SIZE-BIASED RISK MEASURES OF COMPOUND SUMS

M. Denuit

DISCUSSION PAPER | 2019 / 09

SIZE-BIASED RISK MEASURES OF COMPOUND SUMS

MICHEL DENUIT

Institut de statistique, biostatistique et sciences actuarielles - ISBA

UCLouvain

B-1348 Louvain-la-Neuve, Belgium

`michel.denuit@uclouvain.be`

Abstract

The size-biased, or length-biased transform is known to be particularly useful in insurance risk measurement. The case of continuous losses has been extensively considered in the actuarial literature. Given their importance in insurance studies, this paper concentrates on compound sums. The zero-augmented distributions which naturally appear in the individual model of risk theory are obtained as particular cases when the claim frequency distribution is concentrated on $\{0, 1\}$. The general results derived in this paper help actuaries to understand how risk measurement proceeds since the formulas make explicit the loadings corresponding to each source of randomness. Some simple and explicit expressions are obtained when losses are modeled by compound Poisson sums and compound mixed Poisson sums, including the compound Negative Binomial sums. Extensions to conditionally independent risks are discussed. As an illustration, the paper considers the case of compound mixed Poisson sums where the Poisson means are proportional to a random common risk factor.

Keywords: Weighted risk measures, Conditional tail expectation, Panjer recursion, Compound Poisson distribution, Compound mixed Poisson distribution, Infinite divisibility, Common mixture model.

1 Introduction and motivation

Several transforms have been widely used in the actuarial literature. Let us mention the Esscher transform and the stationary-excess operator, for instance. The Esscher transform (or exponential tilting) reduces the probabilities of small values in favor of probabilities of large values and plays an important role in theoretical considerations about premium calculation and risk measurement (see e.g. Chapter 5 in Kaas et al., 2008). The stationary-excess operator (also called equilibrium distribution, see e.g. Chapter 4 in Klugman et al., 2004) appears in the Beekman formula for the ruin probability in the classical risk theory model (see e.g. Chapter 8 in Klugman et al., 2004, or Chapter 4 in Kaas et al., 2008).

Several risk measures for an insurance loss X with distribution function F_X are based on the function

$$t \mapsto \mathbb{E}[X|X > t] = \frac{1}{\mathbb{P}[X > t]} \int_t^\infty x dF_X(x).$$

The Conditional Tail Expectation (CTE) is obtained when t corresponds to the Value-at-Risk (or quantile) of X . Such risk measures are intimately related to the size-biased transform, as shown next. Consider a (measurable) function g and write

$$\begin{aligned} \mathbb{E}[Xg(X)] &= \int_0^\infty xg(x)dF_X(x) \\ &= \mathbb{E}[X] \int_0^\infty g(x) \frac{x dF_X(x)}{\mathbb{E}[X]} \\ &= \mathbb{E}[X] \mathbb{E}[g(\tilde{X})] \end{aligned} \tag{1.1}$$

where the non-negative random variable $\tilde{X}$ with distribution function

$$\mathbb{P}[\tilde{X} \leq t] = \frac{1}{\mathbb{E}[X]} \int_0^t x dF_X(x),$$

is called the size-biased version of X . In some sense, $\tilde{X}$ is larger compared to X , so that $\tilde{X}$ represents a worse loss than X . In order to see why this is true, it suffices to notice that $\tilde{X}$ is distributed as $\max\{X, Z\}$ where the random variable Z is independent of X and has distribution function

$$\mathbb{P}[Z \leq t] = \frac{\mathbb{P}[\tilde{X} \leq t]}{\mathbb{P}[X \leq t]} = \frac{\mathbb{E}[X|X \leq t]}{\mathbb{E}[X]}.$$

Notice that the size-biased version of any constant c leaves it unchanged, that is, $\tilde{c} = c$. This shows that no unjustified loading is induced by the size-biased transform.

Now, inserting the function

$$g(x) = \begin{cases} 1 & \text{if } x > t \\ 0 & \text{otherwise} \end{cases} \tag{1.2}$$

in identity (1.1), we see that the representation

$$\mathbb{E}[X|X > t] = \mathbb{E}[X] \frac{\mathbb{P}[\tilde{X} > t]}{\mathbb{P}[X > t]} \tag{1.3}$$

is valid for any threshold t . As $\tilde{X}$ is larger compared to X , the ratio $P[\tilde{X} > t]/P[X > t]$ appearing in (1.3) exceeds unity so that we recover the classical inequality $E[X|X > t] \geq E[X]$ for all t from elementary probability.

The size-biased transform can be traced back to the late 1960s in the statistical literature. It has proven to be useful in the study of risk measures after the pioneering work by Furman and Landsman (2005, 2008a) and Furman and Zitikis (2008a,b) in connection to (1.3).

The size-biased transform is an example of weighted distribution. Initially developed in order to unify various sampling distributions when the chance of being recorded by an observer varies, weighted distributions are closely related to weighted risk measures and weighted capital allocation rules. See Furman and Zitikis (2009). Among these weighted distributions, the size-biased, or length-biased one corresponds to the identity weight function. It refers to the situation where larger observations are more likely to be recorded. Hence, the available data are of bigger size compared to the actual population values. Translated to an actuarial context, this means that claim amounts are made larger before performing actuarial calculations, which generates a safety loading.

Let us now turn to a collection of n independent insurance losses, denoted as $X_1, X_2, \dots, X_n$. Define the aggregate loss S as $S = \sum_{i=1}^n X_i$. Clearly,

$$E[S|S > t] = \sum_{i=1}^n E[X_i|S > t]$$

so that the total $E[S|S > t]$ can be decomposed into the sum of individual contributions $E[X_i|S > t]$ that are useful in risk management. It is known from Furman and Landsman (2005) that the representation

$$E[X_i|S > t] = E[X_i] \frac{P[S - X_i + \tilde{X}_i > t]}{P[S > t]} \quad (1.4)$$

holds true. In words, the contribution $E[X_i|S > t]$ of risk X_i to $E[S|S > t]$ is equal to its a priori expectation $E[X_i]$ increased by the ratio of the excess probabilities $P[S - X_i + \tilde{X}_i > t]$ and $P[S > t]$, where

$$S - X_i + \tilde{X}_i = \sum_{j \neq i} X_j + \tilde{X}_i$$

is the sum of all risks X_j , except the i th one which is replaced with its size-biased version $\tilde{X}_i$. Hence, $S - X_i + \tilde{X}_i$ tends to be larger compared to S .

Identities (1.3)-(1.4) have been exploited in a number of papers (including those cited above) to derive many useful properties for risk measures, mainly for continuous losses. In this paper, we study the case of compound sums that are particularly useful in insurance applications. When the number of terms is valued in $\{0, 1\}$, we find the zero-augmented distributions which naturally arise in the individual model of risk theory. Particular attention is paid to compound Poisson sums and compound Negative Binomial sums to which Panjer recursive formulas apply.

Notice that some particular cases of compound Poisson sums have been considered previously in the literature. Risk measures for Tweedie distributions have been discussed by Furman and Landsman (2008b, 2010). Recall that Tweedie distributions include compound

Poisson sums with Gamma-distributed claim severities. In this paper, we do not restrict our analysis to Gamma distributed severities but also consider compound Poisson distributions with arbitrary claim severity distribution, as well as compound distributions where the number of terms does not obey the Poisson distribution. An extension to the case of conditional independence is also considered. The present paper complements the results obtained by Cossette et al. (2012) and Kim et al. (2019) who favored a computational approach to the problem.

The remainder of this paper is organized as follows. In Section 2, we recall the main properties of the size-biased transform. Section 3 derives the size-biased transform of compound sums. Section 4 is devoted to the compound Poisson case whereas Section 5 discusses the compound mixed Poisson sums. An extension to conditionally independent random variables is worked out in Section 6. The final Section 7 briefly concludes the paper.

Throughout this paper, we provide elementary proofs of the results. Some of them are available in the literature (and the text refers to the published works if this is the case), but we believe that the simple derivations offered here are both illuminating and easier to follow for an actuarial readership. In addition to making this paper self-contained, the approach adopted here helps actuaries to gain a better understanding of the topic. In particular, some formulas make explicit the loadings corresponding to each source of randomness. This decomposition appears to be useful for risk management.

2 Size-biased transform

2.1 Definition

2.1.1 General case

Consider a risk X (i.e. a non-negative random variable representing a monetary loss) with distribution function F_X . A size-biased version of X is any risk $\tilde{X}$ with distribution function

$$P[\tilde{X} \leq t] = \frac{E[XI[X \leq t]]}{E[X]} = \frac{E[X|X \leq t]}{E[X]}F_X(t) = \frac{1}{E[X]} \int_0^t x dF_X(x),$$

where $I[\cdot]$ denotes the indicator function (which is equal to 1 if the event appearing within brackets is realized, and to 0 otherwise). The operator mapping the distribution function F_X of X to the distribution function $F_{\tilde{X}}$ of $\tilde{X}$ is called the size-biased transform.

A detailed account of the properties of the size-biased transform can be found in Brown (2006) and Arratia et al. (2019). Most parametric models used in actuarial applications are closed under size-biasing, provided the parameters are updated accordingly. See e.g. Table 1 in Patil and Rao (1978).

2.1.2 Continuous case

Consider a claim severity C , that is, a positive random variable C corresponding to the amount of benefit paid in relation to some insurance claim. In general, C is assumed to be

continuous with a probability density function f_C . We then have

$$P[\tilde{C} \leq t] = \int_0^t \frac{x f_C(x)}{E[C]} dx$$

so that the probability density function $f_{\tilde{C}}$ of $\tilde{C}$ is given by

$$f_{\tilde{C}}(x) = \frac{x f_C(x)}{E[C]}, \quad x \geq 0.$$

Brown (2006) explained the relationship of the size-biased transform of a continuous distribution with the waiting time paradox and provided several applications.

2.1.3 Discrete case

Integer-valued random variables are often considered in actuarial applications to model claim frequencies or after discretization of the claim severities to perform computations. Given a counting random variable N valued in $\{0, 1, 2, \dots\}$ with probability mass function p_N , the probability mass function $p_{\tilde{N}}$ of the size-biased version $\tilde{N}$ is obtained as follows:

$$\begin{aligned} p_{\tilde{N}}(k) &= P[\tilde{N} \leq n] - P[\tilde{N} \leq n-1] \\ &= \frac{E[N I[N = k]]}{E[N]} = \frac{k p_N(k)}{E[N]}, \quad k = 0, 1, 2, \dots \end{aligned}$$

Ross (2003) discussed the relationship of the discrete size-biased transform to the inspection paradox. Notice that even if $p_N(0) > 0$, we have

$$p_{\tilde{N}}(0) = 0 \Leftrightarrow \tilde{N} \geq 1 \text{ with probability } 1. \quad (2.1)$$

2.2 Risk measurement based on size-biased transform

It is easy to see that $\tilde{X}$ is stochastically larger compared to X . Indeed, the random variables X and $I[X \leq t]$ are negatively correlated so that

$$P[\tilde{X} \leq t] = \frac{E[X I[X \leq t]]}{E[X]} \leq \frac{E[X] E[I[X \leq t]]}{E[X]} = P[X \leq t]$$

whatever the threshold t . This inequality shows that the probability that X falls below t is larger compared to the corresponding probability for its size-biased version $\tilde{X}$. Hence, replacing the actual loss X with its size-biased transform $\tilde{X}$ can be seen as a safe strategy.

In fact, the stronger likelihood ratio order holds. We refer the reader to Denuit et al. (2005, Chapter 3) for more details concerning this stochastic order relation expressing the idea of “being larger than” for random variables. In the continuous case investigated in Subsection 2.1.2, the ratio

$$\frac{f_{\tilde{C}}(x)}{f_C(x)} = \frac{x}{E[C]}$$

clearly increases in x . In this case, C is said to be smaller than $\tilde{C}$ in the likelihood ratio order. Intuitively speaking, this means that the larger the outcome x , the more likely it corresponds to a realization of $\tilde{C}$ and not of C . A similar interpretation holds in the discrete case considered in Subsection 2.1.3. Specifically, for a counting random variable N with probability mass function p_N , the ratio of the probability mass functions

$$\frac{p_{\tilde{N}}(k)}{p_N(k)} = \frac{k}{\mathbb{E}[N]}$$

clearly increases in k , so that N is smaller than $\tilde{N}$ in the likelihood ratio order.

In general, the size-biased version $\tilde{X}$ of X is larger than X in the likelihood ratio order because $p \mapsto F_{\tilde{X}}(F_X^{-1}(p))$ is convex (see Proposition 3.3.50 in Denuit et al., 2005). This can be seen from the identity

$$F_{\tilde{X}}(F_X^{-1}(p)) = \frac{\mathbb{E}[XI[X \leq F_X^{-1}(p)]]}{\mathbb{E}[X]}$$

which shows that $p \mapsto F_{\tilde{X}}(F_X^{-1}(p))$ is the Lorenz curve of X , which is known to be increasing and convex. The relationship between the size-biased transform and the Lorenz curve has been pointed out by Bartoszewicz and Skolimowska (2006, Section 2.1), after Jain et al. (1989).

2.3 Infinitely divisible case

Assume that X is infinitely divisible. This means that for any positive integer n , X can be represented as the sum of n independent and identically distributed random variables. Infinitely divisible distributions on the half positive real line correspond either to compound Poisson sums if there is a positive probability mass at the origin, or to weak limits of compound Poisson sums. We refer the reader to Theorems 3.2-3.3 in Steutel and van Harn (2003, Chapter 3). The Poisson distribution, the Negative Binomial distribution, the Gamma distribution and the degenerate distribution are examples of infinitely divisible distributions. The uniform distribution and the Binomial distribution are not infinitely divisible, nor are any other distributions with bounded (finite) support.

Pakes et al. (1996, Theorem 2.1) established that the size-biased transform of a random variable X obeying an infinitely divisible distribution can be written as the sum of the initial random variable plus a non-negative, independent one. Precisely, the following distributional equality

$$\tilde{X} \stackrel{d}{=} X + \Delta \text{ for some random variable } \Delta \geq 0, \text{ independent of } X. \quad (2.2)$$

is valid if, and only if the distribution of X is infinitely divisible. We refer the reader to Arratia and Goldstein (2010) for a detailed analysis of various situations where (2.2) is valid. The distributional equality (2.2) shows that $\tilde{X}$ is indeed larger than X , because it corresponds to X increased by Δ in the infinitely divisible case.

3 Size-biased transform of compound sums

Aggregate losses are often represented by means of compound sums in insurance studies. In this section, we consider the size-biased transform of such random variables. Compound sums can be considered as mixtures of sums with a fixed number of terms, letting this number of terms become random. This is why we start with a couple of results of independent interest, dealing with size-biased versions of sums and mixtures.

3.1 Size-biased transform of a sum

The following result can be found in Brown (2006, Example 2) or Arratia et al. (2019, Section 2.4). We provide here an elementary proof of it, for convenience and because it helps to understand how risk measurement proceeds.

Property 3.1. *Consider independent risks $X_1, \dots, X_n$ and let $\tilde{X}_1, \dots, \tilde{X}_n$ be their corresponding size-biased versions, assumed to be independent, and independent of $X_1, \dots, X_n$. Let $S = \sum_{i=1}^n X_i$ be the sum of the risks $X_1, \dots, X_n$. Define the random variable K valued in $\{1, 2, \dots, n\}$, independent of $X_1, \dots, X_n$ and of $\tilde{X}_1, \dots, \tilde{X}_n$ such that*

$$P[K = i] = \frac{E[X_i]}{E[S]}.$$

Then,

$$\tilde{S} =_d S - X_K + \tilde{X}_K = \sum_{j \neq K} X_j + \tilde{X}_K, \quad (3.1)$$

where “ $=_d$ ” means “is distributed as”.

Proof. The random vector $(X_1, \dots, X_{i-1}, \tilde{X}_i, X_{i+1}, \dots, X_n)$ has joint distribution function

$$(x_1, \dots, x_n) \mapsto \frac{E[X_i I[X_i \leq x_i]]}{E[X_i]} \prod_{j \neq i} P[X_j \leq x_j].$$

For any measurable function g , we then have

$$E[g(X_1, \dots, X_{i-1}, \tilde{X}_i, X_{i+1}, \dots, X_n)] = E\left[\frac{X_i}{E[X_i]} g(X_1, \dots, X_n)\right].$$

Let us now consider

$$g(x_1, \dots, x_n) = I\left[\sum_{j=1}^n x_j \leq s\right].$$

We then have

$$P[S - X_i + \tilde{X}_i \leq s] = \frac{E[X_i I[S \leq s]]}{E[X_i]}.$$

Summing this equality over i gives

$$E[SI[S \leq s]] = \sum_{i=1}^n E[X_i] P[S - X_i + \tilde{X}_i \leq s]. \quad (3.2)$$

This can also be rewritten as

$$P[\tilde{S} \leq s] = \frac{E[SI[S \leq s]]}{E[S]} = \sum_{i=1}^n \frac{E[X_i]}{E[S]} P[S - X_i + \tilde{X}_i \leq s].$$

This ends the proof. $\square$

Notice that the proof of Property 3.1 also establishes the validity of the fundamental identity (1.4). Property 3.1 shows that size-biasing a sum of independent risks is equivalent to selecting at random (according to the discrete distribution assigning probability $\frac{E[X_i]}{E[S]}$ to index i) one element appearing in the sum and replacing it with its size-biased version. Thus, the risks X_i with larger means are more likely to be selected.

In case all individual losses X_i share the same distribution, we get the following simpler result.

Corollary 3.2. *Under the conditions of Property 3.1, if $X_1, \dots, X_n$ are identically distributed then*

$$\tilde{S} =_d \sum_{i=1}^{n-1} X_i + \tilde{X}_n. \quad (3.3)$$

Size-biasing a sum of independent and identically distributed risks X_i thus amounts to size-biasing one term and to add it to the remaining $n - 1$ ones that are left unchanged.

Let us now apply these results to risk measurement.

Property 3.3. *Under the conditions of Property 3.1,*

$$E[S|S > t] = E[S] \frac{P\left[\sum_{j \neq K} X_j + \tilde{X}_K > t\right]}{P[S > t]}. \quad (3.4)$$

If $X_1, \dots, X_n$ are identically distributed as in Corollary 3.2 then

$$E[S|S > t] = E[S] \frac{P\left[\sum_{i=1}^{n-1} X_i + \tilde{X}_n > t\right]}{P[S > t]}. \quad (3.5)$$

Notice that the mixture representation involved in (3.4) is in line with (1.4) as it can be rewritten as

$$\begin{aligned} E[S|S > t] &= E[S] \sum_{i=1}^n \frac{E[X_i]}{E[S]} \frac{P\left[\sum_{j \neq i} X_j + \tilde{X}_i > t\right]}{P[S > t]} \\ &= \sum_{i=1}^n E[X_i|S > t]. \end{aligned}$$

3.2 Size-biased transform of a mixture

Arratia et al. (2019, Lemma 2.6) studied the size-biased transform of mixtures. The next result summarizes their findings. We provide the reader with an elementary proof, for convenience and because this result will also be central to the extension to size-biasing conditionally independent risks.

Property 3.4. *Consider a family of non-negative random variables $\{X_\theta, \theta \geq 0\}$ indexed by a single, non-negative parameter θ . Let Θ be a mixing parameter with distribution function F_Θ . The corresponding mixture X_Θ has distribution function*

$$P[X_\Theta \leq x] = \int_0^\infty P[X_\theta \leq x] dF_\Theta(\theta).$$

Then, the size-biased transform of the mixture X_Θ corresponds to the mixture of the non-negative random variables $\{\tilde{X}_\theta, \theta \geq 0\}$ with mixing parameter Θ^ distributed according to*

$$dF_{\Theta^*}(\theta) = \frac{E[X_\theta]}{E[X_\Theta]} dF_\Theta(\theta). \quad (3.6)$$

Proof. This result can easily be obtained as follows:

$$\begin{aligned} P[\tilde{X}_\Theta \leq x] &= \frac{1}{E[X_\Theta]} \int_0^x t dF_{X_\Theta}(t) \\ &= \frac{1}{E[X_\Theta]} \int_0^\infty \int_0^x t dF_{X_\theta}(t) dF_\Theta(\theta) \\ &= \frac{1}{E[X_\Theta]} \int_0^\infty E[X_\theta] F_{\tilde{X}_\theta}(t) dF_\Theta(\theta) \\ &= \int_0^\infty F_{\tilde{X}_\theta}(t) dF_{\Theta^*}(\theta). \end{aligned}$$

This ends the proof. □

Property 3.4 essentially shows that the size-biased version of a mixture corresponds to some mixture of size-biased versions of the mixed components, the mixing distribution being given in (3.6).

Let us now provide a couple of applications of the result stated in Property 3.4. First, we show that size-biasing a product of independent factors amounts to multiply their respective size-biased versions.

Property 3.5. *Consider two independent, positive random variables Z and Θ . Then,*

$$\widetilde{Z\Theta} =_d \tilde{Z}\tilde{\Theta}.$$

Proof. Let us first consider a constant Θ , that is, $\Theta = \theta$ with probability 1. The size-biased transform of θZ has distribution function

$$P[\tilde{Z\theta} \leq t] = \frac{1}{E[Z]} \int_0^t \frac{x}{\theta} dP \left[Z \leq \frac{x}{\theta} \right] = P \left[\tilde{Z} \leq \frac{t}{\theta} \right]$$

so that the distributional equality $\widetilde{Z}\theta =_d \theta\widetilde{Z}$ holds true. Now, define $X_\theta = \theta Z$ so that $\widetilde{X}_\Theta$ corresponds to $\widetilde{Z}\Theta$. Considering Property 3.4, this means that $\widetilde{Z}\Theta$ is a mixture of $\widetilde{Z}\theta =_d \theta\widetilde{Z}$ with mixing parameter Θ^* defined in (3.6), which gives here

$$dP[\Theta^* \leq \theta] = \frac{\theta E[Z]}{E[\Theta]E[Z]} dP[\Theta \leq \theta] \Leftrightarrow \Theta^* =_d \widetilde{\Theta}.$$

This ends the proof. $\square$

The result in Property 3.5 can be traced back to Pakes (1997). If Θ corresponds to the effect of inflation on the nominal amount Z , we see that the size-biased version of the inflated payment ΘZ is obtained by multiplying the size-biased versions of the random inflation rate Θ and of the base payment Z .

Let us now consider an application of Property 3.4 to risk measurement.

Property 3.6. *Under the conditions of Property 3.4, we have*

$$E[X_\Theta | X_\Theta > t] = E[X_\Theta] \frac{\int_0^\infty P[\widetilde{X}_\theta > t] dF_{\Theta^*}(\theta)}{\int_0^\infty P[X_\theta > t] dF_\Theta(\theta)}.$$

Considering the representation in Property 3.6, we see that $E[X_\Theta | X_\Theta > t]$ is obtained by multiplying $E[X_\Theta]$ with a ratio of two excess probabilities. In the numerator, we have the probability that a mixture of $\{\widetilde{X}_\theta, \theta \geq 0\}$ with mixing parameter Θ^* exceeds t whereas in the denominator, the probability that the original mixture $\{X_\theta, \theta \geq 0\}$ with mixing parameter Θ exceeds the same threshold.

3.3 Size-biased transform of a compound sum

We are now ready to derive the size-biased transform of a compound sum. The next result is stated without proof in Brown (2006, Section 2.2) as a direct consequence of the properties of the size-biased transform for sums and mixtures. We provide the reader with a formal derivation and then consider several particular cases of interest in actuarial risk theory.

Proposition 3.7. *Consider the compound sum $X = \sum_{k=1}^N C_k$ where N is valued in $\{0, 1, 2, \dots\}$ and the severities $C_1, C_2, \dots$ are identically distributed, distributed as C , all these random variables being independent. Then,*

$$\widetilde{X} =_d \widetilde{C} + \sum_{k=1}^{\widetilde{N}-1} C_k.$$

Proof. Let $\mu = E[C]$. The compound sum X can be seen as the mixture of $X_n = \sum_{k=1}^n C_k$ with

$$E[X_n] = n\mu \text{ and } E[X] = E[N]\mu.$$

Corollary 3.2 shows that

$$\widetilde{X}_n =_d \widetilde{C} + X_{n-1}.$$

By Property 3.4, we know that $\tilde{X}$ is a mixture of $\tilde{X}_n$ with mixing parameter N^* such that

$$\frac{p_{N^*}(n)}{p_N(n)} = \frac{n}{E[N]} \Rightarrow N^* =_d \tilde{N}.$$

This ends the proof. $\square$

Size-biasing a compound sum thus amounts to size-biasing one term and to add it to a compound sum where the claim severity distribution is left unchanged but the claim frequency becomes $\tilde{N} - 1$. Remember from (2.1) that $\tilde{N} \geq 1$ with probability one, so that $\tilde{N} - 1$ is valued in $\{0, 1, 2, \dots\}$ and thus eligible as a claim frequency. Size-biasing thus impacts on the claim severity distribution through the replacement of one term with its size-biased version $\tilde{C}$ as well as to claim frequency distribution by replacing N with its size-biased version $\tilde{N}$ shifted by 1.

Comparing the result in Proposition 3.7 to Corollary 3.2, we see that the fixed number of terms n has been replaced here with $\tilde{N}$. As size-biasing a constant leaves it unchanged, the result stated in Corollary 3.2 can be recovered as a particular case of Proposition 3.7 when the claim frequency distribution concentrates all its probability on a single value n .

The following result considers the particular case where the claim severities involved in the compound sum obey an infinitely divisible distribution. If C is infinitely divisible then we know from (2.2) that there exists a positive random variable Δ , independent of C such that $\tilde{C} =_d C + \Delta$. This allows us to rewrite $\tilde{C} + \sum_{k=1}^{\tilde{N}-1} C_k$ in Proposition 3.7 as $\Delta + \sum_{k=1}^{\tilde{N}} C_k$, as formally stated next.

Corollary 3.8. *If C is infinitely divisible then the compound sum X in Proposition 3.7 has size-biased transform*

$$\tilde{X} =_d \Delta + \sum_{k=1}^{\tilde{N}} C_k$$

where Δ is the random, independent addition involved in the distributional equality $\tilde{C} =_d C + \Delta$, all the random variables being independent.

When claim severities are infinitely divisible, an extra term can be included in the sum of the severities C_k which now runs to $\tilde{N}$.

Let us now apply these results to risk measurement.

Property 3.9. *Under the conditions of Proposition 3.7,*

$$E[X|X > t] = E[X] \frac{P\left[\tilde{C} + \sum_{k=1}^{\tilde{N}-1} C_k > t\right]}{P[X > t]}.$$

Under the conditions of Corollary 3.8,

$$E[X|X > t] = E[X] \frac{P\left[\tilde{\Delta} + \sum_{k=1}^{\tilde{N}} C_k > t\right]}{P[X > t]}.$$

Replacing N with the constant n , we find the result stated for identically distributed summands in Property 3.3 as a particular case.

Now that general expressions have been obtained for the size-biased version of compound sums, we are ready to study the compound Poisson and compound mixed Poisson sums that are particularly useful in insurance studies. This is the topic of the next two sections.

4 Compound Poisson case

4.1 Size-biased transform of compound Poisson sums

Compound Poisson sums are particular cases of the compound sums considered in Proposition 3.7 where N is Poisson distributed. They have infinitely divisible distributions so that (2.2) applies. The next result identifies the random variable Δ appearing in this distributional equality, which turns out to correspond to the size-biased transform of the generic term C .

Property 4.1. *Consider $X = \sum_{k=1}^N C_k$ as described in Proposition 3.7 with N Poisson distributed with mean λ , which is henceforth denoted as $N \sim \mathcal{Poi}(\lambda)$. Then,*

$$\tilde{X} =_d X + \tilde{C}.$$

Proof. The Poisson distribution can be characterized as follows (Arratia et al., 2019, Proposition 2.1):

$$N \sim \mathcal{Poi}(\lambda) \Leftrightarrow \tilde{N} =_d N + 1.$$

This is easily seen as follows. If $N \sim \mathcal{Poi}(\lambda)$ then $p_{\tilde{N}}(0) = 0$ and for $k \geq 1$,

$$\begin{aligned} p_{\tilde{N}}(k) &= \frac{k \exp(-\lambda) \frac{\lambda^k}{k!}}{\lambda} \\ &= \exp(-\lambda) \frac{\lambda^{k-1}}{(k-1)!} \end{aligned}$$

so that $\tilde{N} =_d N + 1$. The announced result then follows from Proposition 3.7. $\square$

Size-biasing a compound Poisson sum thus amounts to supplement it with the size-biased version of its generic term. As every compound Poisson sum is infinitely divisible, Property 4.1 identifies the random variable Δ involved in (2.2) which corresponds to $\tilde{C}$.

Remark 4.2. In the discrete case, when the random variables C_k are valued in $\{1, 2, 3, \dots\}$, the result stated in Property 4.1 directly follows from Panjer recursion formula

$$\mathbb{P}[X = k] = \frac{\lambda}{k} \sum_{j=1}^k j \mathbb{P}[C = j] \mathbb{P}[X = k - j]$$

starting from $P[X = 0] = \exp(-\lambda)$. Indeed, it suffices to multiply both sides of this recursion formula with $k/E[X] = k/(\lambda E[C])$ to get

$$\begin{aligned} P[\tilde{X} = k] &= \sum_{j=1}^k \frac{jP[C = j]}{E[C]} P[X = k - j] \\ &= \sum_{j=1}^k P[\tilde{C} = j] P[X = k - j] \\ &= P[X + \tilde{C} = k], \end{aligned}$$

which is exactly the result in Property 4.1. Compound Poisson distributions with discrete summands have been extensively considered in Arratia and Goldstein (2010, see formula (22) in that paper) using the standard representation of compound Poisson distributions with discretely distributed summands given e.g. in Theorem 3.4.2 in Kaas et al. (2008), often used in actuarial risk theory. Property 4.1 extends these results to the general case, with arbitrary claim severities.

Let us now consider the particular case where claim severities are infinitely divisible.

Property 4.3. *If C is infinitely divisible then the compound sum X in Property 4.1 has size-biased version*

$$\tilde{X} =_d \Delta + \sum_{k=1}^{N+1} C_k$$

where Δ is the random addition involved in the distributional equality $\tilde{C} =_d C + \Delta$, all the random variables being independent.

4.2 Application to risk measurement

Let us now apply the results obtained for the size-biased transform of compound Poisson sums to risk measurement.

Property 4.4. *Under the conditions of Property 4.1,*

$$E[X|X > t] = E[X] \frac{P[X + \tilde{C} > t]}{P[X > t]} = \frac{E[X]}{P[X > t|X + \tilde{C} > t]}.$$

When the claim severity distribution is infinitely divisible,

$$E[X|X > t] = E[X] \frac{P[X + C + \Delta > t]}{P[X > t]} = \frac{E[X]}{P[X > t|X + C + \Delta > t]}.$$

Let us now consider independent losses $X_1, \dots, X_n$ where each X_i is compound Poisson, that is,

$$X_i = \sum_{k=1}^{N_i} C_{ik}$$

where $N_i \sim \mathcal{Poi}(\lambda_i)$ and the claim severities C_{ik} are positive, continuous, and distributed as C_i , all these random variables being independent. Consider the total amount $S = \sum_{i=1}^n X_i$. According to Property 3.1, we have

$$\tilde{S} =_d S - X_K + \tilde{X}_K \text{ with } P[K = i] = \frac{\lambda_i E[C_i]}{\sum_{j=1}^n \lambda_j E[C_j]}.$$

Property 4.1 shows that the distributional equalities

$$\tilde{X}_i =_d X_i + \tilde{C}_i \text{ hold for all } i \text{ so that } \tilde{X}_K =_d X_K + \tilde{C}_K$$

also holds true. Hence,

$$S - X_K + \tilde{X}_K =_d S + \tilde{C}_K$$

so that we finally obtain

$$E[S|S > t] = E[S] \frac{P[S + \tilde{C}_K > t]}{P[S > t]}$$

with

$$P[S + \tilde{C}_K > t] = \sum_{i=1}^n P[S + \tilde{C}_i > t] \frac{\lambda_i E[C_i]}{\sum_{j=1}^n \lambda_j E[C_j]}. \quad (4.1)$$

Equivalently,

$$E[S|S > t] = \sum_{i=1}^n \lambda_i E[C_i] \frac{P[S + \tilde{C}_i > t]}{P[S > t]}.$$

Also, since

$$S - X_i + \tilde{X}_i =_d S + \tilde{C}_i$$

we get from (1.4)

$$\begin{aligned} E[X_i|S > t] &= E[X_i] \frac{P[S + \tilde{C}_i > t]}{P[S > t]} \\ &= \frac{E[X_i]}{P[S > t|S + \tilde{C}_i > t]}. \end{aligned}$$

The aforementioned results can also be obtained as follows. Because S is a sum of independent compound Poisson sums, it is itself a compound Poisson sum. Precisely, $S =_d \sum_{k=1}^M Z_k$ with $M \sim \mathcal{Poi}(\lambda_\bullet)$ where $\lambda_\bullet = \sum_{i=1}^n \lambda_i$ and $Z_1, Z_2, \dots$ are independent random variables distributed as Z with

$$P[Z \leq t] = \sum_{i=1}^n \frac{\lambda_i}{\lambda_\bullet} P[C_i \leq t],$$

all the random variables being independent. This can be equivalently stated as $Z =_d C_J$ with $P[J = i] = \frac{\lambda_i}{\lambda_\bullet}$. Hence, we can apply Property 4.1 directly to S to get

$$\tilde{S} =_d S + \tilde{Z}.$$

We know from Property 3.4 that $\tilde{Z}$ obeys a mixture of $\tilde{C}_1, \dots, \tilde{C}_n$ with mixing distribution

$$\begin{aligned} P[J^* = i] &= \frac{E[C_i]}{E[C_J]} P[J = i] \\ &= \frac{\lambda_i E[C_i]}{\sum_{j=1}^n \lambda_j E[C_j]} \\ &= P[K = i]. \end{aligned}$$

We thus see that $\tilde{Z} =_d \tilde{C}_K$ and we also recover (4.1) proceeding in this way.

The formulas derived in this section provide the actuary with an effective algorithm to compute $E[X_i|S > s]$. Assume for simplicity that the claim severities C_{ik} are integer-valued (if needed, they can be discretized prior to calculations). Then, S is also compound Poisson distributed and integer-valued. Its distribution can be obtained from Panjer recursion recalled in Remark 4.2. Precisely,

$$P[S = k] = \frac{1}{k} \sum_{j=1}^k j \sum_{i=1}^n \lambda_i P[C_i = j] P[S = k - j].$$

The distribution of $S + \tilde{C}_i$ can then be obtained by convolution and the desired value is obtained by multiplying $E[X_i]$ with the ration $P[S + \tilde{C}_i > s]/P[S > s]$.

Taking for s the Value-at-Risk, or quantile at level p of S , the contribution of X_i to the conditional tail expectation of S at probability level p is given by

$$E[X_i|S > F_S^{-1}(p)] = \frac{E[X_i]}{P[S > F_S^{-1}(p)]} P[S + \tilde{C}_i > F_S^{-1}(p)].$$

In the continuous case, we simply have $P[S > F_S^{-1}(p)] = 1 - p$. Now, if the claim sizes are homogeneous, that is, $C_i =_d C_j$ for all i, j then

$$P[S + \tilde{C}_i > F_S^{-1}(p)] = P[S + \tilde{C}_j > F_S^{-1}(p)]$$

and

$$\frac{E[X_i|S > F_S^{-1}(p)]}{E[X_j|S > F_S^{-1}(p)]} = \frac{\lambda_i}{\lambda_j}.$$

As

$$\sum_{i=1}^n E[X_i|S > F_S^{-1}(p)] = E[S|S > F_S^{-1}(p)]$$

is the conditional tail expectation of S , we finally get

$$E[X_i|S > F_S^{-1}(p)] = \frac{\lambda_i}{\lambda_\bullet} E[S|S > F_S^{-1}(p)].$$

This provides the actuary with a simple formula to distribute the total conditional tail expectation of S , based on the ratios $\frac{\lambda_i}{\lambda_\bullet}$ of expected claim frequencies.

5 Compound mixed Poisson case

5.1 General Poisson mixtures

In this section, we replace the Poisson mean λ with a positive random variable Λ . This construction corresponds to a mixed Poisson distribution with mixing parameter Λ and the corresponding compound sum is a mixed Poisson one. Henceforth, we use subscripts “ Λ ” or “ λ ” to indicate whether we work with the mixture N_Λ or N_λ given $\Lambda = \lambda$, or the corresponding compound sums X_Λ or X_λ .

Property 5.1. *Consider $X_\Lambda = \sum_{k=1}^N C_k$ as described in Proposition 3.7 with N obeying the $\mathcal{Poi}(\Lambda)$ distribution given Λ . Then,*

$$\tilde{X}_\Lambda =_d X_{\tilde{\Lambda}} + \tilde{C}.$$

Proof. Let us apply Property 3.4 with

$$X_\lambda = \sum_{k=1}^{N_\lambda} C_k$$

where $N_\lambda \sim \mathcal{Poi}(\lambda)$ and the C_k are distributed as C , all these random variables being independent. Then, invoking Property 4.1, $\tilde{X}$ appears to be a mixture of

$$\tilde{X}_\lambda =_d X_\lambda + \tilde{C}$$

with mixing

$$\frac{dF_{\Lambda^*}(\lambda)}{dF_\Lambda(\lambda)} = \frac{\lambda}{E[\Lambda]} \Rightarrow \Lambda^* =_d \tilde{\Lambda}.$$

This ends the proof as the distribution of $\tilde{C}$ does not involve the mixing parameter. $\square$

Compared to Property 4.1 applying to homogeneous Poisson claim frequencies, we see that X is replaced with $X_{\tilde{\Lambda}}$ so that the mixing parameter Λ is replaced with its size-biased version $\tilde{\Lambda}$ in the mixed Poisson claim frequency distribution. As we know that size-biasing a constant leaves it unchanged, Property 4.1 can thus be seen as a particular case of Property 5.1.

Coming back to the Poisson process setting, it is often assumed that the mixing distribution involved in a general Poisson mixture is infinitely divisible. This ensures the coherence between the sum of counts recorded over disjoint time intervals and the corresponding total recorded over the entire observation period (i.e., the union of these intervals). If Λ is infinitely divisible then we know from (2.2) that the distributional equality

$$\tilde{\Lambda} =_d \Lambda + \mathcal{E}$$

holds true for some positive random variable $\mathcal{E}$, independent of Λ . Since the Poisson distribution is stable by convolution, we know that

$$N_{\Lambda+\mathcal{E}} =_d N_\Lambda + N_\mathcal{E},$$

all the random variables being independent. In the setting of Property 5.1, we then have

$$\begin{aligned}\tilde{X}_\Lambda &=_{\mathcal{D}} X_{\Lambda+\mathcal{E}} + \tilde{C} \\ &=_{\mathcal{D}} X_\Lambda + X_\mathcal{E} + \tilde{C},\end{aligned}\tag{5.1}$$

all the random variables being independent. Identity (5.1) shows that size biasing a compound mixed Poisson sum X_Λ corresponds to the sum X_Λ itself (as in the homogeneous Poisson case) plus a compound mixed Poisson sum $X_\mathcal{E}$ corresponds to the size-biasing of the mixing parameter to which the size-biased version $\tilde{C}$ of the generic claim severity is added. Thus, $X_\mathcal{E}$ can be considered as an extra loading, compared to the compound Poisson case.

Let us now apply these results to risk measurement.

Property 5.2. *Under the conditions of Property 5.1,*

$$E[X_\Lambda | X_\Lambda > t] = E[X_\Lambda] \frac{P[X_{\tilde{\Lambda}} + \tilde{C} > t]}{P[X_\Lambda > t]}.$$

If the mixing distribution is infinitely divisible then

$$E[X_\Lambda | X_\Lambda > t] = E[X_\Lambda] \frac{P[X_\Lambda + X_\mathcal{E} + \tilde{C} > t]}{P[X_\Lambda > t]}$$

where $\mathcal{E}$ is the positive random variable involved in the distributional equality $\tilde{\Lambda} =_{\mathcal{D}} \Lambda + \mathcal{E}$ known to be valid from (2.2), all the random variables being independent.

5.2 Negative Binomial case

Assume that the Poisson parameter λ is replaced with a random variable Λ obeying the Gamma distribution with mean α/β and variance α/β^2 , which is henceforth denoted as $\Lambda \sim \mathcal{Gam}(\alpha, \beta)$. Then, the counting distribution is the Negative Binomial $\mathcal{NBin}(\alpha, \beta)$ with probability mass function

$$P[N = k] = \frac{\beta^\alpha}{(1 + \beta)^{\alpha+k}} \frac{\Gamma(\alpha + k)}{k! \Gamma(\alpha)}, \quad k = 0, 1, 2, \dots$$

Property 5.3. *Consider $X = \sum_{k=1}^N C_k$ as described in Property 5.1 with $\Lambda \sim \mathcal{Gam}(\alpha, \beta)$ so that N obeys the $\mathcal{NBin}(\alpha, \beta)$ distribution. Then,*

$$\tilde{X}_\Lambda =_{\mathcal{D}} \tilde{C} + X_\Lambda + X_\mathcal{E}$$

where $\mathcal{E} \sim \mathcal{Gam}(1, \beta)$, all the random variables being independent.

Proof. The probability density function of Λ^* obtained from Λ with the transform (3.6) is given by

$$\begin{aligned}f_{\Lambda^*}(\lambda) &= \frac{\lambda}{\alpha/\beta} f_\Lambda(\lambda) \\ &= \frac{\lambda}{\alpha/\beta} \frac{\lambda^{\alpha-1} \beta^\alpha \exp(-\lambda\beta)}{\Gamma(\alpha)} \\ &= \frac{\lambda^\alpha \beta^{\alpha+1} \exp(-\lambda\beta)}{\Gamma(\alpha + 1)}\end{aligned}$$

so that

$$\Lambda^* =_d \tilde{\Lambda} \sim \mathcal{Gam}(\alpha + 1, \beta).$$

Hence, $\Lambda^* =_d \Lambda + \mathcal{E}$ with $\mathcal{E} \sim \mathcal{Gam}(1, \beta)$ and the announced result then follows from (5.1). $\square$

Define N^* as the mixed Poisson distribution, with mixing parameter Λ^* . Then, $N^* \sim \mathcal{NBin}(\alpha + 1, \beta)$. Now, the size-biased transform of the $\mathcal{NBin}(\alpha, \beta)$ distribution has probability mass function

$$\mathbb{P}[\tilde{N} = k] = \frac{k\mathbb{P}[N = k]}{\alpha/\beta} = \frac{\beta^{\alpha+1}}{(1 + \beta)^{\alpha+k}} \frac{\Gamma(\alpha + k)}{(k - 1)!\Gamma(\alpha + 1)}$$

so that $\tilde{N} - 1 =_d \mathcal{NBin}(\alpha + 1, \beta)$. Hence, N^* is distributed as $\tilde{N} - 1$ and we recover a particular case of the general result contained in Proposition 3.7.

5.3 Application to risk measurement

Let us now consider independent losses $X_1, \dots, X_n$ where each X_i is compound mixed Poisson, that is,

$$X_i = \sum_{k=1}^{N_i} C_{ik}$$

where $N_i \sim \mathcal{Poi}(\lambda)$ given $\Lambda_i = \lambda$ and the claim severities C_{ik} are positive, continuous, and distributed as C_i , all these random variables being independent. Consider the total amount $S = \sum_{i=1}^n X_i$.

In the compound mixed Poisson case, the distributional equality

$$S - X_i + \tilde{X}_i =_d \sum_{j \neq i} X_j + X_{\tilde{\Lambda}_i} + \tilde{C}_i$$

holds true for $i = 1, 2, \dots, n$. Assume furthermore that Λ_i is infinitely divisible. Then, $\tilde{\Lambda}_i =_d \Lambda_i + \mathcal{E}_i$ and the distributional equality (5.1) allows us to write

$$S - X_i + \tilde{X}_i =_d S + X_{\mathcal{E}_i} + \tilde{C}_i.$$

Hence, we finally get

$$\mathbb{E}[X_i | S > s] = \mathbb{E}[X_i] \frac{\mathbb{P}[S + X_{\mathcal{E}_i} + \tilde{C}_i > s]}{\mathbb{P}[S > s]}.$$

Comparing this expression to the compound Poisson case, we see that the extra-randomness in the Poisson parameter induces an additional term $X_{\mathcal{E}_i}$ appearing in the excess probability in the numerator.

6 Extension to conditionally independent risks

6.1 Correlated risks

Let us now turn to a collection of n insurance losses, denoted as $X_1, X_2, \dots, X_n$, possibly correlated. Hence, we consider the random vector, or insurance portfolio $\mathbf{X} = (X_1, \dots, X_n)$ with joint distribution function $F_{\mathbf{X}}$. As before, define the aggregate loss S as $S = \sum_{i=1}^n X_i$. Following the same lines as in Section 1, we can write

$$\begin{aligned} \mathbb{E}[X_i g(S)] &= \int_0^\infty \cdots \int_0^\infty x_i g(x_1 + \dots + x_n) dF_{\mathbf{X}}(x_1, \dots, x_n) \\ &= \mathbb{E}[X_i] \int_0^\infty \cdots \int_0^\infty g(x_1 + \dots + x_n) \frac{x_i dF_{\mathbf{X}}(x_1, \dots, x_n)}{\mathbb{E}[X_i]} \\ &= \mathbb{E}[X_i] \mathbb{E}[g(Y_1 + \dots + Y_n)] \end{aligned}$$

where the joint distribution function $F_{\mathbf{Y}}$ of the random vector $\mathbf{Y} = (Y_1, \dots, Y_n)$ is given by

$$F_{\mathbf{Y}}(y_1, \dots, y_n) = \int_0^{y_1} \cdots \int_0^{y_n} \frac{x_i dF_{\mathbf{X}}(x_1, \dots, x_n)}{\mathbb{E}[X_i]}. \quad (6.1)$$

We thus see that Y_i is distributed as $\tilde{X}_i$, marginally. For $j \neq i$, the marginal distribution of Y_j is now given by

$$\mathbb{P}[Y_j \leq y_j] = \int_0^{y_j} \int_0^\infty \frac{x_i dF_{(X_i, X_j)}(x_i, x_j)}{\mathbb{E}[X_i]} = \frac{\mathbb{E}[X_i | X_j \leq y_j]}{\mathbb{E}[X_i]} \mathbb{P}[X_j \leq y_j]. \quad (6.2)$$

Interestingly, some positive dependence is needed among the components of $\mathbf{X}$ to ensure that the random vector $\mathbf{Y}$ is larger compared to $\mathbf{X}$. We refer the reader to Cohen and Sackrowitz (1995) for various conditions (the weakest one being referred to as association; see e.g. Denuit et al., 2005, Chapter 7, and Property 6.1 below). To see why this is actually needed, assume that the individual risks $X_1, \dots, X_n$ are negatively related. Precisely, assume that X_i is negatively expectation dependent on X_j , that is, the inequality

$$\mathbb{E}[X_i | X_j \leq y_j] \geq \mathbb{E}[X_i]$$

holds true for all y_j . In words, this means that the knowledge that X_j is small, that is, X_j falls below the threshold y_j , makes X_i larger on average. We then see from (6.2) that the inequality $\mathbb{P}[Y_j \leq y_j] \geq \mathbb{P}[X_j \leq y_j]$ is valid for all y_j , so that Y_j is smaller than X_j in first-order stochastic dominance: the probability that Y_j falls below any threshold y_j is larger than the corresponding probability for X_j . Switching from $\mathbf{X}$ to $\mathbf{Y}$ is thus not necessarily a conservative strategy in this case. To prevent such a phenomenon to occur, some positive dependence is needed among $X_1, X_2, \dots, X_n$.

We also see that

$$F_{\mathbf{Y}}(y_1, \dots, y_n) = \frac{\mathbb{E}[X_i | X_1 \leq y_1, \dots, X_n \leq y_n]}{\mathbb{E}[X_i]} F_{\mathbf{X}}(y_1, \dots, y_n).$$

Hence, to ensure that $F_{\mathbf{X}}$ dominates $F_{\mathbf{Y}}$, so that $\mathbf{Y}$ is larger than $\mathbf{X}$ (in the sense of the lower orthant order, that is, the joint distribution function $F_{\mathbf{X}}$ dominates $F_{\mathbf{Y}}$ everywhere), the inequality

$$\mathbb{E}[X_i | X_1 \leq y_1, \dots, X_n \leq y_n] \leq \mathbb{E}[X_i]$$

has to be valid for all $y_1, \dots, y_n$. This condition expresses some positive relationship between the individual risks $X_1, \dots, X_n$. It appears to be similar to the one imposed by Guo et al. (2016); see also Denuit and Mesfioui (2017).

Considering independent losses $X_1, \dots, X_n$, we see that $Y_1, \dots, Y_n$ are also independent random variables. Furthermore, Y_j is distributed as X_j for $j \neq i$ whereas Y_i is distributed as $\tilde{X}_i$.

In the remainder of this section, we consider conditionally independent risks. This construction is widely used in actuarial models and allows us to derive extensions of the results obtained earlier in this paper.

6.2 Common mixture model

The common mixture model (in the terminology of Wang, 1998) consists in conditionally independent risks and forms the basis of the credibility approach. The kind of dependence induced by this construction is comprehensively studied in Denuit et al. (2005, Chapter 7). The intuition behind this modeling approach is as follows: an external mechanism, described by the positive random variable Λ , influences several risks $X_1, X_2, \dots, X_n$. Given the environmental parameter Λ , the individual risks are independent. Formally, the joint distribution function of the portfolio vector $\mathbf{X}$ can be written as

$$\begin{aligned} F_{\mathbf{X}}(t_1, \dots, t_n) &= \mathbb{E}[\mathbb{P}[X_1 \leq t_1, \dots, X_n \leq t_n | \Lambda]] \\ &= \int_0^\infty \left(\prod_{i=1}^n \mathbb{P}[X_i \leq t_i | \Lambda = \lambda] \right) dF_\Lambda(\lambda). \end{aligned} \quad (6.3)$$

Notice that this construction is rather general and covers for instance the case of the common shock model, where each risk X_i is obtained as the sum of two independent random variables, where the second one is common to all risks. This common shock then plays the role of Λ in the common mixture model. More examples are considered in the remainder of this section.

It is reasonable to expect that the risks X_i resulting from this construction are positively correlated provided they all move in the same direction with Λ . This makes the random vector $\mathbf{Y}$ with distribution function (6.1) larger so that size-biasing induces a safety margin in this case. This is formally stated in the next result, of which we provide the reader with an elementary proof borrowed from Shaked and Shanthikumar (2007, Theorem 6.B.8).

Property 6.1. *Consider the risks $X_1, X_2, \dots, X_n$ with joint distribution function (6.3). Assume that all the risks X_i become larger when Λ increases, in the sense of first-order stochastic dominance, that is,*

$$\mathbb{P}[X_i \leq t | \Lambda = \lambda] \geq \mathbb{P}[X_i \leq t | \Lambda = \lambda + \delta] \text{ for all } t, \lambda > 0, \delta > 0, \text{ and } i = 1, 2, \dots, n.$$

Then, the random vector $\mathbf{Y}$ with distribution function (6.1) is larger than $\mathbf{X}$ in the sense of multivariate first-order stochastic dominance, that is, the inequality $\mathbb{E}[g(\mathbf{X})] \leq \mathbb{E}[g(\mathbf{Y})]$ holds true for every non-decreasing function g such that the expectations exist.

Proof. We know from Property 7.2.16 in Denuit et al. (2005) that the risks $X_1, \dots, X_n$ are associated random variables, that is, the covariance $C[g_1(\mathbf{X}), g_2(\mathbf{Y})]$ is non-negative for all non-decreasing functions g_1 and g_2 . Now, considering a non-decreasing function g ,

$$\begin{aligned} E[g(\mathbf{Y})] &= \int_0^\infty \cdots \int_0^\infty g(\mathbf{y}) dF_{\mathbf{Y}}(\mathbf{y}) \\ &= \int_0^\infty \cdots \int_0^\infty g(\mathbf{y}) \frac{dF_{\mathbf{Y}}(\mathbf{y})}{dF_{\mathbf{X}}(\mathbf{y})} dF_{\mathbf{X}}(\mathbf{y}) \\ &\geq \int_0^\infty \cdots \int_0^\infty g(\mathbf{y}) dF_{\mathbf{X}}(\mathbf{y}) \int_0^\infty \cdots \int_0^\infty \frac{dF_{\mathbf{Y}}(\mathbf{y})}{dF_{\mathbf{X}}(\mathbf{y})} dF_{\mathbf{X}}(\mathbf{y}) \\ &= E[g(\mathbf{X})] \end{aligned}$$

where the inequality comes from the covariance inequality defining association and the non-decreasingness of the ratio $dF_{\mathbf{Y}}(\mathbf{y})/dF_{\mathbf{X}}(\mathbf{y})$. $\square$

A typical situation where the common mixture construction applies is as follows. Assume that claim amounts are subject to inflation so that $X_i = Z_i \Lambda$ where Z_i is the nominal claim amount and Λ models the impact of inflation. The random variables $Z_1, \dots, Z_n, \Lambda$ are assumed to be independent. Clearly, the risks X_i are positively related so that size-biasing indeed induces the necessary safety loading. Then,

$$S = \Lambda \sum_{i=1}^n Z_i \text{ and } \tilde{S} =_d \widetilde{\Lambda \sum_{i=1}^n Z_i}$$

invoking Property 3.5. Hence,

$$\tilde{S} =_d \tilde{\Lambda} \left(\sum_{i=1}^n Z_i - Z_K + \tilde{Z}_K \right)$$

where the random variable K is defined as in Property 3.1, that is,

$$P[K = i] = \frac{E[X_i]}{E[S]} = \frac{E[Z_i]}{\sum_{j=1}^n E[Z_j]}.$$

6.3 Conditionally independent compound mixed Poisson sums

Let us now assume that given $\Lambda = \lambda$ each X_i is a compound Poisson sum with $N_i \sim \mathcal{Poi}(a_i \lambda)$. Claim severities remain independent but the random vector of claim frequencies $(N_1, \dots, N_n)$ now obeys the common mixture model. In this case, all components X_i increase in Λ in the first-order stochastic dominance (see Chapter 7 in Denuit et al., 2005). This construction is a particular case of the model considered in Kim et al. (2019). The following result gives the size-biased version of S in this case.

Property 6.2. *Assume that $X_i = \sum_{k=1}^{N_i} C_{ik}$ where the claim severities C_{ik} are positive, continuous, and distributed as C_i , all these random variables being independent, $i = 1, 2, \dots, n$,*

and independent of $(N_1, \dots, N_n)$. Each N_i obeys the $\mathcal{Poi}(a_i\Lambda)$ distribution given Λ , for some positive constants $a_1, \dots, a_n$. Then,

$$\tilde{S} =_d S_{\tilde{\Lambda}} + \tilde{C}_K$$

where the random variable K is valued in $\{1, \dots, n\}$ with

$$P[K = i] = \frac{a_i E[C_i]}{\sum_{j=1}^n a_j E[C_j]}, \quad i = 1, \dots, n.$$

Proof. Here, S appears to be a mixture of S_λ where $S_\lambda =_d \sum_{k=1}^M Z_k$, with $M \sim \mathcal{Poi}(a_\bullet \lambda)$ where $a_\bullet = \sum_{i=1}^n a_i$ and $Z_1, Z_2, Z_3, \dots$ are independent and distributed as Z with distribution function

$$P[Z \leq t] = \sum_{i=1}^n \frac{a_i}{a_\bullet} P[C_i \leq t].$$

All the random variables involved in S_λ are independent. Thus, $Z =_d C_J$ with $P[J = i] = \frac{a_i}{a_\bullet}$. We know from Property 4.1 that

$$\tilde{S}_\lambda =_d S_\lambda + \tilde{Z}.$$

Property 3.4 shows that $\tilde{Z}$ obeys a mixture of $\tilde{C}_1, \dots, \tilde{C}_n$ with mixing distribution

$$\begin{aligned} P[J^* = i] &= \frac{E[C_i]}{E[C_J]} P[J = i] \\ &= \frac{a_i E[C_i]}{\sum_{j=1}^n a_j E[C_j]}. \end{aligned}$$

Hence, $\tilde{Z} =_d C_K$ and the distribution of $\tilde{Z}$ does not involve Λ . Now, the size-biased transform of S is a mixture of the size-biased versions $\tilde{S}_\lambda$, with mixing distribution

$$\begin{aligned} dP[\Lambda^* \leq \lambda] &= \frac{E[S_\lambda]}{E[S]} dP[\Lambda \leq \lambda] \\ &= \frac{\lambda}{E[\Lambda]} dP[\Lambda \leq \lambda] \end{aligned}$$

so that $\Lambda^* =_d \Lambda$. This ends the proof. $\square$

Notice that S in Property 6.2 is a compound Poisson sum so that the announced result can also be obtained as a direct consequence of Property 5.1. The results derived in this section remain nevertheless useful to derive the individual contributions $E[X_i | S > t]$ of the conditionally independent compound mixed Poisson sums X_i .

7 Discussion

This paper offers a systematic treatment of the size-biased transform of compound distributions that are particularly relevant for actuarial applications. We consider the cases of

independence and conditional independence, given a common risk factor. The results established in this paper are applied to risk measures derived from the size-biased transform, like the conditional tail expectation.

The formulas established in this paper provide actuaries with effective ways to compute risk measures, mainly based on Panjer recursive formulas in the compound Poisson and Negative Binomial cases. Besides they also allow for a deeper understanding of the way risk measurement proceeds as extra randomness is reflected in additional terms involved in the size-biased version of the initial risk. For instance, switching from a compound homogeneous Poisson sum to a compound mixed Poisson sum is reflected into an additional term involving the size-biased version of the mixing parameter.

References

- Arratia, R., Goldstein, L. (2010). Size bias, sampling, the waiting time paradox, and infinite divisibility: when is the increment independent? arXiv preprint arXiv:1007.3910.
- Arratia, R., Goldstein, L., Kochman, F. (2019). Size bias for one and all. *Probability Surveys* 16, 1-61.
- Bartoszewicz, J., Skolimowska, M. (2006). Preservation of classes of life distributions and stochastic orders under weighting. *Statistics & Probability Letters* 76, 587-596.
- Brown, M. (2006). Exploiting the waiting time paradox: Applications of the size-biasing transformation. *Probability in the Engineering and Informational Sciences* 20, 195-230.
- Cohen, A., Sackrowitz, H.B. (1995). On stochastic ordering of random vectors. *Journal of Applied Probability* 32, 960-965.
- Cossette, H., Mailhot, M., Marceau, E. (2012). TVaR-based capital allocation for multivariate compound distributions with positive continuous claim amounts. *Insurance: Mathematics and Economics* 50, 247-256.
- Denuit, M., Dhaene, J., Goovaerts, M.J., Kaas, R. (2005). *Actuarial Theory for Dependent Risks: Measures, Orders and Models*. Wiley, New York.
- Denuit, M., Mesfioui, M. (2017). Preserving the Rothschild-Stiglitz type increase in risk with background risk: A characterization. *Insurance: Mathematics and Economics* 72, 1-5.
- Furman, E., Landsman, Z. (2005). Risk capital decomposition for a multivariate dependent gamma portfolio. *Insurance: Mathematics and Economics* 37, 635-649.
- Furman, E., Landsman, Z. (2008a). Economic capital allocations for non-negative portfolios of dependent risks. *ASTIN Bulletin* 38, 601-619.
- Furman, E., Landsman, Z. (2008b). Economic capital allocations for non-negative portfolios of dependent risks. *ASTIN Bulletin* 38, 601-619.

- Furman, E., Landsman, Z. (2010). Multivariate Tweedie distributions and some related capital-at-risk analyses. *Insurance: Mathematics and Economics* 46, 351-361.
- Furman, E., Zitikis, R. (2008a). Weighted risk capital allocations. *Insurance: Mathematics and Economics* 43, 263-269.
- Furman, E., Zitikis, R. (2008b). Weighted premium calculation principles. *Insurance: Mathematics and Economics* 42, 459-465.
- Furman, E., Zitikis, R. (2009). Weighted pricing functionals with applications to insurance: An overview. *North American Actuarial Journal* 13, 483-496.
- Guo, X., Li, J., Liu, D., Wang, J. (2016). Preserving the RothschildStiglitz type of increasing risk with background risk. *Insurance: Mathematics and Economics* 70, 144-149.
- Jain, K., Singh, H., Bagai, I. (1989). Relations for reliability measures of weighted distributions. *Communications in Statistics-Theory and Methods* 18, 4393-4412.
- Kaas, R., Goovaerts, M.J., Dhaene, J., Denuit, M. (2008). *Modern Actuarial Risk Theory Using R*. Springer, New York.
- Kim, J. H., Jang, J., Pyun, C. (2019). Capital allocation for a sum of dependent compound mixed Poisson variables: A recursive algorithm approach. *North American Actuarial Journal*, in press.
- Klugman, S.A., Panjer, H.H., Willmot, G.E. (2004). *Loss Models: From Data to Decisions*. Wiley, New York.
- Pakes, A. G. (1997). Characterization by invariance under length-biasing and random scaling. *Journal of Statistical and Inference* 63, 285-310.
- Pakes, A. G., Sapatinas, T., Fosam, E. B. (1996). Characterizations, length-biasing, and infinite divisibility. *Statistical Papers* 37, 53-69.
- Patil, G.P., Rao, C.R. (1978). Weighted distributions and size-biased sampling with applications to wildlife populations and human families. *Biometrics* 34, 179-189.
- Ross, S.M. (2003). The inspection paradox. *Probability in the Engineering and Informational Sciences* 17, 47-51.
- Shaked, M., Shanthikumar, J.G. (2007). *Stochastic Orders*. Springer, New York.
- Steutel, F.W., van Harn, K. (2003). *Infinite Divisibility of Probability Distributions on the Real Line*. CRC Press.
- Wang, S. (1998). Aggregation of correlated risk portfolios: Models and algorithms. *Proceedings of the Casualty Actuarial Society*, 848-939.