Available online at www.sciencedirect.com

SCIENCE DIRECT®

Insurance: Mathematics and Economics 37 (2005) 154–172

INSURANCE
MATHEMATICS & ECONOMICS

www.elsevier.com/locate/ime

Some asymptotic results for sums of dependent random variables, with actuarial applications

Roger J.A. Laeven[a,*], Marc J. Goovaerts[a,b], Tom Hoedemakers[b,c]

[a] *University of Amsterdam, Department of Quantitative Economics, Roetersstraat 11, 1018 WB Amsterdam, The Netherlands*
[b] *Catholic University of Leuven, Department of Applied Economics, Naamsestraat 69, B-3000 Leuven, Belgium*
[c] *Catholic University of Leuven, University Center for Statistics, W. de Croylaan 54, B-3001 Heverlee, Belgium*

Received March 2005; received in revised form May 2005; accepted 26 May 2005

Abstract

This paper establishes some asymptotic results for sums of dependent random variables, in the presence of heavy-tailedness conditions. We demonstrate how the derived results can be used to approximate functionals of sums of dependent random variables for which the analytic expression is too cumbersome to work with and which are of major importance in actuarial applications. Numerical illustrations are provided to assess the quality of the asymptotic approximations.

MSC: 62P05; 62E20; 91B30

Keywords: Asymptotics; Heavy-tailed distributions; Sums of random variables; Stochastic dependence

1. Introduction

Many quantities of relevance in actuarial science involve sums of dependent random variables. For example, one may think of the value-at-risk of a stochastically discounted life annuity, or the stop-loss premium for the aggregate claim amount of a number of interrelated policies. Therefore, distribution functions of sums of dependent random variables are of particular interest. Typically such distribution functions are of a complex form. Consequently, in order to compute functionals of sums of dependent random variables, approximation methods are often indispensable.

In case the dependence structure between the elements of the random sum is known, one could use Monte Carlo simulation to obtain empirical distribution functions. However, this is typically a time-consuming approach, in particular if we want to approximate tail probabilities, which would require an excessive number of simulations. Therefore, alternative methods need to be explored.

* Corresponding author. Tel.: +31 20 525 7317; fax: +31 20 525 4349.
E-mail address: r.j.a.laeven@uva.nl (R.J.A. Laeven).

0167-6687/$ – see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.insmatheco.2005.05.009

By now a rich literature is available on the use of conditional expectations and the concept of *comonotonicity* to obtain bounds in convex order for sums of dependent random variables; the interested reader is referred to Rogers and Shi (1995), Kaas et al. (2000) and Dhaene et al. (2002a,b). While these bounds in convex order have proven to be good approximations in case the distribution of the random sum is light-tailed or moderately heavy-tailed, they perform worse when the heavy-tailedness (the volatility, for the (log)normal case) increases; see e.g., Section 4.4 of Dhaene et al. (2002b).

In actuarial applications, it is often merely the tail of the distribution function that is of interest. Indeed, one may think of (tail-)value-at-risk or expected shortfall estimations. Therefore, approximations for functionals of sums of dependent random variables may alternatively be obtained through the use of asymptotic relations. Though asymptotic results are valid *at* infinity, they may as well serve as approximations *near* infinity.

This paper establishes some asymptotic results for the tail probability of sums of dependent random variables, in the presence of heavy-tailedness conditions. In particular, we establish an asymptotic result for the randomly weighted sum of a sequence of non-negative numbers. Furthermore, we establish, under two different sets of conditions, an asymptotic result for the randomly weighted sum of a sequence of independent random variables that consist of a deterministic and a random component. Throughout, the random weights are products of i.i.d. random variables and thus exhibit an explicit dependence structure. Next, we present three actuarial applications that demonstrate how the derived asymptotic results can be employed to approximate certain functionals of sums of dependent random variables. To explore the quality of the asymptotic approximations, we provide several numerical illustrations that compare the asymptotic approximation values to Monte Carlo simulated values. For one of the illustrations, we have also included the approximation results obtained by two other approximation methods based on comonotonicity and moment matching.

The outline of the paper is as follows: in Section 2, we introduce some notational conventions and provide some preliminaries for heavy-tailed distributions. In Section 3, we present the asymptotic results. Section 4 provides a first application of the obtained asymptotic results, concerning the evaluation of stop-loss premiums and quantiles for general stochastically discounted loss reserves. In Section 5, we present a second application that focuses specifically on IBNR loss reserves. Section 6 presents a third application, which considers the problem of setting an initial provision in such a way that the probabilities of ruin in year i, $i = 1, \ldots, n$, are sufficiently small. Numerical results are presented in Section 7. Proofs of the theorems and the comonotonic and moment matching approximation formulas have been gathered in Appendices A and B.

2. Preliminaries for heavy-tailed distributions

First we introduce some notational conventions. For a random variable (r.v.) X with a distribution function (d.f.) F, we denote its tail probability by $\overline{F}(x) = 1 - F(x) = \mathbb{P}(X > x)$. For two independent r.v.'s X and Y with d.f.'s F and G supported on $(-\infty, +\infty)$, we write by $F * G(x) = \int_{-\infty}^{+\infty} F(x - t)\,\mathrm{d}G(t)$, $-\infty < x < +\infty$, the convolution of F and G, and by $F^{*n} = F * \cdots * F$ the n-fold convolution of F. By $F \otimes G$ we denote the d.f. of XY.

Throughout, unless otherwise stated, all limit relations are for $x \to +\infty$. Let $a(x) \geq 0$ and $b(x) > 0$ be two infinitesimals, satisfying

$$l_1 \leq \liminf_{x\to+\infty} \frac{a(x)}{b(x)} \leq \limsup_{x\to+\infty} \frac{a(x)}{b(x)} \leq l_2.$$

We write $a(x) = O(b(x))$ if $l_2 < +\infty$, $a(x) = o(b(x))$ if $l_2 = 0$ and $a(x) \asymp b(x)$ if both $l_2 < +\infty$ and $l_1 > 0$. We write $a(x) \lesssim b(x)$ if $l_2 = 1$, $a(x) \gtrsim b(x)$ if $l_1 = 1$ and $a(x) \sim b(x)$ if both $l_2 = 1$ and $l_1 = 1$. We say that $a(x)$ and $b(x)$ are weakly equivalent if $a(x) \asymp b(x)$, and say that $a(x)$ and $b(x)$ are (strongly) equivalent if $a(x) \sim b(x)$.

A r.v. X or its d.f. F is said to be *heavy-tailed to the right* if $\mathbb{E}[e^{\gamma X}] = +\infty$ for any $\gamma > 0$. Below we introduce some important classes of heavy-tailed distributions. A d.f. F supported on $[0, +\infty)$ belongs to the *subexponential* class $\mathcal{S}$ if

$$\lim_{x\to+\infty}\frac{\overline{F^{*n}}(x)}{\overline{F}(x)}=n,$$

for any (or equivalently, for some) $n \geq 2$. More generally, a d.f. F supported on $(-\infty, +\infty)$ belongs to the class $\mathcal{S}$ if $\tilde{F}(x) = F(x)I(x \geq 0)$ does, where $I(A)$ denotes the *indicator function*, which equals 1 if event A occurs and 0 otherwise. A d.f. F supported on $(-\infty, +\infty)$ belongs to the *long-tailed* class $\mathcal{L}$ if for any real number y (or equivalently, for $y = 1$) we have that

$$\lim_{x\to+\infty}\frac{\overline{F}(x+y)}{\overline{F}(x)}=1.$$

A class of heavy-tailed distributions that is closely related to the classes $\mathcal{S}$ and $\mathcal{L}$, is the class $\mathcal{D}$ of d.f.'s with *dominatedly varying* tails. A d.f. F supported on $(-\infty, +\infty)$ belongs to the class $\mathcal{D}$ if its tail $\overline{F}$ is of dominated variation in the sense that

$$\limsup_{x\to+\infty}\frac{\overline{F}(xy)}{\overline{F}(x)}<+\infty,$$

for any $0 < y < 1$ (or equivalently for some $0 < y < 1$). It is well-known that

$$\mathcal{D}\cap\mathcal{L}\subset\mathcal{S}\subset\mathcal{L}.$$

We remark that the intersection $\mathcal{D}\cap\mathcal{L}$ contains, many useful heavy-tailed distributions. In particular, the intersection $\mathcal{D}\cap\mathcal{L}$ covers the class $\mathcal{R}$, which consists of all d.f.'s with *regularly varying* tails. A d.f. F supported on $(-\infty, +\infty)$ has a regularly varying tail if there is some $\alpha \geq 0$ such that the relation

$$\lim_{x\to+\infty}\frac{\overline{F}(xy)}{\overline{F}(x)}=y^{-\alpha},$$

holds for any $y > 0$. We write $F \in \mathcal{R}_{-\alpha}$.

In addition to the classes of heavy-tailed distributions introduced above, we introduce the class $\mathcal{R}_{-\infty}$ of d.f.'s with *rapidly varying* tails, containing both heavy-tailed and light-tailed distributions. For a d.f. F supported on $(-\infty, +\infty)$ satisfying $\overline{F}(x) > 0$ for any $x > 0$, F belongs to the class $\mathcal{R}_{-\infty}$ if

$$\lim_{x\to+\infty}\frac{\overline{F}(xy)}{\overline{F}(x)}=\begin{cases}0, & \text{for any } y>1;\\ +\infty, & \text{for any } 0<y<1.\end{cases}$$

We remark that the intersection $\mathcal{S}\cap\mathcal{R}_{-\infty}$ contains, e.g. lognormal distributions and certain Weibull distributions, which are prominent distributions in actuarial applications. For an elaboration on the classes of heavy-tailed distributions and the class of rapidly varying tailed distributions, and their applications in insurance and finance, the interested reader is referred to Bingham et al. (1987) and Embrechts et al. (1997).

3. Asymptotic results

In this section, we derive some asymptotic results for the tail probability of sums of dependent r.v.'s, in the presence of heavy-tailedness conditions. In the following, we let $\{X_n, n = 1, 2, \ldots\}$ and $\{Y_n, n = 1, 2, \ldots\}$ denote two sequences of i.i.d. r.v.'s that are mutually independent. We write F_X for the d.f. of a r.v. X of which $X_n, n = 1, 2, \ldots$, are considered to be independent replicates, and assume it is supported on $(-\infty, +\infty)$. Similarly, we write F_Y for the d.f. of a r.v. Y of which $Y_n, n = 1, 2, \ldots$, are considered to be independent replicates, and assume it is supported on $(0, +\infty)$. For notational convenience, we will use the device of independent replicates throughout.

We state the following theorem:

Theorem 3.1. *Let $Z_i = Y_1 Y_2 \cdots Y_i$ and $0 < a_i < +\infty, i = 1, 2, \ldots$ If $F_Y \in \mathcal{S} \cap \mathcal{R}_{-\infty}$, then for each $n = 1, 2, \ldots$,*

$$\mathbb{P}\left(\sum_{i=1}^{n} a_i Z_i > x\right) \sim \sum_{i=1}^{n} \mathbb{P}(a_i Z_i > x). \tag{1}$$

In an actuarial context, as will become apparent in the applications presented in the next sections, the sequence $\{a_i, i = 1, 2 \ldots\}$ can be regarded as a sequence of deterministic payments. The following theorem applies to the case in which the payments consist of both a deterministic and a random component, and the deterministic component is either an additive or a multiplicative constant. The theorem is an extension of Theorems 5.1 and 5.2 of Tang and Tsitsiashvili (2003).

Theorem 3.2. *Let $Z_i = Y_1 Y_2 \cdots Y_i$ and $0 < a_i < +\infty, i = 1, 2, \ldots$ If the following conditions are valid*:

1. $F_X \in \mathcal{D} \cap \mathcal{L}$,
2. $F_Y \in \mathcal{R}_{-\infty}$,

then for each $n = 1, 2, \ldots$,

$$\mathbb{P}\left(\sum_{i=1}^{n} (a_i + X_i) Z_i > x\right) \sim \sum_{i=1}^{n} \mathbb{P}((a_i + X) Z_i > x). \tag{2}$$

Furthermore, for each $n = 1, 2, \ldots$,

$$\mathbb{P}\left(\sum_{i=1}^{n} (a_i X_i) Z_i > x\right) \sim \sum_{i=1}^{n} \mathbb{P}((a_i X) Z_i > x). \tag{3}$$

Corollary 3.1. *Under the conditions stated in Theorem 3.2, we have for each $n = 1, 2, \ldots$ that*

$$\mathbb{P}\left(\sum_{i=1}^{n} (a_i + X_i) Z_i > x\right) - \mathbb{P}\left(\sum_{i=1}^{n-1} (a_i + X_i) Z_i > x\right) \sim \mathbb{P}((a_n + X) Z_n > x). \tag{4}$$

Furthermore, for each $n = 1, 2, \ldots$,

$$\mathbb{P}\left(\sum_{i=1}^{n} (a_i X_i) Z_i > x\right) - \mathbb{P}\left(\sum_{i=1}^{n-1} (a_i X_i) Z_i > x\right) \sim \mathbb{P}(a_n X Z_n > x). \tag{5}$$

Corollary 3.2. *If condition 1. stated in Theorem 3.2 is replaced by "$F_X \in \mathcal{R}_{-\alpha}$", while the other conditions remain the same, then for each $n = 1, 2, \ldots$,*

$$\mathbb{P}\left(\sum_{i=1}^{n} (a_i + X_i) Z_i > x\right) \sim \sum_{i=1}^{n} \overline{F}_X(x - a_i)(\mathbb{E}[Y^\alpha])^i, \tag{6}$$

and

$$\mathbb{P}\left(\sum_{i=1}^{n}(a_i X_i)Z_i > x\right) \sim \overline{F}_X(x)\sum_{i=1}^{n} a_i^{\alpha}(\mathbb{E}[Y^{\alpha}])^{i}. \tag{7}$$

We remark that the particular case of lognormally distributed payments is not covered by Theorem 3.2, since the lognormal distribution does not belong to the intersection $\mathcal{D} \cap \mathcal{L}$. The lognormal distribution has a moderately heavy tail and has been a popular model for loss severity distributions. Hence, we state the following theorem:

Theorem 3.3. *Relations (2)–(5) remain valid if conditions 1. and 2. stated in Theorem 3.2 are replaced by*

$1'$. *X follows a lognormal* (μ_X, σ_X^2) *law*, $-\infty < \mu_X < +\infty$, $\sigma_X > 0$,
$2'$. *Y follows a lognormal* (μ_Y, σ_Y^2) *law*, $-\infty < \mu_Y < +\infty$, $\sigma_Y > 0$,
$3'$. $\sigma_X > \sigma_Y$.

4. Application 1: general stochastically discounted loss reserves

In this section, we consider the problem of determining stop-loss premiums and quantiles for general discounted loss reserves. We denote by the real-valued r.v. X_i from the i.i.d. sequence $\{X_i, i = 1, \ldots, n\}$ the *net loss* in year i. Furthermore, we denote by the positive r.v. Y_i from the i.i.d. sequence $\{Y_i, i = 1, \ldots, n\}$ the *present value discounting factor* from year i to year $i-1$. The two sequences $\{X_i, i = 1, \ldots, n\}$ and $\{Y_i, i = 1, \ldots, n\}$ are considered to be mutually independent.

The *discounted loss reserve* $\tilde{S}_n$ is given by

$$\tilde{S}_n = \sum_{i=1}^{n} X_i \prod_{j=1}^{i} Y_j. \tag{8}$$

Henceforth, we impose that $\mathbb{E}[\tilde{S}_n I(\tilde{S}_n > 0)] < +\infty$, which is equivalent to the conditions that $\mathbb{E}[XI(X > 0)] < +\infty$ and $\mathbb{E}[Y] < +\infty$. Then the *stop-loss premium* of the discounted loss reserve is

$$\mathbb{E}[(\tilde{S}_n - d)_+] = \int_d^{+\infty} \overline{F}_{\tilde{S}_n}(s)\,\mathrm{d}s, \qquad d \in \mathbb{R}_+. \tag{9}$$

Furthermore, the *p-quantile* of the discounted loss reserve is

$$\inf\{s : F_{\tilde{S}_n}(s) \geq p\}, \qquad p \in (0, 1). \tag{10}$$

Approximate values for expressions (9) and (10) can be obtained by using the previously obtained asymptotic results. In particular, if X and Y satisfy the conditions of Theorems 3.2 or 3.3, then for sufficiently large values of d, expression (9) can be approximated by

$$\mathbb{E}[(\tilde{S}_n - d)_+] \approx \sum_{i=1}^{n} \int_d^{+\infty} \overline{F}_{X\prod_{j=1}^{i} Y_j}(s)\,\mathrm{d}s = \sum_{i=1}^{n} \mathbb{E}\left[\left(X\prod_{j=1}^{i} Y_j - d\right)_+\right]. \tag{11}$$

Since the d.f. of $X\prod_{j=1}^{i} Y_j$ will generally not be analytically tractable, Monte Carlo simulation may still be required. However, the number of simulations has been reduced by a factor $(n+1)/2n$.

In case $F_X \in \mathcal{R}_{-\alpha}$, $0 < \alpha < +\infty$, and $F_Y \in \mathcal{R}_{-\infty}$, the asymptotic approximation for expression (9) reduces to

$$\mathbb{E}[(\tilde{S}_n - d)_+] \approx \int_d^{+\infty} \sum_{i=1}^{n} (\mathbb{E}[Y^\alpha])^i \bar{F}_X(s)\,\mathrm{d}s = \sum_{i=1}^{n} (\mathbb{E}[Y^\alpha])^i \mathbb{E}[(X-d)_+]. \tag{12}$$

Furthermore, in this case for sufficiently large values of p, the asymptotic approximation for expression (10) is given by

$$\inf\{s : F_{\tilde{S}_n}(s) \geq p\} \approx \inf\left\{ s : \sum_{i=1}^{n} (\mathbb{E}[Y^\alpha])^i \bar{F}_X(s) \leq 1 - p \right\}. \tag{13}$$

Under the conditions of Theorem 3.3, for sufficiently large values of p, the asymptotic approximation for expression (10) is given by

$$\inf\{s : F_{\tilde{S}_n}(s) \geq p\} \approx \inf\left\{ s : \sum_{i=1}^{n} \bar{F}_{X \prod_{j=1}^{i} Y_j}(s) \leq 1 - p \right\}. \tag{14}$$

We emphasize that the approximation (14) is not in general valid under the conditions of Theorem 3.2; it requires the additional condition that $F_X \in \mathcal{R}_{-\alpha}$, $0 < \alpha < \infty$, since in that case also $F_{\tilde{S}_n} \in \mathcal{R}_{-\alpha}$ (and in which case (14) reduces to (13)).

As an example, we consider $X_i \sim \text{Pareto}(\alpha, \beta)$ and $Y_i \sim \text{lognormal}(\mu, \sigma^2)$, $i = 1, \ldots, n$, where Pareto(α, β) refers to the Pareto distribution with d.f.

$$F_X(x) = 1 - \left(1 + \frac{x}{\beta}\right)^{-\alpha}, \qquad x > 0,$$

with $\alpha > 0$ and $\beta > 0$. Then clearly $F_X \in \mathcal{R}_{-\alpha}$ and $F_Y \in \mathcal{R}_{-\infty}$. Therefore, the asymptotic approximations (12) and (13) are valid. Notice that for the example under consideration, the asymptotic approximations can even be computed analytically! We performed Monte Carlo simulations for both expressions (9) and (10) to assess the quality of the asymptotic approximations (12) and (13), under various specifications of the parameter n. The results are presented in Table 1.

5. Application 2: IBNR loss reserves

In this section, we focus specifically on the discounted loss reserves for claims already ***i***ncurred ***b***ut ***n***ot yet ***r***eported (IBNR). We let the r.v.'s X_{ij}, $i, j = 1, \ldots, t$, denote the claim figures associated with year of origin i and development year j, which are reported in calendar year $i + j - 1$, and we let t correspond to the last observed calendar year. For (i, j) combinations with $i + j \leq t + 1$, X_{ij} concerns an observation, whereas for (i, j) combinations with $i + j > t + 1$, it concerns a future (incurred but not reported) claim. We consider annual development, although the methods can easily be extended to semi-annual, quarterly or monthly development, and we assume, without loss of generality, that the time it takes for the claims to be completely reported, i.e., the total number of developments period for a complete run-off, is equal to t.

Remark 5.1. IBNR reserves need to be set up for claims that are expected to exist but of which the insurer has not been notified as yet. In the same context, there are also RBNS reserves for claims already ***r***eported ***b***ut ***n***ot yet ***s***ettled. Other acronyms are IBNFR, IBNER and RBNFS, where the "F" is for "Fully" and the "E" for "Enough".

In the following, we assume that the r.v.'s X_{ij}, $i, j = 1, \ldots, t$, can be expressed as products of a deterministic component and an i.i.d. random component. In particular, we consider the following model:

$$X_{ij} = a_{ij} L_{ij}, \qquad i, j = 1, \ldots, t, \tag{15}$$

in which L_{ij}, $i, j = 1, \ldots, t$, are i.i.d. r.v.'s and $a_{ij} > 0$, $i, j = 1, \ldots, t$, are positive numbers.

A simple example of the model in (15) is the well-known loglinear regression model, given by

$$\log \vec{X} = \mathbf{R}\vec{\beta} + \vec{\varepsilon}, \quad \vec{\varepsilon} \sim \mathrm{N}(0, \sigma_\varepsilon^2 \mathbf{I}),$$

where $\vec{X} = (X_{11}, \ldots, X_{1t}, X_{21}, \ldots, X_{2,t-1}, \ldots, X_{t1})$ is the vector of historical claim figures, $\vec{\beta} = (\beta_1, \cdots, \beta_p)'$ is a vector of parameters, $\mathbf{R}$ is the regression matrix of exogenous model variables of dimension $t(t+1)/2 \times p$, $\mathbf{I}$ is the identity matrix, $\mathrm{N}(\cdot, \cdot)$ denotes the normal distribution and $\sigma_\varepsilon^2 > 0$ is unknown but fixed.

A well-known and widely applied linear predictor within the loglinear regression model is of the *chain-ladder* type, given by

$$(\mathbf{R}\vec{\beta})_{ij} = \eta_{ij} = \alpha_i + \beta_j, \tag{16}$$

in which α_i is a parameter for the year of origin i and β_j for development year j. Notice that one of the parameters, for example β_1, must be set equal to zero, in order to have a non-singular regression matrix.

The *accumulated IBNR reserve S_t* is given by

$$S_t = \sum_{i=2}^{t} \sum_{j=t+2-i}^{t} a_{ij} L_{ij}. \tag{17}$$

IBNR reserving entails the determination of the *present value* of the future losses. One of the subproblems in this respect is the discounting of the future losses in the run-off triangle, where interest rates and inflation are not known for certain. As in the previous example, we will incorporate stochastic present value discounting factors. We let the positive r.v. Y_k from the i.i.d. sequence $\{Y_k, k = 1, \ldots, t-1\}$ denote the present value discounting factor from year k to year $k-1$ and consider the two sequences $\{L_{ij}, i = 2, \ldots, t; j = t+2-i, \ldots, t\}$ and $\{Y_k, k = 1, \ldots, t-1\}$ to be mutually independent. For notational convenience, we introduce the positive r.v. $Z_k = Y_1 Y_2 \cdots Y_k, k = 1, \ldots, t-1$. Then the *discounted IBNR reserve $\tilde{S}_t$* is given by

$$\tilde{S}_t = \sum_{i=2}^{t} \sum_{j=t+2-i}^{t} a_{ij} L_{ij} Z_{i+j-t-1}. \tag{18}$$

Henceforth, we impose that $\mathbb{E}[\tilde{S}_t] < +\infty$. Approximate values for stop-loss premiums and quantiles for $\tilde{S}_t$ can be obtained by using our derived asymptotic results. In particular, if $\{L_{ij}, i = 2, \ldots, t; j = t+2-i, \ldots, t\}$ and $\{Y_k, k = 1, \ldots, t-1\}$ satisfy the conditions of Theorems 3.2 or 3.3, then for sufficiently large values of d,

$$\mathbb{E}[(\tilde{S}_t - d)_+] \approx \sum_{i=2}^{t} \sum_{j=t+2-i}^{t} a_{ij} \mathbb{E}\left[\left(L Z_{i+j-t-1} - \frac{d}{a_{ij}}\right)_+\right]. \tag{19}$$

Furthermore, if either $F_X \in \mathcal{R}_{-\alpha}$, $0 < \alpha < +\infty$, and $F_Y \in \mathcal{R}_{-\infty}$, or the conditions of Theorem 3.3 apply, then for sufficiently large values of p,

$$\inf\{s : F_{\tilde{S}_t}(s) \geq p\} \approx \inf\left\{s : \sum_{i=2}^{t} \sum_{j=t+2-i}^{t} \bar{F}_{LZ_{i+j-t-1}}\left(\frac{s}{a_{ij}}\right) \leq 1 - p\right\}. \tag{20}$$

Suppose we consider a loglinear regression model with chain-ladder linear predictor to describe the random claims and we use a geometric Brownian motion with drift to represent the stochastic discount factors. We remark that for this specification Theorem 3.3 applies. Furthermore, for this specification the products $L_{ij}Z_{i+j-t-1}$, $i = 2, \ldots, t$; $j = t+2-i, \ldots, t$ are lognormal and therefore the present value of the IBNR reserve becomes a linear combination of dependent lognormal r.v.'s, given by

$$\tilde{S}_t = \sum_{i=2}^{t} \sum_{j=t+2-i}^{t} \mathrm{e}^{\alpha_i+\beta_j+\varepsilon_{ij}-D(i+j-t-1)}, \tag{21}$$

with $\varepsilon_{ij} \sim$ i.i.d. $\mathrm{N}(0, \sigma_\varepsilon^2)$ and

$$D(i) = D_1 + \cdots + D_i = \mu i + \xi B(i), \tag{22}$$

in which $B(i)$ is a standard Brownian motion, μ is a constant rate of interest and the volatility ξ is a positive constant.

Remark 5.2. In this illustration, we start with a given set of parameters and define the reserve as expressed in (21). In a real reserving exercise, one has to build an appropriate statistical model based on the claims in the run-off triangle and to estimate the parameters from this model. This estimation process would introduce a second component of prediction variance, besides the process variance, namely the estimation variance. The discounted IBNR reserve would then be defined as

$$\sum_{i=2}^{t} \sum_{j=t+2-i}^{t} \mathrm{e}^{\hat{\alpha}_i+\hat{\beta}_j+\varepsilon_{ij}-D(i+j-t-1)},$$

in which $\hat{\alpha}_i + \hat{\beta}_j$ is the maximum likelihood estimate of $\alpha_i + \beta_j$. In the current illustration, for reasons of simplicity, the estimation variance is left out of consideration.

It is well-known that stop-loss premiums and quantiles of a lognormal(μ, σ^2) r.v. X are given by

$$\mathbb{E}[(X-d)_+] = \exp\left(\mu + \frac{\sigma^2}{2}\right)\Phi(d_1) - d\Phi(d_2), \tag{23}$$

and

$$F_X^{-1}(p) = \mathrm{e}^{\mu+\sigma\Phi^{-1}(p)}, \qquad p \in (0,1), \tag{24}$$

respectively, where

$$d_1 = \frac{\mu + \sigma^2 - \log(d)}{\sigma},$$

$$d_2 = d_1 - \sigma,$$

and $\Phi(\cdot)$ denotes the d.f. of a standard normal r.v.

Hence, the asymptotic approximations (19) and (20) can be computed quasi-analytically. We choose $\sigma_\varepsilon = 3$, $\mu = -0.07$, $\xi = 0.2$ and $t = 5$ and use the following chain ladder parameters:

$$\begin{pmatrix}\alpha_1\\ \alpha_2\\ \alpha_3\\ \alpha_4\\ \alpha_5\end{pmatrix} = \begin{pmatrix}1.1\\ 1.6\\ 1.9\\ 2.1\\ 2.2\end{pmatrix}, \qquad \begin{pmatrix}\beta_1\\ \beta_2\\ \beta_3\\ \beta_4\\ \beta_5\end{pmatrix} = \begin{pmatrix}0\\ -0.42\\ -0.38\\ -0.87\\ -0.96\end{pmatrix}.$$

In Table 2 we numerically compare the asymptotic approximations with Monte Carlo simulated values.

In Kaas et al. (2000), conditional expectations and the comonotonic copula are used as tools to construct bounds in *convex* order for sums of dependent r.v.'s, whose marginal distributions are known, but with an unknown or complicated joint distribution. We say that a r.v. X is smaller than a r.v. Y in *convex* order if, for any convex function $f(\cdot)$,

$$\mathbb{E}[f(X)] \leq \mathbb{E}[f(Y)].$$

We write $X \leq_{\mathrm{cx}} Y$. Hoedemakers et al. (2003) extend the approach of Kaas et al. (2000) to sums of scalar products of independent r.v.'s and apply it to the problem of IBNR reserving. In particular, with the notation introduced above, they show that the discounted IBNR reserve $\tilde{S}_t$ can be approximated by a convex lower and a convex upper bound given by

$$\tilde{S}_t^l = \sum_{i=2}^{t} \sum_{j=t+2-i}^{t} \mathbb{E}[a_{ij} Z_{i+j-t-1} | \Lambda] \mathbb{E}[L_{ij}], \tag{25}$$

and

$$\tilde{S}_t^u = \sum_{i=2}^{t} \sum_{j=t+2-i}^{t} F^{-1}_{a_{ij} Z_{i+j-t-1}}(U) F^{-1}_{L_{ij}}(V), \tag{26}$$

respectively, in which Λ is a conditioning r.v. independent of the vector $(Z_1, \ldots, Z_{t-1})$, the r.v.'s U, V are uniformly distributed on (0,1), and U, V and Λ are mutually independent. The following relations hold:

$$\tilde{S}_t^l \leq_{\mathrm{cx}} \tilde{S}_t \leq_{\mathrm{cx}} \tilde{S}_t^u. \tag{27}$$

We will refer to this approximation method as the *comonotonic* approach.

Another approximation method that we will consider is based on moment matching techniques. Moment matching is frequently used to approximate the d.f. of a sum of dependent lognormal r.v.'s; see e.g., Vanduffel et al. (2004). Moment matching means in this case that the unknown d.f. of $\tilde{S}_t$ is approximated by a lognormal d.f. in such a way that the first moments are preserved. The first two moments of a lognormal (μ, σ^2) r.v. X are

$$\mathbb{E}[X] = \mathrm{e}^{\mu+(1/2)\sigma^2} \quad \text{and} \quad \mathbb{E}[X^2] = \mathrm{e}^{2\mu+2\sigma^2}.$$

Expressing the parameters μ and σ^2 of the lognormal distribution in terms of the moments $\mathbb{E}[X]$ and $\mathbb{E}[X^2]$ leads to

$$\mu = \log\left(\frac{\mathbb{E}[X]^2}{\sqrt{\mathbb{E}[X^2]}}\right) \quad \text{and} \quad \sigma^2 = \log\left(\frac{\mathbb{E}[X^2]}{\mathbb{E}[X]^2}\right). \tag{28}$$

Hence, the d.f. of the r.v. $\tilde{S}_t$ can be approximated by a lognormal distribution with first two moments (37) and (38) calculated in Appendix B. The coefficients μ and σ^2 of the lognormal moment matching approximation follow from (28). Approximations for the stop-loss premiums and quantiles are then given by (24) and (23).

Dufresne (2002) obtains a lognormal limit distribution for $\tilde{S}_t$ as the volatility tends to zero, which provides a theoretical justification for the use of the lognormal moment matching approximation.

Numerical results of the comonotonic and moment matching approach have been included in Table 2. We refer to Appendix B for further details concerning the calculation of the stop-loss premiums and quantiles for these two approximation methods. The numerical results demonstrate that the asymptotic approximation values generally outperform the comonotonic upper bound and the lognormal moment matching approximation. Because the comonotonic lower bound performed remarkably bad, its values were left out of the table.

6. Application 3: the hurdle race problem

In this section, we consider the problem of setting the initial provision such that a sequence of future payments can be met for a given investment strategy, under a certain bound on the confidence level. We let S_0 denote the provision at time 0. We assume it is invested such that in year i it generates a continuously compounded stochastic return represented by the r.v. R_i from the i.i.d. sequence $\{R_i, i = 1, \ldots, n\}$ with common d.f. F_R supported on $(-\infty, +\infty)$. Furthermore, we let the r.v. X_i from the i.i.d. sequence $\{X_i, i = 1, \ldots, n\}$ with common d.f. F_X supported on $(-\infty, +\infty)$ represent the total net payment to be met in year i. We consider the two sequences $\{X_i, i = 1, \ldots, n\}$ and $\{R_i, i = 1, \ldots, n\}$ to be mutually independent. Then the provision accumulated until the end of year n can be characterized by the discrete time stochastic process $S_n, n = 1, 2, \ldots$, which satisfies the recurrence equation

$$S_0 = x, \quad S_n = \frac{S_{n-1}}{Y_n} - X_n, \qquad n = 1, 2, \ldots, \tag{29}$$

with $Y_n = \exp(-R_n), n = 1, 2, \ldots$

As a first approach, the initial provision S_0^* could be chosen such that the probability of a non-negative provision at the end of a finite time interval within which all payments take place, is sufficiently large. Formally, this gives

$$S_0^* = \inf\{S_0 : \mathbb{P}(S_n \geq 0 | S_0) \geq 1 - \varepsilon\}, \tag{30}$$

for ε sufficiently small. It is not difficult to verify that expression (30) can be rewritten as

$$S_0^* = \inf\left\{S_0 : \mathbb{P}\left(\sum_{i=1}^{n} X_i \prod_{j=1}^{i} Y_j \leq S_0\right) \geq 1 - \varepsilon\right\}. \tag{31}$$

Notice that expression (31) equals the usual definition of a (pseudo-)inverse distribution function of the random sum $\sum_{i=1}^{n} X_i \prod_{j=1}^{i} Y_j$ evaluated in $1-\varepsilon$. It is obvious that analytic expressions for the distribution of $\sum_{i=1}^{n} X_i \prod_{j=1}^{i} Y_j$ are typically rather involved. We leave it to the reader to verify that approximate values for S_0^* can be obtained by using the asymptotic results of Section 3.

A more stringent approach requires the initial provision to be chosen such that the probability of ruin within a finite time interval is sufficiently small. In that case the initial provision is defined as

$$S_0^* = \inf\{S_0 : \psi(S_0, n) \le \varepsilon\} = \inf\{S_0 : \phi(S_0, 1) + \cdots + \phi(S_0, n) \le \varepsilon\}, \tag{32}$$

for ε sufficiently small, where as usual

$$\psi(x, n) = \mathbb{P}(\inf\{k = 1, 2, \ldots : S_k < 0 | S_0 = x\} \le n), \qquad n = 0, 1, \ldots$$

and

$$\phi(x, n) = \mathbb{P}(\inf\{k = 1, 2, \ldots : S_k < 0 | S_0 = x\} = n), \qquad n = 1, 2, \ldots$$

We will adopt a more sophisticated approach that requires the probability that ruin occurs exactly in year i to be bounded from above by some constant ε_i, $i = 1, \ldots, n$. In fact, the surplus process can be regarded to establish a "hurdle race" of taking successfully the "non-negative surplus hurdle" at the end of each year; see also Vanduffel et al. (2003). Hence, the initial provision is defined as

$$S_0^* = \inf\{S_0 : \phi(S_0, i) \le \varepsilon_i, i = 1, \ldots, n\}, \tag{33}$$

for ε_i sufficiently small. Notice that in the above expression for the initial provision, the bounds on the probabilities of ruin in year i are allowed to be time-dependent. Indeed in practice, when a sequence of payment obligations is to be met over a period of time, the intermediate occurrence of a negative surplus will not in general be fatal, since intermediate adjustments to the provision are typically feasible. It may be natural to let the bounds on the probability of ruin in year i to be decreasing in i.

As an example we consider $X_i \sim \text{lognormal}(\mu_X, \sigma_X^2)$ and $Y_i \sim \text{lognormal}(\mu_Y, \sigma_Y^2)$, $i = 1, \ldots, n$, and consider the insurance risk X to be heavier than the financial risk Y, i.e., $\sigma_X > \sigma_Y$. In that case, approximate values for expression (33) can be obtained by Theorem 3.3 and the asymptotic equality (5). In particular, for sufficiently small values ε_i, the right-hand-side of expression (33) can be approximated by

$$\inf\{S_0 : \phi(S_0, i) \le \varepsilon_i, i = 1, \ldots, n\} \approx \inf\left\{S_0 : \mathbb{P}\left(X \prod_{j=1}^{i} Y_j > S_0\right) \le \varepsilon_i, i = 1, \ldots, n\right\}. \tag{34}$$

We choose $n = 5$ and

$$\begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_5 \end{pmatrix} = \begin{pmatrix} 0.005 \\ 0.004 \\ 0.003 \\ 0.002 \\ 0.001 \end{pmatrix}.$$

Numerical results for the asymptotic approximation and for Monte Carlo simulated values are presented in Table 3.

7. Numerical results

For the various quantities under consideration, the tables that we present below display the Monte Carlo simulated values (“Real”), the approximations obtained by the different approximation methods presented, as well as their absolute deviation (Diff.) and their relative deviation (RDiff.) from the Monte Carlo simulated values.

Table 1
Numerical results of Application 1, Section 4

d	Real	Appr.	Diff.	Rdiff. (%)	p	Real	Appr.	Diff.	Rdiff. (%)
$n=3$									
15	1.50	1.36	0.14	9	0.95	16	14	2	15
20	1.28	1.19	0.09	7	0.975	25	22	3	11
25	1.14	1.07	0.07	6	0.99	44	41	3	7
30	1.03	0.98	0.05	5	0.995	69	66	3	4
35	0.95	0.91	0.04	4	0.999	198	194	4	2
40	0.88	0.85	0.03	4					
50	0.78	0.76	0.02	3					
60	0.71	0.70	0.02	2					
80	0.61	0.61	0.01	1					
100	0.55	0.54	0.00	1					
150	0.44	0.44	0.00	0					
200	0.38	0.38	0.00	0					
$n=5$									
20	2.22	1.89	0.33	15	0.95	24	19	5	22
30	1.75	1.56	0.19	11	0.975	36	30	6	17
40	1.48	1.35	0.13	9	0.99	63	57	6	10
60	1.18	1.11	0.07	6	0.995	96	90	6	6
80	1.01	0.96	0.05	5	0.999	274	265	9	3
100	0.90	0.86	0.04	4					
150	0.72	0.70	0.02	3					
200	0.62	0.61	0.01	2					
250	0.56	0.55	0.01	2					
300	0.51	0.50	0.01	2					
$n=10$									
40	2.91	2.41	0.50	17	0.95	40	28	12	30
60	2.22	1.98	0.24	11	0.975	58	45	13	23
80	1.86	1.72	0.14	7	0.99	98	84	14	14
100	1.62	1.54	0.08	5	0.995	148	133	15	10
150	1.28	1.26	0.02	2	0.999	402	390	12	3
200	1.09	1.09	0.00	0					
300	0.87	0.88	−0.01	−1					
400	0.74	0.75	−0.01	−1					

Notes: “Real” vs. approximate values of stop-loss premiums and quantiles for Pareto claim sizes and lognormal present value discounting factors. Fixed parameter values: $\alpha = 1.5$, $\beta = 1$, $\mu = -0.04$, $\sigma = 0.10$ and 5,000,000 simulations.

Table 2
Numerical results of Application 2, Section 5

	Real	Appr. 1	Appr. 2	Appr. 3	Diff. 1	Diff. 2	Diff. 3	Rdiff. 1 (%)	Rdiff. 2 (%)	Rdiff. 3 (%)
d										
7500	1868.0	1771.6	2541.1	2277.6	96.4	−673.1	−409.6	**5.2**	−36.0	−21.9
10000	1743.5	1658.1	2459.2	2165.8	85.4	−715.7	−422.3	**4.9**	−41.0	−24.2
15000	1568.7	1496.9	2333.6	1998.4	71.8	−764.9	−429.7	**4.6**	−48.8	−27.4
20000	1446.7	1383.1	2237.8	1874.1	63.6	−791.1	−427.4	**4.4**	−54.7	−29.5
25000	1354.0	1295.8	2160.0	1775.2	58.2	−806.0	−421.2	**4.3**	−59.5	−31.1
30000	1279.7	1225.4	2094.4	1693.4	54.3	−814.7	−413.7	**4.2**	−63.7	−32.3
40000	1165.7	1116.7	1987.6	1563.3	49.0	−821.9	−397.6	**4.2**	−70.5	−34.1
50000	1080.4	1034.8	1902.4	1462.2	45.6	−822.0	−381.8	**4.2**	−76.1	−35.3
75000	933.3	892.5	1743.6	1280.5	40.8	−810.3	−347.2	**4.4**	−86.8	−37.2
100000	835.6	797.4	1600.2	1154.9	38.2	−764.6	−319.3	**4.6**	−91.5	−38.2
150000	708.5	673.0	1437.8	985.3	35.5	−729.3	−276.8	**5.0**	−102.9	−39.1
200000	626.2	592.0	1323.5	871.7	34.2	−697.3	−245.5	**5.5**	−111.4	−39.2
250000	566.5	533.4	1260.8	788.2	33.1	−694.3	−221.7	**5.8**	−122.6	−39.1
300000	520.7	488.4	1190.4	723.1	32.2	−669.7	−202.4	**6.2**	−128.6	−38.9
400000	453.9	422.6	1081.8	626.8	31.3	−627.9	−172.9	**6.9**	−138.3	−38.1
500000	406.5	375.9	1000.2	557.7	30.6	−593.7	−151.2	**7.5**	−146.1	−37.2
p										
0.95	8650	7863	4814	7555	787	3836	1095	**9.1**	44.3	12.7
0.975	17000	15868	12436	17296	1132	4564	−296	6.7	26.8	**−1.7**
0.99	38957	37496	37490	45306	1461	1467	−6349	**3.8**	3.8	−16.3
0.995	70795	68885	79477	87283	1910	−8682	−16488	**2.7**	−12.3	−23.3
0.999	257090	253021	374188	337364	4069	−117098	−80274	**1.6**	−45.5	−31.2

Notes: "Real" vs. approximate values of stop-loss premiums and quantiles for chain-ladder claim sizes and lognormal present value discounting factors. Fixed parameter values: $\sigma_\varepsilon = 3$, $\mu = -0.07$, $\xi = 0.2$, $t = 5$ and 50,000,000 simulations. "Appr. 1" refers to the asymptotic approximation, "Appr. 2" to the convex upper bound and "Appr. 3" to the lognormal moment matching approximation. The figures displayed in bold correspond to the best approximations, i.e., the ones with the smallest relative deviation from the Monte Carlo simulated values.

Table 3
Numerical results of Application 3, Section 6

σ_Y	Real	Appr.	Diff.	Rdiff. (%)
0.2	8633	8298	335	4
0.7	25610	24494	1116	4

Notes: "Real" vs. approximate values of initial provisions for lognormal net payments and lognormal present value discounting factors. Fixed parameter values: $\mu_X = 0$, $\sigma_X = 3$, $\mu_Y = -0.07$, $n = 5$ and 10,000,000 simulations.

8. Concluding remarks

We derived some asymptotic results for the tail probability of sums of dependent random variables under specific heavy-tailedness conditions. We showed how to apply the obtained results to approximate certain functionals of sums of dependent random variables. For some cases, the asymptotic results provide us with analytic (!) approximation formulas, while for the other cases they lead to a substantial reduction of the number of simulations to be performed. Our numerical results demonstrate that the asymptotic approximations are typically close to the Monte Carlo value. For the "tail functionals" considered, the asymptotic approximations considerably outperform the approximations obtained by two other approximation methods based on comonotonicity and moment matching.

Acknowledgements

We are grateful to Qihe Tang for valuable discussions on the topic. We thank Rob Kaas and an anonymous referee for comments that improved the paper. Marc Goovaerts and Tom Hoedemakers acknowledge the financial support of the Onderzoeksfonds K.U. Leuven (GOA/02: Actuariële, financiële en statistische aspecten van afhankelijkheden in verzekerings-en financiële portefeuilles).

Appendix A. Proofs

Proof of Theorem 3.1. To prove the theorem, we first restate a result from Tang and Tsitsiashvili (2004).

Lemma A.1. *Let F_1, F_2 and G be three d.f.'s. Suppose that $\overline{F}_i(x) > 0$ for any real number x, $F_i(0)G(0) = 0$, $i = 1, 2$, and $G \in \mathcal{R}_{-\infty}$. If $\overline{F}_1(x) \sim \overline{F}_2(x)$, then*

$$\overline{F_1 \otimes G}(x) \sim \overline{F_2 \otimes G}(x). \tag{35}$$

Then, we proceed with the proof of Theorem 3.1. Clearly,

$$\mathbb{P}\left(\sum_{i=1}^{n} a_i Z_i > x\right) = \mathbb{P}(Y_1(a_1 + Y_2(a_2 + \cdots + Y_{n-1}(a_{n-1} + a_n Y_n))) > x).$$

Since $F_Y \in \mathcal{L}$ and $a_n > 0$, we have that

$$\mathbb{P}(a_{n-1} + a_n Y_n > x) \sim \mathbb{P}(a_n Y_n > x).$$

Hence, applying Lemma A.1 we obtain that

$$\mathbb{P}(Y_{n-1}(a_{n-1} + a_n Y_n) > x) \sim \mathbb{P}(a_n Y_{n-1} Y_n > x).$$

Repeatedly applying Lemma A.1, we finally obtain that

$$\mathbb{P}(Y_1(a_1 + Y_2(a_2 + \cdots + Y_{n-1}(a_{n-1} + a_n Y_n))) > x) \sim \mathbb{P}(a_n Y_1 Y_2 \cdots Y_{n-1} Y_n > x).$$

For the remainder of the proof it suffices to verify that the probabilities $\mathbb{P}(a_i Z_i > x)$, $i = 1, 2, \ldots, n-1$, on the right-hand side of (1) can be neglected when compared with the probability $\mathbb{P}(a_n Z_n > x)$. Since the class $\mathcal{R}_{-\infty}$ is closed under product convolution, the d.f. of the product $\prod_{j=1}^{i} Y_j$ belongs to the class $\mathcal{R}_{-\infty}$ for each $i = 1, 2, \ldots$ Hence, for each $i = 1, 2, \ldots, n-1$ and some $0 < v < 1$,

$$\begin{aligned}\limsup_{x\to+\infty} \frac{\mathbb{P}\left(a_i \prod_{j=1}^{i} Y_j > x\right)}{\mathbb{P}\left(a_n \prod_{j=1}^{n} Y_j > x\right)} &\le \limsup_{x\to+\infty} \frac{\mathbb{P}\left(a_i \prod_{j=1}^{i} Y_j > x\right)}{\mathbb{P}\left(a_i \prod_{j=1}^{i} Y_j > vx, (a_n/a_i) \prod_{j=i+1}^{n} Y_j > 1/v\right)} \\ &= \frac{1}{\mathbb{P}\left((a_n/a_i) \prod_{j=i+1}^{n} Y_j > 1/v\right)} \limsup_{x\to+\infty} \frac{\mathbb{P}\left(a_i \prod_{j=1}^{i} Y_j > x\right)}{\mathbb{P}\left(a_i \prod_{j=1}^{i} Y_j > vx\right)} = 0.\end{aligned}$$

This proves that (1) holds.

Proof of Theorem 3.2. To prove the theorem, we first state three lemmas. The first lemma can easily be proved by Lemmas 3.8 and 3.10 of Tang and Tsitsiashvili (2003).

Lemma A.2. *Let X and Y be two independent r.v.'s, where X is supported on $(-\infty, +\infty)$ with a d.f. F, and Y is strictly positive with a d.f. G. Let $V = XY$ and denote by H the d.f. of V. If $F \in \mathcal{D} \cap \mathcal{L}$ and $G \in \mathcal{R}_{-\infty}$, then $H \in \mathcal{D} \cap \mathcal{L} \subset \mathcal{S}$ and*

$$\bar{H}(x) \asymp \bar{F}(x).$$

The second lemma can be proved by Lemma 3.7 of Tang and Tsitsiashvili (2003).

Lemma A.3. *If $F \in \mathcal{D}$ and $G \in \mathcal{R}_{-\infty}$, there exists some $\varepsilon > 0$ such that*

$$\bar{G}(x^{1-\varepsilon}) = o(\bar{F}(x)).$$

The third lemma can be obtained by fixing $\gamma = 0$ in Lemma 3.2 of Tang and Tsitsiashvili (2003).

Lemma A.4. *Let $F = F_1 * F_2$, where F_1 and F_2 are two d.f.'s supported on $(-\infty, +\infty)$. If $F_1 \in \mathcal{S}$, $F_2 \in \mathcal{L}$, and $\bar{F}_2(x) = O(\bar{F}_1(x))$, then $F \in \mathcal{S}$ and*

$$\bar{F}(x) \sim \bar{F}_1(x) + \bar{F}_2(x).$$

Now to finish the proof of Theorem 3.2, first we prove (2), which says that

$$\begin{aligned}
&\mathbb{P}[(a_1 + X_1)Y_1 + \cdots + (a_{n-1} + X_{n-1})Y_{n-1} \ldots Y_1 + (a_n + X_n)Y_n Y_{n-1} \ldots Y_1 > x] \\
&\sim \mathbb{P}[(a_1 + X_1)Y_1 > x] + \cdots + \mathbb{P}[(a_{n-1} + X_{n-1})Y_{n-1} \ldots Y_1 > x] + \mathbb{P}[(a_n + X_n)Y_n Y_{n-1} \ldots Y_1 > x].
\end{aligned}$$

Applying Lemma A.2, we see that the product $(a_n + X_n)Y_n$ is subexponential and

$$\mathbb{P}[(a_n + X_n)Y_n > x] \asymp \bar{F}(x). \tag{36}$$

Applying Lemma A.4 yields

$$\mathbb{P}[(a_{n-1} + X_{n-1}) + (a_n + X_n)Y_n > x] \sim \mathbb{P}[(a_{n-1} + X_{n-1}) > x] + \mathbb{P}[(a_n + X_n)Y_n > x].$$

Since, by Lemma A.3, there exists some $\varepsilon > 0$ such that $\bar{G}(x^{1-\varepsilon}) = o(\bar{F}(x))$, we have that

$$\begin{aligned}
&\mathbb{P}[(a_{n-1} + X_{n-1})Y_{n-1} + (a_n + X_n)Y_n Y_{n-1} > x] \\
&= \left(\int_0^{x^{1-\varepsilon}} + \int_{x^{1-\varepsilon}}^{+\infty}\right) \mathbb{P}[(a_{n-1} + X_{n-1})y + (a_n + X_n)Y_n y > x]\, \mathrm{d}G(y) \\
&= \int_0^{x^{1-\varepsilon}} \mathbb{P}\left[(a_{n-1} + X_{n-1}) + (a_n + X_n)Y_n > \frac{x}{y}\right] \mathrm{d}G(y) + o(\bar{F}(x)) \\
&\sim \int_0^{x^{1-\varepsilon}} \left(\mathbb{P}\left[(a_{n-1} + X_{n-1}) > \frac{x}{y}\right] + \mathbb{P}\left[(a_n + X_n)Y_n > \frac{x}{y}\right]\right) \mathrm{d}G(y) + o(\bar{F}(x)) \\
&= \left(\int_0^{+\infty} - \int_{x^{1-\varepsilon}}^{+\infty}\right) \left(\mathbb{P}\left[(a_{n-1} + X_{n-1}) > \frac{x}{y}\right] + \mathbb{P}\left[(a_n + X_n)Y_n > \frac{x}{y}\right]\right) \mathrm{d}G(y) + o(\bar{F}(x))
\end{aligned}$$

$$= \mathbb{P}[(a_{n-1}+X_{n-1})Y_{n-1}>x]+\mathbb{P}[(a_n+X_n)Y_nY_{n-1}>x]+o(\overline{F}(x))$$
$$\sim \mathbb{P}[(a_{n-1}+X_{n-1})Y_{n-1}>x]+\mathbb{P}[(a_n+X_n)Y_nY_{n-1}>x].$$

Furthermore, by application of Lemmas A.2 and A.4, it follows that $(a_{n-1}+X_{n-1})Y_{n-1}+(a_n+X_n)Y_nY_{n-1}$ is subexponential and that

$$\mathbb{P}[(a_{n-1}+X_{n-1})Y_{n-1}+(a_n+X_n)Y_nY_{n-1}>x] \asymp \overline{F}(x).$$

Simply repeating the procedure above and observing that

$$\begin{aligned}&(a_{n-2}+X_{n-2})Y_{n-2}+(a_{n-1}+X_{n-1})Y_{n-1}Y_{n-2}+(a_n+X_n)Y_nY_{n-1}Y_{n-2}\\&\quad=[(a_{n-2}+X_{n-2})+(a_{n-1}+X_{n-1})Y_{n-1}+(a_n+X_n)Y_nY_{n-1}]Y_{n-2},\end{aligned}$$

we obtain that

$$\begin{aligned}&\mathbb{P}[(a_{n-2}+X_{n-2})Y_{n-2}+(a_{n-1}+X_{n-1})Y_{n-1}Y_{n-2}+(a_n+X_n)Y_nY_{n-1}Y_{n-2}>x]\\&\quad\sim \mathbb{P}[(a_{n-2}+X_{n-2})Y_{n-2}>x]+\mathbb{P}[(a_{n-1}+X_{n-1})Y_{n-1}Y_{n-2}>x]+\mathbb{P}[(a_n+X_n)Y_nY_{n-1}Y_{n-2}>x].\end{aligned}$$

Hence, repeating the procedure above $n-1$ times yields the announced result (2). The proof of (3) can be given completely analogously to the above, since the distribution of a_iX_i satisfies

$$\mathbb{P}(a_iX_i>x)=\overline{F}(x/a_i)\asymp \overline{F}(x),$$

and is subexponential. This ends the proof of Theorem 3.2.

Proof of Corollary 3.1. Using (36), one can easily verify that

$$\liminf_{x\to+\infty}\frac{\mathbb{P}\left(\sum_{i=1}^{n}(a_i+X_i)Z_i>x\right)}{\mathbb{P}\left(\sum_{i=1}^{n-1}(a_i+X_i)Z_i>x\right)}>1,$$

and that

$$\liminf_{x\to+\infty}\frac{\mathbb{P}\left(\sum_{i=1}^{n}(a_iX_i)Z_i>x\right)}{\mathbb{P}\left(\sum_{i=1}^{n-1}(a_iX_i)Z_i>x\right)}>1.$$

Hence, we can prove (4) and (5) by substituting (2) and (3) into the left-hand-side of (4) and (5), respectively. This ends the proof of Corollary 3.1.

Proof of Corollary 3.2. Given the asymptotic results (2) and (3), the proof of this corollary follows immediately from a well-known result that was referred by Cline (1986) to Proposition 3 of Breiman (1965).

Proof of Theorem 3.3. In case the conditions 1. and 2. of Theorem 3.2 are replaced by the conditions 1′., 2′. and 3′. of Theorem 3.3, the proof of (2) can be established completely analogously to the proof of Theorem 3.2 using the following three lemmas, which are the analogs of Lemmas A.2–A.4, respectively:

Lemma A.5. *Let X and Y be two independent lognormally distributed r.v.'s with* $\sigma_Y<\sigma_X$*. Furthermore, let* $V=XY$ *and denote by H the d.f. of V. Then V follows a lognormal law and* $\overline{F}(x)=o(\overline{H}(x))$*.*

Lemma A.6. *If both F and G are lognormal laws with $\sigma_G < \sigma_F$, then there exists some $\varepsilon > 0$ such that*

$$\bar{G}(x^{1-\varepsilon}) = o(\bar{F}(x)).$$

Lemma A.7. *Let $F = F_1 * F_2$, where F_1 and F_2 are two lognormal laws. Then $F \in \mathcal{S}$ and*

$$\bar{F}(x) \sim \bar{F}_1(x) + \bar{F}_2(x).$$

The results of Lemmas A.5 and A.6 can easily be verified. Lemma A.7 is a special case of Corollary 1 of Cline (1986) and moreover is a special case of Lemma A.4 of the present paper. The proof of (3) can be given analogously, since the distribution of $a_i X_i$ is again lognormal with $\text{Var}[\log(a_i X_i)] = \text{Var}[\log(X_i)] = \sigma_X^2$.

It remains to prove (4) and (5). By application of Lemma A.1 and the same reasoning as in the proof of Theorem 3.1, we have for each $n = 1, 2, \ldots$ and some $0 < v < 1$ that

$$\liminf_{x\to+\infty} \frac{\sum_{i=1}^{n} \mathbb{P}\left((a_i + X)\prod_{j=1}^{i} Y_j > x\right)}{\sum_{i=1}^{n-1} \mathbb{P}\left((a_i + X)\prod_{j=1}^{i} Y_j > x\right)} \geq \liminf_{x\to+\infty} \frac{\mathbb{P}\left((a_n + X)\prod_{j=1}^{n} Y_j > x\right)}{\sum_{i=1}^{n-1} \mathbb{P}\left((a_i + X)\prod_{j=1}^{i} Y_j > x\right)}$$

$$= \frac{1}{\sum_{i=1}^{n-1} \limsup_{x\to+\infty} \frac{\mathbb{P}\left((a_i+X)\prod_{j=1}^{i} Y_j>x\right)}{\mathbb{P}\left((a_n+X)\prod_{j=1}^{n} Y_j>x\right)}} \geq \frac{1}{\sum_{i=1}^{n-1} \limsup_{x\to+\infty} \frac{\mathbb{P}\left((a_i+X)\prod_{j=1}^{i} Y_j>x\right)}{\mathbb{P}\left((a_n+X)\prod_{j=1}^{i} Y_j>vx\right)\mathbb{P}\left(\prod_{j=i+1}^{n} Y_j>1/v\right)}}$$

$$= \frac{1}{\sum_{i=1}^{n-1} \limsup_{x\to+\infty} \frac{\mathbb{P}\left(X\prod_{j=1}^{i} Y_j>x\right)}{\mathbb{P}\left(X\prod_{j=1}^{i} Y_j>vx\right)\mathbb{P}\left(\prod_{j=i+1}^{n} Y_j>1/v\right)}} = +\infty > 1,$$

and

$$\liminf_{x\to+\infty} \frac{\sum_{i=1}^{n} \mathbb{P}\left((a_i X)\prod_{j=1}^{i} Y_j > x\right)}{\sum_{i=1}^{n-1} \mathbb{P}\left((a_i X)\prod_{j=1}^{i} Y_j > x\right)} \geq \liminf_{x\to+\infty} \frac{\mathbb{P}\left((a_n X)\prod_{j=1}^{n} Y_j > x\right)}{\sum_{i=1}^{n-1} \mathbb{P}\left((a_i X)\prod_{j=1}^{i} Y_j > x\right)}$$

$$= \frac{1}{\sum_{i=1}^{n-1} \limsup_{x\to+\infty} \frac{\mathbb{P}\left((a_i X)\prod_{j=1}^{i} Y_j>x\right)}{\mathbb{P}\left((a_n X)\prod_{j=1}^{n} Y_j>x\right)}} = +\infty > 1.$$

Hence, we can prove (4) and (5) by substituting (2) and (3) into the left-hand-side of (4) and (5), respectively. This ends the proof of Theorem 3.3.

Appendix B. Approximation formulas for discounted IBNR loss reserves

In the following, we provide the formulas for the calculation of the IBNR stop-loss premiums and quantiles for the approximation methods based on comonotonicity and moment matching that were briefly described in Section 5. First, we introduce the r.v. W_{ij} defined by

$$W_{ij} = \alpha_i + \beta_j - D(i + j - t - 1),$$

with mean and variance given by

$$\mathbb{E}[W_{ij}] = \alpha_i + \beta_j - (i + j - t - 1)\mu,$$

and

$$\mathrm{Var}[W_{ij}] = \sigma^2_{W_{ij}} = (i + j - t - 1)\xi^2,$$

respectively. For the loglinear regression model with chain-ladder linear predictor and the geometric Brownian discounting process described in Section 5, using the notation introduced above, the asymptotic approximations (19)and (20) become

$$\mathbb{E}[(\tilde{S}_t - d)_+] \approx \sum_{i=2}^{t} \sum_{j=t+2-i}^{t} \mathrm{e}^{\mathbb{E}[W_{ij}]+\frac{1}{2}(\sigma^2_{W_{ij}}+\sigma^2_{\varepsilon_{ij}})} \Phi\left(\frac{\mathbb{E}[W_{ij}] + \sigma^2_{W_{ij}} + \sigma^2_{\varepsilon_{ij}} - \log(d)}{\sqrt{\sigma^2_{W_{ij}} + \sigma^2_{\varepsilon_{ij}}}}\right)$$
$$-d\Phi\left(\frac{\mathbb{E}[W_{ij}] - \log(d)}{\sqrt{\sigma^2_{W_{ij}} + \sigma^2_{\varepsilon_{ij}}}}\right), \qquad d \in \mathbb{R}_+,$$

and

$$F^{-1}_{\tilde{S}_t}(p) \approx \inf\left\{s : \sum_{i=2}^{t} \sum_{j=t+2-i}^{t} \overline{F_L}(s) \leq 1 - p\right\}, \qquad p \in (0, 1),$$

in which F_L is the lognormal$(\mathbb{E}[W_{ij}], \sigma^2_{W_{ij}} + \sigma^2_{\varepsilon_{ij}})$ d.f. Notice that $\mathrm{Var}[\varepsilon_{ij}] = \sigma^2_{\varepsilon_{ij}} = \sigma^2_{\varepsilon}$.

For the convex upper bound $\tilde{S}^u_t$ defined in (26) we have that

$$F_{\tilde{S}^u_t}(x) = \int_0^1 F_{\tilde{S}^u_t|V=v}(x)\,\mathrm{d}v,$$

and

$$\mathbb{E}[(\tilde{S}^u_t - d)_+] = \sum_{i=2}^{t} \sum_{j=t+2-i}^{t} \mathrm{e}^{\mathbb{E}[W_{ij}]+\frac{1}{2}\sigma^2_{W_{ij}}} \int_0^1 \mathrm{e}^{\sigma_\varepsilon \Phi^{-1}(v)} \Phi\left(\sigma_{W_{ij}} - \Phi^{-1}\left(F_{\tilde{S}^u_t|V=v}(d)\right)\right) \mathrm{d}v - d(1 - F_{\tilde{S}^u_t}(d)).$$

Here, $F_{\tilde{S}^u_t|V=v}(x)$ follows from

$$\sum_{i=2}^{t} \sum_{j=t+2-i}^{t} \exp(\mathbb{E}[W_{ij}] + \sigma_{W_{ij}} \Phi^{-1}(F_{\tilde{S}^u_t|V=v}(x)) + \sigma_\varepsilon \Phi^{-1}(v)) = x.$$

Another possibility to derive the d.f. of $\tilde{S}^u_t$ is as follows:

$$F_{\tilde{S}^u_t}(z) = \int_0^1 \Phi_{0,\sigma^2_\varepsilon}\left(\log(z) - \log\left(\sum_{i=2}^{t} \sum_{j=t+2-i}^{t} \exp(\mathbb{E}[W_{ij}] + \sigma_{W_{ij}} \Phi^{-1}(u))\right)\right) \mathrm{d}u.$$

The mean and the variance of $\tilde{S}_t$, to be used in the moment matching approach, are given by

$$\mathbb{E}[\tilde{S}_t] = \sum_{i=2}^{t} \sum_{j=t+2-i}^{t} \mathrm{e}^{\mathbb{E}[W_{ij}]+1/2(\sigma^2_{W_{ij}}+\sigma^2_{\varepsilon})}, \tag{37}$$

and

$$\mathrm{Var}[\tilde{S}_t] = \sum_{i=2}^{t} \sum_{j=t+2-i}^{t} \sum_{k=2}^{t} \sum_{l=t+2-k}^{t} \mathrm{e}^{\sigma^2_{\varepsilon}+(\mathbb{E}[W_{ij}]+\mathbb{E}[W_{kl}])+1/2(\sigma^2_{W_{ij}}+\sigma^2_{W_{kl}})}(\mathrm{e}^{\xi^2 \min(i+j-t-1,k+l-t-1)+\sigma^{2*}_{\varepsilon}} - 1), \tag{38}$$

where

$$\sigma^{2*}_{\varepsilon} = \begin{cases} \sigma^2_{\varepsilon}, & \text{if } i, j = k, l; \\ 0, & \text{if } i, j \neq k, l. \end{cases}$$

References

Bingham, N.H., Goldie, C.M., Teugels, J.L., 1987. Regular Variation. Cambridge University Press, Cambridge.

Breiman, L., 1965. On some limit theorems similar to the arc-sin law. Theory of Probability and its Applications 10, 323–331.

Cline, D.B.H., 1986. Convolution tails, product tails and domains of attraction. Probability Theory and Related Fields 72, 529–557.

Dhaene, J., Denuit, M., Goovaerts, M.J., Kaas, R., Vyncke, D., 2002a. The concept of comonotonicity in actuarial science and finance: theory. Insurance: Mathematics and Economics 31, 3–33.

Dhaene, J., Denuit, M., Goovaerts, M.J., Kaas, R., Vyncke, D., 2002b. The concept of comonotonicity in actuarial science and finance: applications. Insurance: Mathematics and Economics 31, 133–161.

Dufresne, D., 2002. Asian and Basket Asymptotics, Research Paper No. 100. Centre for Actuarial Studies, University of Melbourne.

Embrechts, P., Klüppelberg, C., Mikosch, T., 1997. Modelling Extremal Events for Insurance and Finance. Springer, Berlin.

Hoedemakers, T., Beirlant, J., Goovaerts, M.J., Dhaene, J., 2003. Confidence bounds for discounted loss reserves. Insurance: Mathematics and Economics 33, 297–316.

Kaas, R., Dhaene, J., Goovaerts, M.J., 2000. Upper and lower bounds for sums of random variables. Insurance: Mathematics and Economics 23, 151–168.

Rogers, L.C.G., Shi, Z., 1995. The value of an Asian option. Journal of Applied Probability 32, 1077–1088.

Tang, Q., Tsitsiashvili, G., 2003. Precise estimates for the ruin probability in finite horizon in a discrete-time model with heavy-tailed insurance and financial risks. Stochastic Processes and their Application 108, 299–325.

Tang, Q., Tsitsiashvili, G., 2004. Finite- and infinite-time ruin probabilities in the presence of stochastic returns on investments. Advances in Applied Probability 36, 1278–1299.

Vanduffel, S., Dhaene, J., Goovaerts, M.J., Kaas, R., 2003. The hurdle-race problem. Insurance: Mathematics and Economics 33, 405–413.

Vanduffel, S., Hoedemakers, T., Dhaene, J., 2004. Comparing approximations for risk measures of sums of non-independent lognormal random variables, Research Report OR 0418. Department of Applied Economics, Catholic University of Leuven.