_{t=1}^T \ln x_{i,t} + \ln x_{1,T+1}) - (T \sum_{i=1}^I \mu_i + \mu_1)}{IT+1} - \mu_\lambda \right]^2 \right)} \\
 & \times e^{-\frac{1}{2\sigma_x^2} \left[- \left(\frac{\sum_{i=1}^I \sum_{t=1}^T \ln x_{i,t} + \ln x_{1,T+1}}{IT+1} - \left[T \sum_{i=1}^I \mu_i + \mu_1 \right] \right)^2 \right]} \\
 & \times e^{-\frac{1}{2\sigma_x^2} \left[(\sum_{i=1}^I \sum_{t=1}^T (\ln x_{i,t})^2 + (\ln x_{1,T+1})^2) - 2 \left(\sum_{i=1}^I (\mu_i \sum_{t=1}^T \ln x_{i,t}) + (\ln x_{1,T+1} \cdot \mu_1) \right) \right]} \\
 & \times e^{-\frac{1}{2\sigma_x^2} \left[T \sum_{i=1}^I \mu_i^2 + \mu_1^2 \right]} \times \frac{C_1 \sigma_\lambda}{x_{1,T+1} (\prod_{i=1}^I \prod_{t=1}^T x_{i,t}) (\sigma_x)^{IT} (2\pi)^{\frac{IT+1}{2}} \sqrt{IT+1} \sqrt{\frac{\sigma_\lambda^2(IT+1)}{\sigma_x^2} + 1}}.
 \end{aligned} \tag{3.13}$$

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 substitute the constant back into (3.13), and after extracting and grouping the terms containing $(\ln x_{1,T+1})^2$ and $\ln x_{1,T+1}$ in this equation, we obtain the result as presented in (3.11).

STEP 3. Now we prove the Theorem 3.1. By applying the square operation:

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

into (3.11) then we have

$$f_{X_{1,T+1}|\vec{X}}(x_{1,T+1}|\vec{x}) = \frac{C_2}{x_{1,T+1}} e^{-\frac{1}{2} \left[\frac{(\ln x_{1,T+1} - \frac{B}{A})^2}{\frac{1}{A}} \right]} \times e^{-\frac{1}{2} [K - \frac{B^2}{A}]}. \tag{3.14}$$

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

$$\begin{aligned}
 E[X_{1,T+1}|\vec{X} = \vec{x}] &= e^{\mu_{1,T+1} + \frac{\sigma_{1,T+1}^2}{2}} \\
 &= e^{\left[\frac{\sigma_\lambda^2 (\sum_{i=1}^I \sum_{t=1}^T \ln x_{i,t} - T \sum_{i=1}^I \mu_i + \mu_1 IT) + \sigma_x^2 (\mu_\lambda + \mu_1)}{\sigma_x^2 + \sigma_\lambda^2 IT} \right]} \\
 & \times e^{\left[\frac{(\sigma_\lambda^2 (IT+1) + \sigma_x^2) \sigma_x^2}{2(\sigma_x^2 + \sigma_\lambda^2 IT)} \right]}.
 \end{aligned} \tag{3.15}$$

This concludes the proof. \square

3.2 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 3.2. *Suppose the random variable Λ and the random vector \vec{X} satisfy all assumptions as in Lemma 2.1. Assume further that $X_{j,t}|\lambda$ and common effect λ satisfy **N1** and **L2**, respectively. Then the Bayesian premium can be written as*

$$\begin{aligned} E[X_{j,T+1}|\vec{X} = \vec{x}] &= \frac{\sigma_\lambda^2 IT \left(\frac{\sum_{i=1}^I \sum_{t=1}^T x_{i,t}}{IT} \right) + \sigma_x^2 (\mu_\lambda + \mu)}{\sigma_\lambda^2 IT + \sigma_x^2} \\ &= w_1 \bar{\bar{X}} + (1 - w_1)(\mu_\lambda + \mu) \end{aligned} \quad (3.16)$$

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

Proof. The proof for this theorem is similar that for Theorem (3.1). By just substituting normal density for lognormal density in equation (3.2) and (3.3). Thus continuing the proof in the same manner as in Theorem (3.1) one gets the credibility formula (3.16) \square

4 Conclusions

This article proposes a model for claim dependence across insured individuals by using the common effect (in the terminology of Yeo and Valdez, 2006[5]). This model is built within the framework for calculating the Bayesian premium. We are able to derive an explicit form of the Bayesian premium in the case where the common effect follows normal distribution and claim amounts are lognormally distributed. Moreover, the Bayesian premium can be expressed in the credibility form when claim amounts follow normal distribution.

It is worth mentioning that the Bayesian premium with lognormal claims (see, assumption **L1**) needs to choose the value μ_j that depend on individual j . To obtain a suitable value is still an interesting problem needing further investigation. However, we can use claims experience including other factors from their portfolios to justify the appropriacy of this value. Moreover, 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[7]).

In a future, follow-up work towards applications of the theoretical results in this paper, the effect of prior distributions, as well as the sampling models should be addressed, when real data are available. This is a mandate to make sure data fit well the model, and robustness of prior models are checked. For these purposes, the Bayesian model checking procedures outlined in Chapter 6 of the Text by Gelman et al. (2014)[8], could be used

Acknowledgements : The authors wish to acknowledge the referees of the paper for careful reading and several suggestions which improved this paper. This work is partially supported by the Centre of Excellence in Mathematics in Thailand.

References

- [1] H. Gerber, D. Jones, Credibility formula of the updating type, *Transactions of the Society of Actuaries* 27 (1975) 31-52.
- [2] E.W. Free, V.R. Young, Y. Lou, A longitudinal data analysis interpretation of credibility models, *Insurance: Mathematics and Economics* 24 (1999) 229-248.
- [3] S. Wang, Aggregation of correlated risk portfolios, *Proceedings of the Casualty Actuarial Society* 85 (1998) 848-939.
- [4] O. Purcaru, M. Denuit, On the dependence induced by frequency credibility models, *Belgian Actuarial Bulletin* 2 (2002) 73-79.
- [5] K.L. Yeo, E.A. Valdez, Claim dependence with common effects in credibility models, *Insurance: Mathematics and Economics* 38 (2006) 609-629.
- [6] E.A. Valdez, Some less-known but useful results for normal distribution, Working paper, UNSW, Sydney, Australia, 2004.
- [7] L. Wen, W. Xianyi, Z. Xian, The credibility premiums for models with dependence induced by common effects, *Insurance: Mathematics and Economics* 44 (2009) 19-25.
- [8] A. Gelman, J.B. Carlin, H.S. Stern, D.B. Dunson, A. Vehtari, D.B. Rubin, *Bayesian Data Analysis*, 3rd Edn, CRC Press, Boca Raton, 2014.

(Received 1 September 2016)

(Accepted 10 November 2016)