

Claim Dependence with Common Effects in Credibility Models*

Yeo Keng Leong and Emiliano A. Valdez
School of Actuarial Studies
Faculty of Commerce & Economics
University of New South Wales
Sydney, AUSTRALIA

Latest Draft: 18 October 2004[†]

Abstract

Several credibility models found in published literature have largely been single dimensional in the sense that the observable claims are derived from a single individual risk or a single group of homogeneous risks over a period of time. In the case where the additional dimension of observing different individual risks or different groups of risks are allowed for, the assumption of independence across the observable claims is often made. This is a matter of convenience and mathematical tractability, though in general, everyone agrees this may seem unrealistic. As such, dependence must be taken into account when modelling risks for assessing credibility premiums. In this paper, we introduce the notion of modelling claim dependence across individuals and simultaneously across time within individuals using common effects. The resulting model is then 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 in the mean-squared error sense. We express this conditional expectation in the form of a credibility premium. We further give illustrative examples to demonstrate the ideas.

*Keywords: claim dependence, insurance premium, credibility models, and common effects.

[†]The authors wish to acknowledge financial support from the Australian Research Council Discovery Grant DP0345036 and the UNSW Actuarial Foundation of The Institute of Actuaries of Australia. Please do not quote without authors' permission. Corresponding author's email address: yeoster28@pacific.net.sg.

1 Introduction

Consider the problem of pricing an insurance contract. The premium is determined by assessing the observable claims from a portfolio of such homogeneous contracts. Here, we shall denote the observable claims by a random variable $X_{i,t}$ where $i = 1, 2, \dots, I$ denotes the individual risk and $t = 1, 2, \dots, T$ denotes the time period. Clearly, we are assuming I contracts in the portfolio for a total of T time periods. The same time period applies to all individuals although the modelling aspects can easily be adjusted to accommodate differing time periods for different individuals.

In insurance premium determination, it is a common practice to group individual risks so that the risks within each group are as homogeneous as possible in terms of certain observable risk characteristics. A collective premium, also known as the manual premium, is then calculated and charged for this group. The grouping is made primarily to reach a fair and equitable premium across all individuals. It also helps to accomplish isolating a large group of independent and identical risks so that the law of large numbers can be invoked in the claims prediction and therefore, minimizing variability in the claims experience. However, there will always be imprecision in the grouping of the individual risks making the risks within each group not completely homogeneous. A limited number of unobservable traits will always contribute to the possible presence of heterogeneity among the individuals.

Past claims experience provides an invaluable insight into the unobservable characteristics of the individual risks. Certainly in pricing for general (or casualty) insurance products, there is the practice of accounting for the past experience of the insured individual in claims prediction and in premium calculation. Experience rating is the term used in this exercise and is generally viewed to reach a fair and equitable insurance price. For example, in motor insurance, a driver may have had a number of years of experience available to the insurer. Some drivers with little or no claims experience will simply be assessed an additional risk premium, and prediction of its own future claims may largely be based on the claims experience of the group it belongs to.

These considerations naturally point to some sort of a compromise between the two sets of experience in claims prediction: the group's claims experience and the individual's claims experience, if any. This has led actuaries to use a pricing formula of the form

$$\text{Premium} = Z \cdot \text{Own Experience} + (1 - Z) \cdot \text{Group Experience}, \quad (1)$$

where Z , a value between 0 and 1 (inclusive), is well-known to be the "credibility factor". The credibility factor in (1) is a weight assigned to the individual's own claims experience, and most credibility models developed for almost a century now in turn leads to the calculation of this weight. It is to be noted that on one hand, the group's collective experience is extensive enough so that the law of large numbers is applicable and therefore, ignores the presence of heterogeneity. On the other hand, the individual's own experience contains valuable information about the risk characteristics of the individual but may not be complete due to lack of volume. Credibility models must therefore be able to reach an intuitively appealing formula allowing for larger credibility for larger number of years of individual experience, for example.

Such a credibility factor should be close to unity in the case of abundant individual risk experience or where there is a high degree of heterogeneity in the overall experience. It should be close to nil if individual risk experience is lacking or unreliable, or where there is a high degree of homogeneity in overall experience.

The development of credibility models is believed to be first studied by Mowbray (1914) and a few years later, by Whitney (1918) where he suggested using a weighted average between the individual and the collective experience. Bailey (1950) later examined the several ways to develop a credibility formula using the ideas of classical statistics. The study of the numerous ways of deriving or estimating the credibility factor subsequently became known as credibility theory. Today, several textbooks cover this material and many of which provide introduction to students who wish to learn the material for actuarial examinations. Excellent sources include Klugman, et al. (1998), Kaas, et al. (2001) and Mahler and Dean (2001).

Because of tractability, it has been the traditional practice to assume independence of claims in several of these credibility models. Using our notation introduced earlier, claims $X_{1,t}, X_{2,t}, \dots, X_{I,t}$ are assumed to be independent across the individuals for a fixed time t , stating that claims of one insured individual do not directly impact those of other insured individuals. In some situations, this may be unrealistic as for example, in house insurance, geographic proximity of insureds may result in exposure to a common catastrophe, and in motor insurance, accidents may involve several insureds at once in a collision. For a fixed individual $i = 1, 2, \dots, I$, claims $X_{i,1}, X_{i,2}, \dots, X_{i,T}$ are also often assumed independent across different time periods. Again, this may seem unrealistic because for example, in a motor insurance, accident proneness may be present for an individual. Modelling the time dependence is a more common practice when developing credibility models, but not dependence across individuals. The early paper by Gerber and Jones (1975) and the more recent ones by Frees, et al. (1999) are examples of credibility models with time dependence of claims.

In this paper, we address a simultaneous dependence of claims across individuals for a fixed time period and across time periods for a fixed individual. We accomplish this by introducing the notion of a common effect affecting all individuals and another common effect affecting a fixed individual over time. As used in this paper, a random variable Λ has been used to describe the common dependence across the insured individuals, and for a fixed individual i , the random variable Θ_i has been used to describe the common dependence across the time periods. In statistics, such dependence has sometimes been called “common effects”, “latent or unobservable variables”, and the term “frailty” variables is more often used in the biostatistics and survival models literature. See, for example, Vaupel, et al. (1979) and Oakes (1989).

In the actuarial and insurance literature, the notion of claims dependence is increasingly becoming an important part of the modelling process. In Wang (1998), a set of statistical tools for modelling dependencies of risks in an insurance portfolio has been suggested. Using copula framework, Valdez and Mo (2002) and Albrecher and Kantor (2002) have both examined the impact of claim dependencies on the probability of ruin. The works of Heilmann (1986) and Hürlimann (1993) have investigated the effect of dependencies of risks on stop-loss premiums. Several general-

izations and alternative models of dependence have since followed including, Dhaene and Goovaerts (1996, 1997) and Müller (1997), addressing their impact on stop-loss premiums. Other models have included the works of Genest, et al. (1999) and Cossette, et al. (2002) where claim dependence have been addressed in the framework of individual risk models. Furthermore, using the notion of a stochastic order, the recent papers by Purcaru and Denuit (2002, 2003) provide excellent discussion of dependencies in claim frequency for credibility models.

Our paper offers additional insight into the modelling of claim dependencies, and within the framework of developing credibility premiums. We address modelling the possible dependence present across the insured individuals and simultaneously across the time periods by introducing common effects. Nothing similar has been addressed in the actuarial and insurance literature, and we believe this provides a valuable contribution to the already increasing literature on the issue of claim dependence. The motivation for this paper is in contrast to the papers by Purcaru and Denuit (2002, 2003) where they address dependencies only for the frequency of claims. In addition, this paper allows for developing a credibility premium which has the advantage of being expressed as in (1). We compare numerical results of our model with that of the ordinary Bayesian framework commonly used in developing credibility premiums.

The structure for the rest of the paper has been made as follows. First in Section 2, we discuss the construction of the two-level common effects model. We call this two-level because one common effect is to describe, as already alluded in the early paragraphs, the dependency across insured individuals and another common effect to describe the time dependence. In subsequent sections, we discuss the results of a specific model based on the Normal distribution assumptions. In particular, section 3 explores the resulting model when the risks, $X_{i,t}$, $i = 1, 2, \dots, I$ and $t = 1, 2, \dots, T$, and risk parameters, Λ and Θ_i , $i = 1, 2, \dots, I$, are all Normally distributed. The credibility premium formula is derived and some asymptotic properties of the formula are explored providing us intuitive insights into the formulas developed. Section 4 demonstrates the application and usefulness of the model using a set of simulated claim observations. We then compare the results of this application to that of the case of the more well-known Normal credibility model that allows for a single common effect only. We conclude in section 5. The appendix provides some of the detailed proofs and calculations for the credibility premium formula derived in section 3.

2 The Two-Level Common Effects Model of Dependence

Consider a portfolio of insurance contracts consisting of I insured individuals and suppose that each individual has available a history of a total of T time periods. Denote by $X_{i,t}$ the claim amount for individual i during period t . For convenience, we shall sometimes use the random vector

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

to denote the vector of claims for a particular individual $i = 1, 2, \dots, I$. Our primary interest is to predict the next period claim for each individual based on all the observed claims $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_I$. This will be denoted by the random variable $X_{i,T+1}$. It is well-known in statistics that the best predictor of $X_{i,T+1}$, based on all the observed claims $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_I$, in the sense of the “mean squared prediction error” is the conditional expectation

$$E(X_{i,T+1} | \mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_I). \quad (2)$$

In other words, the predictor in (2) is the required functional $g(\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_I)$ that minimizes the following mean squared prediction error:

$$E[X_{i,T+1} - g(\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_I)]^2.$$

See, for example, Shao (2003) for a nice proof on page 40. Thus, our objective is to evaluate this claims prediction based on the model of dependence described in the ensuing paragraphs. In effect, we end up evaluating the distribution of the random variable $X_{i,T+1} | \mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_I$.

As already mentioned in the introduction section, the model of dependence being proposed in this paper will allow for both the dependence among the individual risks as well as the dependence of experience for a particular individual risk over time. The dependence among individual risks will be described by a common effect random variable Λ whose probability function will be assumed to be known and denoted by $f_\Lambda(\lambda)$. Realizations of this common effect is denoted by λ . Conditionally on this common effect, the random vectors \mathbf{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. Thus, we have the following two assumptions, labelled A1 and A2, respectively.

- A1. The random variable Λ has known probability function $f_\Lambda(\lambda)$.
- A2. The random vectors $\mathbf{X}_i | \Lambda$, $i = 1, \dots, I$, where $\mathbf{X}_i = (X_{i,1}, \dots, X_{i,T})'$, are conditionally independent.

Now, for a fixed individual say i , the dependence of claims across time will be described by another common effect random variable denoted by Θ_i whose probability function is assumed known and denoted by $f_{\Theta_i}(\theta_i)$, with θ_i unambiguously denoting its realizations. This assumption we list as A3 below. This random variable could either be a discrete, continuous, or a mixed random variable. Additional assumptions regarding this common effect are listed below.

- A3. For a fixed $i = 1, \dots, I$, the random variable Θ_i has known probability function $f_{\Theta_i}(\theta_i)$.
- A4. The random variables $\Theta_1, \Theta_2, \dots, \Theta_I$ are pairwise independent, that is, Θ_i is independent of Θ_j , $\forall i \neq j$, $i, j = 1, \dots, I$.
- A5. For a fixed $i = 1, \dots, I$, the random variable Θ_i is independent of Λ .

A6. For a fixed $i = 1, \dots, I$ and a fixed $t = 1, \dots, T$, the conditional random variable $X_{i,t}|\Theta_i, \Lambda$ has known probability function denoted by $f_{X_{i,t}|\Theta_i, \Lambda}(x_{i,t}|\theta_i, \lambda)$.

As is usual with common effects models, the assumptions of conditional independence are imposed. Assumption A7 below merely asserts that given the overall risk parameter, all individuals are independent. Finally, assumption A8 is merely stating that given the overall risk parameter and the individual's risk parameter, the individual risk's experience at a particular time period is independent of that of all other individuals as well as the individual risk's experience of other time periods.

A7. The random vectors $\mathbf{X}_i, \Theta_i|\Lambda, i = 1, \dots, I$, are conditionally independent.

A8. The random variables $X_{i,t}|\Theta_i, \Lambda, i = 1, \dots, I$ and $t = 1, \dots, T$, are conditionally independent.

For convenience, we shall sometimes write the random vector $\Theta = (\Theta_1, \Theta_2, \dots, \Theta_I)'$ with realizations denoted by the vector $\theta = (\theta_1, \theta_2, \dots, \theta_I)'$. As in the Bayesian premium setting, it should be noted that in the strictest sense, the distributions of Θ and Λ are not prior distributions but are probability distributions describing some unknown processes. These processes model the level of risk for each individual, $X_{i,t}$, with λ and θ_i as the realizations. Thus, assumption A8 states that given the knowledge of λ and θ_i , each individual risk's experience at any time period is independent of all other experience, including those of the same individual but at a different time period. This is reasonable as knowledge of both λ and θ_i removes the sources of dependence between risks and across time. One can think of Λ as the variable inducing the dependence of claims among the individuals, such as in the case of an epidemic in life insurance, a catastrophe in general insurance, or simply, a bad weather condition in a day where accident collisions become common. On the other hand, one can think of Θ_i as the variable inducing claims for an individual across different time periods, such as in the case of accident proneness of individuals, risk characteristics like lifestyles often unobservable by the insurer or risk characteristics which may not be legally used for pricing discrimination by the insurer.

Assumption A5 states that the overall risk parameter is independent of all individual risk parameters. This appears to be a reasonable assumption and is best explained by a simple example in the case of a motor insurance. A typical driver has certain individual driving characteristics unobservable to the insurer such as propensity to exceed speed limits. As already alluded in the previous paragraph, these factors will be described by Θ_i . This same driver may live in a neighborhood with certain levels of crime rates and traffic congestion contributing to increasing probability of a claim. These factors will in turn be included in Λ as they are experienced by other drivers living in the same region. These two sets of factors logically does not affect each other and hence, it is reasonable to assume independence between Λ and Θ_i .

Assumption A4 states that individual risk parameters are independent of one another. Again, this is reasonable because of the existence of Λ . The random variable Λ already takes into account all the common characteristics across individual risks so that Θ_i are truly peculiar individual characteristics.

It is important to compare these assumptions with that of the ordinary Bayesian model. In the Bayesian model, there is no overall risk parameter Λ and thus, its impact has been implicitly included in each individual risk parameter Θ_i . Yet, it is still assumed that each Θ_i is independent of one another. This apparent contradiction deems the ordinary Bayesian model unsuitable and incomplete for our purposes. It is for this reason that Λ is introduced into the model set-up described in this section.

Proceeding now, the conditional density, $f_{X_{j,T+1}|\mathbf{X}_1, \dots, \mathbf{X}_I}(x_{j,T+1}|\mathbf{x}_1, \dots, \mathbf{x}_I)$ is desired as the conditional expectation in (2) gives our best estimate of next period's claims and also gives our desired premium. For convenience, denote the random vector

$$\mathbf{X} = (\mathbf{X}'_1, \dots, \mathbf{X}'_I)'$$

which gives all the observable claims from all individuals and across all available time periods. The conditional expectation can then be conveniently expressed as

$$E(X_{j,T+1}|\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_I) = E(X_{j,T+1}|\mathbf{X}) = \int x_{j,T+1} \cdot f_{X_{j,T+1}|\mathbf{X}}(x_{j,T+1}|\mathbf{x}) dx_{j,T+1}, \quad (3)$$

where the integral is the Riemann-Stieltjes integral. Consider first the joint density of the individual risks \mathbf{X} , the individual risk parameters Θ , and the overall risk parameter Λ . We give the following result as a lemma.

Lemma 1 *Consider the two-level common effects model satisfying assumptions A1 to A8 described in this section. The joint density of the individual risks \mathbf{X} , the individual risk parameters Θ , and the overall risk parameter Λ can be expressed as*

$$f_{\mathbf{X}, \Theta, \Lambda}(\mathbf{x}, \boldsymbol{\theta}, \lambda) = f_{\Lambda}(\lambda) \times \prod_{i=1}^I f_{\Theta_i}(\theta_i) \times \prod_{i=1}^I f_{\mathbf{X}_i|\Theta_i, \Lambda}(\mathbf{x}_i|\theta_i, \lambda). \quad (4)$$

Proof. First we observe that the conditional density of $\mathbf{X}_1, \dots, \mathbf{X}_I|\Theta, \Lambda$ can be written as

$$f_{\mathbf{X}|\Theta, \Lambda}(\mathbf{x}|\boldsymbol{\theta}, \lambda) = \frac{f_{\mathbf{X}, \Theta, \Lambda}(\mathbf{x}, \boldsymbol{\theta}, \lambda)}{f_{\Theta, \Lambda}(\boldsymbol{\theta}, \lambda)}$$

and applying assumptions A4 and A5, we indeed have

$$f_{\mathbf{X}|\Theta, \Lambda}(\mathbf{x}|\boldsymbol{\theta}, \lambda) = \frac{f_{\mathbf{X}, \Theta, \Lambda}(\mathbf{x}, \boldsymbol{\theta}, \lambda)}{f_{\Lambda}(\lambda) \left[\prod_{i=1}^I f_{\Theta_i}(\theta_i) \right]} = \frac{f_{\mathbf{X}, \Theta|\Lambda}(\mathbf{x}, \boldsymbol{\theta}|\lambda)}{\prod_{i=1}^I f_{\Theta_i}(\theta_i)}.$$

Assumption A7 further leads us to

$$\begin{aligned} f_{\mathbf{X}|\Theta, \Lambda}(\mathbf{x}|\boldsymbol{\theta}, \lambda) &= \frac{\prod_{i=1}^I f_{\mathbf{X}_i, \Theta_i|\Lambda}(\mathbf{x}_i, \theta_i|\lambda)}{\prod_{i=1}^I f_{\Theta_i}(\theta_i)} = \frac{\prod_{i=1}^I [f_{\mathbf{X}_i, \Theta_i|\Lambda}(\mathbf{x}_i, \theta_i|\lambda) f_{\Lambda}(\lambda)]}{[f_{\Lambda}(\lambda)]^I \prod_{i=1}^I f_{\Theta_i}(\theta_i)} \\ &= \prod_{i=1}^I \frac{f_{\mathbf{X}_i, \Theta_i, \Lambda}(\mathbf{x}_i, \theta_i, \lambda)}{f_{\Theta_i, \Lambda}(\theta_i, \lambda)} = \prod_{i=1}^I f_{\mathbf{X}_i|\Theta_i, \Lambda}(\mathbf{x}_i|\theta_i, \lambda). \end{aligned}$$

Next, we observe that the desired joint density can be expressed as

$$\begin{aligned}
f_{\mathbf{X}, \Theta, \Lambda}(\mathbf{x}, \boldsymbol{\theta}, \lambda) &= f_{\Theta, \Lambda}(\boldsymbol{\theta}, \lambda) \cdot f_{\mathbf{X}|\Theta, \Lambda}(\mathbf{x}|\boldsymbol{\theta}, \lambda) \\
&= f_{\Lambda}(\lambda) \cdot f_{\Theta}(\boldsymbol{\theta}) \cdot f_{\mathbf{X}|\Theta, \Lambda}(\mathbf{x}|\boldsymbol{\theta}, \lambda) \\
&= f_{\Lambda}(\lambda) \times \prod_{i=1}^I f_{\Theta_i}(\theta_i) \times f_{\mathbf{X}|\Theta, \Lambda}(\mathbf{x}|\boldsymbol{\theta}, \lambda) \\
&= f_{\Lambda}(\lambda) \times \prod_{i=1}^I f_{\Theta_i}(\theta_i) \times \prod_{i=1}^I f_{\mathbf{X}_i|\Theta_i, \Lambda}(\mathbf{x}_i|\theta_i, \lambda),
\end{aligned}$$

where the second equality follows from assumption A5 and the third equality follows from assumption A4. The last step follows from the result of the previous intermediate result. ■

Using definition of conditional density, we have

$$f_{\Theta, \Lambda|\mathbf{X}}(\boldsymbol{\theta}, \lambda|\mathbf{x}) = \frac{f_{\mathbf{X}, \Theta, \Lambda}(\mathbf{x}, \boldsymbol{\theta}, \lambda)}{f_{\mathbf{X}}(\mathbf{x})},$$

and applying the result in the previous lemma, we have

$$f_{\Theta, \Lambda|\mathbf{X}}(\boldsymbol{\theta}, \lambda|\mathbf{x}) = c \times f_{\Lambda}(\lambda) \times \prod_{i=1}^I f_{\Theta_i}(\theta_i) \times \prod_{i=1}^I f_{\mathbf{X}_i|\Theta_i, \Lambda}(\mathbf{x}_i|\theta_i, \lambda),$$

where c is a normalizing constant and can be expressed as

$$c = [f_{\mathbf{X}}(\mathbf{x})]^{-1} = \left[\int \int \cdots \int f_{\mathbf{X}, \Theta, \Lambda}(\mathbf{x}, \boldsymbol{\theta}, \lambda) \left(\prod_{i=1}^I d\theta_i \right) d\lambda \right]^{-1},$$

where $\prod_{i=1}^I d\theta_i = d\theta_1 d\theta_2 \cdots d\theta_I$. We now state the result for the desired conditional density of $X_{j, T+1}|\mathbf{X}$.

Theorem 1 Consider the two-level common effects model satisfying assumptions A1 to A8 described in this section. For a fixed individual $j = 1, 2, \dots, I$, the conditional density of $X_{j, T+1}|\mathbf{X}$ can be written as

$$\begin{aligned}
&f_{X_{j, T+1}|\mathbf{X}}(x_{j, T+1}|\mathbf{x}) \\
&= \int \int \cdots \int f_{X_{j, T+1}|\Theta_j, \Lambda}(x_{j, T+1}|\theta_j, \lambda) f_{\Theta, \Lambda|\mathbf{X}}(\boldsymbol{\theta}, \lambda|\mathbf{x}) \left(\prod_{i=1}^I d\theta_i \right) d\lambda. \quad (5)
\end{aligned}$$

Proof. Consider a fixed individual j , $1 \leq j \leq I$. It is clear that by definition of conditional density, we have

$$f_{X_{j, T+1}|\mathbf{X}}(x_{j, T+1}|\mathbf{x}) = \frac{f_{X_{j, T+1}, \mathbf{X}}(x_{j, T+1}, \mathbf{x})}{f_{\mathbf{X}}(\mathbf{x})}$$

and notice that the numerator above can be written as

$$\begin{aligned}
& f_{X_{j,T+1}, \mathbf{x}}(x_{j,T+1}, \mathbf{x}) \\
&= \int \int \cdots \int f_{X_{j,T+1}, \mathbf{x} | \Theta, \Lambda}(x_{j,T+1}, \mathbf{x} | \boldsymbol{\theta}, \lambda) f_{\Theta, \Lambda}(\boldsymbol{\theta}, \lambda) d\theta_1 \cdots d\theta_I d\lambda \\
&= \int \int \cdots \int f_{X_{j,T+1} | \Theta_j, \Lambda}(x_{j,T+1} | \theta_j, \lambda) f_{\mathbf{x} | \Theta, \Lambda}(\mathbf{x} | \boldsymbol{\theta}, \lambda) f_{\Theta, \Lambda}(\boldsymbol{\theta}, \lambda) d\theta_1 \cdots d\theta_I d\lambda.
\end{aligned}$$

To see the last step, observe that, from assumption A6 and the previous results, we have

$$\begin{aligned}
f_{X_{j,T+1}, \mathbf{x} | \Theta, \Lambda}(x_{j,T+1}, \mathbf{x} | \boldsymbol{\theta}, \lambda) &= f_{X_{j,T+1}, \mathbf{x}_j | \Theta_j, \Lambda}(x_{j,T+1}, \mathbf{x}_j | \theta_j, \lambda) \prod_{\substack{i=1 \\ i \neq j}}^I f_{\mathbf{x}_i | \Theta, \Lambda}(\mathbf{x}_i | \boldsymbol{\theta}, \lambda) \\
&= f_{X_{j,T+1} | \Theta_j, \Lambda}(x_{j,T+1} | \theta_j, \lambda) \prod_{i=1}^I f_{\mathbf{x}_i | \Theta, \Lambda}(\mathbf{x}_i | \boldsymbol{\theta}, \lambda) \\
&= f_{X_{j,T+1} | \Theta_j, \Lambda}(x_{j,T+1} | \theta_j, \lambda) f_{\mathbf{x} | \Theta, \Lambda}(\mathbf{x} | \boldsymbol{\theta}, \lambda).
\end{aligned}$$

Thus,

$$\begin{aligned}
& f_{X_{j,T+1} | \mathbf{x}}(x_{j,T+1} | \mathbf{x}) \\
&= \frac{1}{f_{\mathbf{x}}(\mathbf{x})} \int \int \cdots \int f_{X_{j,T+1} | \Theta_j, \Lambda}(x_{j,T+1} | \theta_j, \lambda) f_{\mathbf{x} | \Theta, \Lambda}(\mathbf{x} | \boldsymbol{\theta}, \lambda) f_{\Theta, \Lambda}(\boldsymbol{\theta}, \lambda) d\theta_1 \cdots d\theta_I d\lambda \\
&= \int \int \cdots \int f_{X_{j,T+1} | \Theta_j, \Lambda}(x_{j,T+1} | \theta_j, \lambda) \frac{f_{\mathbf{x} | \Theta, \Lambda}(\mathbf{x} | \boldsymbol{\theta}, \lambda) f_{\Theta, \Lambda}(\boldsymbol{\theta}, \lambda)}{f_{\mathbf{x}}(\mathbf{x})} d\theta_1 \cdots d\theta_I d\lambda,
\end{aligned}$$

and the desired result in (5) immediately follows. ■

The purpose of the theorem above is to derive an explicit expression for the conditional density in terms of all the available or given information. First, notice from this theorem that this conditional density involves the multiple integration (or summation in the case of discrete) of the product of the conditional density

$$f_{X_{j,T+1} | \Theta_j, \Lambda}(x_{j,T+1} | \theta_j, \lambda)$$

which according to assumption A6 is known and given, and that of

$$f_{\Theta, \Lambda | \mathbf{x}}(\boldsymbol{\theta}, \lambda | \mathbf{x}) = \frac{f_{\mathbf{x} | \Theta, \Lambda}(\mathbf{x} | \boldsymbol{\theta}, \lambda) f_{\Theta, \Lambda}(\boldsymbol{\theta}, \lambda)}{f_{\mathbf{x}}(\mathbf{x})}$$

for which the numerator can be evaluated using lemma 1 together with the independence of all the common effects. 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.

3 Normal Common Effects

In this and subsequent sub-sections, we derive the credibility premium as expressed in (3) for the case where the common effects follow Normal distributions. As we will notice in the results, we are able to derive interesting, explicit expressions for the predicted claim amount, or for the credibility premium.

3.1 Credibility Premium Formula

Before we consider special cases of the Normal common effects 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:

- N1. The random variables $X_{i,t}|\theta_i, \lambda$ are Normally distributed with mean $(\theta_i + \lambda)$ and common variance σ_x^2 , that is $X_{i,t}|\theta_i, \lambda \sim \text{Normal}(\theta_i + \lambda, \sigma_x^2)$, for $i = 1, 2, \dots, I$;
- N2. The 'individual' common effects θ_i are Normally distributed with common mean μ_θ and common variance σ_θ^2 ; and
- N3. The 'overall' common effects λ are also Normally distributed with mean μ_λ and variance σ_λ^2 .

It follows therefore that we have

$$\begin{aligned} f_{X_{i,t}|\Theta_i,\Lambda}(x_{i,t}|\theta_i, \lambda) &= \frac{1}{\sqrt{2\pi}\sigma_x} e^{-\frac{1}{2}\left[\frac{x_{i,t}-(\theta_i+\lambda)}{\sigma_x}\right]^2}, \\ f_{\Theta_i}(\theta_i) &= \frac{1}{\sqrt{2\pi}\sigma_\theta} e^{-\frac{1}{2}\left(\frac{\theta_i-\mu_\theta}{\sigma_\theta}\right)^2} \text{ and} \\ f_\Lambda(\lambda) &= \frac{1}{\sqrt{2\pi}\sigma_\lambda} e^{-\frac{1}{2}\left(\frac{\lambda-\mu_\lambda}{\sigma_\lambda}\right)^2}. \end{aligned}$$

First we derive the density of $X_{j,T+1}|\mathbf{X}_1, \mathbf{X}_2$ using Theorem 1. To do so, we fix the individual j , so that without loss of generality, we can assume $j = 1$. Similar forms of expressions will follow for all the other individuals. In the appendix, we show that $X_{1,T+1}|\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_I$ is Normally distributed where for convention, we write the mean as $\mu_{1,T+1}$ and the variance as $\sigma_{1,T+1}^2$. In short, we have that

$$X_{1,T+1}|\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_I \sim \text{Normal}(\mu_{1,T+1}, \sigma_{1,T+1}^2).$$

The appendix also shows the explicit expressions for the mean and the variance. In particular, we have

$$E(X_{1,T+1}|\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_I) = \mu_{1,T+1} = w_1\bar{X}_1 + w_{i \neq 1}\bar{\bar{X}}_{i \neq 1} + w_{\theta,\lambda}(\mu_\theta + \mu_\lambda), \quad (6)$$

where $\bar{X}_1 = \frac{1}{T} \sum_t X_{1,t}$ is the observed sample mean of the individual concerned (which is $j = 1$ for ease of exposition) and $\bar{X}_{i \neq 1} = \frac{1}{(I-1)T} \sum_{i \neq 1} \sum_t X_{i,t}$ is the observed sample mean of the rest of the individuals. Thus, we see from (6) that the credibility premium is a weighted average of these observed sample means together with the aggregated means of the common effects. As demonstrated in the appendix, the weights turned out to have the following expressions:

W1. the weight attached to individual's own experience:

$$w_1 = \frac{T [(\sigma_\lambda^2 I + \sigma_\theta^2) \sigma_\theta^2 T + \sigma_x^2 (\sigma_\lambda^2 + \sigma_\theta^2)]}{[(\sigma_\lambda^2 I + \sigma_\theta^2) T + \sigma_x^2] (\sigma_\theta^2 T + \sigma_x^2)},$$

W2. the weight attached to the rest of the group's experience:

$$w_{i \neq 1} = \frac{T (I - 1) \sigma_\lambda^2 \sigma_x^2}{[(\sigma_\lambda^2 I + \sigma_\theta^2) T + \sigma_x^2] (\sigma_\theta^2 T + \sigma_x^2)}; \text{ and}$$

W3. the weight attached to prior beliefs:

$$w_{\theta, \lambda} = \frac{\sigma_x^2 (\sigma_\theta^2 T + \sigma_x^2)}{[(\sigma_\lambda^2 I + \sigma_\theta^2) T + \sigma_x^2] (\sigma_\theta^2 T + \sigma_x^2)}.$$

It is straightforward to show that the credibility premium in (6) is actually a weighted average in the sense that the weights above sum to one:

$$w_1 + w_{i \neq 1} + w_{\theta, \lambda} = 1.$$

Although it has very little value for our purpose, for completeness purposes, we also give the variance of $X_{1,T+1} | \mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_I$. We have, as shown in the appendix, that

$$\begin{aligned} \text{Var}(X_{1,T+1} | \mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_I) &= \sigma_{1,T+1}^2 \\ &= \frac{\sigma_x^2 \left\{ \begin{array}{l} (\sigma_\lambda^2 I + \sigma_\theta^2) \sigma_\theta^2 T^2 \\ + [\sigma_\lambda^2 I (\sigma_\theta^2 + \sigma_x^2) + \sigma_\theta^2 (\sigma_\theta^2 + 2\sigma_x^2)] T \\ + \sigma_x^2 (\sigma_\lambda^2 + \sigma_\theta^2 + \sigma_x^2) \end{array} \right\}}{[(\sigma_\lambda^2 I + \sigma_\theta^2) T + \sigma_x^2] (\sigma_\theta^2 T + \sigma_x^2)}. \end{aligned}$$

Furthermore, it is also interesting to notice that the credibility premium is in fact an unbiased predictor for next period claim. It is straightforward process to demonstrate that

$$\begin{aligned} E(\mu_{1,T+1}) &= E \left[w_1 \bar{X}_1 + w_{i \neq 1} \bar{X}_{i \neq 1} + w_{\theta, \lambda} (\mu_\theta + \mu_\lambda) \right] \\ &= w_1 E(\bar{X}_1) + w_{i \neq 1} E(\bar{X}_{i \neq 1}) + w_{\theta, \lambda} (\mu_\theta + \mu_\lambda) \\ &= \mu_\theta + \mu_\lambda. \end{aligned}$$

Some additional interesting asymptotic properties of the credibility formula are deferred to the next sub-section.

It should be noted also that the credibility premium may be alternatively represented in the form:

$$\mu_{1,T+1} = w_1^* \overline{X}_1 + w_\Sigma \overline{\overline{X}} + w_{\theta,\lambda} (\mu_\theta + \mu_\lambda), \quad (7)$$

where \overline{X}_1 , as before, is the observed sample mean of the individual concerned and $\overline{\overline{X}} = \frac{1}{IT} \sum_i \sum_t X_{i,t}$ is now the observed sample mean of all the individuals in the group, including this one individual $j = 1$. This is the rather conventional form of writing the credibility premium in (7) as a weighted average of the observed sample means for the individuals and the groups, together with the prior means. In this re-expression, the weights turned out to have the following forms:

W1*. the weight attached to individual's own experience:

$$w_1^* = \frac{T [(\sigma_\lambda^2 I + \sigma_\theta^2) \sigma_\theta^2 T + \sigma_x^2 \sigma_\theta^2]}{[(\sigma_\lambda^2 I + \sigma_\theta^2) T + \sigma_x^2] (\sigma_\theta^2 T + \sigma_x^2)},$$

W2*. the weight attached to the group's experience, including the individual of interest:

$$w_\Sigma = \frac{IT \sigma_\lambda^2 \sigma_x^2}{[(\sigma_\lambda^2 I + \sigma_\theta^2) T + \sigma_x^2] (\sigma_\theta^2 T + \sigma_x^2)}; \text{ and}$$

W3*. the weight attached to prior beliefs:

$$w_{\theta,\lambda} = \frac{\sigma_x^2 (\sigma_\theta^2 T + \sigma_x^2)}{[(\sigma_\lambda^2 I + \sigma_\theta^2) T + \sigma_x^2] (\sigma_\theta^2 T + \sigma_x^2)}.$$

In this alternative representation, the credibility premium is also decomposed into three components. These are the individual's own experience, the overall average experience which includes the experience of the individual of interest, and prior beliefs respectively. The difference between the two forms of representation of the credibility premium lies in the second term. In the latter representation, it is the mean of all individuals' experience, $\overline{\overline{X}}$, whereas in the first representation, the individual of interest's experience is excluded when computing this mean, thereby giving only the term $\overline{\overline{X}}_{i \neq 1}$.

3.2 Asymptotic Properties

There are some interesting asymptotic properties that can be derived from the credibility premium formulas developed in the previous sub-section. For purposes of developing these asymptotic properties, we focus on the credibility premium formula in (6) where this premium is the weighted sum of the observed sample means of the individual and the rest of the individuals. Similar interesting observations can be made if one focus on the credibility premium formula in (7).

These asymptotic properties are summarized below:

P1. Lack of past claims experience. If we let $T \rightarrow 0$, that is, past experience is lacking for all individuals, then $w_1 \rightarrow 0$, $w_{i \neq 1} \rightarrow 0$ and $w_{\theta, \lambda} \rightarrow 1$. The fewer experience available for the insurer to assess future claims experience, the more weight it will attach to what it believes (that is, the prior) it should be.

P2. Abundant past individual experience. This is the complete opposite of the previous property. Here, if we let $T \rightarrow \infty$, that is, there is abundance of past experience for all individuals, then we can show that

$$\begin{aligned}
w_1 &= \frac{(\sigma_\lambda^2 I + \sigma_\theta^2) \sigma_\theta^2 T^2 + (\sigma_\lambda^2 + \sigma_\theta^2) \sigma_x^2 T}{(\sigma_\lambda^2 I + \sigma_\theta^2) \sigma_\theta^2 T^2 + (\sigma_\lambda^2 I + 2\sigma_\theta^2) \sigma_x^2 T + \sigma_x^4} \\
&= \frac{(\sigma_\lambda^2 I + \sigma_\theta^2) \sigma_\theta^2 + (\sigma_\lambda^2 + \sigma_\theta^2) \sigma_x^2 \frac{1}{T}}{(\sigma_\lambda^2 I + \sigma_\theta^2) \sigma_\theta^2 + (\sigma_\lambda^2 I + 2\sigma_\theta^2) \sigma_x^2 \frac{1}{T} + \sigma_x^4 \frac{1}{T^2}} \\
&\rightarrow \frac{(\sigma_\lambda^2 I + \sigma_\theta^2) \sigma_\theta^2}{(\sigma_\lambda^2 I + \sigma_\theta^2) \sigma_\theta^2} = 1,
\end{aligned}$$

and that

$$\begin{aligned}
w_{i \neq 1} &= \frac{\sigma_\lambda^2 \sigma_x^2 (I - 1) \frac{1}{T}}{(\sigma_\lambda^2 I + \sigma_\theta^2) \sigma_\theta^2 + (\sigma_\lambda^2 I + 2\sigma_\theta^2) \sigma_x^2 \frac{1}{T} + \sigma_x^4 \frac{1}{T^2}} \\
&\rightarrow \frac{0}{(\sigma_\lambda^2 I + \sigma_\theta^2) \sigma_\theta^2} = 0
\end{aligned}$$

and also that

$$\begin{aligned}
w_{\theta, \lambda} &= \frac{\sigma_x^2 \sigma_\theta^2 \frac{1}{T} + \sigma_x^4 \frac{1}{T^2}}{(\sigma_\lambda^2 I + \sigma_\theta^2) \sigma_\theta^2 + (\sigma_\lambda^2 I + 2\sigma_\theta^2) \sigma_x^2 \frac{1}{T} + \sigma_x^4 \frac{1}{T^2}} \\
&\rightarrow \frac{0}{(\sigma_\lambda^2 I + \sigma_\theta^2) \sigma_\theta^2} = 0.
\end{aligned}$$

This is intuitively appealing as one would expect to attach more weight to individual's own experience as there are more experience available about that individual's claims experience.

P3. Abundant group experience. If we now let $I \rightarrow \infty$, that is, there is abundance of group experience, then we find that

$$\begin{aligned}
w_1 &= \frac{\sigma_\lambda^2 \sigma_\theta^2 T^2 I + \sigma_\theta^4 T^2 + (\sigma_\lambda^2 + \sigma_\theta^2) \sigma_x^2 T}{\sigma_\lambda^2 T I (\sigma_\theta^2 T + \sigma_x^2) + (\sigma_\theta^2 T + \sigma_x^2)^2} \\
&= \frac{\sigma_\lambda^2 \sigma_\theta^2 T^2 + [\sigma_\theta^4 T^2 + (\sigma_\lambda^2 + \sigma_\theta^2) \sigma_x^2 T] \frac{1}{I}}{\sigma_\lambda^2 T (\sigma_\theta^2 T + \sigma_x^2) + (\sigma_\theta^2 T + \sigma_x^2)^2 \frac{1}{I}} \\
&\rightarrow \frac{\sigma_\theta^2 T}{\sigma_\theta^2 T + \sigma_x^2},
\end{aligned}$$

and that

$$\begin{aligned} w_{i \neq 1} &= \frac{\sigma_\lambda^2 \sigma_x^2 T - \sigma_\lambda^2 \sigma_x^2 T \frac{1}{I}}{\sigma_\lambda^2 T (\sigma_\theta^2 T + \sigma_x^2) + (\sigma_\theta^2 T + \sigma_x^2)^2 \frac{1}{I}} \\ &\longrightarrow \frac{\sigma_x^2}{\sigma_\theta^2 T + \sigma_x^2} \end{aligned}$$

and also that

$$w_{\theta, \lambda} = \frac{\sigma_x^2 (\sigma_\theta^2 T + \sigma_x^2) \frac{1}{I}}{\sigma_\lambda^2 T (\sigma_\theta^2 T + \sigma_x^2) + (\sigma_\theta^2 T + \sigma_x^2)^2 \frac{1}{I}} \longrightarrow 0.$$

First observe that there is lesser weight attached to the prior belief or experience. We can then re-write the credibility premium as

$$\frac{\sigma_\theta^2 T}{\sigma_\theta^2 T + \sigma_x^2} \cdot \bar{X}_1 + \frac{\sigma_x^2}{\sigma_\theta^2 T + \sigma_x^2} \cdot \bar{\bar{X}}_{i \neq 1}.$$

Here we observe that the credibility factor has the usual conventional form

$$z = \frac{\sigma_\theta^2 T}{\sigma_\theta^2 T + \sigma_x^2} = \frac{T}{T + (\sigma_x^2 / \sigma_\theta^2)}$$

and no longer depends on the variability of the 'overall' common effects. This overall common effects have been removed because of the abundance of group experience. In addition, if on one hand, we have that $\sigma_\theta^2 \longrightarrow \infty$, that is, 'individual' common effects become large in variation, then we would simply attach full credibility to the individual. That is, $w_1 \longrightarrow 1$ whenever $\sigma_\theta^2 \longrightarrow \infty$ and $I \longrightarrow \infty$. If on the other hand, we have $\sigma_x^2 \longrightarrow \infty$ together with $I \longrightarrow \infty$, that is, the variability of the individual's claims are too large to be able to draw any meaningful prediction, then we would simply attach all the weights to the experience of the rest of the individuals in the group.

P4. Large variation of individual claims. We find that as the variability of individual claims increases, that is, $\sigma_x^2 \longrightarrow \infty$, we have

$$\begin{aligned} w_1 &= \frac{(\sigma_\lambda^2 + \sigma_\theta^2) T \sigma_x^2 + (\sigma_\lambda^2 I + \sigma_\theta^2) \sigma_\theta^2 T^2}{\sigma_x^4 + (\sigma_\lambda^2 I + 2\sigma_\theta^2) T \sigma_x^2 + (\sigma_\lambda^2 I + \sigma_\theta^2) \sigma_\theta^2 T} \\ &= \frac{(\sigma_\lambda^2 + \sigma_\theta^2) T \frac{1}{\sigma_x^2} + (\sigma_\lambda^2 I + \sigma_\theta^2) \sigma_\theta^2 T^2 \frac{1}{\sigma_x^4}}{1 + (\sigma_\lambda^2 I + 2\sigma_\theta^2) T \frac{1}{\sigma_x^2} + (\sigma_\lambda^2 I + \sigma_\theta^2) \sigma_\theta^2 T \frac{1}{\sigma_x^4}} \longrightarrow 0, \end{aligned}$$

and that

$$w_{i \neq 1} = \frac{\sigma_\lambda^2 (I - 1) T \frac{1}{\sigma_x^2}}{1 + (\sigma_\lambda^2 I + 2\sigma_\theta^2) T \frac{1}{\sigma_x^2} + (\sigma_\lambda^2 I + \sigma_\theta^2) \sigma_\theta^2 T \frac{1}{\sigma_x^4}} \longrightarrow 0,$$

and also that

$$w_{\theta,\lambda} = \frac{1 + \sigma_\theta^2 T \frac{1}{\sigma_x^2}}{1 + (\sigma_\lambda^2 I + 2\sigma_\theta^2) T \frac{1}{\sigma_x^2} + (\sigma_\lambda^2 I + \sigma_\theta^2) \sigma_\theta^2 T \frac{1}{\sigma_x^4}} \longrightarrow 1.$$

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 of individual's risk parameters. As $\sigma_\theta^2 \longrightarrow \infty$, that is, a larger variability in individual risk parameters, then we have

$$\begin{aligned} w_1 &= \frac{T^2 \sigma_\theta^4 + (\sigma_\lambda^2 IT^2 + \sigma_x^2 T) \sigma_\theta^2 + \sigma_x^2 \sigma_\lambda^2 T}{T^2 \sigma_\theta^4 + (\sigma_\lambda^2 IT^2 + 2\sigma_x^2 T) \sigma_\theta^2 + \sigma_\lambda^2 \sigma_x^2 IT + \sigma_x^4} \\ &= \frac{T^2 + (\sigma_\lambda^2 IT^2 + \sigma_x^2 T) \frac{1}{\sigma_\theta^2} + \sigma_x^2 \sigma_\lambda^2 T \frac{1}{\sigma_\theta^4}}{T^2 + (\sigma_\lambda^2 IT^2 + 2\sigma_x^2 T) \frac{1}{\sigma_\theta^2} + [\sigma_\lambda^2 \sigma_x^2 IT + \sigma_x^4] \frac{1}{\sigma_\theta^4}} \longrightarrow \frac{T^2}{T^2} = 1, \end{aligned}$$

$$w_{i \neq 1} = \frac{\sigma_\lambda^2 \sigma_x^2 (I - 1) T \frac{1}{\sigma_\theta^4}}{T^2 + (\sigma_\lambda^2 IT^2 + 2\sigma_x^2 T) \frac{1}{\sigma_\theta^2} + [\sigma_\lambda^2 \sigma_x^2 IT + \sigma_x^4] \frac{1}{\sigma_\theta^4}} \longrightarrow \frac{0}{T^2} = 0,$$

and

$$w_{\theta,\lambda} = \frac{\sigma_x^2 T \frac{1}{\sigma_\theta^2} + \sigma_x^4 \frac{1}{\sigma_\theta^4}}{T^2 + (\sigma_\lambda^2 IT^2 + 2\sigma_x^2 T) \frac{1}{\sigma_\theta^2} + [\sigma_\lambda^2 \sigma_x^2 IT + \sigma_x^4] \frac{1}{\sigma_\theta^4}} \longrightarrow \frac{0}{T^2} = 0.$$

All the weights are assigned to the individual's own experience. This appears intuitively as a highly variable θ diminishes the reliability of both $\bar{X}_{i \neq 1}$ and μ_θ as a measure of the extent of the risk due to individual 1.

P6. Large variation in overall risk parameter. As $\sigma_\lambda^2 \longrightarrow \infty$, that is, a larger variability in overall risk parameter, then we have

$$\begin{aligned} w_1 &= \frac{(\sigma_\theta^2 IT^2 + \sigma_x^2 T) \sigma_\lambda^2 + \sigma_\theta^4 T^2 + \sigma_x^2 \sigma_\theta^2}{(\sigma_\theta^2 T + \sigma_x^2) IT \sigma_\lambda^2 + (\sigma_\theta^2 T + \sigma_x^2)^2} \\ &= \frac{\sigma_\theta^2 IT^2 + \sigma_x^2 T + (\sigma_\theta^4 T^2 + \sigma_x^2 \sigma_\theta^2) \frac{1}{\sigma_\lambda^2}}{(\sigma_\theta^2 T + \sigma_x^2) IT + (\sigma_\theta^2 T + \sigma_x^2)^2 \frac{1}{\sigma_\lambda^2}} \longrightarrow \frac{\sigma_\theta^2 IT + \sigma_x^2}{\sigma_\theta^2 IT + \sigma_x^2 I}, \end{aligned}$$

and

$$w_{i \neq 1} = \frac{\sigma_x^2 (I - 1) T}{(\sigma_\theta^2 T + \sigma_x^2) IT + (\sigma_\theta^2 T + \sigma_x^2)^2 \frac{1}{\sigma_\lambda^2}} \longrightarrow \frac{(I - 1) \sigma_x^2}{\sigma_\theta^2 IT + \sigma_x^2 I}$$

and

$$w_{\theta,\lambda} = \frac{\sigma_x^2 (\sigma_\theta^2 T + \sigma_x^2) \frac{1}{\sigma_\lambda^2}}{(\sigma_\theta^2 T + \sigma_x^2) IT + (\sigma_\theta^2 T + \sigma_x^2)^2 \frac{1}{\sigma_\lambda^2}} \longrightarrow \frac{0}{\sigma_\theta^2 IT + \sigma_x^2 I} = 0.$$

- P6a. If additionally, $T \rightarrow \infty$ or $\sigma_\theta^2 \rightarrow \infty$, then $w_1 \rightarrow 1$ and $w_{i \neq 1} \rightarrow 0$; or
P6b. If additionally, $I \rightarrow \infty$, then $w_1 \rightarrow \frac{\sigma_\theta^2 T}{\sigma_\theta^2 T + \sigma_x^2}$ and $w_{i \neq 1} \rightarrow \frac{\sigma_x^2}{\sigma_\theta^2 T + \sigma_x^2}$;
or
P6c. If additionally, $\sigma_x^2 \rightarrow \infty$, then $w_1 \rightarrow \frac{1}{I}$ and $w_{i \neq 1} \rightarrow \frac{I-1}{I}$. For this last scenario, observe that

$$\begin{aligned}
\mu_{1,T+1} &= w_1 \bar{X}_1 + w_{i \neq 1} \bar{\bar{X}}_{i \neq 1} + w_{\theta, \lambda} (\mu_\theta + \mu_\lambda) \\
&\rightarrow \frac{1}{I} \bar{X}_1 + \frac{I-1}{I} \bar{\bar{X}}_{i \neq 1} \\
&= \frac{1}{I} \left(\frac{1}{T} \sum_{t=1}^T X_{1,t} \right) + \frac{I-1}{I} \left(\frac{1}{(I-1)T} \sum_{i=2}^I \sum_{t=1}^T X_{i,t} \right) \\
&= \frac{1}{IT} \left(\sum_{t=1}^T X_{1,t} + \sum_{i=2}^I \sum_{t=1}^T X_{i,t} \right) = \frac{1}{IT} \sum_{i=1}^I \sum_{t=1}^T X_{i,t} = \bar{\bar{X}},
\end{aligned}$$

which is the overall aggregated mean.

Once again, all these additional observations are properties expected of a realistic credibility factor.

3.3 Two Individual Risks with Unit Variance

If in the previous sub-sections we let $I = 2$ and $\sigma_\theta^2 = \sigma_\lambda^2 = \sigma_x^2 = 1$, then we could further simplify and derive some interesting results. So essentially, we are now working with just two individual risks where the common effects, both individual and overall, have unit variances. It is easy to show that the credibility premium, in this case, simplifies to

$$E(X_{1,T+1} | \mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_I) = \mu_{1,T+1} = w_1 \bar{X}_1 + w_2 \bar{X}_2 + w_{\theta, \lambda} (\mu_\theta + \mu_\lambda), \quad (8)$$

where \bar{X}_1 and \bar{X}_2 are respectively, the observed sample means from the first and second individuals. The credibility weights are therefore

$$w_1 = \frac{T(3T+2)}{(3T+1)(T+1)} = 1 - \frac{1}{2(T+1)} - \frac{1}{2(3T+1)},$$

$$w_2 = \frac{T}{(3T+1)(T+1)} = \frac{1}{2(T+1)} - \frac{1}{2(3T+1)},$$

and

$$w_{\theta, \lambda} = \frac{T+1}{(3T+1)(T+1)} = \frac{1}{3T+1}.$$

In addition, the variance can be expressed as

$$\begin{aligned}
Var(X_{1,T+1} | \mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_I) &= \sigma_{1,T+1}^2 = \frac{3T^2 + 7T + 3}{(3T+1)(T+1)} \\
&= 1 + \frac{3}{2(3T+1)} + \frac{1}{2(T+1)}.
\end{aligned}$$

Some intuitive and useful observations can be gleaned from these results. Firstly, observe that the credibility premium is linear in \bar{X}_1 , \bar{X}_2 and $(\mu_\theta + \mu_\lambda)$. The weights, namely w_1 , w_2 and $w_{\theta,\lambda}$, also sum up to unity. It is also noted that when individual experience is abundant, i.e. as $T \rightarrow \infty$, we have $w_1 \rightarrow 1$, $w_2 \rightarrow 0$ and $w_{\theta,\lambda} \rightarrow 0$, which implies increasing reliance on individual experience at the expense of ignoring both the past knowledge of claims experience and experience of the other individual. On the other hand, when individual experience is lacking, i.e. as $T \rightarrow 0$, we have $w_1 \rightarrow 0$, $w_2 \rightarrow 0$ and $w_{\theta,\lambda} \rightarrow 1$, which implies increasing reliance on past knowledge of claims experience and individual experiences are ignored. The latter two observations are properties expected of a realistic credibility factor.

3.4 I Individual Risks with Unit Variance

In the special case where we have $I > 2$ individuals, but still with unit variances, we can show that the credibility premium can be expressed as

$$E(X_{1,T+1} | \mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_I) = \mu_{1,T+1} = w_1 \bar{X}_1 + w_{i \neq 1} \bar{\bar{X}}_{i \neq 1} + w_{\theta,\lambda} (\mu_\theta + \mu_\lambda),$$

where \bar{X}_1 and $\bar{\bar{X}}_{i \neq 1} = \frac{1}{(I-1)T} \sum_{i \neq 1}^I \sum_{t=1}^T X_{i,t}$ are the usual observed sample means. Interestingly, the weights are

$$w_1 = \frac{T[(I+1)T+2]}{[(I+1)T+1](T+1)} = 1 - \frac{I-1}{I(T+1)} - \frac{1}{I[(I+1)T+1]},$$

$$w_{i \neq 1} = \frac{T(I-1)}{[(I+1)T+1](T+1)} = \frac{I-1}{I(T+1)} - \frac{I-1}{I[(I+1)T+1]},$$

and

$$w_{\theta,\lambda} = \frac{T+1}{[(I+1)T+1](T+1)} = \frac{1}{[(I+1)T+1]}.$$

In addition, the variance can be expressed as

$$\begin{aligned} \text{Var}(X_{1,T+1} | \mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_I) &= \sigma_{1,T+1}^2 = \frac{(I+1)T^2 + (2I+3)T + 3}{[(I+1)T+1](T+1)} \\ &= 1 + \frac{I+1}{I[(I+1)T+1]} + \frac{I-1}{I(T+1)}. \end{aligned}$$

Firstly, as before, observe that the credibility premium is linear in \bar{X}_1 , $\bar{\bar{X}}_{i \neq 1}$ and $(\mu_\theta + \mu_\lambda)$. The weights, namely w_1 , $w_{i \neq 1}$ and $w_{\theta,\lambda}$, also sum up to unity. In the presence of abundant individual experience, i.e. as $T \rightarrow \infty$, we have $w_1 \rightarrow 1$, $w_{i \neq 1} \rightarrow 0$ and $w_{\theta,\lambda} \rightarrow 0$, which implies increasing reliance to individual experience at the expense of both the past knowledge of claims experience and experience of the other individuals.

On the other hand, when individual experience is lacking, i.e. as $T \rightarrow 0$, we have $w_1 \rightarrow 0$, $w_2 \rightarrow 0$ and $w_{\theta,\lambda} \rightarrow 1$, which implies increasing reliance on past knowledge of claims experience and individual experience is ignored.

As the number of individual risks increases, i.e. as $I \rightarrow \infty$, we have $w_1 \rightarrow \frac{T}{T+1}$, $w_{i \neq 1} \rightarrow \frac{1}{T+1}$ and $w_{\theta, \lambda} \rightarrow 0$, which implies increasing reliance will be placed on individual experience, especially that of the own individual's, and prior knowledge of claims is ignored.

4 Numerical Example

The two-level common effects model of claim dependence described in the previous sections can be applied in many forms of insurance. Consider, for example, private household or motor insurance. In a household insurance, coverage is provided for a private home structure against fire and theft, as well as other possible catastrophic events such as typhoons or floods. One can view the two levels of common effects as those affected by the ordinary, non-catastrophic events that is usually individualistic in nature, and those affected by presumably the less common, but catastrophic events usually affecting an aggregated portion of the insureds in the portfolio.

As with other types of developing statistical models for other purposes, one prime concern always is the parameter estimates to use for the model. The convention is to examine historical claims experience, to the extent reliable, as a source of knowledge for parameter estimates. In implementing the model suggested in this paper, better parameter estimates can be obtained if the historical claims data can be separated into categories of “catastrophe” and “non-catastrophe”. This categorization information should be readily available for most insurance companies that keeps track of claims, especially the cause of claims, as they come in reported.

For purposes of numerical illustration, we produce some simulated claims data to examine what the effect there might be from assuming some level of dependence between individuals as well as across time, using the two-level common effects framework suggested in this paper, and compare it with the case of the ordinary Bayesian Normal model. The assumption of dependence only across individuals in the Bayesian Normal model, for simplicity and mathematical tractability, is always made without regard as to whether it is violated or not. But here we find that dependence across time can have an effect on the level of credibility premium that may be charged for insurance.

4.1 Assumptions and Simulated Data

This sub-section briefly describes the nature of the assumptions used to draw numerical results. First, we simulated observations assuming the two-level common effects model assumptions hold in reality and then compare the results based on two different models: the two-level common effects model described in this paper, using the Normal distribution assumptions of the common effects, and the conventional Bayesian Normal model which allows for only one level of common effect. We are interested in a numerical comparison of the resulting credibility premiums predicted under each model.

A summary of the specification, description, as well as the parameter values used in the simulation is found in Table 1. In order to allow for a meaningful comparison

between these two models, we have chosen the parameters in the two respective models to be consistent with that of the other. More precisely, for example, we have that $\mu = \mu_\theta + \mu_\lambda$ and $\sigma^2 = \sigma_\theta^2 + \sigma_\lambda^2$. The variances are additive in the Bayesian Normal model because it has been assumed that θ_i and λ are independent of each other.

We generated $n = 1,000$ different 10-year paths of claims for 10 different individuals assuming the two-level common effects model is the true model. Hence, we are saying that in reality, there are two common effects as described in this paper that are inducing the claims. Here we can view these common effects as either catastrophic or non-catastrophic events. These simulations are performed in an Excel spreadsheet where we also stored the observations. Recall that it is assumed that for each individual, the claims amount for each time period conditional on θ_i and λ , i.e. $X_{i,t}|\theta_i, \lambda$, is Normally distributed with mean $\theta_i + \lambda$ and variance σ_x^2 . θ_i and λ are in turn also Normally distributed, with means μ_θ and μ_λ and variances σ_θ^2 and σ_λ^2 respectively.

Thus, these mechanics give the values of θ_i and λ which are simulated first and later stored. From these stored values, the claim amounts, conditional on the common effects, $X_{i,t}|\theta_i, \lambda$ for each i and t , are later simulated and also stored. These simulated paths of claims which are then stored are used to predict the claims that are expected, for each individual i and for the next time period $T + 1$. Figure 1 shows a sample path for each of the ten insured individuals.

Table 1: Summary of Model Assumptions and Parameters used in Simulation

Specification	Description
<i>Model I: Two-Level Normal Common Effects Model</i>	
Conditional density	$X_{i,t} \theta_i, \lambda \sim N(\theta_i + \lambda, \sigma_x^2)$, for $i = 1, 2, \dots, I$ and $t = 1, \dots, T$
'Individual' common effects	$\theta_i \sim N(\mu_\theta, \sigma_\theta^2)$, for $i = 1, 2, \dots, I$
'Overall' common effect	$\lambda \sim N(\mu_\lambda, \sigma_\lambda^2)$
Assumptions	$I = 10$ individuals, $T = 10$ years
Parameter values	$\sigma_x^2 = 6,084$ $\mu_\theta = 100, \sigma_\theta^2 = 1,024$ $\mu_\lambda = 200, \sigma_\lambda^2 = 4,096$
<i>Model II: Bayesian Normal Model</i>	
Conditional density	$X_{i,t} \theta \sim N(\theta, \sigma_x^2)$, for $i = 1, 2, \dots, I$ and $t = 1, \dots, T$
Single common effect	$\theta \sim N(\mu, \sigma^2)$
Assumptions	$I = 10$ individuals, $T = 10$ years
Parameter values	$\sigma_x^2 = 6,084, \mu = 300, \sigma^2 = 5,120$

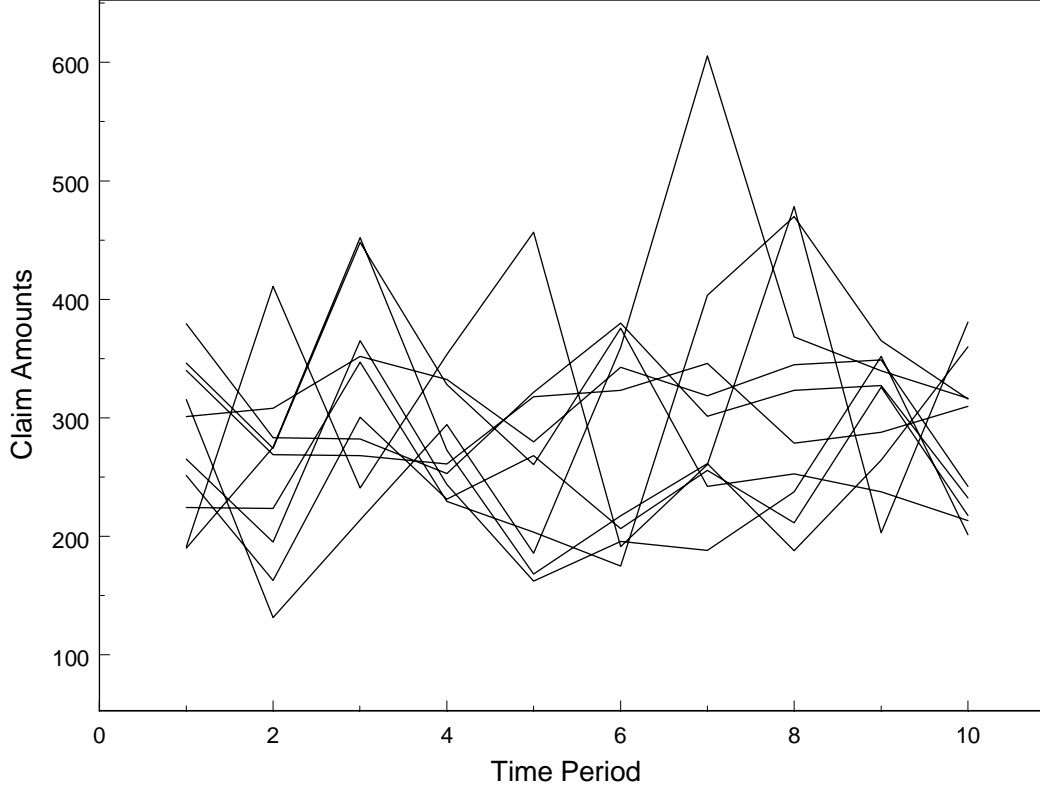


Figure 1: Sample Paths of Claims from 10 Individuals.

4.2 Results and Discussion

We shall call the two-level Normal common effects model our Model I and the Bayesian Normal Model our Model II. For each one of sample paths of claims from the 10 individuals, we computed the credibility premium for year 11 (the next period) for individual 1.

For Model I, we use the formula derived in Section (3) which again we summarize below:

$$E(X_{1,11} | \mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_{10}) = w_1 \bar{X}_1 + w_{i \neq 1} \bar{\bar{X}}_{i \neq 1} + w_{\theta, \lambda} (\mu_\theta + \mu_\lambda), \quad (9)$$

where the weight values used can be computed using the parameter values assumed in the model. Specifically, we have for the credibility weights the following values for Model I:

$$w_1 = \frac{T [(\sigma_\lambda^2 I + \sigma_\theta^2) \sigma_\theta^2 T + \sigma_x^2 (\sigma_\lambda^2 + \sigma_\theta^2)]}{[(\sigma_\lambda^2 I + \sigma_\theta^2) T + \sigma_x^2] (\sigma_\theta^2 T + \sigma_x^2)} = 0.663139$$

for the weight attached to individual 1's own experience,

$$w_{i \neq 1} = \frac{T (I - 1) \sigma_\lambda^2 \sigma_x^2}{[(\sigma_\lambda^2 I + \sigma_\theta^2) T + \sigma_x^2] (\sigma_\theta^2 T + \sigma_x^2)} = 0.322577$$

for the weight attached to the rest of the group's experience and

$$w_{\theta,\lambda} = \frac{\sigma_x^2 (\sigma_\theta^2 T + \sigma_x^2)}{[(\sigma_\lambda^2 I + \sigma_\theta^2) T + \sigma_x^2] (\sigma_\theta^2 T + \sigma_x^2)} = 0.14284$$

for the weight attached to prior beliefs. Table 2 demonstrates a detailed calculation of the credibility premium for individual 1 under Model I for one set of simulations. This table gives the values of the parameter generated for the common effects for each individual. Notice the common value for all insured individuals for the parameter that describes the 'overall' common effects.

Table 2: Detailed Computation of the Credibility Premium for Model I

Individual j	Observed θ_j	Observed λ	$X_{j,1}$	$X_{j,2}$...	$X_{j,10}$
1	120.2717	158.1960	262.1506	131.1447	...	110.6222
2	95.2689	158.1960	262.1677	388.9029	...	149.1350
3	73.5339	158.1960	268.4374	273.5788	...	234.8364
4	117.9636	158.1960	369.7292	249.6484	...	329.5931
5	125.3547	158.1960	325.7415	138.3715	...	253.5196
6	114.6384	158.1960	255.1586	249.5788	...	277.0331
7	101.3216	158.1960	256.3065	250.8617	...	236.5538
8	113.8054	158.1960	101.9994	146.6685	...	128.1672
9	80.5072	158.1960	302.3823	353.8572	...	245.7404
10	118.0959	158.1960	318.2558	203.8160	...	274.6407

The credibility premium for Model II is well-known and the formula, using the notations introduced for this model (see Table 1), is given by

$$E(X_{1,11} | \mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_{10}) = Z \cdot \overline{\overline{X}} + (1 - Z) \cdot \mu, \quad (10)$$

where the credibility factor can be expressed as

$$Z = \frac{T}{T + (\sigma_x^2 / \sigma^2)}.$$

See, for example, Kaas, et al. (2001) and Klugman, et al. (1998) for discussion of these credibility formulas.

As a convention, we shall use $\mu_{j,T+1}^{\text{Model I}}$ to denote the credibility premium in Model I as given in (9) and use $\mu_{j,T+1}^{\text{Model II}}$ to denote that of Model II as given in (10). Now for comparison purposes, we then computed the percentage difference in credibility premiums for individual j between these two models using:

$$\begin{aligned} \Delta_j &= \frac{\text{Model I Premium} - \text{Model II Premium}}{\text{Model II Premium}} \\ &= \frac{\mu_{j,T+1}^{\text{Model I}} - \mu_{j,T+1}^{\text{Model II}}}{\mu_{j,T+1}^{\text{Model II}}}, \end{aligned} \quad (11)$$

For each simulation, we can compute this percentage premium difference Δ_j and examine the resulting distribution of these premium differences for the entire 1,000 simulations. In order not to overwhelm the reader with lots of statistics, we chose to present the results only in terms of Individual 1 and the aggregate of all the ten individuals. The aggregate percentage difference has been computed using:

$$\Delta = \frac{\sum_{j=1}^{10} \mu_{j,T+1}^{\text{Model I}} - \sum_{j=1}^{10} \mu_{j,T+1}^{\text{Model II}}}{\sum_{j=1}^{10} \mu_{j,T+1}^{\text{Model II}}}. \quad (12)$$

Some summary of the resulting percentage differences are given in Table 3 below, together with Figures 2a and 2b that provide histograms of the distribution of the percentage differences. Figure 2a provides the histogram of the percentage difference for the case of individual 1 only, and Figure 2b for the aggregate portfolio (that is, sum of all the individuals). For convenience, the dotted lines in the histograms show the zero reference point. We also show in Figures 3a and 3b the Normal Q-Q plots of these respective distributions to show the skewness, or the non-symmetry, observed from these resulting differences. According to these figures and statistics, the ordinary Bayesian Normal model tends to overstate the credibility premium from its true value, supposedly that of the premium calculated based on the two-level common effects model.

Table 3: Some Descriptive Statistics of the Percentage Difference between the Credibility Premiums in Models I and II

Statistic	Individual 1	Aggregate
Mean	-0.0076298	-0.0048645
Median	-0.0025542	-0.0021824
Variance	0.0081391	0.0061402
Standard Deviation	0.0902168	0.0247795
Minimum	-0.5921914	-0.1584046
Maximum	0.2585304	0.0348486

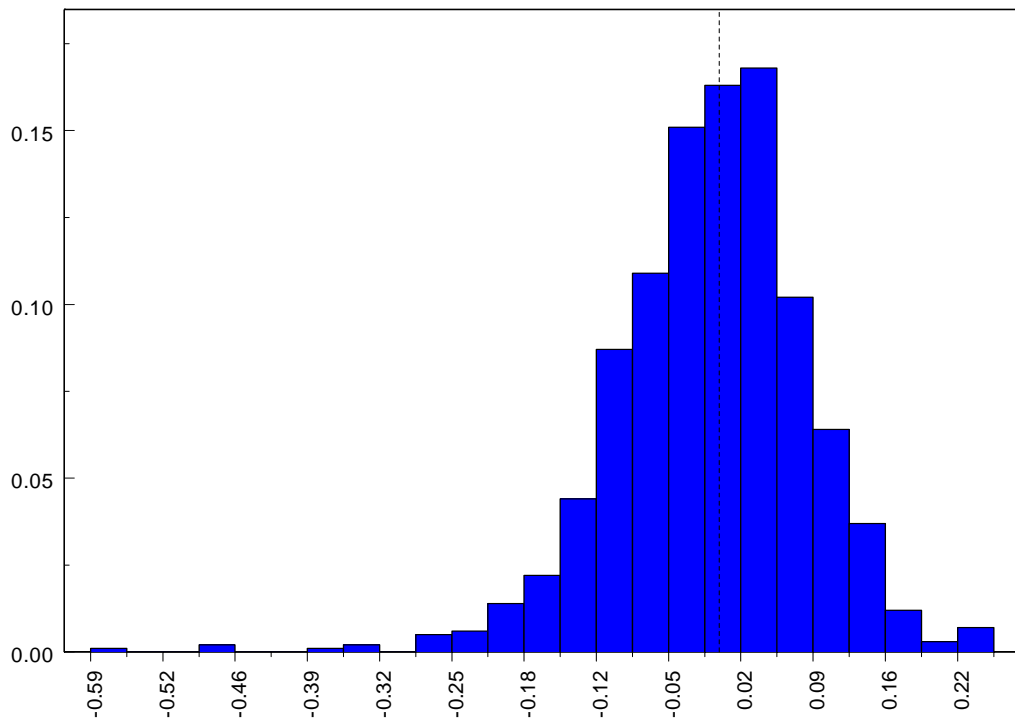


Figure 2a: Histogram of the Percentage Premium Differences for Individual 1

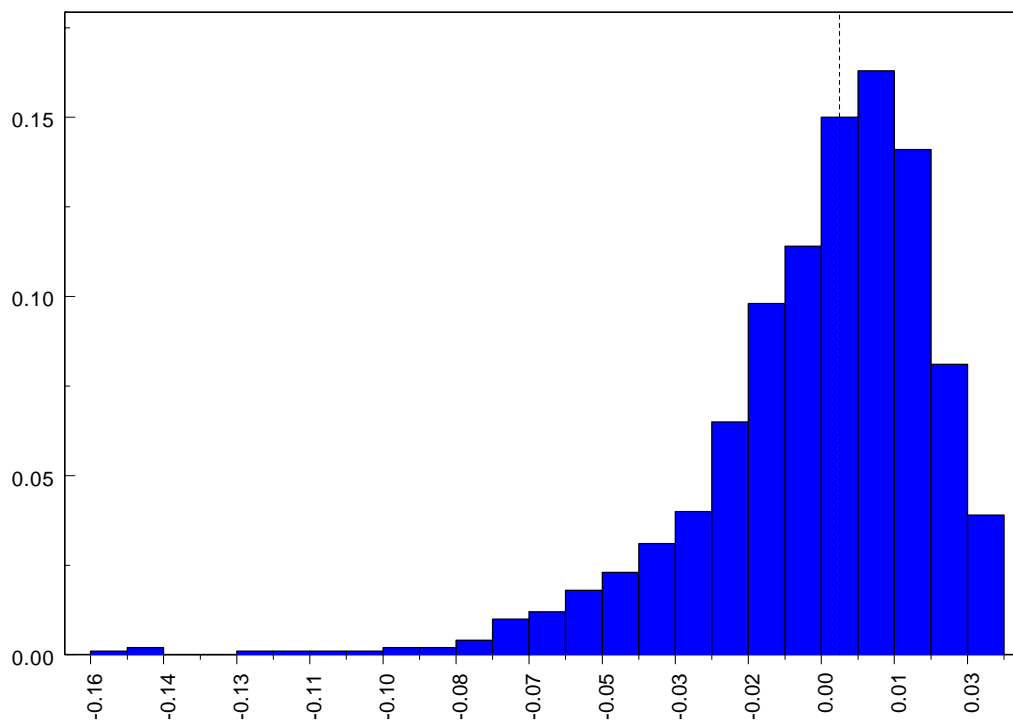


Figure 2b: Histogram of the Percentage Premium Differences for the Aggregate Portfolio

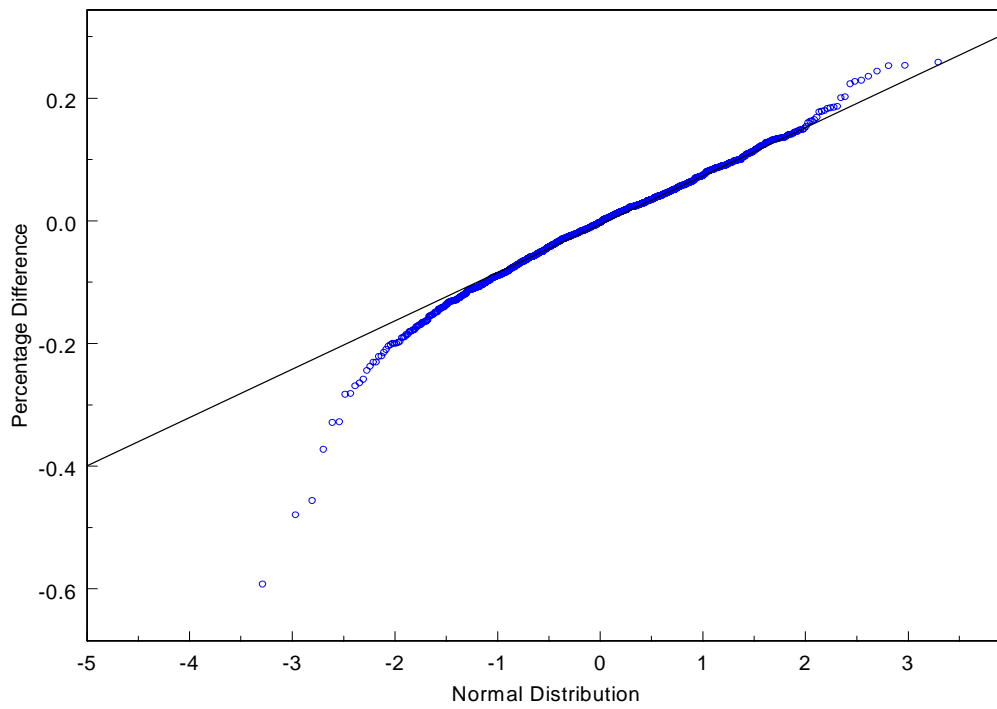


Figure 3a: A Normal Q-Q Plot of these Differences for Individual 1

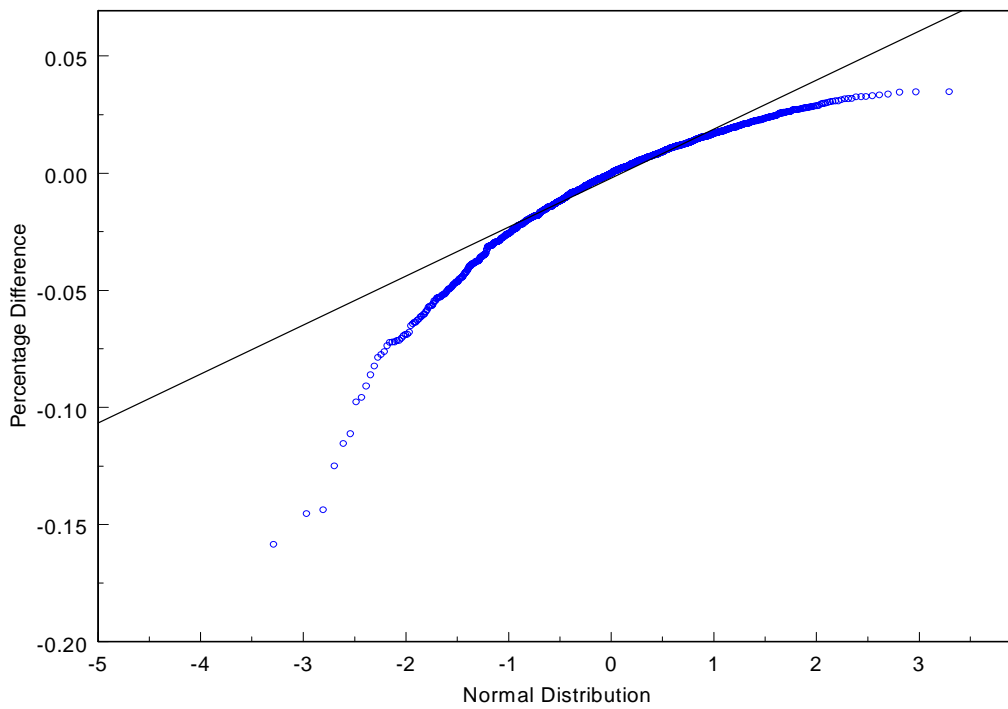


Figure 3b: A Normal Q-Q Plot of these Differences for the Aggregate Portfolio

5 Conclusion

It is a common practice to group individual risks in an insurance portfolio to achieve homogeneity in reaching a fair and equitable premium across the individuals. However, a limited number of observable traits will always contribute to the possible presence of heterogeneity among the individuals. One way to correct for this is to introduce claim dependencies, and it is the purpose of this paper to introduce and offer additional insights into the modelling of claim dependencies. We developed a so-called two-level common effects model within the framework of calculating credibility premiums. These common effects are introduced to account for the possible dependence that may be present across the insured individuals and simultaneously across the time periods. Using this framework, we demonstrate how to compute the required premium for the next period, following a series of observed claims from the individuals in the portfolio for a period of years. We are able to show, as a specific example, that when these common effects follow Normal distributions, that this required premium can be expressed in the form of a credibility-type formula. This formula contain a credibility factor attached to an individual's own experience. We find that the formula also gives weight to the claims experience of the rest of the individuals together with a prior-type claims that account for the common effects. We also find that when asymptotics are considered, we are able to derive explicitly and intuitively appealing results.

A more general framework of modelling claim dependencies could be to use the concept of copulas. Copulas offer the flexibility of modelling dependent random variables, but offer very limited mathematical tractability. This paper provides a framework that can allow, besides the intuitive appeal, mathematical tractability in modelling claim dependencies. However, copula models are still appropriate tools and these can be explored in future work.

Appendix A

From Section 3.1,

$$\begin{aligned} f_{X_{i,t}|\Theta_i,\Lambda}(x_{i,t}|\theta_i, \lambda) &= \frac{1}{\sqrt{2\pi}\sigma_x} e^{-\frac{1}{2}\left[\frac{x_{i,t}-(\theta_i+\lambda)}{\sigma_x}\right]^2}, \\ f_{\Theta_i}(\theta_i) &= \frac{1}{\sqrt{2\pi}\sigma_\theta} e^{-\frac{1}{2}\left(\frac{\theta_i-\mu_\theta}{\sigma_\theta}\right)^2} \text{ and} \\ f_\Lambda(\lambda) &= \frac{1}{\sqrt{2\pi}\sigma_\lambda} e^{-\frac{1}{2}\left(\frac{\lambda-\mu_\lambda}{\sigma_\lambda}\right)^2}. \end{aligned}$$

The objective is to derive the density of $X_{j,T+1}|\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_I$ where without loss of generality, we fix $j = 1$. Thus,

$$\begin{aligned} & f_{X_{1,T+1}|\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_I}(x_{1,T+1}|\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_I) \\ &= \frac{\int \int \cdots \int f_{X_{1,T+1}|\Theta_1,\Lambda}(x_{1,T+1}|\theta_1, \lambda) f_{\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_I, \Theta, \Lambda}(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_I, \boldsymbol{\theta}, \lambda) d\boldsymbol{\theta} d\lambda}{f_{\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_I}(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_I)} \\ &= C_1 \int \int \cdots \int f_{X_{1,T+1}|\Theta_1,\Lambda}(x_{1,T+1}|\theta_1, \lambda) f_{\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_I, \Theta, \Lambda}(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_I, \boldsymbol{\theta}, \lambda) d\boldsymbol{\theta} d\lambda, \end{aligned} \tag{13}$$

where $C_1 = \frac{1}{f_{\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_I}(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_I)}$ is just a normalizing constant and do 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 $f_{X_{1,T+1}|\Theta_1,\Lambda}(x_{1,T+1}|\theta_1, \lambda)$ is already known to be

$$f_{X_{1,T+1}|\Theta_1,\Lambda}(x_{1,T+1}|\theta_1, \lambda) = \frac{1}{\sqrt{2\pi}\sigma_x} e^{-\frac{1}{2}\left[\frac{x_{1,T+1}-(\theta_1+\lambda)}{\sigma_x}\right]^2}. \tag{14}$$

The joint density $f_{\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_I, \Theta, \Lambda}(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_I, \boldsymbol{\theta}, \lambda)$ can be derived by utilizing (4), giving

$$\begin{aligned} & f_{\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_I, \Theta, \Lambda}(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_I, \boldsymbol{\theta}, \lambda) \\ &= \frac{1}{\sqrt{2\pi}\sigma_\lambda} e^{-\frac{1}{2}\left(\frac{\lambda-\mu_\lambda}{\sigma_\lambda}\right)^2} \prod_{i=1}^I \left(\frac{1}{\sqrt{2\pi}\sigma_\theta} e^{-\frac{1}{2}\left(\frac{\theta_i-\mu_\theta}{\sigma_\theta}\right)^2} \left\{ \prod_{t=1}^T \frac{1}{\sqrt{2\pi}\sigma_x} e^{-\frac{1}{2}\left[\frac{x_{i,t}-(\theta_i+\lambda)}{\sigma_x}\right]^2} \right\} \right). \end{aligned} \tag{15}$$

Substituting (14) and (15) into (13), we have

$$\begin{aligned} & f_{X_{1,T+1}|\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_I}(x_{1,T+1}|\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_I) \\ &= \int \int \cdots \int \frac{C_1}{(2\pi)^{\frac{I(T+1)+2}{2}} \sigma_x^{IT+1} \sigma_\theta^I \sigma_\lambda} e^{-\frac{1}{2}\left\{ \sum_{t=1}^{T+1} \left[\frac{x_{1,t}-(\theta_1+\lambda)}{\sigma_x}\right]^2 + \sum_{i=2}^I \sum_{t=1}^T \left[\frac{x_{i,t}-(\theta_i+\lambda)}{\sigma_x}\right]^2 \right\}} \\ & e^{-\frac{1}{2}\left[\sum_{i=1}^I \left(\frac{\theta_i-\mu_\theta}{\sigma_\theta}\right)^2 + \left(\frac{\lambda-\mu_\lambda}{\sigma_\lambda}\right)^2 \right]} d\boldsymbol{\theta} d\lambda. \end{aligned} \tag{16}$$

Considering only the terms containing θ_1 of (16) and after simplifying, we have

$$\begin{aligned}
& \int \frac{1}{2\pi} e^{-\frac{1}{2} \left\{ \sum_{t=1}^{T+1} \left[\frac{x_{1,t} - (\theta_1 + \lambda)}{\sigma_x} \right]^2 + \left(\frac{\theta_1 - \mu_\theta}{\sigma_\theta} \right)^2 \right\}} d\theta_1 \\
&= \int \frac{1}{2\pi} e^{-\frac{T+1}{2\sigma_x^2} \left[\theta_1 - \frac{\sum_{t=1}^{T+1} x_{1,t} - \lambda(T+1)}{T+1} \right]^2} e^{-\frac{1}{2} \left(\frac{\theta_1 - \mu_\theta}{\sigma_\theta} \right)^2} \\
& \quad e^{-\frac{1}{2\sigma_x^2} \left\{ \frac{-\left[\sum_{t=1}^{T+1} x_{1,t} - \lambda(T+1) \right]^2}{T+1} + \sum_{t=1}^{T+1} x_{1,t}^2 + (T+1)\lambda^2 - 2\lambda \sum_{t=1}^{T+1} x_{1,t} \right\}} d\theta_1. \tag{17}
\end{aligned}$$

Notice that part of (17) can be simplified as follows:

$$\begin{aligned}
& \frac{1}{2\pi} e^{-\frac{T+1}{2\sigma_x^2} \left[\theta_1 - \frac{\sum_{t=1}^{T+1} x_{1,t} - \lambda(T+1)}{T+1} \right]^2} e^{-\frac{1}{2} \left(\frac{\theta_1 - \mu_\theta}{\sigma_\theta} \right)^2} \\
&= \frac{\sigma_x \sigma_\theta}{\sqrt{T+1}} \varphi \left\{ \frac{\sqrt{T+1}}{\sigma_x} \left[\theta_1 - \frac{\sum_{t=1}^{T+1} x_{1,t} - \lambda(T+1)}{T+1} \right] \right\} \varphi \left(\frac{\theta_1 - \mu_\theta}{\sigma_\theta} \right),
\end{aligned}$$

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

At this point, a useful result from Valdez (2004) will be used to further simplify (17). This result states that for $\varphi(z)$ and any constants a and b , the following is true:

$$\int_{-\infty}^{\infty} \varphi(z) \varphi(a - bz) dz = \frac{1}{\sqrt{b^2 + 1}} \varphi \left(\sqrt{\frac{a^2}{b^2 + 1}} \right). \tag{18}$$

Thus, by letting $z = \frac{\theta_1 - \mu_\theta}{\sigma_\theta}$ so that $dz = \frac{1}{\sigma_\theta} d\theta_1$, then applying (18) to (17) and simplifying, we have

$$\begin{aligned}
& \int \frac{1}{2\pi} e^{-\frac{1}{2} \left\{ \sum_{t=1}^{T+1} \left[\frac{x_{1,t} - (\theta_1 + \lambda)}{\sigma_x} \right]^2 + \left(\frac{\theta_1 - \mu_\theta}{\sigma_\theta} \right)^2 \right\}} d\theta_1 \\
&= \frac{\sigma_x \sigma_\theta^2}{\sqrt{T+1} \sqrt{\frac{\sigma_\theta^2(T+1)}{\sigma_x^2} + 1}} e^{-\frac{1}{2\sigma_x^2} \left\{ \frac{-\left[\sum_{t=1}^{T+1} x_{1,t} - \lambda(T+1) \right]^2}{T+1} + \sum_{t=1}^{T+1} x_{1,t}^2 + (T+1)\lambda^2 - 2\lambda \sum_{t=1}^{T+1} x_{1,t} \right\}} \\
& \quad \varphi \left\{ \sqrt{\frac{T+1}{\sigma_\theta^2(T+1) + \sigma_x^2}} \left[\frac{\sum_{t=1}^{T+1} x_{1,t} - \lambda(T+1)}{T+1} - \mu_\theta \right] \right\}. \tag{19}
\end{aligned}$$

Considering only the terms containing θ_i , $i = 2, 3, \dots, I$, of (16), applying (18) and simplifying in the same manner as above, we have

$$\begin{aligned}
& \int \frac{1}{2\pi} e^{-\frac{1}{2} \left\{ \sum_{t=1}^T \left[\frac{x_{i,t} - (\theta_i + \lambda)}{\sigma_x} \right]^2 + \left(\frac{\theta_i - \mu_\theta}{\sigma_\theta} \right)^2 \right\}} d\theta_i \\
&= \frac{\sigma_x \sigma_\theta^2}{\sqrt{T} \sqrt{\frac{\sigma_\theta^2 T}{\sigma_x^2} + 1}} e^{-\frac{1}{2\sigma_x^2} \left[\frac{-\left(\sum_{t=1}^T x_{i,t} - \lambda T \right)^2}{T} + \sum_{t=1}^T x_{i,t}^2 + \lambda^2 T - 2\lambda \sum_{t=1}^T x_{i,t} \right]} \\
& \quad \varphi \left[\sqrt{\frac{T}{\sigma_\theta^2 T + \sigma_x^2}} \left(\frac{\sum_{t=1}^T x_{i,t} - \lambda T}{T} - \mu_\theta \right) \right]. \tag{20}
\end{aligned}$$

Continuing from (16) by substituting (19) and (20) and then collecting the terms containing λ , we have

$$\begin{aligned}
& f_{X_{1,T+1}|\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_I}(x_{1,T+1}|\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_I) \\
&= \int C_2 e^{-\frac{1}{2}\left(\frac{\lambda-\mu_\lambda}{\sigma_\lambda}\right)^2} \\
& \quad e^{-\frac{1}{2}\left\{ \begin{aligned} & \frac{-(\sum_{t=1}^{T+1} x_{1,t})^2}{(T+1)\sigma_x^2} + \frac{\sum_{t=1}^{T+1} x_{1,t}^2}{\sigma_x^2} - \frac{\sum_{i=2}^I (\sum_{t=1}^T x_{i,t})^2}{T\sigma_x^2} + \frac{\sum_{i=2}^I \sum_{t=1}^T x_{i,t}^2}{\sigma_x^2} \\ & + \frac{(\sum_{t=1}^{T+1} x_{1,t})^2}{(T+1)[\sigma_\theta^2(T+1)+\sigma_x^2]} + \frac{(T+1)\mu_\theta^2}{\sigma_\theta^2(T+1)+\sigma_x^2} - \frac{2\mu_\theta \sum_{t=1}^{T+1} x_{1,t}}{\sigma_\theta^2(T+1)+\sigma_x^2} \\ & + \frac{\sum_{i=2}^I (\sum_{t=1}^T x_{i,t})^2}{T(\sigma_\theta^2 T + \sigma_x^2)} + \frac{(I-1)T\mu_\theta^2}{\sigma_\theta^2 T + \sigma_x^2} - \frac{2\mu_\theta \sum_{i=2}^I \sum_{t=1}^T x_{i,t}}{\sigma_\theta^2 T + \sigma_x^2} \end{aligned} \right\}} \\
& \quad e^{-\frac{1}{2}\left[\begin{aligned} & \frac{2\lambda \sum_{t=1}^{T+1} x_{1,t}}{\sigma_x^2} - \frac{\lambda^2(T+1)}{\sigma_x^2} + \frac{\lambda^2(T+1)}{\sigma_x^2} - \frac{2\lambda \sum_{t=1}^{T+1} x_{1,t}}{\sigma_x^2} \\ & + \frac{2\lambda \sum_{i=2}^I \sum_{t=1}^T x_{i,t}}{\sigma_x^2} - \frac{\lambda^2(I-1)T}{\sigma_x^2} + \frac{\lambda^2(I-1)T}{\sigma_x^2} - \frac{2\lambda \sum_{i=2}^I \sum_{t=1}^T x_{i,t}}{\sigma_x^2} \\ & - \frac{2\lambda \sum_{t=1}^{T+1} x_{1,t}}{\sigma_\theta^2(T+1)+\sigma_x^2} + \frac{(T+1)\lambda^2}{\sigma_\theta^2(T+1)+\sigma_x^2} + \frac{2\mu_\theta(T+1)\lambda}{\sigma_\theta^2(T+1)+\sigma_x^2} \\ & - \frac{2\lambda \sum_{i=2}^I \sum_{t=1}^T x_{i,t}}{\sigma_\theta^2 T + \sigma_x^2} + \frac{(I-1)T\lambda^2}{\sigma_\theta^2 T + \sigma_x^2} + \frac{2\mu_\theta(I-1)T\lambda}{\sigma_\theta^2 T + \sigma_x^2} \end{aligned} \right]} d\lambda, \quad (21)
\end{aligned}$$

where $C_2 = \frac{C_1 \sigma_\theta^I}{(2\pi)^{\frac{IT+1}{2}} \sigma_x^{I(T-1)+1} \sigma_\lambda T^{\frac{I-1}{2}} (T+1) \left(\frac{\sigma_\theta^2 T}{\sigma_x^2} + 1\right)^{\frac{I-1}{2}} \left(\frac{\sigma_\theta^2(T+1)}{\sigma_x^2} + 1\right)^{\frac{1}{2}}}$.

Now consider only the terms containing λ of (21). Notice that these terms can be simplified as follows:

$$\begin{aligned}
& \frac{1}{2\pi} e^{-\frac{1}{2}\left[-\frac{2\lambda \sum_{t=1}^{T+1} x_{1,t}}{\sigma_\theta^2(T+1)+\sigma_x^2} + \frac{(T+1)\lambda^2}{\sigma_\theta^2(T+1)+\sigma_x^2} + \frac{2\mu_\theta(T+1)\lambda}{\sigma_\theta^2(T+1)+\sigma_x^2} - \frac{2\lambda \sum_{i=2}^I \sum_{t=1}^T x_{i,t}}{\sigma_\theta^2 T + \sigma_x^2} + \frac{(I-1)T\lambda^2}{\sigma_\theta^2 T + \sigma_x^2} + \frac{2\mu_\theta(I-1)T\lambda}{\sigma_\theta^2 T + \sigma_x^2} \right]} \\
& e^{-\frac{1}{2}\left(\frac{\lambda-\mu_\lambda}{\sigma_\lambda}\right)^2} \\
&= \frac{\sigma_\lambda}{\sqrt{\frac{(I-1)T}{\sigma_\theta^2 T + \sigma_x^2} + \frac{T+1}{\sigma_\theta^2(T+1)+\sigma_x^2}}} \\
& \quad \varphi\left(\sqrt{\frac{(I-1)T}{\sigma_\theta^2 T + \sigma_x^2} + \frac{T+1}{\sigma_\theta^2(T+1)+\sigma_x^2}} \left\{ \begin{aligned} & \lambda - \frac{1}{\frac{(I-1)T}{\sigma_\theta^2 T + \sigma_x^2} + \frac{T+1}{\sigma_\theta^2(T+1)+\sigma_x^2}} \\ & \frac{\sum_{t=1}^{T+1} x_{1,t}}{\sigma_\theta^2(T+1)+\sigma_x^2} + \frac{\sum_{i=2}^I \sum_{t=1}^T x_{i,t}}{\sigma_\theta^2 T + \sigma_x^2} \\ & - \left(\frac{(I-1)T}{\sigma_\theta^2 T + \sigma_x^2} + \frac{T+1}{\sigma_\theta^2(T+1)+\sigma_x^2} \right) \mu_\theta \end{aligned} \right\} \right) \\
& \quad \varphi\left(\frac{\lambda-\mu_\lambda}{\sigma_\lambda}\right). \quad (22)
\end{aligned}$$

Thus, by letting $z = \frac{\lambda-\mu_\lambda}{\sigma_\lambda}$ so that $dz = \frac{1}{\sigma_\lambda} d\lambda$, then applying (18) to (22) and simplifying, we have

$$\begin{aligned}
& \int \frac{1}{2\pi} e^{-\frac{1}{2} \left(\frac{\lambda - \mu_\lambda}{\sigma_\lambda} \right)^2} \\
& e^{-\frac{1}{2} \left[-\frac{2\lambda \sum_{t=1}^{T+1} x_{1,t}}{\sigma_\theta^2(T+1) + \sigma_x^2} + \frac{(T+1)\lambda^2}{\sigma_\theta^2(T+1) + \sigma_x^2} + \frac{2\mu_\theta(T+1)\lambda}{\sigma_\theta^2(T+1) + \sigma_x^2} - \frac{2\lambda \sum_{i=2}^I \sum_{t=1}^T x_{i,t}}{\sigma_\theta^2 T + \sigma_x^2} + \frac{(I-1)T\lambda^2}{\sigma_\theta^2 T + \sigma_x^2} + \frac{2\mu_\theta(I-1)T\lambda}{\sigma_\theta^2 T + \sigma_x^2} \right]} d\lambda \\
& = \frac{\sigma_\lambda}{\sqrt{\frac{(I-1)T}{\sigma_\theta^2 T + \sigma_x^2} + \frac{T+1}{\sigma_\theta^2(T+1) + \sigma_x^2}}} \\
& e^{-\frac{1}{2} \left(\frac{-1}{\frac{(I-1)T}{\sigma_\theta^2 T + \sigma_x^2} + \frac{T+1}{\sigma_\theta^2(T+1) + \sigma_x^2}} \left\{ \frac{\sum_{t=1}^{T+1} x_{1,t}}{\sigma_\theta^2(T+1) + \sigma_x^2} + \frac{\sum_{i=2}^I \sum_{t=1}^T x_{i,t}}{\sigma_\theta^2 T + \sigma_x^2} - \left[\frac{T+1}{\sigma_\theta^2(T+1) + \sigma_x^2} + \frac{(I-1)T}{\sigma_\theta^2 T + \sigma_x^2} \right] \mu_\theta \right\}^2 \right)} \\
& \frac{1}{\sqrt{\sigma_\lambda^2 \left[\frac{(I-1)T}{\sigma_\theta^2 T + \sigma_x^2} + \frac{T+1}{\sigma_\theta^2(T+1) + \sigma_x^2} \right] + 1}} \\
& \varphi \left[\left(\frac{1}{\frac{(I-1)T}{\sigma_\theta^2 T + \sigma_x^2} + \frac{T+1}{\sigma_\theta^2(T+1) + \sigma_x^2}} \left\{ \frac{\sum_{t=1}^{T+1} x_{1,t}}{\sigma_\theta^2(T+1) + \sigma_x^2} + \frac{\sum_{i=2}^I \sum_{t=1}^T x_{i,t}}{\sigma_\theta^2 T + \sigma_x^2} \right\} - \mu_\lambda \right) \right. \\
& \left. \sqrt{\frac{\frac{(I-1)T}{\sigma_\theta^2 T + \sigma_x^2} + \frac{T+1}{\sigma_\theta^2(T+1) + \sigma_x^2}}{\sigma_\lambda^2 \left(\frac{(I-1)T}{\sigma_\theta^2 T + \sigma_x^2} + \frac{T+1}{\sigma_\theta^2(T+1) + \sigma_x^2} \right) + 1}} \right] \right]. \tag{23}
\end{aligned}$$

Continuing from (21) and substituting (23), we obtain

$$\begin{aligned}
& f_{X_{1,T+1} | \mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_I} (x_{1,T+1} | \mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_I) \\
& = C_3 e^{-\frac{1}{2} \left(\frac{-\left(\frac{\sum_{t=1}^{T+1} x_{1,t}}{(T+1)\sigma_x^2} + \frac{\sum_{t=1}^{T+1} x_{1,t}^2}{\sigma_x^2} - \frac{\sum_{i=2}^I \left(\frac{\sum_{t=1}^T x_{i,t}}{T\sigma_x^2} \right)^2}{T\sigma_x^2} + \frac{\sum_{i=2}^I \sum_{t=1}^T x_{i,t}^2}{\sigma_x^2} \right)}{\frac{(T+1)\sigma_x^2}{\sigma_\theta^2(T+1) + \sigma_x^2} + \frac{\sum_{t=1}^{T+1} x_{1,t}^2}{\sigma_\theta^2(T+1) + \sigma_x^2} + \frac{(T+1)\mu_\theta^2}{\sigma_\theta^2(T+1) + \sigma_x^2} - \frac{2\mu_\theta \sum_{t=1}^{T+1} x_{1,t}}{\sigma_\theta^2(T+1) + \sigma_x^2} + \frac{\sum_{i=2}^I \left(\frac{\sum_{t=1}^T x_{i,t}}{\sigma_\theta^2 T + \sigma_x^2} \right)^2}{T(\sigma_\theta^2 T + \sigma_x^2)} + \frac{(I-1)T\mu_\theta^2}{\sigma_\theta^2 T + \sigma_x^2} - \frac{2\mu_\theta \sum_{i=2}^I \sum_{t=1}^T x_{i,t}}{\sigma_\theta^2 T + \sigma_x^2} - \frac{1}{\frac{(I-1)T}{\sigma_\theta^2 T + \sigma_x^2} + \frac{T+1}{\sigma_\theta^2(T+1) + \sigma_x^2}} \left\{ \frac{\sum_{t=1}^{T+1} x_{1,t}}{\sigma_\theta^2(T+1) + \sigma_x^2} + \frac{\sum_{i=2}^I \sum_{t=1}^T x_{i,t}}{\sigma_\theta^2 T + \sigma_x^2} \right\}^2 - \left[\frac{T+1}{\sigma_\theta^2(T+1) + \sigma_x^2} + \frac{(I-1)T}{\sigma_\theta^2 T + \sigma_x^2} \right] \mu_\theta} \right)^2} \\
& = C_3 e^{-\frac{1}{2} \left(\frac{\frac{(I-1)T}{\sigma_\theta^2 T + \sigma_x^2} + \frac{T+1}{\sigma_\theta^2(T+1) + \sigma_x^2}}{\sigma_\lambda^2 \left[\frac{(I-1)T}{\sigma_\theta^2 T + \sigma_x^2} + \frac{T+1}{\sigma_\theta^2(T+1) + \sigma_x^2} \right] + 1} \right)} \\
& \left[* \left(\frac{1}{\frac{(I-1)T}{\sigma_\theta^2 T + \sigma_x^2} + \frac{T+1}{\sigma_\theta^2(T+1) + \sigma_x^2}} \left\{ \frac{\sum_{t=1}^{T+1} x_{1,t}}{\sigma_\theta^2(T+1) + \sigma_x^2} + \frac{\sum_{i=2}^I \sum_{t=1}^T x_{i,t}}{\sigma_\theta^2 T + \sigma_x^2} \right\} - \mu_\lambda \right) \right]^2, \tag{24}
\end{aligned}$$

$$\text{where } C_3 = \frac{C_2 \sigma_\lambda}{\sqrt{2\pi} \sqrt{\frac{(I-1)T}{\sigma_\theta^2 T + \sigma_x^2} + \frac{T+1}{\sigma_\theta^2(T+1) + \sigma_x^2}} \sqrt{\sigma_\lambda \left[\frac{(I-1)T}{\sigma_\theta^2 T + \sigma_x^2} + \frac{T+1}{\sigma_\theta^2(T+1) + \sigma_x^2} \right] + 1}}.$$

Extracting the terms containing $x_{1,T+1}$ in (24) and further simplification yields

$$\begin{aligned}
& f_{X_{1,T+1}|\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_I}(x_{1,T+1}|\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_I) \\
&= C_3 e^{-\frac{1}{2} \left(\left\{ \frac{-1}{\sigma_x^2(T+1)} + \frac{1}{(T+1)[\sigma_\theta^2(T+1)+\sigma_x^2]} \right\} \left(x_{1,T+1}^2 + 2x_{1,T+1} \sum_{t=1}^T x_{1,t} \right) \right.} \\
&\quad \left. + \frac{x_{1,T+1}^2}{\sigma_x^2} - \frac{2\mu_\theta x_{1,T+1}}{\sigma_\theta^2(T+1)+\sigma_x^2} \right. \\
&\quad \left. + \left(\frac{\frac{(I-1)T}{\sigma_\theta^2 T + \sigma_x^2} + \frac{T+1}{\sigma_\theta^2(T+1)+\sigma_x^2}}{\frac{(I-1)T}{\sigma_\theta^2 T + \sigma_x^2} + \frac{T+1}{\sigma_\theta^2(T+1)+\sigma_x^2}} \right) \left\{ \sigma_\lambda^2 \left[\frac{(I-1)T}{\sigma_\theta^2 T + \sigma_x^2} + \frac{T+1}{\sigma_\theta^2(T+1)+\sigma_x^2} \right] + 1 \right\} \right. \\
&\quad \left. - \frac{\frac{(I-1)T}{\sigma_\theta^2 T + \sigma_x^2} + \frac{T+1}{\sigma_\theta^2(T+1)+\sigma_x^2}}{\frac{(I-1)T}{\sigma_\theta^2 T + \sigma_x^2} + \frac{T+1}{\sigma_\theta^2(T+1)+\sigma_x^2}} \right. \\
&\quad \left. * \left(\frac{x_{1,T+1}^2 + 2x_{1,T+1} \sum_{t=1}^T x_{1,t}}{[\sigma_\theta^2(T+1)+\sigma_x^2]^2} \right. \right. \\
&\quad \left. \left. + \frac{2x_{1,T+1}}{\sigma_\theta^2(T+1)+\sigma_x^2} \left\{ \frac{\sum_{i=2}^I \sum_{t=1}^T x_{i,t}}{\sigma_\theta^2 T + \sigma_x^2} - \left[\frac{(I-1)T}{\sigma_\theta^2 T + \sigma_x^2} + \frac{T+1}{\sigma_\theta^2(T+1)+\sigma_x^2} \right] \mu_\theta \right\} \right) \right. \\
&\quad \left. - \frac{2\mu_\lambda x_{1,T+1}}{(\sigma_\theta^2(T+1)+\sigma_x^2) \left\{ \sigma_\lambda^2 \left[\frac{(I-1)T}{\sigma_\theta^2 T + \sigma_x^2} + \frac{T+1}{\sigma_\theta^2(T+1)+\sigma_x^2} \right] + 1 \right\}} \right) + K \right), \tag{25}
\end{aligned}$$

where

$$\begin{aligned}
K &= \left\{ \frac{-1}{\sigma_x^2(T+1)} + \frac{1}{(T+1)[\sigma_\theta^2(T+1)+\sigma_x^2]} \right\} \left(\sum_{t=1}^T x_{1,t} \right)^2 + \frac{\sum_{t=1}^T x_{1,t}^2}{\sigma_x^2} \\
&\quad - \frac{2\mu_\theta \sum_{t=1}^T x_{1,t}}{\sigma_\theta^2(T+1)+\sigma_x^2} + \left[\frac{-1}{\sigma_x^2 T} + \frac{1}{T(\sigma_\theta^2 T + \sigma_x^2)} \right] \sum_{i=2}^I \left(\sum_{t=1}^T x_{i,t} \right)^2 \\
&\quad + \frac{\sum_{i=2}^I \sum_{t=1}^T x_{i,t}^2}{\sigma_x^2} - \frac{2\mu_\theta \sum_{i=2}^I \sum_{t=1}^T x_{i,t}}{\sigma_\theta^2 T + \sigma_x^2} + \left[\frac{(I-1)T}{\sigma_\theta^2 T + \sigma_x^2} + \frac{T+1}{\sigma_\theta^2(T+1)+\sigma_x^2} \right] \mu_\theta^2 \\
&\quad + \left(\frac{\frac{(I-1)T}{\sigma_\theta^2 T + \sigma_x^2} + \frac{T+1}{\sigma_\theta^2(T+1)+\sigma_x^2}}{\frac{(I-1)T}{\sigma_\theta^2 T + \sigma_x^2} + \frac{T+1}{\sigma_\theta^2(T+1)+\sigma_x^2}} \right) \left\{ \sigma_\lambda^2 \left[\frac{(I-1)T}{\sigma_\theta^2 T + \sigma_x^2} + \frac{T+1}{\sigma_\theta^2(T+1)+\sigma_x^2} \right] + 1 \right\} \\
&\quad - \frac{\frac{(I-1)T}{\sigma_\theta^2 T + \sigma_x^2} + \frac{T+1}{\sigma_\theta^2(T+1)+\sigma_x^2}}{\frac{(I-1)T}{\sigma_\theta^2 T + \sigma_x^2} + \frac{T+1}{\sigma_\theta^2(T+1)+\sigma_x^2}} \\
&\quad * \left(\frac{(\sum_{t=1}^T x_{1,t})^2}{[\sigma_\theta^2(T+1)+\sigma_x^2]^2} + \left\{ \frac{\sum_{i=2}^I \sum_{t=1}^T x_{i,t}}{\sigma_\theta^2 T + \sigma_x^2} - \left[\frac{(I-1)T}{\sigma_\theta^2 T + \sigma_x^2} + \frac{T+1}{\sigma_\theta^2(T+1)+\sigma_x^2} \right] \mu_\theta \right\}^2 \right) \\
&\quad + \frac{2 \sum_{t=1}^T x_{1,t}}{\sigma_\theta^2(T+1)+\sigma_x^2} \left\{ \frac{\sum_{i=2}^I \sum_{t=1}^T x_{i,t}}{\sigma_\theta^2 T + \sigma_x^2} - \left[\frac{(I-1)T}{\sigma_\theta^2 T + \sigma_x^2} + \frac{T+1}{\sigma_\theta^2(T+1)+\sigma_x^2} \right] \mu_\theta \right\} \\
&\quad + \frac{\frac{(I-1)T}{\sigma_\theta^2 T + \sigma_x^2} + \frac{T+1}{\sigma_\theta^2(T+1)+\sigma_x^2}}{\sigma_\lambda^2 \left[\frac{(I-1)T}{\sigma_\theta^2 T + \sigma_x^2} + \frac{T+1}{\sigma_\theta^2(T+1)+\sigma_x^2} \right] + 1} \left(\mu_\lambda^2 - \frac{2\mu_\lambda}{\sigma_\theta^2 T + \sigma_x^2 + \frac{T+1}{\sigma_\theta^2(T+1)+\sigma_x^2}} \right. \\
&\quad \left. * \left\{ \frac{\sum_{t=1}^T x_{1,t}}{\sigma_\theta^2(T+1)+\sigma_x^2} + \frac{\sum_{i=2}^I \sum_{t=1}^T x_{i,t}}{\sigma_\theta^2 T + \sigma_x^2} - \left[\frac{(I-1)T}{\sigma_\theta^2 T + \sigma_x^2} + \frac{T+1}{\sigma_\theta^2(T+1)+\sigma_x^2} \right] \mu_\theta \right\} \right).
\end{aligned}$$

Finally, from (25), we group the terms containing $x_{1,T+1}^2$ and $x_{1,T+1}$, giving

$$\begin{aligned}
& f_{X_{1,T+1}|\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_I} (x_{1,T+1}|\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_I) \\
&= C_3 e^{-\frac{1}{2}} \left\{ \begin{aligned} & \left[\begin{aligned} & \frac{-1}{\sigma_x^2(T+1)} + \frac{1}{(T+1)[\sigma_\theta^2(T+1)+\sigma_x^2]} + \frac{1}{\sigma_x^2} \\ & + \left(\frac{1}{\left[\frac{(I-1)T}{\sigma_\theta^2 T + \sigma_x^2} + \frac{T+1}{\sigma_\theta^2(T+1)+\sigma_x^2} \right] \left\{ \sigma_\lambda^2 \left[\frac{(I-1)T}{\sigma_\theta^2 T + \sigma_x^2} + \frac{T+1}{\sigma_\theta^2(T+1)+\sigma_x^2} \right] + 1 \right\}} \right. \\ & \quad \left. - \frac{(I-1)T}{\sigma_\theta^2 T + \sigma_x^2} + \frac{T+1}{\sigma_\theta^2(T+1)+\sigma_x^2} \right. \\ & \quad \left. * \frac{1}{[\sigma_\theta^2(T+1)+\sigma_x^2]^2} \right) \end{aligned} \right] x_{1,T+1}^2 \\ & - 2 \left[\begin{aligned} & \left\{ \frac{-1}{\sigma_x^2(T+1)} + \frac{1}{(T+1)[\sigma_\theta^2(T+1)+\sigma_x^2]} \right\} \sum_{t=1}^T x_{1,t} + \frac{\mu_\theta}{\sigma_\theta^2(T+1)+\sigma_x^2} \\ & - \left(\frac{1}{\left[\frac{(I-1)T}{\sigma_\theta^2 T + \sigma_x^2} + \frac{T+1}{\sigma_\theta^2(T+1)+\sigma_x^2} \right] \left\{ \sigma_\lambda^2 \left[\frac{(I-1)T}{\sigma_\theta^2 T + \sigma_x^2} + \frac{T+1}{\sigma_\theta^2(T+1)+\sigma_x^2} \right] + 1 \right\}} \right. \\ & \quad \left. - \frac{(I-1)T}{\sigma_\theta^2 T + \sigma_x^2} + \frac{T+1}{\sigma_\theta^2(T+1)+\sigma_x^2} \right. \\ & \quad \left. * \left(\frac{\sum_{t=1}^T x_{1,t}}{[\sigma_\theta^2(T+1)+\sigma_x^2]^2} + \frac{\sum_{i=2}^I \sum_{t=1}^T x_{i,t}}{\sigma_\theta^2 T + \sigma_x^2} \right) \right. \\ & \quad \left. * \left(\frac{1}{\sigma_\theta^2(T+1)+\sigma_x^2} \left\{ - \left[\frac{(I-1)T}{\sigma_\theta^2 T + \sigma_x^2} + \frac{T+1}{\sigma_\theta^2(T+1)+\sigma_x^2} \right] \mu_\theta \right\} \right) \right) \end{aligned} \right] x_{1,T+1} \\ & + \frac{\mu_\lambda}{[\sigma_\theta^2(T+1)+\sigma_x^2] \left\{ \sigma_\lambda^2 \left[\frac{(I-1)T}{\sigma_\theta^2 T + \sigma_x^2} + \frac{T+1}{\sigma_\theta^2(T+1)+\sigma_x^2} \right] + 1 \right\}} + K \end{aligned} \right\}. \tag{26}
\end{aligned}$$

From (26), we can simplify the coefficients of $x_{1,T+1}^2$ and $x_{1,T+1}$ and write them as A and B respectively where

$$A = \frac{(\sigma_\lambda^2 I T + \sigma_\theta^2 T + \sigma_x^2)(\sigma_\theta^2 T + \sigma_x^2)}{C_4 \sigma_x^2}$$

and

$$\begin{aligned}
B &= \frac{[(\sigma_\lambda^2 I + \sigma_\theta^2) \sigma_\theta^2 T + \sigma_x^2 (\sigma_\lambda^2 + \sigma_\theta^2)] \sum_{t=1}^T x_{1,t}}{C_4 \sigma_x^2} \\ &+ \frac{\sigma_\lambda^2 \sigma_x^2 \sum_{i=2}^I \sum_{t=1}^T x_{i,t}}{C_4 \sigma_x^2} + \frac{\sigma_x^2 (\sigma_\theta^2 T + \sigma_x^2) (\mu_\theta + \mu_\lambda)}{C_4 \sigma_x^2}
\end{aligned}$$

with $C_4 = (\sigma_\lambda^2 I + \sigma_\theta^2) \sigma_\theta^2 T^2 + [\sigma_\lambda^2 I (\sigma_\theta^2 + \sigma_x^2) + \sigma_\theta^2 (2\sigma_\theta^2 + \sigma_x^2)] T + \sigma_x^2 (\sigma_\lambda^2 + \sigma_\theta^2 + \sigma_x^2)$.

By performing a completing the squares operation, and the density of $X_{1,T+1}|\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_I$ can therefore be simplified to

$$\begin{aligned}
& f_{X_{1,T+1}|\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_I} (x_{1,T+1}|\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_I) \\
&= C_3 e^{-\frac{K}{2}} e^{-\frac{A}{2} \left[\left(x_{1,T+1} - \frac{B}{A} \right)^2 - \left(\frac{B}{A} \right)^2 \right]} \\
&= C_3 e^{-\frac{1}{2} \left(K - \frac{B^2}{A} \right)} e^{-\frac{1}{2} \left[\frac{\left(x_{1,T+1} - \frac{B}{A} \right)^2}{\frac{1}{A}} \right]}.
\end{aligned}$$

Observe that $e^{-\frac{1}{2} \left[\frac{(x_{1,T+1} - \frac{B}{A})^2}{\frac{1}{A}} \right]}$ is the kernel of a normal distribution. Therefore, it can be concluded that $X_{1,T+1} | \mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_I \sim \text{Normal}(\mu_{1,T+1}, \sigma_{1,T+1}^2)$, where $\mu_{1,T+1} = \frac{B}{A}$ and $\sigma_{1,T+1}^2 = \frac{1}{A}$. Further simplification gives

$$\mu_{1,T+1} = w_1 \overline{X}_1 + w_{i \neq 1} \overline{X}_{i \neq 1} + w_{\theta, \lambda} (\mu_\theta + \mu_\lambda),$$

with

$$\begin{aligned} w_1 &= \frac{[(\sigma_\lambda^2 I + \sigma_\theta^2) \sigma_\theta^2 T + \sigma_x^2 (\sigma_\lambda^2 + \sigma_\theta^2)] T}{(\sigma_\lambda^2 I T + \sigma_\theta^2 T + \sigma_x^2) (\sigma_\theta^2 T + \sigma_x^2)}, \\ w_2 &= \frac{\sigma_\lambda^2 \sigma_x^2 (I - 1) T}{(\sigma_\lambda^2 I T + \sigma_\theta^2 T + \sigma_x^2) (\sigma_\theta^2 T + \sigma_x^2)}, \\ w_{\theta, \lambda} &= \frac{\sigma_x^2 (\sigma_\theta^2 T + \sigma_x^2)}{(\sigma_\lambda^2 I T + \sigma_\theta^2 T + \sigma_x^2) (\sigma_\theta^2 T + \sigma_x^2)}, \end{aligned}$$

and

$$\sigma_{1,T+1}^2 = \frac{C_4 \sigma_x^2}{(\sigma_\lambda^2 I T + \sigma_\theta^2 T + \sigma_x^2) (\sigma_\theta^2 T + \sigma_x^2)}.$$

References

- [1] Albers, W., 1999. Stop-loss Premiums under Dependence. *Insurance: Mathematics & Economics*, **24**, 173-85.
- [2] Albrecher, H., J. Kantor, 2002. Simulation of Ruin Probabilities for Risk Processes of Markovian Type. *Monte Carlo Methods & Application*, **8**, 111-127.
- [3] Bailey, A.L., 1950. Credibility Procedures - La Place's Generalization of Bayes' Rule and the Combination of Collateral Knowledge with Observed Data. *Proceedings of the Casualty Actuarial Society*, **37**, 7-23.
- [4] Bühlmann, H., 1967. Experience Rating and Credibility. *ASTIN Bulletin*, **4**, 199-207.
- [5] Bühlmann, H., E. Straub, 1970. Glaubwürdigkeit für Schadensätze (Credibility for Loss Ratios). *Mitteilungen der Vereinigung Schweizerischer*, **70**, 111-33.
- [6] Cossette, H., P. Gaillardetz, E. Marceau, J. Rioux, 2002. On Two Dependent Individual Risk Models. *Insurance: Mathematics & Economics*, **30**, 153-66.
- [7] Dhaene, J., M.J. Goovaerts, 1996. Dependency of Risks and Stop-Loss Order. *ASTIN Bulletin* **26**, 201-212.
- [8] Dhaene, J., M.J. Goovaerts, 1997. On the Dependency of Risks in the Individual Life Model. *Insurance: Mathematics & Economics* **19**, 243-253.
- [9] De Pril, N., L. D'Hooge, M. J. Goovaerts, 1976. A Bibliography on Credibility Theory and its Applications. *Journal of Computational and Applied Mathematics*, **2**, 55-62.
- [10] Frees, E.W., V.R. Young, Y. Luo, 1999. A Longitudinal Data Analysis Interpretation of Credibility Models. *Insurance: Mathematics & Economics* **24**, 229-247.
- [11] Genest, C., É Marceau, M. Mesfioui, 2003. Compound Poisson approximations for individual models with dependent risks. *Insurance: Mathematics and Economics*, **32**, 73-91.
- [12] Gerber, H., D. Jones, 1975. Credibility Formulas of the Updating Type. *Transactions of the Society of Actuaries* **27**, 31-52.
- [13] Hachemeister, C.A., 1975. Credibility for Regression Models with Application to Trend. In *Credibility: Theory and Applications* (P. Kahn ed.), Academic Press, New York.
- [14] Heilmann, W.-R., 1986. On the Impact of Independence of Risks on Stop-Loss Transforms. *Insurance: Mathematics and Economics*, **5**, 197-99.
- [15] Hürlimann, W., 1993. Bivariate Distributions with Diatomic Conditionals and Stop-Loss Transforms of Random Sums. *Statistics and Probability Letters*, **17**, 329-35.

- [16] Kaas, R., M.J. Goovaerts, J. Dhaene, M. Denuit, 2001. *Modern Actuarial Risk Theory*. Kluwer Academic Publishers, Dordrecht.
- [17] Klugman, S.A., 1992. *Bayesian Statistics in Actuarial Science with Emphasis on Credibility*. Kluwer Academic Publishers, Boston/Dordrecht/London.
- [18] Klugman, S.A., H.H. Panjer, G.E. Willmot, 1998. *Loss Models: From Data to Decisions*. John Wiley & Sons, New York.
- [19] Longley-Cook, L. H., 1962. An Introduction to Credibility Theory. *Proceedings of the Casualty Actuarial Society*, **49**, 194-221.
- [20] Mahler, H. C., Dean, C. G., 2001. "Credibility," *Foundations of Casualty Actuarial Science* (4th Ed.), Casualty Actuarial Society, Chapter 8.
- [21] Mowbray, A.H., 1914. How Extensive a Payroll Exposure is Necessary to Give Dependable Pure Premiums. *Proceedings of the Casualty Actuarial Society*, **1**, 24-30.
- [22] Müller, A., 1997. Stop-Loss Order for Portfolios of Dependent Risks. *Insurance: Mathematics and Economics*, **21**, 219-23.
- [23] Norberg, R., 1979. The Credibility Approach to Experience Rating. *Scandinavian Actuarial Journal*, 181-221.
- [24] Oakes, D., 1989. Bivariate Survival Models Induced by Frailties. *Journal of the American Statistical Association*, **84**, 487-93.
- [25] Philbrick, S.W., 1981. An Examination of Credibility Concepts. *Proceedings of the Casualty Actuarial Society*, **68**, 195-219.
- [26] Purcaru, O., M. Denuit, 2002. On the dependence induced by frequency credibility models. *Belgian Actuarial Bulletin*, **2**, 73-79.
- [27] Purcaru, O., M. Denuit, 2003. Dependence in Dynamic Claim Frequency Credibility Models. *ASTIN Bulletin*, **33**, 23-40.
- [28] Shao, J., 2003. *Mathematical Statistics*. Springer-Verlag, New York.
- [29] Sundt, B., 1986. Special Issue on Credibility Theory. *Insurance: Abstracts and Reviews*, **2**.
- [30] Taylor, G.C., 1979. Credibility Analysis of a General Hierarchical Model. *Scandinavian Actuarial Journal*, 1-12.
- [31] Valdez, E.A., 2004. Some Less-Known but Useful Results for Normal Distribution. Working paper, UNSW, Sydney, Australia.
- [32] Valdez, E.A., K. Mo, 2002. Ruin Probabilities with Dependent Claims. Working paper, UNSW, Sydney, Australia.

- [33] Vaupel, J.W., K.G. Manton, E. Stallard, 1979. The Impact of Heterogeneity in Individual Frailty and the Dynamics of Mortality. *Demography*, **16**, 439-54.
- [34] Wang, S., 1998. Aggregation of Correlated Risk Portfolios. *Proceedings of the Casualty Actuarial Society*, **85**, 848-939.
- [35] Whitney, A.W., 1918. The Theory of Experience Rating. *Proceedings of the Casualty Actuarial Society*, **4**, 274-92.