One-year premium risk and emergence pattern of ultimate loss based on conditional distribution *

Łukasz Delong[1], Marcin Szatkowski[1,2]

[1]SGH Warsaw School of Economics, Institute of Econometrics

Niepodległości 162, Warsaw 02-554, Poland,

lukasz.delong@sgh.waw.pl

[2]STU ERGO Hestia SA, Risk Department

Hestii 1, Sopot 81-731, Poland

marcin.szatkowski@ergohestia.pl

Abstract: We study the relation between one-year premium risk and ultimate premium risk. In practice, the one-year risk is sometimes related to the ultimate risk by using a so-called emergence pattern formula which postulates a linear relation between both risks. We define the true emergence pattern of the ultimate loss for the one-year premium risk based on a conditional distribution of the ultimate loss derived from a multivariate distribution of the claims development process. We investigate three models commonly used in claims reserving and prove that the true emergence pattern formulas are different from the linear emergence pattern formula used in practice. We show that the one-year risk, when measured by VaR, can be under and overestimated if the linear emergence pattern formula is applied. We present two modifications of the linear emergence pattern formula. These modifications allow us to go beyond the claims development models investigated in the first part and work with an arbitrary distribution of the ultimate loss.

Keywords: One-year risk, ultimate risk, premium risk, emergence pattern, Solvency II, claims development process.

1 Introduction

Insurance companies are exposed to *premium and reserve risk* and they quantify so-called *one-year and ultimate risk*. By one-year premium risk we understand the risk that the premiums earned in a given year are not sufficient to cover the payments made in the year and the reserve set at the end of the year for the losses incurred in this accident year (the one-year loss). By ultimate premium risk we understand the risk that the premiums earned in a given year are not sufficient to cover all the payments made in infinite time horizon for the losses incurred in this accident year (the ultimate loss). One-year and ultimate reserve risks are defined analogously, but the reserve risk is related to the adequacy of the current volumes of the claims reserves (the reserve risk concerns payments for the losses incurred in the past accident years). One-year risk is investigated for Solvency II risk capital, whereas ultimate risk is investigated for pricing and business strategies. The natural question is what is a relation between the one-year risk and the ultimate risk and how to model this relation. We try to answer this question in this paper. We focus on the premium risk, i.e. on losses from a new, single accident year.

We work in a discrete time setting. Let $X = (X_0, X_1, X_2, ..., X_n)$ denote a stochastic process on a probability space $(\Omega, \mathbb{F}, \mathbb{P})$ with a filtration $\mathbb{F} = (\mathcal{F}_k)_{k=0,1,...,n}$. We set $X_0 = 0$. We assume that $\mathbb{F}$ is the natural filtration generated by X. Consequently, the process X is $\mathbb{F}$-adapted. We use X_0 to introduce an explicit notation for the information at time 0, which we later need to correctly define the true emergence pattern as a conditional distribution. The random variables $(X_1, X_2, ..., X_n)$ describe the claims development process for an accident year: X_1 denotes the claims paid in the first development year and X_n denotes the aggregate value of all claims paid. The ultimate premium risk is modelled with X_n. The one-year premium risk is modelled with $BE_1 = \mathbb{E}[X_n|\mathcal{F}_1]$. The conditional expectation $\mathbb{E}[X_n|\mathcal{F}_1]$ is an $\mathcal{F}_1$-measurable random variable. The expected value $BE_1 = \mathbb{E}[X_n|\mathcal{F}_1] = \mathbb{E}[X_n|X_1]$ is called the best estimate of the ultimate loss after the first year. For simplicity, we neglect discounting in the definition of the best estimate of the ultimate loss. We can see that the relation between the one-year premium risk and the ultimate premium risk depends on the relation between X_1 and X_n. The relation between X_1 and X_n is usually specified in a claims reserving model which specifies the claims development process $(X_1, ..., X_n)$. In claims reserving models, see e.g. Radtke et al. (2016), Hertig (1985), Wüthrich, Merz (2008), we choose the distribution X_1 and the conditional distributions $(X_{i+1}|X_i, ..., X_1)_{i=1}^{n-1}$. We can model the one-year risk and the ultimate risk (BE_1, X_n) by generating them in a forward way: we generate the first year loss X_1, we calculate the best estimate $\mathbb{E}[X_n|X_1]$, using the development factors $(X_{i+1}|X_i, ..., X_1)_{i=1}^{n-1}$ we generate the losses in the next development years and we get the ultimate loss X_n. Such a forward simulation scheme is investigated e.g. in Wüthrich, Merz (2015) where the relation between the one year reserve risk and the

ultimate reserve risk is discussed in details. The new problem which we study in this paper is how to model the one-year premium risk and the ultimate premium risk (BE_1, X_n) by generating them in a backward way starting with the ultimate loss X_n for the new accident year under consideration.

This paper is motivated by a real problem faced by insurance companies in Solvency II regime. From business point of view, the unconditional distribution of X_n is well-understood by decision makers, used in all planning reports, is the basis of pricing and long-term risk analysis. At the same time, the premium risk module in Solvency II focuses on the risk in the first year for the premiums earned, hence the distribution of X_1 must be used. Due to the business reasons mentioned above, insurance companies are less willing to use the distribution of X_1. They would usually prefer to perform a backward simulation of (X_1, X_n), i.e simulate the ultimate loss from a well-understood unconditional distribution of X_n and allocate the simulated ultimate loss X_n to the first development year. This allocation should be done based on the conditional distribution of $X_1|X_n = x$, or based on the conditional distribution of $BE_1|X_n = x$ since BE_1 describes the one-year premium risk in Solvency II.

A practical approach to a backward simulation of the cumulative payments $(X_1, ..., X_n)$ starting from the ultimate loss X_n is discussed in England et al. (2012) and Bird, Cairns (2011), who introduce concepts of an *emergence pattern* and *emergence factors* in reserve risk. These concepts can also be applied to premium risk. The authors suggest that the ultimate loss X_n can be mapped into the best estimate of the ultimate loss BE_1 by using a linear function. To our knowledge, this allocation rule is used in practice by many insurance companies to model one-year premium and reserve risks. We postulate a new approach to define the emergence pattern of the ultimate loss. Since we investigate the one-year premium risk, we only have to define the emergence pattern of X_n in the first development year, which is different to the reserve risk where we have to define the emergence patterns of X_n in all development years except the first one. In this paper, the emergence pattern formula for the one-year premium risk is based on the conditional distribution of the best estimate of the ultimate loss given the ultimate loss $BE_1|(\mathcal{F}_0, X_n)$ where the conditional distribution is derived from the multivariate distribution of the claims development process $(X_0, X_1, ..., X_n)$ and the definition of the best estimate of the ultimate loss after the first year BE_1. The definition $BE_1|(\mathcal{F}_0, X_n)$ emphasises that we should not view the emergence pattern of the ultimate loss as $BE_1|\mathcal{F}_n$. In the sequel, we use $BE_1|X_n$ as a simplified notation for $BE_1|(\mathcal{F}_0, X_n)$. The conditional distribution of $BE_1|X_n$ is interpreted as the *true* emergence pattern of the ultimate loss X_n. From the conditional distribution of $BE_1|X_n$, we next derive the unconditional distribution of the best estimate of the ultimate loss BE_1 used for quantifying the *true* (unconditional) one-year premium risk.

We consider three common claims development models used for claims reserving (Gaus-

sian Incremental Loss Ratio, Hertig's Lognormal and Over-Dispersed Poisson models) and we establish the conditional distributions of $BE_1|X_n$ and the unconditional distributions of BE_1 in these models. We prove that the true emergence patterns in these three models are different from the emergence pattern postulated by England et al. (2012), Bird, Cairns (2011). We identify that the true one-year premium risk, when measured with Value-at-Risk, can be under and overestimated if the emergence pattern formula from England et al. (2012), Bird, Cairns (2011) is applied, and the true one-year premium risk, when measured with Value-at-Risk, can be higher than the ultimate premium risk at some low confidence levels. Each of the three investigated claims development models assumes a particular distribution of the ultimate loss and a particular claims development process (respectively: normal, lognormal, and Poisson). In order to investigate possible relations between the one-year premium risk and the ultimate premium risk in models with various distributions of the ultimate loss and various claims development processes, we keep the conditional distributions $BE_1|X_n$ from the well-known claims development models but use an arbitrary distribution X_n. The key conclusions from this part of the paper are that the distribution of the one-year premium risk predicted by the emergence pattern formula by England et al. (2012), Bird, Cairns (2011) can have incorrect tail behaviour and the true one-year premium risk, when measured with Value-at-Risk, can be higher than the ultimate premium risk at all high confidence levels. The last result contradicts the common belief among actuaries that the one-year risk is always lower than the ultimate risk if we go far in the right tail, see e.g. Lloyd's (2014) or AISAM-ACME (2007). Yet, if standard deviation is used to measure the risk, then the true emergence pattern formulas from this paper give the same one-year premium risk as the emergence pattern formula from England et al. (2012), Bird, Cairns (2011) and the one-year premium risk is always lower than the ultimate premium risk in our models.

We consider light-tailed, subexponential and Pareto-type distributions. Unless otherwise stated, we assume that the distribution F of the ultimate loss X_n is absolutely continuous, supported on $(0, \infty)$ with infinite right-end point and has finite second moment. By a light-tailed loss distribution we mean a distribution F such that

$$\lim_{x \to \infty} e^{tx}(1 - F(x)) = 0, \quad for \ some \ t > 0. \tag{1.1}$$

By a subexponential loss distribution we mean a distribution F such that for independent copies of $X \sim F$ we have

$$Pr(X_1 + ... + X_n > x) \sim nPr(X > x) \quad x \to \infty. \tag{1.2}$$

It is known that $\lim_{x \to \infty} e^{tx}(1 - F(x)) = \infty$ for any $t > 0$. By a Pareto-type distribution with a tail index θ we mean a distribution F such that

$$1 - F(x) = L(x)x^{-\theta}, \quad x > 0, \tag{1.3}$$

where L denotes a slowly-varying function, i.e. a function which satisfies $\lim_{x \to \infty} \frac{L(tx)}{L(x)} = 1$ for any $t > 0$. Since subexponential distributions include Pareto-type distributions, we focus on subexponential distributions with all moments finite.

We always assume that the risk is measured with Value-at-Risk, unless otherwise stated. We quantify the (unconditional) ultimate premium risk and the (unconditional) one-year premium risk with $VaR_\gamma[X_n - \mathbb{E}[X_n]]$ and $VaR_\gamma[BE_1 - \mathbb{E}[BE_1]]$ at some confidence level γ. In reserve risk models, so-called one-year claims development results $CDRs$ are investigated, see Wüthrich, Merz (2015). The counterpart of the one-year CDR, which we want to use for the one-year premium risk, is the technical result for the new accident year, defined as the difference between the premiums earned in the year and the best estimate of the ultimate loss at the end of the year for the losses incurred in this accident year. Since the premiums include an expected profit margin, we replace the premiums with the expected loss $\mathbb{E}[X_n]$ in our definition of the premium risk. From insurance point of view, we only consider confidence levels $\gamma > 0.5$ such that $VaR_\gamma[X_n] > \mathbb{E}[X_n]$ (but we may have $VaR_\gamma[BE_1] < \mathbb{E}[BE_1]$).

In Section 2 we describe the emergence pattern formula from England et al. (2012), Bird, Cairns (2011). In Section 3 we investigate the true emergence patterns and the true one-year risks in Gaussian Incremental Loss Ratio, Hertig's Lognormal and Over-Dispersed Poisson models. In Section 4 we study the one-year premium risk if we use an arbitrary distribution of the ultimate loss and allocate this loss to the first year by using the conditional distributions from Section 3. All proofs are presented in the last section.

2   The linear emergence pattern

England et al. (2012) and Bird, Cairns (2011) introduced the concept of an emergence pattern of the ultimate loss for one-year reserve risk. They postulate a linear relation between the best estimate of the ultimate loss and the ultimate loss. If their approach is applied to one-year premium risk, then we have the relation:

$$BE_1^{ep} = \alpha X_n + (1 - \alpha)\mathbb{E}[X_n], \tag{2.1}$$

where α is called an emergence factor, and $\alpha \in (0, 1)$. England et al. (2012), Bird, Cairns (2011) propose to use

$$\alpha = \frac{SD[BE_1]}{SD[X_n]} = \frac{SD\big[BE_1 - \mathbb{E}[BE_1]\big]}{SD\big[X_n - \mathbb{E}[X_n]\big]}. \tag{2.2}$$

The emergence factor α in (2.2) measures the relation between the one-year premium risk and the ultimate premium risk for the new accident year under consideration, where the risk is measured with standard deviation.

We would like to remark that in this paper we don't consider estimation error and we assume that all parameters of the claims development process are given. Consequently, the emergence factor α can be directly derived from the multivariate distribution of $(X_1, ..., X_n)$ assumed for the accident year. In practice, assuming a Markovian claims development process or a claims development process with independent increments, we would have to estimate the distribution of the loss in the first year X_1 and the distributions of the development factors $(X_{i+1}|X_i)_{i=1}^{n-1}$, or $(X_{i+1} - X_i)_{i=1}^{n-1}$, in consecutive years and, next, compare the deviations of BE_1 and X_n calculated/simulated in the forward way. The idea of estimating α for the one-year premium risk is the same as for the one-year reserve risk, see England et al. (2012) and Bird, Cairns (2011). The distributions of the development factors $(X_{i+1}|X_i)_{i=1}^{n-1}$, or $(X_{i+1} - X_i)_{i=1}^{n-1}$, are estimated from the historical losses in a run-off triangle in a claims reserving model (e.g. Chain-Ladder, incremental loss ratio, additive model). However, in order to derive (2.2) for the one-year premium risk, we also have to establish the unconditional distribution of X_1, which is usually related to the exposure (premium) in the accident year, see incremental loss ratio and additive models in Chapter 8.3 in Wüthrich, Merz (2008). The parameters of the claims development process estimated from historical data in a claims reserving model can be used in the new accident year. In particular, X_1 can be simulated using the planned volume of the premiums to be earned in the new accident year, coming from the financial plan of the company. Due to the assumption that we know all parameters of the claims development process, the one-year premium risk is, and can be, investigated independently of the one-year reserve risk in the sense that the development factors observed in the new calendar year in the loss run-off triangle don't impact the best estimate of the ultimate loss for the losses incurred in the new accident year, i.e. we don't have to consider so-called re-reserving/bootstraping techniques for $\mathbb{E}[X_n|X_1]$. We are aware of the simplifications, however, the key goal of this paper is to analyse probabilistic properties of the one-year premium risk and the ultimate premium risk implied by various claims development processes.

We have a simple algorithm (2.1) for the simulation of BE_1 starting from X_n. Moreover, the parametrisation of the formula (2.1) is very appealing. We point out that we can use any distribution of the ultimate loss X_n parameterised in the premium risk model and we can switch from the ultimate premium risk to the one-year premium risk using a single scaling factor α parameterised in the claims reserving model. The distribution of X_n used in (2.1) does not have to coincide with the distribution of the ultimate loss from the claims reserving model where the parameter α is calibrated, see England et al. (2012) and Bird, Cairns (2011).

Let us investigate theoretical foundations of the emergence pattern formula (2.1). Since we want to model a pair of dependent random variables (BE_1, X_n), the simplest approach would be to use a linear factor model for (BE_1, X_n). It is natural to assume that $\mathbb{E}[BE_1] =$

$\mathbb{E}[X_n]$ to guarantee that the estimation of the ultimate loss is unbiased. We get the relation:

$$BE_1^f = \eta X_n + (1 - \eta)\mathbb{E}[X_n] + \epsilon, \quad \eta = \rho(BE_1, X_n)\frac{SD(BE_1)}{SD(X_n)}, \tag{2.3}$$

where $\mathbb{E}[\epsilon] = 0, Var[\epsilon] = \left(1 - \rho^2(BE_1, X_n)\right)Var[BE_1]$ and ϵ is independent of X_n. We can see that the linear factor model (2.3) leads to the emergence pattern (2.1) with α from (2.2) only if $\rho(BE_1, X_n) = 1$. We can immediately derive the following conclusions.

Corollary 2.1.   *1. The emergence pattern formula (2.1) is built on the assumption that the best estimate of the ultimate loss BE_1 is perfectly linearly correlated with the ultimate loss X_n,*

2. The emergence pattern formula (2.1) cannot describe the true emergence pattern of the ultimate loss, in the sense of the conditional distribution of $BE_1|X_n = x$, and the true joint distribution of the one-year risk and the ultimate risk (BE_1, X_n) in non-trivial models of claims development, including Gaussian Incremental Loss Ratio, Hertig's Lognormal, and Over-Dispersed Poisson models.

The emergence pattern formula (2.1) implies that the conditional distribution $BE_1^{ep}|X_n = x$ is degenerate. However, the true conditional distribution $BE_1|X_n = x$ and the true conditional emergence pattern formula for the ultimate loss cannot be degenerate in a non-trivial claims development model, since we expect that there are many scenarios with many different possible values of BE_1 which can lead to the same value of X_n. Hence, there is a clear need for an improvement in the emergence pattern formula.

In many applications, including Solvency II calculation of the capital requirement, we only need an unconditional sample of the best estimate of the ultimate loss BE_1, instead of a joint sample (BE_1, X_n) or a conditional sample $BE_1|X_n = x$. Let us use the formula (2.1) to generate the unconditional sample of BE_1. We are interested in properties of the (unconditional) one-year risk vs. the ultimate risk.

Theorem 2.1. *Let us consider the emergence pattern formula (2.1) with an emergence factor $\alpha \in (0, 1)$.*

1. $\mathbb{E}[BE_1^{ep}] = \mathbb{E}[X_n]$ and $Var[BE_1^{ep}] = \alpha^2 Var[X_n] < Var[X_n]$,

2. If X_n has a light-tailed distribution (subexponential with all moments finite), then BE_1^{ep} has a light-tailed distribution (subexponential with all moments finite),

3. If X_n has a Pareto-type distribution with tail index θ, then BE_1^{ep} has a Pareto-type distribution with tail index θ,

4. $VaR_\gamma\left[BE_1^{ep} - E[BE_1^{ep}]\right] = \alpha VaR_\gamma\left[X_n - E[X_n]\right] < VaR_\gamma\left[X_n - E[X_n]\right].$

7

If we use (2.1), then the unconditional expected value of the best estimate of the ultimate loss is equal to the unconditional expected value of the ultimate loss. This property is desired since otherwise the estimation of the ultimate loss would be biased. If the emergence factor α is set in accordance with condition (2.2), then $Var[BE_1^{ep}] = Var[BE_1]$, which is again a desired property. Next, the one-year premium risk is a fraction of order α of the ultimate premium risk, if the risk is measured with standard deviation and Value-at-Risk. This property is less clear but it is in accordance with a common belief among actuaries that the one-year risk is lower than the ultimate risk, see e.g. Lloyd's (2014) or AISAM-ACME (2007). Finally, the distributions of the one-year premium risk and the ultimate premium risk have the same tail behaviour. Theorem 2.1 shows that the formula (2.1) has some desirable properties if applied to simulate the unconditional samples of the best estimates of the ultimate loss BE_1 from the ultimate losses X_n.

3   The true emergence patterns in claims development models

In the next subsections, we derive the conditional distributions of $BE_1|X_n$ and the unconditional distributions of BE_1 in three claims development models used for claims reserving.

3.1   Gaussian Incremental Loss Ratio model

We consider Incremental Loss Ratio model (ILR) with Gaussian incremental losses, see e.g. Radtke et al. (2016) and Wüthrich, Merz (2008). We investigate cumulative payments given by

$$X_j = \sum_{i=1}^{j} \epsilon_i, \qquad \text{where:} \quad \epsilon_i \sim N(E\mu_i; E\sigma_i^2) \quad \text{for } i \in \{1, ..., n\}, \tag{3.1}$$

and $(\epsilon_i)_{i=1}^{n}$ is a sequence of independent random variables. By E we denote the exposure in the accident year, and ϵ_i represents the incremental loss in development year i. We assume that $\mu_i \in \mathbb{R}, \sigma_i > 0, E > 0$. We denote

$$\mu = \sum_{i=1}^{n} \mu_i, \quad \sigma^2 = \sum_{i=1}^{n} \sigma_i^2.$$

The best estimate of the ultimate loss is defined by

$$BE_1 = \mathbb{E}[X_n|X_1] = X_1 + E(\mu - \mu_1). \tag{3.2}$$

The next theorem describes the distributions in which we are interested. The distributions have special parameterisations which mimic the emergence pattern formula (2.1).

Theorem 3.1. *Let us consider the model (3.1) of claims development and the multivariate distribution of the claims development process* $(X_1, ..., X_n)$. *Let*

$$\mu_{X_n} = \mathbb{E}[X_n], \quad \sigma_{X_n}^2 = Var[X_n], \quad \alpha = \frac{SD[BE_1]}{SD[X_n]}.$$

We have the following loss distributions:

$$X_n \quad \sim \quad N\left(\mu_{X_n}; \sigma_{X_n}^2\right), \tag{3.3}$$

$$BE_1|X_n = x \quad \sim \quad N\left(\alpha^2 x + (1-\alpha^2)\mu_{X_n}; \alpha^2(1-\alpha^2)\sigma_{X_n}^2\right), \tag{3.4}$$

$$BE_1 \quad \sim \quad N\left(\mu_{X_n}; \alpha^2\sigma_{X_n}^2\right), \tag{3.5}$$

with the parameter α *which satisfies* $\alpha \in (0, 1)$.

Formula (3.4) establishes the true emergence pattern of the ultimate loss X_n in our claims development model. Moreover, it describes a backward simulation scheme for BE_1, starting from X_n, where we can switch (in accordance with the underlying probabilistic model of claims development) from the ultimate premium risk to the one-year premium risk. As in (2.1), we only use the distribution of the ultimate loss X_n from the ultimate premium risk and a single emergence factor α which summarises the information from the reserve risk module about the claims development process required for the one-year premium risk. We follow the main idea behind the emergence pattern formula (2.1). However, we are able to improve the emergence pattern formula from England et al. (2012), Bird, Cairns (2011) so that it yields the correct conditional distribution of $BE_1|X_n = x$ and, consequently, the correct unconditional distribution of BE_1 in our claims development model, given by (3.5).

If we assume, as in Theorem 3.1, that the ultimate loss $X_n \sim N(\mu_{X_n}, \sigma_{X_n}^2)$, then the emergence pattern formula (2.1) yields the following unconditional distribution of the best estimate of the ultimate loss:

$$BE_1^{ep} \quad \sim \quad N\left(\mu_{X_n}, \alpha^2\sigma_{X_n}^2\right). \tag{3.6}$$

Interestingly, the unconditional distributions of the best estimate of the ultimate loss (3.5) and (3.6) are the same. Consequently, there exists a claims development model where the unconditional distribution of the true one-year premium risk can be characterised with the emergence pattern formula from England et al. (2012), Bird, Cairns (2011). However, as we see in the next subsections, this is an exceptional case.

We conclude with properties of the unconditional distribution of BE_1, which are special cases of the properties for BE_1^{ep} from Theorem 2.1.

Theorem 3.2. *Let us consider the model and the assumptions from Theorem 3.1.*

1. $\mathbb{E}[BE_1] = \mathbb{E}[X_n]$ and $Var[BE_1] = \alpha^2 Var[X_n] < Var[X_n]$,

2. $VaR_\gamma[BE_1 - E[BE_1]] = \alpha VaR_\gamma[X_n - E[X_n]] < VaR_\gamma[X_n - E[X_n]]$, where

$$VaR_\gamma[BE_1 - E[BE_1]] = \alpha \Phi^{-1}(\gamma) SD[X_n].$$

3.2   Hertig's Lognormal model

We consider a multiplicative loss model where the development factors are modelled with lognormal distributions, see e.g. Hertig (1985) or Wüthrich, Merz (2008). We investigate cumulative payments given by

$$X_1 = \epsilon_1, \quad X_i = X_{i-1} \cdot \epsilon_i, \qquad \text{where: } \epsilon_i \sim LogN(m_i, s_i^2) \quad \text{for } i \in \{1, ..., n\}, \tag{3.7}$$

and $(\epsilon_i)_{i=1}^n$ is a sequence of independent random variables. We assume that $m_i \in \mathbb{R}, s_i > 0$. We denote

$$m = \sum_{i=1}^n m_i, \quad s^2 = \sum_{i=1}^n s_i^2.$$

The best estimate of the ultimate loss is defined by

$$BE_1 = \mathbb{E}[X_n | X_1] = X_1 e^{m - m_1 + \frac{1}{2}(s^2 - s_1^2)}. \tag{3.8}$$

Using the same reasoning as in the previous section, we can derive the true emergence pattern for the ultimate loss which comes from the claims development process (3.7).

Theorem 3.3. *Let us consider the model* (3.7) *of claims development and the multivariate distribution of the claims development process* $(X_1, ..., X_n)$. *Let*

$$\mu_{X_n} = \mathbb{E}[X_n], \quad \psi_{X_n} = \frac{SD[X_n]}{\mathbb{E}[X_n]}, \quad \alpha = \frac{SD[BE_1]}{SD[X_n]}.$$

We have the following loss distributions:

$$X_n \sim LogN\left(\tilde{m}; \tilde{s}^2\right), \tag{3.9}$$

$$BE_1 | X_n = x \sim LogN\left(\tilde{\alpha}^2 \log(x) + \left(1 - \tilde{\alpha}^2\right)\left(\tilde{m} + \frac{\tilde{s}^2}{2}\right); \tilde{\alpha}^2(1 - \tilde{\alpha}^2)\tilde{s}^2\right), \tag{3.10}$$

$$BE_1 \sim LogN\left(\tilde{m} + (1 - \tilde{\alpha}^2)\frac{\tilde{s}^2}{2}; \tilde{\alpha}^2 \tilde{s}^2\right), \tag{3.11}$$

where

$$\tilde{m} = log(\mu_{X_n}) - \frac{1}{2}\log\left(1 + \psi_{X_n}^2\right), \quad \tilde{s}^2 = \log\left(1 + \psi_{X_n}^2\right),$$

$$\tilde{\alpha}^2 = \frac{\log\left(1 + \alpha^2 \psi_{X_n}^2\right)}{\log\left(1 + \psi_{X_n}^2\right)},$$

with the parameter α which satisfies $\alpha \in (0, 1)$.

10

Theorem 3.4. *Let us consider the model and the assumptions from Theorem 3.3.*

1. $\mathbb{E}[BE_1] = \mathbb{E}[X_n]$ *and* $Var[BE_1] = \alpha^2 Var[X_n] < Var[X_n]$,

2. $VaR_\gamma[BE_1 - E[BE_1]] < VaR_\gamma[X_n - E[X_n]]$ *for* $\gamma > \gamma^*$, $VaR_\gamma[BE_1 - E[BE_1]] > VaR_\gamma[X_n - E[X_n]]$ *for* $\gamma < \gamma^*$, *where*

$$VaR_\gamma[BE_1] = \mu_{X_n}\left(1 + \alpha^2\psi_{X_n}^2\right)^{-1/2} e^{\sqrt{\log\left(1+\alpha^2\psi_{X_n}^2\right)}\,\Phi^{-1}(\gamma)},$$

$$\gamma^* = \Phi\left(\frac{1}{2}\sqrt{\log(1+\alpha^2\psi_{X_n}^2)} + \frac{1}{2}\sqrt{\log(1+\psi_{X_n}^2)}\right),$$

3. *We have the limit*

$$\lim_{\gamma\to 1} \frac{VaR_\gamma[BE_1 - \mathbb{E}[BE_1]]}{VaR_\gamma[X_n - \mathbb{E}[X_n]]} = 0.$$

If we make the assumptions from Theorem 3.3 and apply the emergence pattern formula (2.1) to the claims development process (3.7), then we can deduce the distribution:

$$BE_1^{ep} \sim \alpha \cdot LogN(\tilde{m}, \tilde{s}^2) + (1-\alpha)\mu_{X_n}. \tag{3.12}$$

Let us compare Theorem 3.4 with 2.1 and the emergence pattern formula (3.10) with (2.1). Recalling the discussion from the previous sections, it is clear that the conditional distributions of the best estimate of the ultimate loss given the ultimate loss (3.10) and (2.1) are different. We can now observe that the unconditional distributions of the best estimate of the ultimate loss (3.11) and (3.12) are also different if derived from the emergence pattern formula and the conditional distribution approach.

We now focus on the premium risk in the claims development model (3.7). We can notice that the supports of BE_1 and BE_1^{ep} generated with (3.11) and (3.12) are different. From point 2 from Theorem 3.4 we deduce that the true one-year risk can be higher than the ultimate risk if the confidence level γ is too low. We point out that a confidence level γ at which we may observe that $VaR_\gamma[BE_1] > VaR_\gamma[X_n]$ is above the lowest confidence level we allow for, i.e. is above $\tilde{\gamma} = \Phi\left(\frac{1}{2}\sqrt{\log(1+\psi_{X_n}^2)}\right)$ at which $VaR_{\tilde{\gamma}}[X_n] = E[X_n]$. For a skewed distribution of X_n with high ψ_{X_n}, a *low* confidence level γ at which we may observe that $VaR_\gamma[BE_1] > VaR_\gamma[X_n]$ can be very high, if α is also high. However, at all sufficiently high confidence levels the true one-year risk is lower then the ultimate risk. This property stands in contrast with the property of the emergence pattern formula (2.1) where the one-year risk is lower than the ultimate risk at all confidence levels. From point 3 from Theorem 3.4 we can conclude that the ratio between the Value-at-Risk measures for the one-year risk BE_1 and the ultimate risk X_n at a sufficiently high confidence level must be lower than α (whereas the ratio between the VaR for BE_1^{ep} and the VaR for X_n is α). Summing up,

we can conclude that the emergence pattern formula (2.1) underestimates the true one-year premium risk at low confidence levels and overestimates the true one-year premium risk at high confidence levels in the claims development model (3.7). The precise values of low/high confidence levels depend on α and ψ_{X_n}.

Example 3.1. We choose $\mu_{X_n} = 1$ since μ_{X_n} is scaling parameter for X_n. We calculate the ratios of $\frac{VaR_\gamma[BE_1 - \mathbb{E}[BE_1]]}{VaR_\gamma[X_n - \mathbb{E}[X_n]]}$ as a function of α, for different ψ_{X_n} and γ. If $\frac{VaR_\gamma[BE_1 - \mathbb{E}[BE_1]]}{VaR_\gamma[X_n - \mathbb{E}[X_n]]} > 1$, then the true one-year risk is higher than the ultimate risk. Let us recall that $\frac{VaR_\gamma[BE_1^{ep} - \mathbb{E}[BE_1^{ep}]]}{VaR_\gamma[X_n - \mathbb{E}[X_n]]} = \alpha$ for any ψ_{X_n} and γ. Hence, if $\frac{VaR_\gamma[BE_1 - \mathbb{E}[BE_1]]}{VaR_\gamma[X_n - \mathbb{E}[X_n]]} > \alpha$, the emergence pattern formula (2.1) underestimates the true one-year risk, and if $\frac{VaR_\gamma[BE_1 - \mathbb{E}[BE_1]]}{VaR_\gamma[X_n - \mathbb{E}[X_n]]} < \alpha$, then the emergence pattern formula (2.1) overestimates the true one-year risk. The results are presented in Figure 1. They confirm that the emergence pattern formula (2.1) can under and overestimate the true one-year premium risk in the claims development model (3.7). E.g. for $\psi_{X_n} = 1$ and $\alpha = 50\%$, the emergence pattern formula (2.1) overestimates the true one year risk at $\gamma = 99.5\%$ by about 25%. For $\psi_{X_n} = 3$ and $\alpha = 50\%$, the emergence pattern formula (2.1) underestimates the true one year risk at $\gamma = 99.5\%$ by about 9%. We can see that the discrepancies between the linear emergence pattern formula and the true one-year risk can be much larger in both directions for different parameters. In practice, we would rarely expect $\psi_{X_n} > 3$. We also observe that the true one-year premium risk can be higher than the ultimate premium risk at low confidence levels γ. Even though point 3 from Theorem 3.4 shows that $VaR_\gamma[BE_1 - \mathbb{E}[BE_1]] = o\big(VaR_\gamma[X_n - \mathbb{E}[X_n]]\big)$ for $\gamma \to 1$, we end up with $VaR_\gamma[BE_1 - \mathbb{E}[BE_1]] \approx \alpha^2 VaR_\gamma[X_n - \mathbb{E}[X_n]]$ for γ in the range $99.99\% - 99.999\%$ in our experiments. $\qquad\square$

3.3 Over-Dispersed Poisson model

Finally, we consider Incremental Loss Ratio model with Over-Dispersed Poisson incremental losses, see e.g. Wüthrich, Merz (2008) or Taylor (2016). We investigate cumulative payments given by

$$X_j = \sum_{i=1}^{j} \epsilon_i, \quad \text{where:} \quad \epsilon_i \sim ODP\big(\mu\omega_i, \psi\big) \quad \text{for } i \in \{1, ..., n\}, \tag{3.13}$$

and $(\epsilon_i)_{i=1}^n$ is a sequence of independent random variables. We assume that $\mu > 0, \omega_i > 0, \psi > 0$. The parameter μ is interpreted as the expected ultimate loss, and ω_i is interpreted as the proportion of μ which is expected to be paid in development year i. Hence, we assume that $\sum_{i=1}^n \omega_i = 1$. We assume that the dispersion ψ is independent of the development year so that we can derive analytical results. The assumption that $\epsilon \sim ODPoiss\big(\mu, \psi\big)$ means
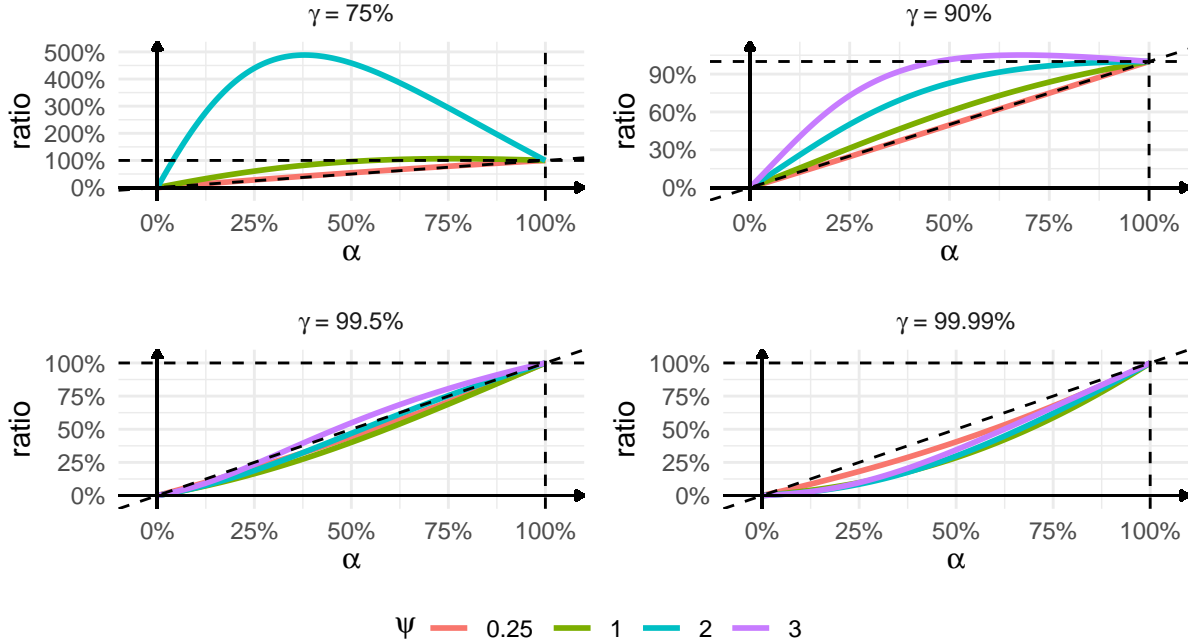
Figure 1: The ratios $\frac{VaR_\gamma[BE_1-\mathbb{E}[BE_1]]}{VaR_\gamma[X_n-\mathbb{E}[X_n]]}$ in the Hertig's Lognormal model.

that $\epsilon/\psi \sim Poiss(\mu/\psi)$. The best estimate of the ultimate loss is defined by

$$BE_1 = \mathbb{E}[X_n|X_1] = X_1 + \mu(1-\omega_1). \tag{3.14}$$

Theorem 3.5. *Let us consider the model* (3.13) *of claims development and the multivariate distribution of the claims development process* $(X_1, ..., X_n)$. *Let*

$$\mu_{X_n} = \mathbb{E}[X_n], \quad \psi_{X_n} = \frac{Var[X_n]}{\mathbb{E}[X_n]}, \quad \alpha = \frac{SD[BE_1]}{SD[X_n]}.$$

We have the following loss distributions:

$$X_n \quad \sim \quad \psi_{X_n} \cdot Poiss(\mu_{X_n}/\psi_{X_n}), \tag{3.15}$$

$$BE_1|X_n = x \quad \sim \quad \psi_{X_n} \cdot Bin\left(\frac{x}{\psi_{X_n}}; \alpha^2\right) + (1-\alpha^2)\mu_{X_n}, \tag{3.16}$$

$$BE_1 \quad \sim \quad \psi_{X_n} \cdot Poiss(\alpha^2\mu_{X_n}/\psi_{X_n}) + (1-\alpha^2)\mu_{X_n}, \tag{3.17}$$

with the parameter α *which satisfies* $\alpha \in (0,1)$.

Theorem 3.6. *Let us consider the model and the assumptions from Theorem 3.5.*

1. $\mathbb{E}[BE_1] = \mathbb{E}[X_n]$ *and* $Var[BE_1] = \alpha^2 Var[X_n] < Var[X_n]$,

13

2. $VaR_\gamma\big[BE_1 - E[BE_1]\big] < VaR_\gamma[X_n - E[X_n]]$ *for* $\gamma > \gamma^*$ *where* $\gamma^* < 1$ *is the last point where the distribution functions of* BE_1 *and* X_n *intersect, or* $VaR_\gamma\big[BE_1 - E[BE_1]\big] < VaR_\gamma[X_n - E[X_n]]$ *for all* γ *if the distribution functions of* BE_1 *and* X_n *don't intersect. The relation between* $VaR_\gamma\big[BE_1 - E[BE_1]\big]$ *and* $VaR_\gamma[X_n - E[X_n]]$ *can change for* $\gamma < \gamma^*$ *since the distribution functions of* BE_1 *and* X_n *can intersect more than once (if they intersect),*

3. *We have the limit*

$$\lim_{\gamma \to 1} \frac{VaR_\gamma\big[BE_1 - \mathbb{E}[BE_1]\big]}{VaR_\gamma\big[X_n - \mathbb{E}[X_n]\big]} = 1.$$

If we make the assumptions from Theorem 3.5 and we apply the emergence pattern formula (2.1) to the claims development process (3.13), then we can derive the distribution:

$$BE_1^{ep} \sim \alpha\psi_{X_n} \cdot Poiss\big(\mu_{X_n}/\psi_{X_n}\big) + (1-\alpha)\mu_{X_n}. \tag{3.18}$$

By point 2 from Theorem 3.6, there exists a confidence level $\gamma^* < 1$ such that the true one-year risk is lower than the ultimate risk at all sufficiently high confidence levels $\gamma > \gamma^*$. At low confidence levels $\gamma < \gamma^*$, the true one-year risk can be lower or higher than the ultimate risk. This property makes the interpretation of the one-year premium risk in relation to the ultimate premium risk difficult in the claims development model (3.13) since the relation between the true one-year premium risk and the ultimate premium risk can change multiple times as we change the confidence level. From point 3 from Theorem 3.6 we can deduce that the ratio between the Value-at-Risk measures for the one-year risk BE_1 and the ultimate risk X_n at a sufficiently high confidence level must be higher than α. Consequently, the emergence pattern formula (2.1) underestimates the true one-year premium risk at all high confidence levels. The emergence pattern formula (2.1) can also underestimate the true one-year premium risk at some low confidence levels e.g. when the true one-year risk is above the ultimate risk.

Example 3.2. We choose $\psi_{X_n} = 1$ and we consider Poisson distributions of X_n with different μ_{X_n}. The Over-Dispersed Poisson model approaches the Gaussian Incremental Loss Ratio model for $\mu_{X_n} \to \infty$, hence we are interested in low μ_{X_n}. Our key results are presented in Figure 2. We remark that some ratios are negative since $VaR_\gamma[BE_1] < \mathbb{E}[BE_1]$ (but $VaR_\gamma[X_n] > \mathbb{E}[X_n]$ by our assumption). We can confirm that the emergence pattern formula (2.1) can under and overestimate the true one-year premium risk in the claims development model (3.13). E.g. for $\mu_{X_n} = 1.5$ and $\alpha = 85\%$, the emergence pattern formula (2.1) underestimates the true one-year risk at $\gamma = 99.5\%$ by 24% and the true one-year risk is above the ultimate risk by 12%. For $\mu_{X_n} = 25$ and $\alpha = 15\%$, the emergence pattern formula (2.1) underestimates the true one-year risk at $\gamma = 99.5\%$ by 14%. For $\mu_{X_n} = 1.5$
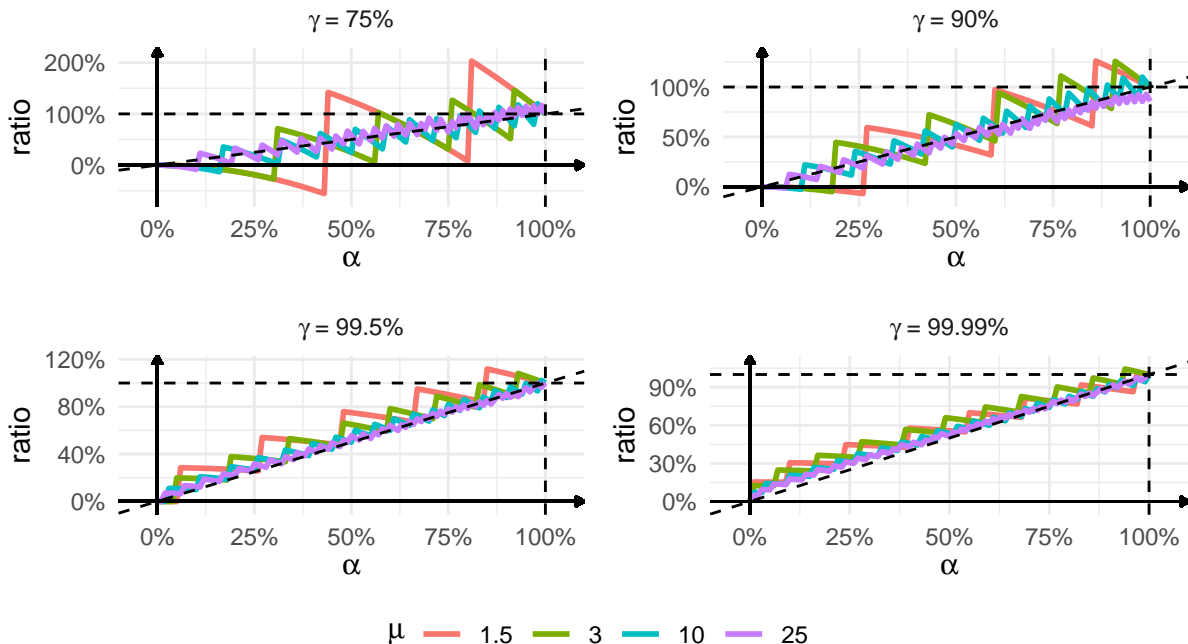
Figure 2: The ratios $\frac{VaR_\gamma[BE_1 - \mathbb{E}[BE_1]]}{VaR_\gamma[X_n - \mathbb{E}[X_n]]}$ in the Poisson model.

and $\alpha = 85\%$, the true one-year risk is higher than the ultimate risk at $\gamma = 75\%$ and 99.5%, but the true one-year risk is lower than the ultimate risk for $\gamma = 90\%$ and 99.99% - in this example the distribution functions of BE_1 and X_n intersect 9 times in the range from the median up to the quantile of order $\gamma^* = 99.555\%$. Moreover, the emergence pattern formula (2.1) overestimates the true one-year risk at $\gamma = 90\%$ by 39%. Finally, even though point 3 from Theorem 3.6 shows that $VaR_\gamma[BE_1 - \mathbb{E}[BE_1]] \sim VaR_\gamma[X_n - \mathbb{E}[X_n]]$ for $\gamma \to 1$, the convergence for γ in the range $99.99\% - 99.999\%$ is slow in our experiments, yet it can be observed. $\square$

4 One-year premium risk with an arbitrary ultimate loss distribution

The models which we investigated in the previous section assume a particular form of the claims development process and lead to a particular distribution of the ultimate loss. What is important for practical applications in Solvency II in premium risk module is the possibility of using any distribution of the ultimate loss. Yet, it may be difficult to specify a priori the joint multivariate distribution for cumulative payments $(X_1, ..., X_n)$ which lead to a pre-specified distribution of the ultimate loss X_n. What we suggest is to use an arbitrary unconditional

distribution of the ultimate loss X_n and a conditional distribution of $BE_1|X_n$, extracted from the well-known claims development model, as a mechanism for allocating the ultimate loss X_n to the best estimate BE_1. We focus on the conditional distributions $BE_1|X_n$ from the Gaussian Incremental Loss Ratio model and the Hertig's Lognormal model.

We use the following construction. Let us consider a vector of dependent random variables $(BE_1, X_n) \sim f$ with a joint density function of f. The joint density f implies the marginal and conditional densities:

$$X_n \sim f_{X_n}, \quad X_n|BE_1 = x \sim f_{X_n|BE_1=x}, \quad BE_1|X_n = x \sim f_{BE_1|X_n=x}.$$

We define a new joint density of $(BE_1, X_n) \sim f^{new}$ by choosing $BE_1|X_n = x \sim f_{BE_1|X_n=x}$ and $X_n \sim f_{X_n}^{new}$ where $f_{X_n}^{new}$ is a density absolutely continuous with respect to f_{X_n}. We have the new marginal and conditional densities:

$$f_{BE_1}^{new}(z) = \int f_{BE_1|X_n=x}(z) f_{X_n}^{new}(x) dx, \tag{4.1}$$

$$f_{X_n|BE_1=x}^{new}(z) = C \cdot f_{X_n|BE_1=x}(z) \frac{f_{X_n}^{new}(z)}{f_{X_n}(z)}, \tag{4.2}$$

where C denotes a normalising constant which depends on x but is independent of z.

We can construct new models of claims development by using the distribution of the ultimate loss X_n and the emergence pattern of X_n characterized with the conditional distribution of $BE_1|X_n$. We recall that the emergence factor α needed for $BE_1|X_n$ is discussed in Section 2. Our approach has two advantages. Firstly, we have a flexible and interpretable probabilistic model where we can freely choose the distribution of the ultimate loss and switch from the ultimate premium risk to the one-year premium risk (the backward simulation scheme). This is a desired feature in Solvency II modelling. Secondly, we can investigate properties of the one-year premium risk vs. the ultimate premium risk in various claims development models, beyond the models we know from the claims reserving literature. This is the topic of the next two subsections.

4.1   The conditional distribution from Gaussian ILR model

From (3.4) we can deduce that the unconditional distribution of BE_1 is a mixture of normal distributions. We can state the representation:

$$BE_1 = \alpha^2 X_n + (1 - \alpha^2)\mu_{X_n} + \sqrt{\alpha^2(1 - \alpha^2)}\sigma_{X_n}\xi, \tag{4.3}$$

where $\xi \sim N(0,1)$, and ξ is independent of X_n. The formula (4.3) is our new emergence pattern formula which allocates the ultimate loss with an arbitrary distribution to the best estimate of the ultimate loss. We can call (4.3) *the additive normal emergence pattern*

formula. The additive normal emergence pattern can be seen as a version of the classical emergence pattern formula (2.1) where we simply add a Gaussian noise in order to have a non-degenerate distribution of $BE_1 | X_n = x$. If we choose a normal distribution for X_n in (4.3), then we are in the framework of the Gaussian ILR model and the results from Section 3.1 apply.

Theorem 4.1. *Let us assume that the ultimate loss is modelled with $X_n \sim F_{X_n}$. Let $\mu_{X_n} = \mathbb{E}[X_n]$, $\sigma^2_{X_n} = Var[X_n]$ and choose an emergence factor $\alpha \in (0, 1)$. The emergence pattern of the ultimate loss X_n is described with the formula (4.3). We have the following properties:*

1. *$\mathbb{E}[BE_1] = \mathbb{E}[X_n]$ and $Var[BE_1] = \alpha^2 Var[X_n] < Var[X_n]$,*

2. *If X_n has a light-tailed distribution (subexponential with all moments finite), then BE_1 has a light-tailed distribution (subexponential with all moments finite). Moreover, if X_n has a light-tailed distribution such that $\lim_{x \to \infty}(1 - F_{X_n}(x))e^{vx^2} = 0$ for all $v > 0$, then $VaR_\gamma\big[BE_1 - \mathbb{E}[BE_1]\big] > VaR_\gamma\big[X_n - \mathbb{E}[X_n]\big]$ for $\gamma > \gamma^*$, with some $\gamma^* < 1$,*

3. *If X_n has a Pareto-type distribution with tail index θ, then BE_1 has a Pareto-type distribution with tail index θ, and we have the limit*

$$\lim_{\gamma \to 1} \frac{VaR_\gamma\big[BE_1 - \mathbb{E}[BE_1]\big]}{VaR_\gamma\big[X_n - \mathbb{E}[X_n]\big]} = \alpha^2.$$

From point 2 we deduce a new and interesting property that the true one-year premium risk can be higher than the ultimate premium risk at all high confidence levels. Points 2 and 3 indicate that the distributions of the true one-year premium risk and the ultimate premium risk have the same tail behaviour. This property holds in Gaussian Incremental Loss Ratio, Hertig's Lognormal and Over-Dispersed Poisson models where the distribution of X_n is uniquely specified by the underlying claims development model, as well as if we model the one-year risk with the emergence pattern formula (2.1). Finally, point 3 gives a new possible value of $\lim_{\gamma \to 1} \frac{VaR_\gamma\big[BE_1 - \mathbb{E}[BE_1]\big]}{VaR_\gamma\big[X_n - \mathbb{E}[X_n]\big]}$.

Example 4.1. First, we assume that $X_n \sim Exp(\lambda)$ with $1 - F_{X_n}(x) = e^{-\lambda x}$. In this case, the claims development process for the pair (BE_1, X_n) with the ultimate loss $X_n \sim Exp(\lambda)$ and the emergence pattern of the ultimate loss given by the conditional distribution (3.4) can be fully characterised. From (4.1)-(4.2), using classical calculus, we can derive the distributions:

$$Pr(BE_1 \le x) = \Phi\Big(\frac{\lambda x - (1 - \alpha^2)}{\sqrt{\alpha^2(1 - \alpha^2)}}\Big) - exp\Big(\frac{3(1 - \alpha^2) - 2\lambda x}{2\alpha^2}\Big)\Phi\Big(\frac{\lambda x - 2(1 - \alpha^2)}{\sqrt{\alpha^2(1 - \alpha^2)}}\Big),$$

$$X_n | BE_1 = x \sim N_0\Big(\frac{\lambda x - 2(1 - \alpha^2)}{\alpha^2 \lambda}, \frac{1 - \alpha^2}{\alpha^2 \lambda^2}\Big),$$

where N_0 denotes normal distribution left truncated at 0 and Φ denotes the standard normal distribution function. To have more flexibility in our numerical study, we consider $X_n \sim Weibull(\lambda, \tau)$ with $1 - F_{X_n}(x) = e^{-\lambda x^\tau}$. We choose $\lambda = 1$ since λ is scaling parameter for the distribution of X_n. If $\tau = 1$, then $X_n \sim Exp(1)$. If $\tau > 2$, then $\lim_{x \to \infty}(1 - F_{X_n}(x))e^{vx^2} = 0$ for any $v > 0$. In order to calculate $VaR_\gamma[BE_1]$, we use Monte Carlo simulations and we generate a sample of size 10^6. The results are presented in Figure 3. For $\tau < 2$, the one-year risk is lower than the ultimate risk at sufficiently high confidence levels but the one-year risk can be higher than the ultimate risk at low confidence levels. For $\tau > 2$, we observe the new property: the one-year risk can be higher than the ultimate risk at all high confidence levels. E.g. for $\alpha = 85\%$ and $\tau = 20$, we can observe that the one-year risk is higher than the ultimate risk at $\gamma = 99.5\%$ and 99.99%, and at all higher confidence levels. The higher α and τ, the lower the minimal confidence level γ^* from which the one-year risk dominates the ultimate risk. If $\tau > 3.6$, then the Weibull distribution is negatively skewed. Negatively skewed distributions can appear in situations where data are clustered near an upper limit, and the most well-known case where we use a negatively skewed Weibull distribution is modelling of the future lifetime in developed countries where people tend to survive to older ages. $\qquad \Box$
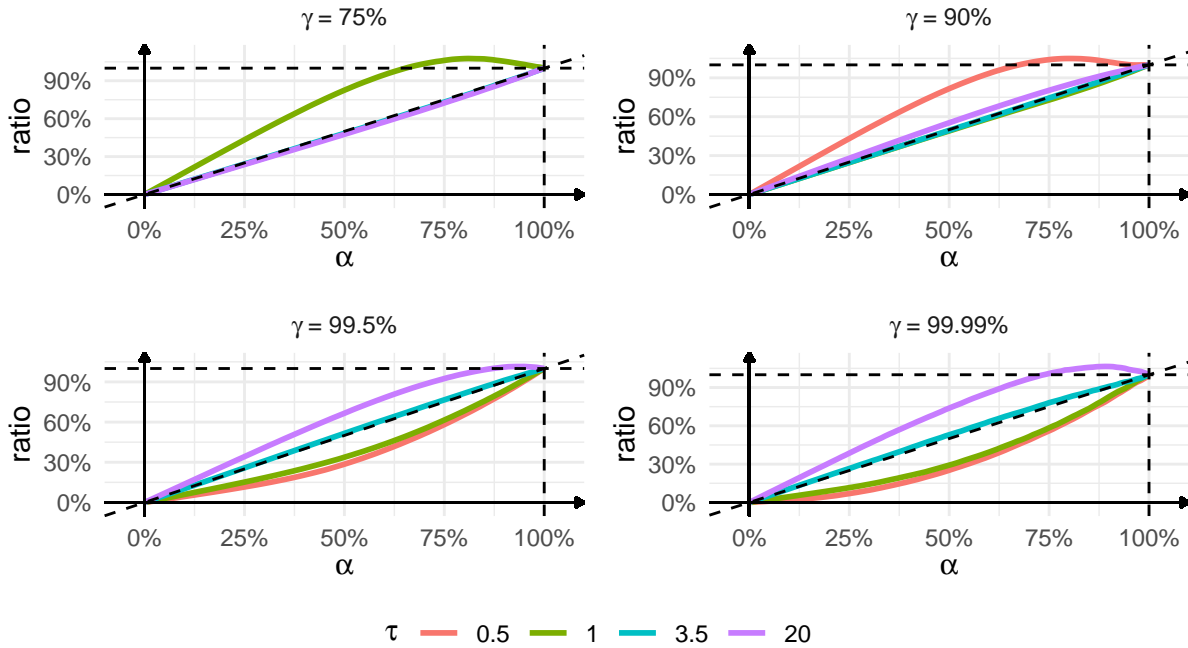

Figure 3: The ratios $\frac{VaR_\gamma[BE_1 - \mathbb{E}[BE_1]]}{VaR_\gamma[X_n - \mathbb{E}[X_n]]}$ in the model where $X_n \sim Weibull(\tau)$ and $BE_1|X_n$ comes from the Gaussian ILR model.

4.2 The conditional distribution from Hertig's Lognormal model

From (3.10) we can deduce that the unconditional distribution of BE_1 is a mixture of lognormal distributions. We can state the representation:

$$BE_1 = (X_n)^{\tilde{\alpha}^2} e^{(1-\tilde{\alpha}^2)\left(\tilde{m}+\frac{\tilde{s}^2}{2}\right)+\sqrt{\tilde{\alpha}^2(1-\tilde{\alpha}^2)}\tilde{s}\xi}, \tag{4.4}$$

where $\xi \sim N(0,1)$, and ξ is independent of X_n. The formula (4.4) is our new emergence pattern formula which allocates the ultimate loss with an arbitrary distribution to the best estimate of the ultimate loss. We can call (4.4) *the multiplicative lognormal emergence pattern formula*. The multiplicative lognormal emergence pattern can be seen as a version of the classical emergence pattern formula (2.1) where we allocate X_n to BE_1 with a random scaling factor in order to have a non-degenerate distribution of $BE_1|X_n = x$. If we choose a lognormal distribution for X_n in (4.4), then we are in the framework of the Hertig's model and the results from Section 3.2 apply.

Theorem 4.2. *Let us assume that the ultimate loss is modelled with $X_n \sim F_{X_n}$. Let $\mu_{X_n} = \mathbb{E}[X_n]$, $\psi_{X_n} = \frac{SD[X_n]}{\mathbb{E}[X_n]}$ and choose an emergence factor $\alpha \in (0,1)$. We assume that there exists a unique solution $\tilde{\alpha} \in (0,1)$ to the equation*

$$\frac{\mathbb{E}[X_n^{2\tilde{\alpha}^2}]}{(\mathbb{E}[X_n^{\tilde{\alpha}^2}])^2}\left(1+\psi_{X_n}^2\right)^{\tilde{\alpha}^2(1-\tilde{\alpha}^2)} = 1 + \alpha^2\psi_{X_n}^2. \tag{4.5}$$

We set

$$\begin{aligned}
\tilde{m} &= \frac{\log(\mu_{X_n}) - \log\left(\mathbb{E}[X_n^{\tilde{\alpha}^2}]\right) - (1-\tilde{\alpha}^4)\frac{\tilde{s}^2}{2}}{1-\tilde{\alpha}^2}, \\
\tilde{s}^2 &= \log\left(1+\psi_{X_n}^2\right).
\end{aligned}$$

The emergence pattern of the ultimate loss X_n is described with the formula (4.4). We have the following properties:

1. *$\mathbb{E}[BE_1] = \mathbb{E}[X_n]$ and $Var[BE_1] = \alpha^2 Var[X_n] < Var[X_n]$,*

2. *If X_n has a light-tailed distribution, then BE_1 has a subexponential distribution with all moments finite. Moreover, $VaR_\gamma[BE_1 - \mathbb{E}[BE_1]] > VaR_\gamma[X_n - \mathbb{E}[X_n]]$ for $\gamma > \gamma^*$, with some $\gamma^* < 1$,*

3. *If X_n has a subexponential distribution with all moments finite such that $\limsup_{x\to\infty} \frac{1-F_{X_n}(vx)}{1-F_{X_n}(x)} < 1$ for some $v > 1$ and $X_n^{\tilde{\alpha}^2}$ is light-tailed or subexponential, then BE_1 has a subexponential distribution with all moments finite,*

19

4. *If X_n has a Pareto-type distribution with tail index θ and regularly varying function L such that $L(x^{\frac{1}{\tilde{\alpha}^2}}) \sim C \cdot L(x)$, $x \to \infty$, then BE_1 has a Pareto-type distribution with tail index $\theta/\tilde{\alpha}^2 > \theta$, and*

$$\lim_{\gamma \to 1} \frac{VaR_\gamma[BE_1 - \mathbb{E}[BE_1]]}{VaR_\gamma[X_n - \mathbb{E}[X_n]]} = 0.$$

Remark 4.1. *Point 3 is satisfied e.g. for lognormal, Weibull, Benktander type I and type II distributions, see e.g. the remark before Theorem 2.1 in Tang (2006). Point 4 is satisfied e.g. if $L(x) \sim a(\log(x))^b$, $x \to \infty$, which is the case for Pareto, Burr and loggamma distributions.*

Lemma 4.1. *Let us consider the equation* (4.5) *with $\alpha \in (0, 1)$.*

1. *There exists at least one solution $\tilde{\alpha} \in (0, 1)$ to* (4.5)*,*

2. *There exists a unique solution $\tilde{\alpha} \in (0, 1)$ to* (4.5) *for Gamma, Weibull, Pareto (for shape parameter greater than 2.15) and loggamma (for shape parameter greater than 2.15) distributions of X_n.*

Remark 4.2. *For Pareto and loggamma distributions with shape parameters close to 2, the equation* (4.5) *may have several solutions. Low values of shape parameter are unlikely to be used in practice in our model since the emergence factor α is based on variance and variance of Pareto or loggamma explodes when the shape parameter approaches 2.*

We have found another interesting property of the one-year premium risk vs. the ultimate premium risk: the distribution of the true one-year premium risk can have a different tail behaviour than the distribution of the ultimate premium risk. If the tail of the one-year risk is heavier than the tail of the ultimate risk, as in point 2 from Theorem 4.2, then we expect that the one-year risk is higher than the ultimate risk at all high confidence levels, and we have already observed such a property in the previous section. If the tail of the one-year risk is lighter than the tail of the ultimate risk, as in point 4 from Theorem 4.2, then we expect that the one-year risk is lower than the ultimate risk at all high confidence levels.

Example 4.3. We consider $X_n \sim Weibull(\tau)$. We want to illustrate point 2 from Theorem 4.2, hence we are interested in $\tau \geq 1$. In order to calculate $VaR_\gamma[BE_1]$, we use Monte Carlo simulations and we generate a sample of size 10^6. The results are presented in Figure 4. For $\tau = 10$ and $\alpha = 85\%$, the one-year risk is higher than the ultimate risk at $\gamma = 99.5\%$ by about 0.5%, and at $\gamma = 99.99\%$ by about 10%. If the emergence pattern formula (2.1) is applied in this claims development model to measure the one-year premium risk, then the true one-year risk is underestimated, respectively, by about 15% and 23% at $\gamma = 99.5\%$ and 99.99%. For $\alpha = 85\%$, the one-year risk is higher than the ultimate risk at the confidence level $\gamma = 99.5\%$, and also at all confidence levels $\gamma > 99.5\%$, if we consider Weibull distributions

with $\tau > 7.5$. As discussed in Example 4.2, negatively skewed Weibull distributions are not common in applications but they should not be excluded a priori from considerations.
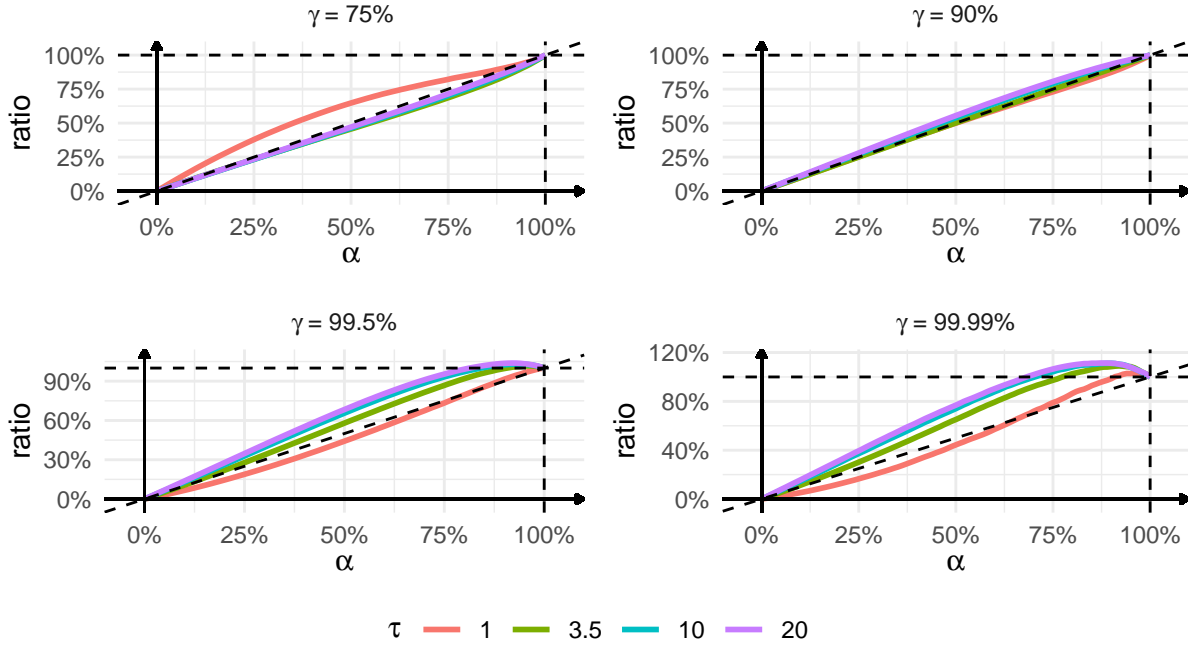

Figure 4: The ratios $\frac{VaR_\gamma[BE_1 - \mathbb{E}[BE_1]]}{VaR_\gamma[X_n - \mathbb{E}[X_n]]}$ in the model where $X_n \sim Weibull(\tau)$ and $BE_1|X_n$ comes from the Hertig's Lognormal model.

We illustrate consequences of point 4 from Theorem 4.2. We consider $X_n \sim Pareto(1, \theta)$ with $1 - F_{X_n}(x) = x^{-\theta}, x > 1$. Then, $\log(X_n)$ has exponential distribution with parameter θ and we can use the results from Example 4.1 to calculate $VaR_\gamma[BE_1]$. The results are presented in Figure 5. For $\theta = 5$ and $\alpha = 50\%$, the one-year risk is lower than the ultimate risk at $\gamma = 99.5\%$ by about 68% and at $\gamma = 99.99\%$ by about 81%. If the emergence pattern formula (2.1) is applied in this claims development model to measure the one-year premium risk, then the true one-year risk is overestimated, respectively, by about 58% and 165% at $\gamma = 99.5\%$ and 99.99%. We can also observe that the one-year risk can be higher than the ultimate risk at low confidence levels. Finally, even though point 4 from Theorem 4.2 shows that $VaR_\gamma[BE_1 - \mathbb{E}[BE_1]] = o(VaR_\gamma[BE_1 - \mathbb{E}[BE_1]])$ for $\gamma \to 1$, we end up with $VaR_\gamma[BE_1 - \mathbb{E}[BE_1]] \approx \alpha^3 VaR_\gamma[BE_1 - \mathbb{E}[BE_1]]$ for γ in the range 99.99% − 99.999% in our experiments. □

Theorems 4.1 and 4.2 show that the common belief among actuaries that the one-year risk is always lower than the ultimate risk does not always hold to be true. We point out that the relation between the one-year risk and the ultimate risk also depends on the risk measure, not only on the claims development process. If standard deviation is used to measure the
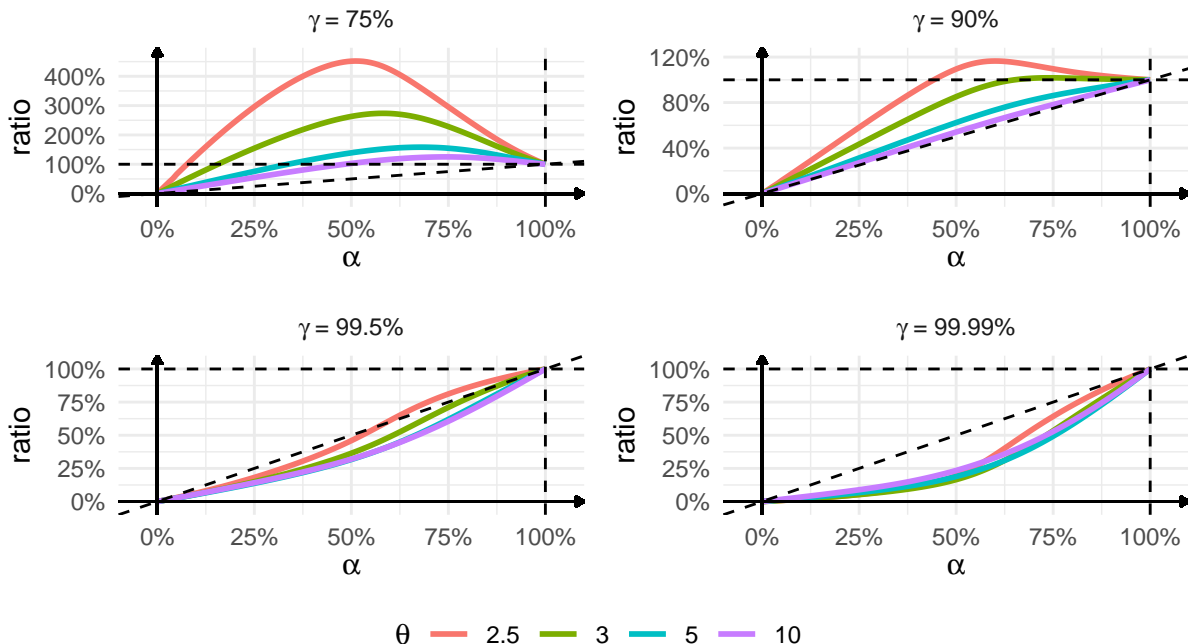
Figure 5: The ratios $\frac{VaR_\gamma[BE_1 - \mathbb{E}[BE_1]]}{VaR_\gamma[X_n - \mathbb{E}[X_n]]}$ in the model where $X_n \sim Pareto(\theta)$ and $BE_1|X_n$ comes from the Hertig's Lognormal model.

risk, then point 1 in Theorems 3.2, 3.4, 3.6, 4.1, 4.2 shows that the true one-year premium risk is always lower than the ultimate premium risk in all our claims development models (in Theorems 3.2, 3.4, 3.6, we prove that $\alpha \in (0,1)$ and, consequently, we assume that $\alpha \in (0,1)$ in Theorems 4.1, 4.2). This conclusion agrees with the results from Wüthrich, Merz (2015) where the authors use mean square error of prediction to compare the true one-year reserve risk with the ultimate reserve risk. However, in this paper, we focus on Value-at-Risk. If Value-at-Risk is used to measure the risk, then the relation between the one-year risk and the ultimate risk depends on the confidence level and the claims development process. In many claims development models, including the models presented in Section 3 which are commonly used in claims reserving, if a sufficiently high confidence level for Value-at-Risk is chosen, then the true one-year premium risk is lower than the ultimate premium risk, which agrees with the common belief. However, there are claims development models, and examples of such model are discussed in Sections 4.1-4.2, where the true one-year premium risk can be higher than the ultimate premium risk at all high confidence levels. Actuaries should be aware of this property in order not to misunderstand and underestimate the one-year risk. On the other side, the true one-year premium risk can be much less dangerous than it is believed by users of the linear emergence pattern formula (2.1). Point 4 from Theorem 4.2 proves that, in a claims development model with the emergence pattern of the ultimate

loss given by (4.4), the tail index of the Pareto-type distribution of the true one-year risk is higher than the tail index of the Pareto-type distribution of the ultimate risk, whereas from Theorem 2.1 we know that the tail index of the Pareto-type distribution of the one-year risk predicted with (2.1) remains equal to the tail index of the distribution of the ultimate loss.

5   Proofs

By $f(x) \sim Cg(x)$ we mean the limit $\lim_{x \to \infty} \frac{f(x)}{g(x)} = C$, by $f(x) = o(g(x))$ we mean the limit $\lim_{x \to \infty} \frac{f(x)}{g(x)} = 0$, unless different limits are specified.

First, we collect some results which we use in the proofs below.

Lemma 5.1. *We consider two independent r.v. Z_1 and Z_2. Let $\mathcal{S}$ denote the class of subexponential distributions, and $\mathcal{P}(\theta)$ denotes the class of Pareto-type distributions with tail index θ.*

1. *If Z_1 and Z_2 are light-tailed, then $aZ_1 + bZ_2$ is light-tailed, for any $a, b \in \mathbb{R}$,*

2. *If $Z_1 \in \mathcal{S}$ and $1 - F_{Z_2}(x) = o\bigl(1 - F_{Z_1}(x)\bigr)$, then $Z_1 + Z_2 \in \mathcal{S}$ and $1 - F_{Z_1+Z_2}(x) \sim 1 - F_{Z_1}(x)$ [Lemma 9 in Geluk, De Vries (2006)],*

3. *If $Z_1 \in \mathcal{S}$ and Z_2 is a positive, bounded r.v., then $Z_1 \cdot Z_2 \in \mathcal{S}$ [Corollary 2.5 in Cline, Samorodnitsky (1994)]. If $Z_1 \in \mathcal{S}$ such that $\limsup_{x \to \infty} \frac{1 - F_{Z_1}(vx)}{1 - F_{Z_1}(x)} < 1$ for some $v > 1$ and Z_2 is a positive r.v. such that $1 - F_{Z_2}(vx) = o\bigl(1 - F_{Z_1}(x)\bigr)$ for some $v > 0$, or $1 - F_{Z_2}(vx) = o\bigl(1 - F_{Z_2}(x)\bigr)$ for some $v > 1$, then $Z_1 \cdot Z_2 \in \mathcal{S}$ [Theorem 2.1 and Corollary 2.1.2 in Tang (2006)],*

4. *If $Z_1 \in \mathcal{P}(\theta)$ is a positive r.v. and Z_2 is a positive r.v. with a light-tailed distribution, then $Z_1 \cdot Z_2 \in \mathcal{P}(\theta)$ and $1 - F_{Z_1 \cdot Z_2}(x) \sim \mathbb{E}[Z_2^\theta]\bigl(1 - F_{Z_1}(x)\bigr)$ [Proposition 1.3.9.b in Mikosch (1999)],*

5. *If Z_1, Z_2 have all moments finite, then $Z_1 + Z_2$ and $Z_1 \cdot Z_2$ have all moments finite.*

The proof of Theorem 2.1. <u>Points 1 and 4</u> are obvious. <u>Points 2 and 3</u> follow from Lemma 5.1 if we choose a degenerate distribution for Z_2. $\qquad\square$

The proofs of Theorems 3.1, 3.3, 3.5. We use the probabilistic assumptions behind the claims development processes, derive the distributions and use appropriate parameterisations.

<u>Gaussian ILR.</u> By Result 4.6 in Johnson-Wichern (2007), if (Z_1, Z_2) has a bivariate normal distribution with $Z_1 \sim N(\mu_{Z_1}, \sigma_{Z_1}^2)$, $Z_2 \sim N(\mu_{Z_2}, \sigma_{Z_2}^2)$ and zero correlation, then

$$Z_1 | Z_1 + Z_2 = z \sim N\left(\mu_{Z_1} \frac{\sigma_{Z_2}^2}{\sigma_{Z_1}^2 + \sigma_{Z_2}^2} + \frac{\sigma_{Z_1}^2}{\sigma_{Z_1}^2 + \sigma_{Z_2}^2}(z - \mu_{Z_2}); \frac{\sigma_{Z_1}^2 \sigma_{Z_2}^2}{\sigma_{Z_1}^2 + \sigma_{Z_2}^2}\right). \tag{5.1}$$

We set $Z_1 = X_1 = \epsilon_1 \sim N(\mu_1, E\sigma_1^2)$, $Z_2 = X_n - X_1 = \sum_{i=2}^{n} \epsilon_i \sim N(\mu - \mu_1, E(\sigma^2 - \sigma_1^2))$, $Z = Z_1 + Z_2 = X_n \sim N(E\mu, E\sigma^2)$. The pair (Z_1, Z_2) has a bivariate normal distribution with zero correlation. We apply (5.1) and get the conditional distribution:

$$X_1|X_n = x \sim N\left(E\mu_1 + \frac{\sigma_1^2}{\sigma^2}(x - E\mu); E\frac{\sigma_1^2(\sigma^2 - \sigma_1^2)}{\sigma^2}\right). \tag{5.2}$$

Since $BE_1 = X_1 + E(\mu - \mu_1)$, we derive our key conditional distribution:

$$BE_1|X_n = x \sim N\left(\frac{\sigma_1^2}{\sigma^2}x + \left(1 - \frac{\sigma_1^2}{\sigma^2}\right)E\mu; E\frac{\sigma_1^2(\sigma^2 - \sigma_1^2)}{\sigma^2}\right) \tag{5.3}$$

We note that $X_n \sim N(E\mu, E\sigma^2)$ and $BE_1 \sim N(E\mu, E\sigma_1^2)$. We set

$$\mu_{X_n} = \mathbb{E}[X_n] = E\mu, \quad \sigma_{X_n}^2 = Var[X_n] = E\sigma^2, \quad \alpha^2 = \frac{Var[BE_1]}{Var[X_n]} = \frac{\sigma_1^2}{\sigma^2}, \tag{5.4}$$

and we can now prove the formulas for the distributions of X_n, BE_1 and $BE_1|X_n$ by using the new parameters from (5.4). Since $0 < \sigma_1^2 < \sigma^2$, we conclude that $\alpha \in (0, 1)$.

Hertig's model. The development process (3.7) satisfies the following dynamics

$$\log(X_i) = \log(X_{i-1}) + \log(\epsilon_i) = \log(X_{i-1}) + \eta_i, \quad \text{where: } \eta_i \sim N(m_i, s_i^2), \quad i = 2, ..., n,$$

which fits the Gaussian ILR model. By (5.2), we state the conditional distribution:

$$\log(X_1)|\log(X_n) = x \sim N\left(m_1 + \frac{s_1^2}{s^2}(x - m); \frac{s_1^2(s^2 - s_1^2)}{s^2}\right),$$

and, using the definition of BE_1, see (3.8), we can derive our key conditional distribution:

$$BE_1|X_n = x \sim LogN\left(\frac{s_1^2}{s^2}\log(x) + \left(1 - \frac{s_1^2}{s^2}\right)\left(m + \frac{s^2}{2}\right); \frac{s_1^2(s^2 - s_1^2)}{s^2}\right). \tag{5.5}$$

We deduce that $X_n \sim LogN(m, s^2)$ and $BE_1 \sim LogN(m + \frac{1}{2}(s^2 - s_1^2), s_1^2)$. We can set

$$\mu_{X_n} = \mathbb{E}[X_n] = e^{m + \frac{1}{2}s^2}, \quad \psi_{X_n}^2 = \frac{Var[X_n]}{\left(\mathbb{E}[X_n]\right)^2} = e^{s^2} - 1, \quad \alpha^2 = \frac{Var[BE_1]}{Var[X_n]} = \frac{e^{s_1^2} - 1}{e^{s^2} - 1}, \tag{5.6}$$

and we solve (5.6) to represent the parameters:

$$s^2 = \log(1 + \psi_{X_n}^2), \quad s_1^2 = \log(1 + \alpha^2\psi_{X_n}^2), \quad m = \log(\mu_{X_n}) - \frac{1}{2}\log(1 + \psi_{X_n}^2). \tag{5.7}$$

Finally, we introduce three new parameters: $\tilde{s} = s$, $\tilde{m} = m$ and $\tilde{\alpha}^2 = \frac{s_1^2}{s^2}$. We can now deduce the formulas for the distributions of X_n, BE_1 and $BE_1|X_n$. Since $0 < s_1^2 < s^2$, we conclude that $\alpha \in (0, 1)$.

Over-Dispersed Poisson: It is a known result that if $Z_1 \sim Poiss(\mu_{Z_1})$, $Z_2 \sim Poiss(\mu_{Z_2})$ and Z_1 is independent of Z_2, then

$$Z_1|Z_1 + Z_2 = z \sim Bin\left(z, \frac{\mu_{Z_1}}{\mu_{Z_1} + \mu_{Z_2}}\right), \tag{5.8}$$

Using the definitions of the Over-Dispersed Poisson distribution and the best estimate of the ultimate loss in this model, we can deduce the conditional distributions:

$$X_1|X_n = x \sim \psi \cdot Bin\left(\frac{x}{\psi}; \omega_1\right), \quad BE_1|X_n = x \sim \psi \cdot Bin\left(\frac{x}{\psi}; \omega_1\right) + \mu(1 - \omega_1). \tag{5.9}$$

Moreover, $X_n \sim \psi Poiss(\mu/\psi)$ and $BE_1 \sim \psi Poiss(\mu\omega_1/\psi) + \mu(1 - \omega_1)$. We set

$$\mu_{X_n} = \mathbb{E}[X_n] = \mu, \quad \psi_{X_n} = \frac{Var[X_n]}{\mathbb{E}[X_n]} = \psi, \quad \alpha^2 = \frac{Var[BE_1]}{Var[X_n]} = \omega_1. \tag{5.10}$$

The formulas for the distributions X_n, BE_1 and $BE_1|X_n$ can now be established by using the new parameters from (5.10). Since $0 < \omega_1 < 1$, $\alpha \in (0, 1)$. $\qquad\square$

The proofs of Theorems 3.2, 3.4, 3.6. <u>Point 1.</u> The results can be deduced from the distributions derived in Theorems 3.1, 3.3, 3.5 and the property that $\alpha \in (0, 1)$.

<u>Point 2 from Theorem 3.2.</u> We use the well-known formula for the quantile of normal distribution together with $\alpha \in (0, 1)$ and $VaR_\gamma[X_n] > \mathbb{E}[X_n]$.

<u>Points 2-3 from Theorem 3.4.</u> We use the distribution (3.11) and the formulas

$$VaR_\gamma[X_n] = e^{\tilde{m} + \tilde{s}\Phi^{-1}(\gamma)}, \quad VaR_\gamma[BE_1] = e^{\tilde{m} + (1-\tilde{\alpha}^2)\frac{\tilde{s}^2}{2} + \tilde{\alpha}\tilde{s}\Phi^{-1}(\gamma)}, \tag{5.11}$$

which describe the quantiles of the lognormal distributions of BE_1 and X_n. We can now solve the inequality $VaR_\gamma[BE_1] < VaR_\gamma[X_n]$, which is equivalent to $VaR_\gamma[BE_1 - \mathbb{E}[BE_1]] < VaR_\gamma[X_n - \mathbb{E}[X_n]]$, and find γ^*. The limit $\lim_{\gamma \to 1} \frac{VaR_\gamma[BE_1 - \mathbb{E}[BE_1]]}{VaR_\gamma[X_n - \mathbb{E}[X_n]]}$ can be directly calculated using the formula (5.11).

<u>Point 2 from Theorem 3.6.</u> Let $Y \sim Poiss(\alpha^2\mu)$ and $X \sim Poiss(\mu)$. It is known that $F_X(x) \leq F_Y(x)$ for all x, see e.g. the remark at the end of section 1.2 in Klenke and Mattner (2010). Let us choose arbitrary $k > 0$. Our first goal is to show that $F_X(x + k) < F_Y(x)$ for all sufficiently high x. We only consider $x > \mu$. By Theorem 2 from Short (2013), we have the bounds:

$$\Phi\left(\sqrt{2H(\mu, x)}\right) < F_X(x) < \Phi\left(\sqrt{2H(\mu, x + 1)}\right),$$

where $H(\mu, x) = \mu - x + x\log\left(\frac{x}{\mu}\right)$. Consequently, our assertion holds if $H(\alpha^2\mu, x) > H(\mu, x + k + 1)$ for all sufficiently high x. Using the definition of H, we have to investigate the inequality:

$$x\log\left(\frac{x}{x + k + 1}\right) + k + 1$$
$$- (k + 1)\log(x + k + 1) - x\log(\alpha^2) > \mu(1 - \alpha^2) - (k + 1)\log\mu. \tag{5.12}$$

25

We can prove that $\lim_{x\to+\infty}\left(x\log\left(\frac{x}{x+k+1}\right)+k+1\right)=0$. Moreover, $\lim_{x\to+\infty}\left(-(k+1)\log(x+k+1)-x\log\left(\alpha^2\right)\right)=+\infty$, since $\alpha\in(0,1)$. Hence, the inequality (5.12) is satisfied for all sufficiently high x.

Let us consider our claims development model. For simplicity we assume that $\psi_{X_n}=1$. Let $Y=BE_1-(1-\alpha^2)\mu$. By the arguments from above, we know that $F_Y(x)>F_{X_n}\left(x+(1-\alpha^2)\mu\right)$ for all sufficiently high x. Since the distribution F_{BE_1} is the distribution F_Y shifted to the right by $(1-\alpha^2)\mu$, we deduce that $F_{BE_1}(x)>F_{X_n}(x)$ for all sufficiently high x. The result is proved, and can be extended to arbitrary ψ_{X_n}.

<u>Point 3 from Theorem 3.6.</u> By Lemma 9.1 from Steutel and Van Harn (2004), if $X\sim Poiss(\mu)$, then $-\log(1-F_X(x))\sim x\log(x)$. Consequently, we can derive the limit:

$$\lim_{t\to\infty}\frac{\log(1-F_{BE_1}(tx))}{\log(1-F_{X_n}(t))}=x,\quad x>0.\tag{5.13}$$

We use the arguments from Propositions 2.2 and 2.6.vi from Resnick (2007). Since the limit (5.13) holds for a sequence of functions, the sequence of inverse functions also converges. We can establish the limit:

$$\lim_{t\to\infty}\frac{F_{BE_1}^{-1}\left(1-(1-F_{X_n}(t))^y\right)}{t}=y,\quad y>0.\tag{5.14}$$

By the property that $F_{X_n}(F_{X_n}^{-1}(t))\sim t, t\to 1$, we end up with

$$\lim_{t\to\infty}\frac{F_{BE_1}^{-1}(1-1/t^y)}{F_{X_n}^{-1}(1-1/t)}=y,\quad y>0.\tag{5.15}$$

Finally, we choose $y=1$ in (5.15). We conclude that $VaR_\gamma[BE_1-\mathbb{E}[BE_1]]\sim VaR_\gamma[X_n-\mathbb{E}[X_n]]$ for $\gamma\to 1$. $\qquad\square$

The proof Theorem 4.1. <u>Point 1</u> is obvious.

<u>Point 2.</u> The first statement follows from Lemma 5.1. The result for light-tailed distributions is obvious. We consider a subexponential distribution with all moments finite. We know that $Z_1=\alpha^2 X_n+(1-\alpha^2)\mu_{X_n}$ is subexponential with all moments finite. The random variable $Z_2=\sqrt{\alpha^2(1-\alpha^2)}\sigma_{X_n}\xi$ has normal distribution and is light-tailed. We observe that

$$\lim_{x\to\infty}\frac{1-F_{Z_2}(x)}{1-F_{Z_1}(x)}=\lim_{x\to\infty}\frac{e^{tx}(1-F_{Z_2}(x))}{e^{tx}(1-F_{Z_1}(x))}=0,\quad for\ some\ t>0,\tag{5.16}$$

since the numerator converges to zero (Z_2 is light-tailed) and the denominator converges to infinity (Z_1 is subexponential). We deduce that $BE_1=Z_1+Z_2$ has a subexponential distribution with all moments finite.

We prove the second statement. We can derive the lower bound for the convolution:

$$\begin{aligned}1-F_{BE_1}(x) &= Pr\left(\alpha^2 X_n+(1-\alpha^2)\mu_{X_n}+Z_2>x\right)\geq Pr\left((1-\alpha^2)\mu_{X_n}+Z_2>x,\alpha^2 X_n\leq x\right)\\ &= Pr\left((1-\alpha^2)\mu_{X_n}+Z_2>x\right)F_{X_n}\left(\frac{x}{\alpha^2}\right),\quad x>0.\end{aligned}$$

Using a bound for the tail of normal distribution, see Gordon (1941), the assumption that $\lim_{x\to\infty}(1-F_{X_n}(x))e^{vx^2}=0$ and $\lim_{x\to\infty}\frac{e^{wx}}{x}=+\infty$ for all $w>0, v>0$, we can show that

$$
\begin{aligned}
\lim_{x\to\infty}\frac{1-F_{BE_1}(x)}{1-F_{X_n}(x)} &\geq \lim_{x\to\infty}\frac{Pr\big((1-\alpha^2)\mu_{X_n}+Z_2>x\big)F_{X_n}(\frac{x}{\alpha^2})}{1-F_{X_n}(x)}\\
&\geq \frac{1}{\sqrt{2\pi}}\lim_{x\to\infty}\frac{e^{-\frac{(u(x))^2}{2}}F_{X_n}(\frac{x}{\alpha^2})}{(u(x)+\frac{1}{u(x)})(1-F_{X_n}(x))}\\
&\geq K\lim_{x\to\infty}\frac{e^{wx}}{e^{vx^2}(u(x)+\frac{1}{u(x)})(1-F_{X_n}(x))}=+\infty,
\end{aligned}
$$

where $u(x)=\frac{x-(1-\alpha^2)\mu_{X_n}}{\sqrt{\alpha^2(1-\alpha^2)\sigma_{X_n}^2}}$ and K,v,w denote some positive constants. We can deduce that $1-F_{BE_1}(x)>1-F_{X_n}(x)$ for all sufficiently high x, and $VaR_\gamma[BE_1]>VaR_\gamma[X_n]$ for all sufficiently high γ.

<u>Point 3.</u> Z_1 and Z_2 are still defined as in point 1. By Lemma 5.1 and similar arguments as in (5.16), $1-F_{BE_1}(x)\sim 1-F_{Z_1}(x)\sim\alpha^{2\theta}\big(1-F_{X_n}(x)\big)$, and $1-F_{Z_2}(x)=o\big(1-F_{Z_1}(x)\big)$. We can deduce that BE_1 inherits the tail from X_n. Using Lemma 5.1 again, we prove that $1-F_{BE_1-\mu}(x)\sim\alpha^{2\theta}(1-F_{X_n-\mu}(x))$. The result for VaR follows from the arguments used in Proposition 2.6.vi from Resnick (2007). $\qquad\square$

The proof Theorem 4.2. <u>Points 1.</u> Using independence between X_n and $\xi\sim N(0,1)$, as well as moments for log-normal distribution, we can calculate the moments:

$$
\mathbb{E}\big[BE_1^k\big]=\mathbb{E}\Big[(X_n)^{k\tilde\alpha^2}\Big]e^{k(1-\tilde\alpha^2)\big(\tilde m+\frac{\tilde s^2}{2}\big)+\frac{1}{2}k^2\tilde\alpha^2(1-\tilde\alpha^2)\tilde s^2},\quad k=1,2. \tag{5.17}
$$

We require that $\mathbb{E}[BE_1]=\mathbb{E}[X_n]$ and $Var[BE_1]=\alpha^2 Var[X_n]$. If we use (5.17), then we derive the equation:

$$
\frac{\mathbb{E}\big[(X_n)^{2\tilde\alpha^2}\big]}{\big(\mathbb{E}\big[(X_n)^{\tilde\alpha^2}\big]\big)^2}e^{\tilde\alpha^2(1-\tilde\alpha^2)\tilde s^2}=1+\alpha^2\psi_{X_n}^2. \tag{5.18}
$$

We can see that we can choose any $\tilde s^2>0$ and look for $\tilde\alpha$ which satisfies (5.18). The most obvious choice is $\tilde s^2$ from Theorem 3.3 so that our new emergence pattern formula coincides with the true emergence pattern formula in the Hertig's model if X_n has lognormal distribution. The parameter $\tilde m$ is next derived from $\mathbb{E}[BE_1]=\mathbb{E}[X_n]$.

<u>Point 2.</u> Since X_n has a moment generating function and $\tilde\alpha\in(0,1)$, then $Z_1=X_n^{\tilde\alpha^2}$ has a moment generating function and is light-tailed. The lognormal distribution $Z_2=e^{(1-\tilde\alpha^2)\big(\tilde m+\frac{\tilde s^2}{2}\big)+\sqrt{\tilde\alpha^2(1-\tilde\alpha^2)}\tilde s\xi}$ is subexponential and satisfies $\limsup_{x\to\infty}\frac{1-F_{Z_2}(vx)}{1-F_{Z_2}(x)}<1$ for some $v>1$, see the remark before Theorem 2.1 in Tang (2006). Moreover, we observe that

$$
\lim_{x\to\infty}\frac{1-F_{Z_1}(vx)}{1-F_{Z_2}(x)}=\lim_{x\to\infty}\frac{e^{tvx}(1-F_{Z_1}(vx))}{e^{t(v-1)x}e^{tx}(1-F_{Z_2}(x))}=0,\quad for\ some\ t>0,\ for\ all\ v\geq 1, \tag{5.19}
$$

27

since the numerator converges to zero and the denominator converges to infinity. By Lemma 5.1, $BE_1 = Z_1 Z_2$ is subexponential with all moments finite. Finally, we prove that

$$\lim_{x \to \infty} \frac{1 - F_{X_n}(x)}{1 - F_{BE_1}(x)} = \lim_{x \to \infty} \frac{e^{tx}(1 - F_{X_n}(x))}{e^{tx}(1 - F_{BE_1}(x))} = 0, \quad for \ some \ t > 0,$$

since the numerator converges to zero and the denominator converges to infinity. We can deduce that $1 - F_{BE_1}(x) > 1 - F_{X_n}(x)$ for all sufficiently high x, and $VaR_\gamma[BE_1] > VaR_\gamma[X_n]$ for all sufficiently high γ.

<u>Point 3.</u> If the distribution of X_n satisfies the assumption that $\limsup_{x \to \infty} \frac{1 - F_{X_n}(vx)}{1 - F_{X_n}(x)} < 1$ for some $v > 1$, then the distribution of $X_n^{\tilde{\alpha}^2}$, given by $F_{X_n^{\tilde{\alpha}}}(x) = F_{X_n}(x^{\frac{1}{\tilde{\alpha}}})$, also satisfies this assumption. If $X_n^{\tilde{\alpha}}$ has a light-tailed distribution, then the result follows from point 2. Let $X_n^{\tilde{\alpha}}$ be subexponential. We recall that we put $Z_2 = e^{(1-\tilde{\alpha}^2)(\tilde{m} + \frac{\tilde{s}^2}{2}) + \sqrt{\tilde{\alpha}^2(1-\tilde{\alpha}^2)}\tilde{s}\xi}$. We show that

$$\lim_{x \to \infty} \frac{1 - F_{Z_2}(vx)}{1 - F_{Z_2}(x)} \leq \lim_{x \to \infty} \frac{(\log(x) - a)^2 + b^2}{(\log(vx) - a)(\log(x) - a)} \frac{e^{-\frac{(\log(vx)-a)^2}{2b^2}}}{e^{-\frac{(\log(x)-a)^2}{2b^2}}}$$

$$= \lim_{x \to \infty} \frac{(\log(x) - a)^2 + b^2}{(\log(vx) - a)(\log(x) - a)} e^{\frac{-\log(v)}{2b^2}(2(\log(x)-a) + \log(v))} = 0, \quad for \ all \ v > 1,$$

where a, b are the parameters of Z_2, and we use bounds for the tail of normal distribution, see Gordon (1941). The statement now follows from Lemma 5.1.

<u>Point 4.</u> We can immediately prove the relation:

$$1 - F_{X_n^{\tilde{\alpha}^2}}(x) = L(x^{\frac{1}{\tilde{\alpha}^2}})x^{-\frac{\theta}{\tilde{\alpha}^2}} \sim CL(x)x^{-\frac{\theta}{\tilde{\alpha}^2}}.$$

By Lemma 5.1, we derive the asymptotic formula:

$$1 - F_{BE_1}(x) \sim C\mathbb{E}\big[Z_2^{\frac{\theta}{\tilde{\alpha}^2}}\big]L(x)x^{-\frac{\theta}{\tilde{\alpha}^2}}. \tag{5.20}$$

Since X_n has a Pareto-type distribution $1 - F(x) = L(x)x^{-\theta}$, we have the asymptotic formula:

$$VaR_\gamma[X_n] \sim (1 - \gamma)^{-1/\theta}\tilde{L}\Big(\frac{1}{1 - \gamma}\Big), \quad \gamma \to 1, \tag{5.21}$$

where $\tilde{L}$ denotes the de Bruyn conjugate of L, see (2.7) and Proposition 2.5 from Berlaint et al. (2004). Using (5.20)-(5.21) and the arguments from Proposition 2.6.vi from Resnick (2007), we can establish that

$$VaR_\gamma[BE_1] \sim \Big(C\mathbb{E}\big[Z_1^{\frac{\theta}{\tilde{\alpha}^2}}\big]\Big)^{\frac{\tilde{\alpha}^2}{\theta}}(1 - \gamma)^{-\frac{\tilde{\alpha}^2}{\theta}}\tilde{L}\Big(\frac{1}{1 - \gamma}\Big)$$

$$\sim \Big(C\mathbb{E}\big[Z_1^{\frac{\theta}{\tilde{\alpha}^2}}\big]\Big)^{\frac{\tilde{\alpha}^2}{\theta}}(1 - \gamma)^{\frac{1-\tilde{\alpha}^2}{\theta}}VaR_\gamma[X_n], \quad \gamma \to 1,$$

and the result is proved. $\qquad\square$

The proof of Lemma 4.1. Let $k = \tilde{\alpha}^2$. We simply denote $X = X_n, \psi = \psi_{X_n}$. We choose $\alpha \in (0,1)$. We introduce the functions:

$$f(k) = \mathbb{E}[X^k] = \mathbb{E}[e^{k \log(X)}], \quad h(k) = \log f(2k) - 2 \log f(k) + k(1-k) \log(1+\psi^2).$$

It is easy to show that $h(k)$ is continuous on $(0,1)$. We look for $k \in (0,1)$ such that

$$h(k) = \log(1 + \alpha^2 \psi^2). \tag{5.22}$$

We have $h(0) = 0$ and $h(1) = \log(1+\psi^2) > \log(1+\alpha^2\psi^2)$. It is known that $k \mapsto f(k)$ is log-convex. Hence $h'(k) = 2\left(\frac{d}{dx} \log f(x)|_{x=2k} - \frac{d}{dx} \log f(x)|_{x=k}\right) + (1-2k)\log(1+\psi^2) > 0$, and $k \mapsto h(k)$ is increasing on $(0, \frac{1}{2})$. We conclude that there exists at least one solution to (5.22).

<u>Gamma distribution.</u> Let $X \sim Gamma(a,b)$. By straightforward calculations, we show that

$$f(k) = \mathbb{E}[X^k] = b^k \frac{\Gamma(k+a)}{\Gamma(a)}, \quad \frac{d^2}{dk^2} \log f(k) = \frac{d^2}{dk^2} \log \Gamma(k+a),$$

and

$$\begin{aligned} \frac{d^2}{dk^2}h(k) &= 4\frac{d^2}{dx^2} \log \Gamma(x+a)|_{x=2k} - 2\frac{d^2}{dx^2} \log \Gamma(x+a)|_{x=k} - 2\log(1+\psi^2) \\ &= \frac{d^2}{dk^2} \log \left(\frac{\Gamma(2k+a)}{\left(\Gamma(k+a)\right)^2} \right) - 2\log(1+\psi^2) = \frac{d^2}{dk^2}g(k) - 2\log(1+\psi^2), \end{aligned}$$

where we define

$$g(k) = \log \left(\frac{\Gamma(2k+a)}{\left(\Gamma(k+a)\right)^2} \right).$$

Let $a \geq \frac{1}{2}$. By Theorem 5 from Qi et al. (2008), the function g is 3-log-concave, which means that $\frac{d^3}{dk^3} \log g(k) \leq 0$. We can conclude that the function $k \mapsto \frac{d^2}{dk^2} \log g(k)$ is decreasing. We have to consider three cases: 1) $\frac{d^2}{dk^2}g(k) > 2\log(1+\psi^2)$ for all $k \in (0,1)$: the function $k \mapsto h(k)$ is convex on $(0,1)$; 2) $\frac{d^2}{dk^2}g(k) < 2\log(1+\psi^2)$ for all $k \in (0,1)$: the function $k \mapsto h(k)$ is concave on $(0,1)$; 3) There exists $k_0 \in (0,1)$ such that $\frac{d^2}{dk^2}g(k) > 2\log(1+\psi^2)$ for $0 < k < k_0$ and $\frac{d^2}{dk^2}g(k) < 2\log(1+\psi^2)$ for $1 > k > k_0$: the function $k \mapsto h(k)$ is convex on $(0, k_0)$ and concave on $(k_0, 1)$. In all three cases, the properties imply, together with the properties mentioned in the beginning, that there exists a unique solution to (5.22).

Let $a < \frac{1}{2}$. From series representation of trigamma function, we deduce that the second derivative of loggamma function is positive. By inequality (1.10) from Guo et al. (2015), which gives upper and lower bounds on the second derivative of loggamma function, we get

$$\frac{d^2}{dk^2}g(k) \leq \frac{2a+3}{(2k+a+\frac{1}{2})(k+a+1)} + \frac{2a^2 - 4k^2}{(2k+a)^2(k+a)^2}. \tag{5.23}$$

Since $k \mapsto h(k)$ is increasing on $(0, \frac{1}{2})$, we only have to consider $k \in (\frac{1}{2}, 1)$. Using (5.23) and the assumption that $k \in (\frac{1}{2}, 1)$, we derive the inequality

$$\frac{d^2}{dk^2}h(k) = \frac{d^2}{dk^2}g(k) - 2\log\left(1 + \frac{1}{a}\right) \leq 2\left(\frac{1}{a + \frac{3}{2}} - \log\left(1 + \frac{1}{a}\right) + \frac{a^2 - 2k^2}{(2k + a)^2(k + a)^2}\right). \quad (5.24)$$

Let us introduce the function $p(a) = \frac{1}{a + \frac{3}{2}} - \log(1 + \frac{1}{a})$. We have $p(0^+) = -\infty$ and $p(\frac{1}{2}) = \frac{1}{2} - \log 3 < 0$. Moreover, by straightforward calculation, we can check that $p'(a) > 0$ for $a \in (0, \frac{1}{2})$. Consequently, the first two terms on the right hand side of (5.24) are negative. It is clear that the last term in (5.24) is also negative. Hence, we have proved that the function $k \mapsto h(k)$ is concave on $(\frac{1}{2}, 1)$. This property implies, together with the properties mentioned in the beginning, that there exists a unique solution to (5.22).

<u>Pareto distribution.</u> Let $X \sim Pareto(x_0, \theta)$ with $1 - F(x) = x_0^\theta x^{-\theta}$ for $x > x_0$. We assume that $\theta > 2$ so that $Var[X] < \infty$. By straightforward calculations, we can show that

$$f(k) = \mathbb{E}[X^k] = \frac{\theta x_0^k}{\theta - k}, \quad \frac{d}{dk}\log f(k) = \log(x_0) + \frac{1}{\theta - k}.$$

We can also calculate that

$$\frac{d}{dk}h(k) = \frac{2}{\theta - 2k} - \frac{2}{\theta - k} - (2k - 1)\log(1 + \psi^2).$$

Since $\psi^2 = \frac{1}{\theta(\theta - 2)}$ and $\log(1 + x) < x$, we can prove the inequality:

$$\frac{d}{dk}h(k) > \frac{\theta^2 + \theta(6k^2 - 7k) + 2k^2 - 4k^3}{(\theta - 2)\theta(\theta - 2k)(\theta - k)}. \quad (5.25)$$

The denominator of the fraction on the right hand side of (5.25) is always positive and the numerator is a quadratic function in θ. The larger root of this quadratic function is given by $\tilde{\theta}(k) = \frac{1}{2}k(\sqrt{36k^2 - 68k + 41} - 6k + 7)$. One can check that the maximal value of the function $k \mapsto \tilde{\theta}(k)$ on $(0, 1)$ is equal to $\theta^* \approx 2.15$. Consequently, we can conclude that for $\theta > \theta^*$ the first derivative (5.25) is positive for any $k \in (0, 1)$, meaning that the function $k \mapsto h(k)$ is increasing on $(0, 1)$. This property implies, together with the properties mentioned in the beginning, that there exists a unique solution to (5.22).

<u>Weibull distribution.</u> The proof is analogous as for Gamma distribution.

<u>Loggamma distribution.</u> The proof is analogous as for Pareto distribution. $\square$

References

AISAM-ACME. (2007) AISAM-ACME study on non-life long tail liabilities: Reserve risk and risk margin assessment under Solvency II. https://www.amice-eu.org/

Berlaint, J., Goegebeur, Y., Segers, J., Teugels, J. (2004). Statistics of Extremes: Theory and applications. *Wiley Series.*

Bird, C., Cairns, M. (2011). Practical experiences of modelling one-year risk emergence. *GIRO Conference and Exhibition 2011.*

Cline, D., Samorodnitsky, G. (1994). Subexponentiality of the product of independent random variables. *Stochastic Processes and their Applications* 49, 75-98.

England, P., Cairns, M., Scarth, R. (2012). The 1 year view of reserving risk: The "actuary-in-the-box" vs. emergence patterns. *GIRO Conference and Exhibition 2012.*

Geluk, J., De Vries, C. (2006). Weighted sums of subexponential random variables and asymptotic dependence between returns on reinsurance equities. *Insurance: Mathematics and Economics* 38, 39-56.

Gordon, R. D. (1941). Values of Mills' ratio of area to bounding ordinate of the normal probability integral for large values of the argument. *Annals of Mathematical Statistics* 12, 364-366.

Guo, B., Feng, Q., Zhao, J., Luo, Q. (2015). Sharp Inequalities for Polygamma Functions. *Mathematica Slovaca* 65.

Hertig, J. (1985). A Statistical Approach to IBNR-reserves in Marine Reinsurance. *ASTIN Bulletin* 15, 171-183.

Johnson, R. A., Wichern D. W. (2007) Applied Multivariate Statistical Analysis. *Pearson Education, Inc.* Sixth edition.

Klenke, A., Mattner, L. (2010). Stochastic ordering of classical discrete distributions. *Advances in Applied Probability* 42, 392-410.

Lloyd's. (2014). Solvency II Model Validation Guidance. https://www.lloyds.com/∼/media/Files/The-Market/Operating-at-Lloyds/Solvency-II/2014-Guidance/Model-Validation-Guidance-April-2014-v1.pdf

Mikosch, T. (1999). Regular Variation, Subexponentiality and Their Applications in Probability Theory. *Eindhoven University of Technology.*

Qi, F., Guo, S., Gou, B., Chen, S. (2008). A class of k-log-convex functions and their applications to some special functions. *Integral Transforms and Special Functions* 19, 195-200

Radtke, M., Schmidt K. D., Schnaus, A. (2016). Handbook on Loss Reserving *Springer.*

Resnick, S. (2007). Heavy-Tail Phenomena. Probabilistic and Statistical Modeling. *Springer.*

Rolski, T., Schmidli, H., Schmidt, V., Teugels, J. (2001). Stochastic Processes for Insurance and Finance. *Wiley Series.*

Short, M. (2013). Improved inequalities for the Poisson and Binomial distribution and upper tail quantile functions. *International Scholarly Research Notices: Probability and Statistics* 2013, 1-6.

Steutel, F., Van Harn, K. (2004). Infinite Divisibility of Probability Distributions on the Real Line. *Marcel and Dekker.*

Tang, Q. (2006). The subexponentiality of products revisited. *Extremes* 9, 231-241.

Taylor, G., McGuire, G. (2016). Stochastic Loss Reserving Using Generalized Linear Models. *Casualty Actuarial Society* CAS Monograph Series Number 3.

Wüthrich, M. V., Merz, M. (2008). Stochastic Claims Reserving Methods in Insurance. *John Wiley & Sons Ltd, Chichester.*

Wüthrich, M. V., Merz, M. (2015). Claims run-off uncertainty: the full picture. Swiss Finance Institute Research Paper No. 14-69. Available at SSRN: https://ssrn.com/abstract=2524352 or http://dx.doi.org/10.2139/ssrn.2524352.