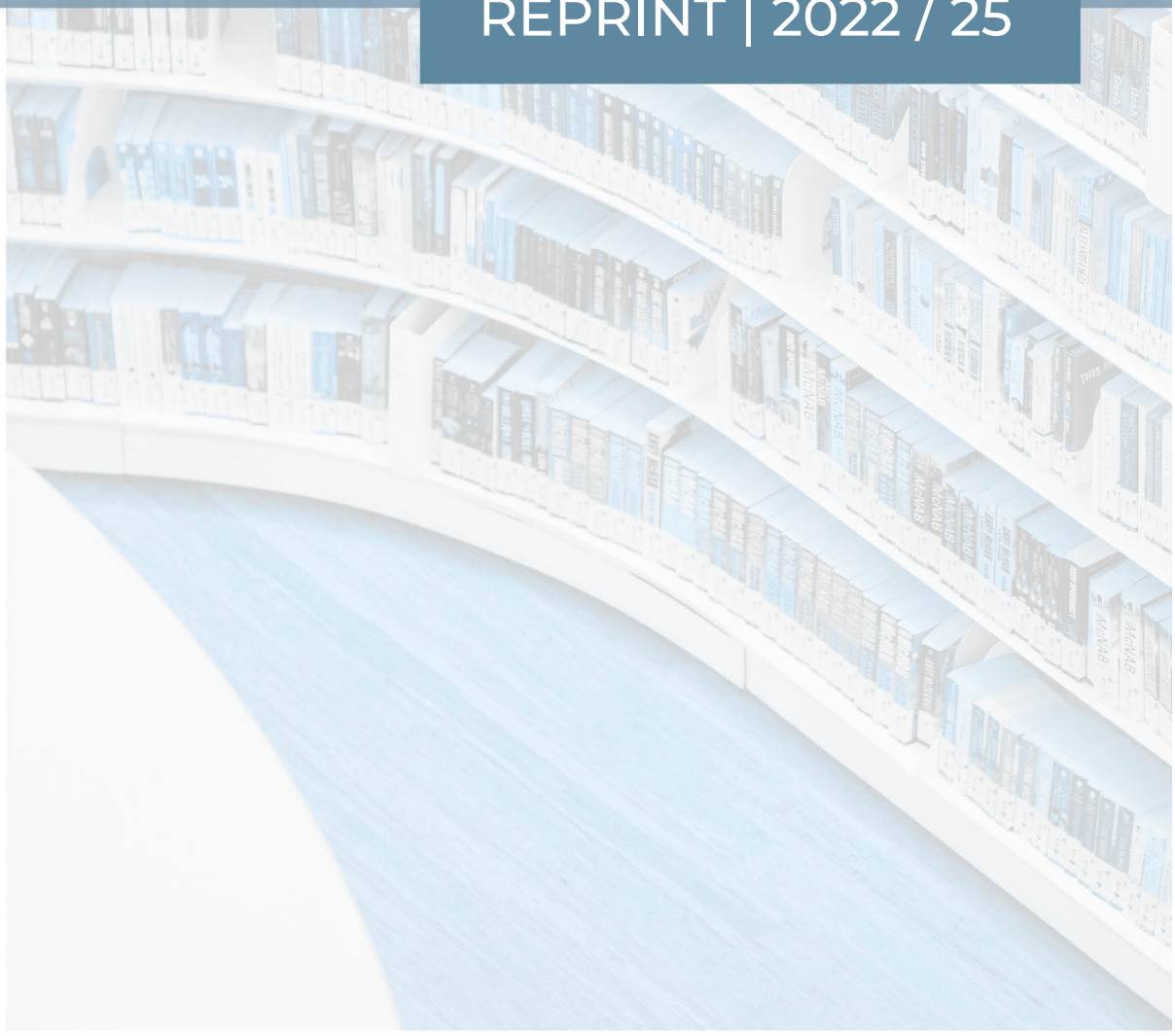
CONDITIONAL TAIL EXPECTATION DECOMPOSITION AND CONDITIONAL MEAN RISK SHARING FOR DEPENDENT AND CONDITIONALLY INDEPENDENT LOSSES

Michel Denuit, Christian Y. Robert

REPRINT | 2022 / 25



UCLouvain

ISBA | LIDAM Louvain Institute of Data Analysis and Modeling in economics and statistics

ISBA

Voie du Roman Pays 20 - L1.04.01

B-1348 Louvain-la-Neuve

Email : lidam-library@uclouvain.be

https://uclouvain.be/en/research-institutes/lidam/isba/publication.html

CONDITIONAL TAIL EXPECTATION DECOMPOSITION AND CONDITIONAL MEAN RISK SHARING FOR DEPENDENT AND CONDITIONALLY INDEPENDENT RISKS

Michel Denuit
Institute of Statistics, Biostatistics and Actuarial Science - ISBA
Louvain Institute of Data Analysis and Modeling - LIDAM
UCLouvain
Louvain-la-Neuve, Belgium
michel.denuit@uclouvain.be

Christian Y. Robert
Laboratory in Finance and Insurance - LFA
CREST - Center for Research in Economics and Statistics
ENSAE
Paris, France
chrobert@ensae.fr

June 8, 2020

Abstract

Conditional tail expectations are often used in risk measurement and capital allocation. Conditional mean risk sharing appears to be effective in collaborative insurance, to distribute total losses among participants. This paper develops analytical results for risk allocation among different, correlated units based on conditional tail expectations and conditional mean risk sharing. Results available in the literature for independent risks are extended to correlated ones, in a unified way. The approach is applied to mixture models with correlated latent factors that are often used in insurance studies.

Keywords: Weighted distributions, size-biased transform, mixture models.

1 Introduction and motivation

The paper aims at contributing to the rich literature on risk allocation and risk sharing that are both core topics in actuarial science. In the former case, actuaries often need to allocate the available amount of capital across various entities, such as business lines, territories or even individual products. Costs of holding capital can then be distributed among entities so that it is fairly supported by policyholders, on the one hand, and the respective performances of each entity can be assessed, on the other hand. There are many ways to allocate capital, most of them being based on risk measures such as Conditional Tail Expectation (CTE) considered in this paper.

Risk sharing is a related, though distinct problem. In that context, economic agents hand their individual losses over to a pool and agree on some rule as to how the total pooled loss has to be divided among them. In the present paper, we focus on the conditional mean risk sharing, as defined by Denuit and Dhaene (2012). According to this rule, each participant to an insurance pool contributes the conditional expectation of the loss brought to the pool, given the total loss experienced by the entire pool.

In both cases, we wish to determine how to split the risk measure of the total risk, or the total risk itself, among individual entities. Precisely, the contribution of this paper is as follows. Representations for capital allocation based on CTE and conditional mean risk sharing derived in the literature for independent risks are extended here to correlated ones. To this end, we use a multivariate weighted distribution considered by Arratia et al. (2019) in the context of size-biasing sums of random variables. Several examples are discussed, including Liouville and infinitely divisible multivariate distributions. Mixture models where correlation arises from latent variables are carefully studied because of their wide applicability to insurance studies.

The remainder of this paper is organized as follows. Section 2 establishes the decomposition of CTE among correlated business lines whereas Section 3 considers the conditional mean risk sharing of correlated losses. As an application, Section 4 deals with multivariate mixture models.

2 General CTE decomposition formulas

2.1 CTE and size-biasing

Consider an insurance loss modeled as a non-negative random variable X with distribution function F_X and expected value $\mathrm{E}[X]$ such that $0 < \mathrm{E}[X] < \infty$. Several risk measures for X are based on the function

$$t \mapsto \mathrm{E}[X|X > t] = \frac{1}{\mathrm{P}[X > t]} \int_t^\infty x \mathrm{d}F_X(x)$$

where t is such that $\mathrm{P}[X > t] > 0$. The Conditional Tail Expectation (CTE) is obtained when t corresponds to some 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

$$\begin{aligned} \mathrm{E}[Xg(X)] &= \int_0^\infty xg(x)\mathrm{d}F_X(x) \\ &= \mathrm{E}[X]\int_0^\infty g(x)\frac{x\mathrm{d}F_X(x)}{\mathrm{E}[X]} \\ &= \mathrm{E}[X]\mathrm{E}[g(\widetilde{X})] \end{aligned} \tag{2.1}$$

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

$$\mathrm{P}[\widetilde{X} \leq t] = \frac{1}{\mathrm{E}[X]}\int_0^t x\mathrm{d}F_X(x),$$

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

$$\mathrm{P}[Z \leq t] = \frac{\mathrm{P}[\widetilde{X} \leq t]}{\mathrm{P}[X \leq t]} = \frac{\mathrm{E}[X|X \leq t]}{\mathrm{E}[X]}. \tag{2.2}$$

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

Now, inserting the function

$$g(x) = \mathrm{I}[x > t] = \begin{cases} 1 \text{ if } x > t \\ \\ 0 \text{ otherwise} \end{cases} \tag{2.3}$$

in identity (2.1), we see that the representation

$$\mathrm{E}[X|X > t] = \mathrm{E}[X]\frac{\mathrm{P}[\widetilde{X} > t]}{\mathrm{P}[X > t]} \tag{2.4}$$

is valid for any threshold t. The ratio $\mathrm{P}[\widetilde{X} > t]/\mathrm{P}[X > t]$ appearing in (2.4) exceeds unity so that we recover the classical inequality $\mathrm{E}[X|X > t] \geq \mathrm{E}[X]$ for all t from elementary probability.

The size-biased transform can be traced back to the late 1960s in the statistical literature. It 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. The size-biased transform has proven to be useful in the study of risk measures after the pioneering work by Furman and Landsman (2005, 2008) and Furman and Zitikis (2008a,b) in connection to (2.4).

2.2 Independent losses

Consider n insurance losses modeled as non-negative, mutually independent random variables $X_1, X_2, \ldots, X_n$, each with $0 < \mathrm{E}[X_i] < \infty$. Define the aggregate loss S as $S = \sum_{i=1}^n X_i$. Clearly,

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

so that the total $\mathrm{E}[S|S > t]$ can be decomposed into the sum of individual contributions $\mathrm{E}[X_i|S > t]$ that are useful in risk management to allocate the CTE across lines of business, for instance (taking for t a quantile of S). It is known from Furman and Landsman (2005) that the representation

$$\mathrm{E}[X_i|S > t] = \mathrm{E}[X_i]\frac{\mathrm{P}[S - X_i + \widetilde{X}_i > t]}{\mathrm{P}[S > t]} \tag{2.5}$$

holds true, where the size-biased version $\widetilde{X}_i$ of X_i is assumed to be independent of $X_1, X_2, \ldots, X_n$. In words, the contribution $\mathrm{E}[X_i|S > t]$ of risk X_i to $\mathrm{E}[S|S > t]$ is equal to its expected value $\mathrm{E}[X_i]$ increased by the ratio of the excess probabilities $\mathrm{P}[S - X_i + \widetilde{X}_i > t]$ and $\mathrm{P}[S > t]$, where

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

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

Identities (2.4)-(2.5) have been exploited in a number of papers (including those cited in the preceding section) to derive many useful properties for risk measures, mainly for continuous losses assumed to be mutually independent. In the next section, we extend these results to general losses, possibly correlated.

2.3 Dependent losses

Let us now extend the representation (2.5) recalled in the preceding section to correlated risks. To this end, consider a collection of n insurance losses, modeled as n non-negative, possibly correlated random variables $X_1, X_2, \ldots, X_n$. Hence, we consider the random vector, or insurance portfolio $\boldsymbol{X} = (X_1, \ldots, X_n)$ with joint distribution function $F_{\boldsymbol{X}}$. As before, define the aggregate loss S as $S = \sum_{i=1}^n X_i$.

For each $i \in \{1, 2, \ldots, n\}$, define the random vector $\boldsymbol{Y}^{[i]} = (Y_1^{[i]}, \ldots, Y_n^{[i]})$ with joint distribution function $F_{\boldsymbol{Y}^{[i]}}$ given by

$$\begin{aligned} F_{\boldsymbol{Y}^{[i]}}(y_1, \ldots, y_n) &= \int_0^{y_1} \cdots \int_0^{y_n} \frac{x_i \mathrm{d}F_{\boldsymbol{X}}(x_1, \ldots, x_n)}{\mathrm{E}[X_i]} \\ &= \frac{\mathrm{E}[X_i|X_1 \leq y_1, \ldots, X_n \leq y_n]}{\mathrm{E}[X_i]} F_{\boldsymbol{X}}(y_1, \ldots, y_n). \end{aligned} \tag{2.6}$$

The random vector $\boldsymbol{Y}^{[i]}$ with distribution function (2.6) plays a central role in the extensions of the results presented in the preceding section to correlated risks. Distributions (2.6)

have been considered by Arratia et al. (2019) in relation to size-biasing sums of random variables. See Section 2.4 below for details. The distribution function (2.6) is a multivariate weighted version of the distribution function for $\boldsymbol{X}$. Multivariate weighted distributions have been reviewed by Navarro et al. (2006). The weight function w corresponding to (2.6) is $w(x_1, \ldots, x_n) = x_i/\mathrm{E}[X_i]$ that has been considered in Jain and Nanda (1995).

We see from (2.6) that $Y_i^{[i]}$ is distributed as $\widetilde{X}_i$, marginally. The marginal distribution of $Y_j^{[i]}$ for $j \neq i$ is given by

$$\begin{aligned}
\mathrm{P}[Y_j^{[i]} \leq y_j] &= \int_0^{y_j} \int_0^{\infty} \frac{x_i \mathrm{d}F_{(X_i,X_j)}(x_i, x_j)}{\mathrm{E}[X_i]} \\
&= \frac{\mathrm{E}\big[X_i \mathrm{I}[X_j \leq y_j]\big]}{\mathrm{E}[X_i]} \\
&= \mathrm{P}[X_j \leq y_j] + \frac{\mathrm{Cov}\big[X_i, \mathrm{I}[X_j \leq y_j]\big]}{\mathrm{E}[X_i]} \qquad (2.7) \\
&= \frac{\mathrm{E}[X_i | X_j \leq y_j]}{\mathrm{E}[X_i]} \mathrm{P}[X_j \leq y_j], \qquad (2.8)
\end{aligned}$$

where $\mathrm{I}[\cdot]$ is the indicator function, equal to 1 if the condition appearing between the brackets is satisfied, and to 0 otherwise. Interestingly, some positive dependence is needed among the components of $\boldsymbol{X}$ to ensure that the components of the random vector $\boldsymbol{Y}^{[i]}$ are "larger" compared to those of $\boldsymbol{X}$. If X_i and X_j are negatively related then it can be expected that X_i and $\mathrm{I}[X_j \leq y_j]$ are positively correlated (since $x_j \mapsto \mathrm{I}[x_j \leq y_j]$ is a decreasing function) so that we see from (2.7) that the inequality $\mathrm{P}[Y_j^{[i]} \leq y_j] \geq \mathrm{P}[X_j \leq y_j]$ holds true. If this is the case for all y_j then $Y_j^{[i]}$ tends to be smaller compared to X_j, in the sense that $Y_j^{[i]}$ is dominated by X_j in (first-order) stochastic dominance. This is for instance the case when X_i is negatively expectation dependent on X_j, that is, when the inequality

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

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 (2.8) that the inequality $\mathrm{P}[Y_j^{[i]} \leq y_j] \geq \mathrm{P}[X_j \leq y_j]$ is valid for all y_j, so that $Y_j^{[i]}$ is smaller than X_j in (first-order) stochastic dominance, as mentioned previously. Switching from $\boldsymbol{X}$ to $\boldsymbol{Y}^{[i]}$ is thus not necessarily a conservative strategy in this case. To prevent such a phenomenon to occur, some positive dependence is needed among $X_1, X_2, \ldots, X_n$.

Considering (2.6), in order to ensure that $F_{\boldsymbol{X}}$ dominates $F_{\boldsymbol{Y}^{[i]}}$, so that $\boldsymbol{Y}^{[i]}$ is larger than $\boldsymbol{X}$ (in the sense of the lower orthant order, that is, the joint distribution function $F_{\boldsymbol{X}}$ dominates $F_{\boldsymbol{Y}^{[i]}}$ everywhere), the inequality

$$\mathrm{E}[X_i | X_1 \leq y_1, \ldots, X_n \leq y_n] \leq \mathrm{E}[X_i]$$

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

We know from Proposition 6.1(v) in Navarro et al. (2006) that provided $\boldsymbol{X}$ possesses positively associated components, $\boldsymbol{Y}^{[i]}$ is larger than $\boldsymbol{X}$ in multivariate stochastic dominance. Notice that

$$\mathrm{d}F_{\boldsymbol{Y}^{[i]}}(y_1,\ldots,y_n) = \frac{y_i}{\mathrm{E}[X_i]}\mathrm{d}F_{\boldsymbol{X}}(y_1,\ldots,y_n)$$

so that the ratio $\mathrm{d}F_{\boldsymbol{Y}^{[i]}}/\mathrm{d}F_{\boldsymbol{X}}$ is non-decreasing. This corresponds to condition (1.8) in Cohen and Sackrowitz (1995). Together with association, that condition is known to imply multivariate (first-order) stochastic dominance. 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 2.1. *Assume that the risks $X_1,\ldots,X_n$ are associated random variables, that is, the covariance $\mathrm{Cov}[g_1(\boldsymbol{X}),g_2(\boldsymbol{X})]$ is non-negative for all non-decreasing functions g_1 and g_2. Then, the random vector $\boldsymbol{Y}^{[i]}$ with distribution function (2.6) is larger than $\boldsymbol{X}$ in the sense of multivariate first-order stochastic dominance, that is, the inequality $\mathrm{E}[g(\boldsymbol{X})] \leq \mathrm{E}[g(\boldsymbol{Y}^{[i]})]$ holds true for every non-decreasing function g such that the expectations exist.*

Proof. Considering a non-decreasing function g,

$$\begin{aligned}
\mathrm{E}[g(\boldsymbol{Y}^{[i]})] &= \int_0^\infty \cdots \int_0^\infty g(\boldsymbol{y})\mathrm{d}F_{\boldsymbol{Y}^{[i]}}(\boldsymbol{y}) \\
&= \int_0^\infty \cdots \int_0^\infty g(\boldsymbol{y})\frac{\mathrm{d}F_{\boldsymbol{Y}^{[i]}}(\boldsymbol{y})}{\mathrm{d}F_{\boldsymbol{X}}(\boldsymbol{y})}\mathrm{d}F_{\boldsymbol{X}}(\boldsymbol{y}) \\
&\geq \int_0^\infty \cdots \int_0^\infty g(\boldsymbol{y})\mathrm{d}F_{\boldsymbol{X}}(\boldsymbol{y}) \int_0^\infty \cdots \int_0^\infty \frac{\mathrm{d}F_{\boldsymbol{Y}^{[i]}}(\boldsymbol{y})}{\mathrm{d}F_{\boldsymbol{X}}(\boldsymbol{y})}\mathrm{d}F_{\boldsymbol{X}}(\boldsymbol{y}) \\
&= \mathrm{E}[g(\boldsymbol{X})]
\end{aligned}$$

where the inequality comes from the covariance inequality defining association and the non-decreasingness of the ratio $\mathrm{d}F_{\boldsymbol{Y}^{[i]}}(\boldsymbol{y})/\mathrm{d}F_{\boldsymbol{X}}(\boldsymbol{y})$. This ends the proof. □

Property 2.1 considers risks $X_1,\ldots,X_n$ that are associated random variables. Association is a positive dependence concept reviewed in Section 7.2.3 of Denuit et al. (2005). It is known to increase many risk measures compared to the independent case.

The next result gives a general representation formula for $\mathrm{E}[X_i|S>s]$ based on the random vector $\boldsymbol{Y}^{[i]}$. As a corollary, we obtain the existing result for independent risks recalled in Section 2.2.

Proposition 2.2. *Consider insurance losses $X_1,\ldots,X_n$ with joint distribution function $F_{\boldsymbol{X}}$ and expected values $\mathrm{E}[X_i]$ such that $0<\mathrm{E}[X_i]<\infty$. Let $\boldsymbol{Y}^{[i]}=(Y_1^{[i]},\ldots,Y_n^{[i]})$ be a random vector with joint distribution function $F_{\boldsymbol{Y}^{[i]}}$ given by (2.6). Then for any $s\geq 0$ and for any $i\in\{1,2,\ldots,n\}$, we have*

$$\mathrm{E}[X_i|S>s] = \mathrm{E}[X_i]\,\frac{\mathrm{P}[Y_1^{[i]}+\ldots+Y_n^{[i]}>s]}{\mathrm{P}[X_1+\ldots+X_n>s]}.$$

Proof. Following the lines leading to (2.1), we obtain

$$\begin{aligned}
\mathrm{E}[X_i g(S)] &= \int_0^\infty \cdots \int_0^\infty x_i g(x_1 + \ldots + x_n) \mathrm{d}F_{\boldsymbol{X}}(x_1, \ldots, x_n) \\
&= \mathrm{E}[X_i] \int_0^\infty \cdots \int_0^\infty g(x_1 + \ldots + x_n) \frac{x_i \mathrm{d}F_{\boldsymbol{X}}(x_1, \ldots, x_n)}{\mathrm{E}[X_i]} \\
&= \mathrm{E}[X_i] \mathrm{E}[g(Y_1^{[i]} + \ldots + Y_n^{[i]})]. \qquad (2.9)
\end{aligned}$$

Now, let us apply (2.9) to the function g given by

$$g(x_1, \ldots, x_n) = \mathrm{I}\left[\sum_{j=1}^n x_j > s\right]$$

to obtain the identity

$$\mathrm{P}\left[Y_1^{[i]} + \ldots + Y_n^{[i]} > s\right] = \frac{\mathrm{E}\left[X_i \mathrm{I}\left[S > s\right]\right]}{\mathrm{E}[X_i]}.$$

This ends the proof. □

Corollary 2.3. *Consider independent risks $X_1, X_2, \ldots, X_n$. Let $\widetilde{X}_1, \widetilde{X}_2, \ldots, \widetilde{X}_n$ be their corresponding size-biased versions, assumed to be independent and independent of $X_1, X_2, \ldots, X_n$. If $X_1, \ldots, X_n$ are mutually independent then $Y_1^{[i]}, \ldots, Y_n^{[i]}$ are also independent random variables. Furthermore,* (2.8) *shows that $Y_j^{[i]}$ is distributed as X_j for $j \neq i$, and $Y_i^{[i]} \overset{\mathrm{d}}{=} \widetilde{X}_i$, where "$\overset{\mathrm{d}}{=}$" means "is distributed as". The distributional equality*

$$Y_1^{[i]} + \ldots + Y_n^{[i]} \overset{\mathrm{d}}{=} S - X_i + \widetilde{X}_i$$

thus holds true when $X_1, \ldots, X_n$ are mutually independent and we recover (2.5) *as a particular case.*

2.4 Size-biased transform of a sum of dependent risks

Arratia et al. (2019, Section 2.4) studied the effect of size-biasing on a sum of random variables. Their result can also be obtained with the help of the CTE decomposition derived in Proposition 2.2. This can be shown as follows. On the one hand, the identity appearing in equation (2.4) allows us to write

$$\mathrm{E}[S|S > t] = \mathrm{E}[S] \frac{\mathrm{P}[\widetilde{S} > t]}{\mathrm{P}[S > t]}$$

for any threshold t. On the other hand, Proposition 2.2 allows us to write

$$\begin{aligned}
\mathrm{E}[S|S > s] &= \sum_{i=1}^n \mathrm{E}[X_i|S > s] \\
&= \sum_{i=1}^n \mathrm{E}\left[X_i\right] \frac{\mathrm{P}[Y_1^{[i]} + \ldots + Y_n^{[i]} > s]}{\mathrm{P}[X_1 + \ldots + X_n > s]}.
\end{aligned}$$

Combining these identities, we get

$$\mathrm{P}[\widetilde{S} > t] = \sum_{i=1}^{n} \frac{\mathrm{E}\left[X_i\right]}{\mathrm{E}[S]} \mathrm{P}[Y_1^{[i]} + \ldots + Y_n^{[i]} > t].$$

This shows that $\widetilde{S}$ obeys a discrete mixture with mixing distribution assigning probability $\frac{\mathrm{E}[X_i]}{\mathrm{E}[S]}$ on index i. Precisely, define the random variable K valued in $\{1, 2, \ldots, n\}$, independent of $\boldsymbol{X}$ and of $\boldsymbol{Y}^{[1]}, \ldots, \boldsymbol{Y}^{[n]}$ such that $\mathrm{P}[K = k] = \frac{\mathrm{E}[X_k]}{\mathrm{E}[S]}$ for $k \in \{1, 2, \ldots, n\}$. Then,

$$\widetilde{S} \overset{\mathrm{d}}{=} \sum_{i=1}^{n} Y_i^{[K]} = \begin{cases} Y_1^{[1]} + \ldots + Y_n^{[1]} \text{ with probability } \frac{\mathrm{E}[X_1]}{\mathrm{E}[S]} \\ Y_1^{[2]} + \ldots + Y_n^{[2]} \text{ with probability } \frac{\mathrm{E}[X_2]}{\mathrm{E}[S]} \\ \vdots \\ Y_1^{[n]} + \ldots + Y_n^{[n]} \text{ with probability } \frac{\mathrm{E}[X_n]}{\mathrm{E}[S]} \end{cases}.$$

This shows that size-biasing a sum of correlated risks is equivalent to selecting at random (according to the discrete distribution assigning probability $\frac{\mathrm{E}[X_k]}{\mathrm{E}[S]}$ to index k) one element among $\{\boldsymbol{Y}^{[1]}, \ldots, \boldsymbol{Y}^{[n]}\}$. Thus, the elements $\boldsymbol{Y}^{[i]}$ corresponding to risks X_i with larger expected values are more likely to be selected.

As a particular case, consider independent risks $X_1, \ldots, X_n$. Let $\widetilde{X}_1, \ldots, \widetilde{X}_n$ be their corresponding size-biased versions, assumed to be independent, and independent of $X_1, \ldots, X_n$. Let the random variable K valued in $\{1, 2, \ldots, n\}$ be defined as above, and independent of $X_1, \ldots, X_n$ and of $\widetilde{X}_1, \ldots, \widetilde{X}_n$. We then have $\widetilde{S} \overset{\mathrm{d}}{=} S - X_K + \widetilde{X}_K = \sum_{j \neq K} X_j + \widetilde{X}_K$.

3 Conditional mean risk sharing

3.1 General representation formula

Consider the random vector, or insurance portfolio $\boldsymbol{X} = (X_1, \ldots, X_n)$ with joint distribution function $F_{\boldsymbol{X}}$. The next result gives a general representation formula for $\mathrm{E}[X_i | S = s]$ in terms of the random vectors $\boldsymbol{Y}^{[i]}$.

Proposition 3.1. *Consider insurance losses $X_1, \ldots, X_n$ with joint distribution function $F_{\boldsymbol{X}}$ and expected values $\mathrm{E}[X_i]$ such that $0 < \mathrm{E}[X_i] < \infty$. Let $\boldsymbol{Y}^{[i]} = (Y_1^{[i]}, \ldots, Y_n^{[i]})$ be a random vector with joint distribution function $F_{\boldsymbol{Y}^{[i]}}$ given by (2.6). The following results hold:*

(i) if $X_1, \ldots, X_n$ are continuous random variables with joint probability density function $f_{\boldsymbol{X}}$ then for any $s \geq 0$,

$$\mathrm{E}[X_i | S = s] \;=\; \mathrm{E}\left[X_i\right] \frac{f_{Y_1^{[i]} + \ldots + Y_n^{[i]}}(s)}{f_{X_1 + \ldots + X_n}(s)}.$$

(ii) if $X_1, \ldots, X_n$ are valued in $\{0, 1, 2, \ldots\}$ then for any $s \in \{0, 1, 2, \ldots\}$,

$$\mathrm{E}[X_i | S = s] \;=\; \mathrm{E}\left[X_i\right] \frac{\mathrm{P}[Y_1^{[i]} + \ldots + Y_n^{[i]} = s]}{\mathrm{P}[X_1 + \ldots + X_n = s]}.$$

(iii) if $X_1, \ldots, X_n$ are zero-augmented random variables with positive probability masses at the origin and probability density functions over $(0, \infty)$ then $\mathrm{E}[X_i|S=0]=0$ and for any $s>0$, the representation in (i) holds true.

Proof. Let us apply (2.9) to the function g given by

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

to get the identity

$$\mathrm{P}\left[Y_1^{[i]} + \ldots + Y_n^{[i]} \leq s\right] = \frac{\mathrm{E}\big[X_i \mathrm{I}\left[S \leq s\right]\big]}{\mathrm{E}[X_i]} \Leftrightarrow \mathrm{E}\big[X_i \mathrm{I}\left[S \leq s\right]\big] = \mathrm{E}[X_i]\mathrm{P}\left[Y_1^{[i]} + \ldots + Y_n^{[i]} \leq s\right].$$

To establish the validity of statement (i), notice that we can also write

$$\mathrm{E}\big[X_i \mathrm{I}[S \leq s]\big] = \int_0^s \mathrm{E}\big[X_i \big| S=t\big] f_S(t) \mathrm{d}t.$$

Taking the derivative of these expressions with respect to s gives the identity stated in (i). Turning to (ii), we can write

$$\begin{aligned} \mathrm{E}\big[X_i \mathrm{I}\left[S=s\right]\big] &= \mathrm{E}\big[X_i \mathrm{I}\left[S \leq s\right]\big] - \mathrm{E}\big[X_i \mathrm{I}\left[S \leq s-1\right]\big] \\ &= \mathrm{E}[X_i]\mathrm{P}\left[Y_1^{[i]} + \ldots + Y_n^{[i]} = s\right], \end{aligned}$$

as announced. This ends the proof. □

It is worth to stress that the representation in Proposition 3.1(i) is related to Bartlett's formula for conditional expectations $\mathrm{E}[U|V=v]$ in terms of characteristic functions (Zabell, 1979). Essentially, the densities are replaced with their inverse transform to obtain Bartlett's formula.

When $X_1, \ldots, X_n$ are mutually independent, we recover as a particular case the result obtained by Denuit (2019, Proposition 2.3), as formally stated next.

Corollary 3.2. *Consider independent risks $X_1, X_2, \ldots, X_n$. Let $\widetilde{X}_1, \widetilde{X}_2, \ldots, \widetilde{X}_n$ be their corresponding size-biased versions, assumed to be independent and independent of $X_1, X_2, \ldots, X_n$. If $X_1, \ldots, X_n$ are mutually independent then $Y_1^{[i]}, \ldots, Y_n^{[i]}$ are also independent random variables. Furthermore, $Y_j^{[i]}$ is distributed as X_j for $j \neq i$, while $Y_i^{[i]} \overset{\mathrm{d}}{=} \widetilde{X}_i$. The distributional equality*

$$Y_1^{[i]} + \ldots + Y_n^{[i]} \overset{\mathrm{d}}{=} S - X_i + \widetilde{X}_i$$

thus holds true when $X_1, \ldots, X_n$ are mutually independent and

(i) if $X_1, X_2, \ldots, X_n$ are absolutely continous random variables with respective probability density functions $f_{X_1}, f_{X_2}, \ldots, f_{X_n}$, then for any $s>0$

$$\mathrm{E}[X_i|S=s] = \mathrm{E}[X_i] \frac{f_{S-X_i+\widetilde{X}_i}(s)}{f_S(s)}. \tag{3.1}$$

(ii) if $X_1, X_2, \ldots, X_n$ are valued in $\{0, 1, 2, \ldots\}$, then for any $s \in \{0, 1, 2, \ldots\}$

$$\mathrm{E}[X_i|S = s] = \mathrm{E}[X_i]\frac{\mathrm{P}[S - X_i + \widetilde{X}_i = s]}{\mathrm{P}[S = s]}. \tag{3.2}$$

(iii) if $X_1, \ldots, X_n$ are zero-augmented random variables with positive probability masses at the origin and probability density functions over $(0, \infty)$ then $\mathrm{E}[X_i|S = 0] = 0$ and for any $s > 0$, (3.1) holds true.

3.2 Linear case

3.2.1 Characterization

With the help of multivariate Laplace transforms, Furman et al. (2018) characterized the case where the conditional expectations are linear in s, that is, the situation where the identity $\mathrm{E}[X_i|S = s] = \frac{\mathrm{E}[X_i]}{\mathrm{E}[S]}s$ holds true. Considering Theorem 3.2 in that paper, denote as $X = X_i$ and $Y = \sum_{j\neq i} X_j$. Let $L(u, v) = \mathrm{E}[\exp(-uX - vY)]$ be the Laplace transform of the random couple (X, Y). Then,

$$\mathrm{E}[X_i|S = s] = \frac{\mathrm{E}\left[X_i\right]}{\mathrm{E}[S]}s \Leftrightarrow \left.\frac{\frac{\mathrm{d}}{\mathrm{d}u}L(u, v)}{\frac{\mathrm{d}}{\mathrm{d}v}L(u, v)}\right|_{(u,v)=(t,t)} = \frac{\mathrm{E}\left[X_i\right]}{\mathrm{E}\left[\sum_{j\neq i} X_j\right]} \text{ for all } t. \tag{3.3}$$

This further constrains the distribution of the random vectors $\boldsymbol{Y}^{[i]}$, as shown next. Considering Proposition 3.1(i), we see that under (3.3), we must have

$$\frac{\mathrm{E}\left[X_i\right]}{\mathrm{E}[S]}s = \mathrm{E}\left[X_i\right]\frac{f_{Y_1^{[i]}+\ldots+Y_n^{[i]}}(s)}{f_S(s)} \Rightarrow f_{\widetilde{S}}(s) = f_{Y_1^{[i]}+\ldots+Y_n^{[i]}}(s). \tag{3.4}$$

Thus, (3.3) implies that all sums $Y_1^{[i]} + \ldots + Y_n^{[i]}$ are identically distributed, as $\widetilde{S}$.

3.2.2 Inverted Dirichlet distributions

Assume that $\boldsymbol{X}$ obeys the Inverted Dirichlet$(a_1, \ldots, a_{n+1})$ distribution, for some positive parameters $a_1, \ldots, a_{n+1}$, that is, $\boldsymbol{X}$ possesses the joint probability density function

$$f_{\boldsymbol{X}}\left(x_1, \ldots, x_n\right) = \frac{\Gamma\left(\sum_{i=1}^{n+1} a_i\right)}{\prod_{i=1}^{n+1}\Gamma\left(a_i\right)}\prod_{i=1}^{n} x_i^{a_i-1}\left(1 + \sum_{i=1}^{n} x_i\right)^{-\sum_{i=1}^{n+1} a_i} \tag{3.5}$$

for positive $x_1, x_2, \ldots, x_n$. The random vector $\boldsymbol{X}$ admits the representation

$$\boldsymbol{X} \overset{\mathrm{d}}{=} \frac{1}{Z_{n+1}}\left(Z_1, \ldots, Z_n\right)$$

where $Z_1, \ldots, Z_n, Z_{n+1}$ are independent random variables obeying Chi-Square distributions with $2a_j$ degrees of freedom, $j \in \{1, \ldots, n, n + 1\}$.

If $a_{n+1} > 1$, we have $\mathrm{E}[X_i] = \frac{a_i}{a_{n+1}-1}$ and the joint probability density function of $\boldsymbol{Y}^{[i]}$ is obtained from

$$\begin{aligned}\frac{x_i}{\mathrm{E}[X_i]} f_{\mathbf{X}}(x_1, \ldots, x_n) &= \frac{\Gamma\left(\sum_{j=1}^{n+1} a_j\right)}{\Gamma(a_i+1)\,\Gamma(a_{n+1}-1)\prod_{i=1,\ldots,n,j\neq i}\Gamma(a_j)} \\ &\quad x_i^{a_i} \prod_{i=1,\ldots,n,j\neq i} x_j^{a_j-1} \left(1+\sum_{j=1}^{n} x_j\right)^{-\sum_{j=1}^{n+1} a_j}.\end{aligned}$$

We deduce that $\boldsymbol{Y}^{[i]}$ follows the Inverted Dirichlet$(a_1, \ldots, a_i+1, \ldots, a_{n+1}-1)$ distribution.

Since

$$S = \sum_{j=1}^{n} X_j = \frac{\sum_{j=1}^{n} Z_j}{Z_{n+1}},$$

the probability density function of the sum S is given by

$$f_S(s) = \frac{\Gamma\left(\sum_{i=1}^{n+1} a_i\right)}{\Gamma\left(\sum_{i=1}^{n} a_i\right)\Gamma(a_{n+1})} s^{\sum_{i=1}^{n} a_i} (1+s)^{-\sum_{i=1}^{n+1} a_i}.$$

For the same reason, we also get

$$f_{\sum_{j=1}^{n} Y_j^{[i]}}(s) = \frac{\Gamma\left(\sum_{i=1}^{n+1} a_i\right)}{\Gamma\left(\sum_{i=1}^{n} a_i+1\right)\Gamma(a_{n+1}-1)} s^{\sum_{i=1}^{n} a_i+1} (1+s)^{-\sum_{i=1}^{n+1} a_i}.$$

It follows that from Proposition 3.1(i) that

$$\begin{aligned}\mathrm{E}[X_i|S=s] &= \mathrm{E}[X_i] \frac{f_{\sum_{j=1}^{n} Y_j^{[i]}}(s)}{f_S(s)} \\ &= \frac{a_i}{a_{n+1}-1} \frac{\Gamma\left(\sum_{j=1}^{n} a_j\right)}{\Gamma\left(\sum_{j=1}^{n} a_j+1\right)} \frac{\Gamma(a_{n+1})}{\Gamma(a_{n+1}-1)} s \\ &= \frac{a_i}{\sum_{j=1}^{n} a_j} s.\end{aligned}$$

The conditional expectation is thus linear in s, with a slope equal to the ratio of a_i over the sum of $a_1, \ldots, a_n$. As it will be explained below for the whole class of Liouville distributions, this result can also be deduced from (3.3).

3.2.3 Liouville distributions

The result established in the preceding section is just a particular case of a more general statement applying to the whole family of Liouville distributions. Recall that an absolutely continuous random vector $\boldsymbol{X}$ obeys a multivariate Liouville distribution if its joint probability density function is proportional to

$$g\left(\sum_{j=1}^{n} x_j\right) \prod_{i=1}^{n} x_i^{a_i-1}$$

where $x_i > 0$, $a_i > 0$ and the function g is positive, continuous and appropriately integrable. Gupta and Richards (1987) have used the notation $L_n[g; a_1, ..., a_n]$ to refer to this distribution.

We only consider Liouville distributions of the first kind for which the support is non-compact. In this case

$$\boldsymbol{X} \overset{\mathrm{d}}{=} R\boldsymbol{Z},$$

where $R = \sum_{j=1}^{n} X_j$ obeys the univariate Liouville distribution $L_1[g; a]$ with $a = \sum_{i=1}^{n} a_i$ and $\boldsymbol{Z}$ is independent of R and possesses the Dirichlet distribution with joint probability density function

$$f_{\boldsymbol{Z}}(z_1, \ldots, z_n) = \frac{\Gamma\left(\sum_{i=1}^{n} a_i\right)}{\prod_{i=1}^{n} \Gamma(a_i)} \prod_{i=1}^{n-1} z_i^{a_i - 1} \left(1 - \sum_{i=1}^{n-1} z_i\right)^{a_n - 1}, \quad z_i > 0, \quad \sum_{i=1}^{n} z_i = 1.$$

We recover as a particular case the Inverted Dirichlet distribution considered in the preceding section with $g(t) = (1+t)^{-\sum_{i=1}^{n+1} a_i}$ for $t > 0$, $a_{n+1} > 0$. The corresponding probability density function is proportional to

$$\frac{\prod_{i=1}^{n} x_i^{a_i - 1}}{\left(1 + \sum_{i=1}^{n} x_i\right)^{\sum_{i=1}^{n+1} a_i}}$$

which corresponds to (3.5). As another example, the multivariate correlated Gamma distribution is obtained with $g(t) = t^{c-1} e^{-bt}$ for $t > 0$, $c > 0$, $b > 0$. The corresponding probability density function is proportional to

$$\left(\sum_{i=1}^{n} x_i\right)^{c-1} \prod_{i=1}^{n} \left(e^{-b x_i} x_i^{a_i - 1}\right).$$

Let us now derive the distribution of $\boldsymbol{Y}^{[i]}$. Since $x_i f_{\boldsymbol{X}}(x_1, \ldots, x_n)$ is proportional to

$$g\left(\sum_{j=1}^{n} x_j\right) x_i^{a_i} \prod_{i=1, \ldots, n, j \neq i} x_j^{a_j - 1},$$

we deduce that $\boldsymbol{Y}^{[i]}$ follows the $L_n[g; a_1, ..., a_i + 1, ..., a_n]$ distribution as soon as $\mathrm{E}[X_i] < \infty$. We have $f_S(s) \propto g(s) s^{a-1}$ where "$\propto$" means "is proportional to", and therefore, $f_{\sum_{j=1}^{n} Y_j^{[i]}}(s) \propto g(s) s^a$. It follows that $\mathrm{E}[X_i | S_n = s] \propto s$ which implies

$$\mathrm{E}[X_i | S_n = s] = \frac{\mathrm{E}[X_i]}{\mathrm{E}[S_n]} s.$$

Since $\mathrm{E}[X_i] = a_i \mathrm{E}[R] / a$, we finally get

$$\mathrm{E}[X_i | S_n = s] = \frac{a_i}{\sum_{j=1}^{n} a_j} s.$$

This result can also be obtained as a consequence of (3.3). To this end, consider

$$X = X_i = RZ_i \text{ and } Y = \sum_{j \neq i} X_j = R \sum_{j \neq i} Z_j = R(1 - Z_i).$$

Then,

$$L(u,v) = \mathrm{E}\left[\exp\left(-R\left(uZ_i + v\sum_{j\neq i} Z_j\right)\right)\right]$$

and the ratio appearing in (3.3) writes

$$\begin{aligned}
\left.\frac{\frac{\mathrm{d}}{\mathrm{d}u}L(u,v)}{\frac{\mathrm{d}}{\mathrm{d}v}L(u,v)}\right|_{(u,v)=(t,t)} &= \frac{\mathrm{E}[RZ_i\exp(-tR)]}{\sum_{j\neq i}\mathrm{E}[RZ_j\exp(-tR)]} \\
&= \frac{\mathrm{E}[Z_i]}{\sum_{j\neq i}\mathrm{E}[Z_j]} \\
&= \frac{\mathrm{E}\left[X_i\right]}{\mathrm{E}\left[\sum_{j\neq i}X_j\right]}.
\end{aligned}$$

This confirms that the conditional mean risk sharing is linear in the case of Liouville distributions.

In Section 4, 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 the independent case. This is also another approach to deal with Liouville distributions which are correlated by the common factor R that plays the role of the latent variable Λ inducing correlation between the conditionally independent risks.

3.2.4 Infinitely divisible distributions

According to Corollary 2.5 in Horn and Steutel (1978), a positive random vector $\boldsymbol{X}$ with infinitely divisible distribution on $\mathbb{R}^n_+$ can be characterized by its Laplace transform of the form

$$\mathrm{E}[e^{-\langle \boldsymbol{t},\boldsymbol{X}\rangle}] = \exp\left(\int_{\mathbb{R}^n_+}\frac{e^{-\langle \boldsymbol{t},\boldsymbol{x}\rangle}-1}{\|\boldsymbol{x}\|}\nu\left(\mathrm{d}\boldsymbol{x}\right)\right)$$

where $\langle \boldsymbol{x},\boldsymbol{y}\rangle$ stands for the inner product $\sum_{i=1}^n x_i y_i$, $\|\boldsymbol{x}\| = \sqrt{\langle \boldsymbol{x},\boldsymbol{x}\rangle}$ for the Euclidean norm and where ν is a measure such that $\nu\left(\{0\}\right)=0$ and $\int_{\boldsymbol{y}>\boldsymbol{x}}\frac{\nu(\mathrm{d}\boldsymbol{y})}{\|\mathbf{y}\|}<\infty$ for all $\boldsymbol{x}$ in the interior of $\mathbb{R}^n_+$. The expectation of X_i is given by

$$\mathrm{E}[X_i] = \int_{\mathbb{R}^n_+}\frac{x_i}{\|\boldsymbol{x}\|}\nu\left(\mathrm{d}\boldsymbol{x}\right).$$

The random vector $\boldsymbol{Y}^{[i]}$ possesses the Laplace transform

$$\begin{aligned}
\mathrm{E}[e^{-\left\langle \boldsymbol{t},\boldsymbol{Y}^{[i]}\right\rangle}] &= \frac{\mathrm{E}[X_i e^{-\langle \boldsymbol{t},\boldsymbol{X}\rangle}]}{\mathrm{E}[X_i]} = -\frac{1}{\mathrm{E}[X_i]}\frac{\partial}{\partial t_i}\mathrm{E}[e^{-\langle \boldsymbol{t},\boldsymbol{X}\rangle}] \\
&= \frac{1}{\int_{\mathbb{R}^n_+}\left(x_i/\left\|\boldsymbol{x}\right\|\right)\nu\left(\mathrm{d}\boldsymbol{x}\right)}\int_{\mathbb{R}^n_+}\frac{x_i}{\|\boldsymbol{x}\|}e^{-\langle \boldsymbol{t},\boldsymbol{x}\rangle}\nu\left(\mathrm{d}\boldsymbol{x}\right)\int_{[0,\infty)}e^{itx}\frac{\nu\left(\mathrm{d}x\right)}{\nu\left([0,\infty)\right)}\mathrm{E}[e^{-\langle \boldsymbol{t},\boldsymbol{X}\rangle}] \\
&= \mathrm{E}[e^{-\left\langle \boldsymbol{t},\boldsymbol{Z}^{[i]}\right\rangle}]\mathrm{E}[e^{-\langle \boldsymbol{t},\boldsymbol{X}\rangle}]
\end{aligned}$$

where $\boldsymbol{Z}^{[i]}$ is a random variable with probability density function proportional to $(x_i/\left\|\boldsymbol{x}\right\|)\,\nu\left(\mathrm{d}\boldsymbol{x}\right)$, and therefore $\boldsymbol{Y}^{[i]} \overset{\mathrm{d}}{=} \boldsymbol{X} + \boldsymbol{Z}^{[i]}$ where $\boldsymbol{X}$ and $\boldsymbol{Z}^{[i]}$ are independent. See also item (b) in Theorem 2.4 by Horn and Steutel (1978).

Assume that there exist positive constants λ_{ij} such that for all $i \neq j$

$$\frac{\int_{\mathbb{R}_+^n} \frac{x_i}{\left\|\boldsymbol{x}\right\|} e^{-\langle \boldsymbol{t},\boldsymbol{x}\rangle} \nu\left(\mathrm{d}\boldsymbol{x}\right)}{\int_{\mathbb{R}_+^n} \frac{x_j}{\left\|\boldsymbol{x}\right\|} e^{-\langle \boldsymbol{t},\boldsymbol{x}\rangle} \nu\left(\mathrm{d}\boldsymbol{x}\right)} = \lambda_{ij}, \text{ for all } \boldsymbol{t}.$$

In this case, we have

$$\lambda_{ij} = \frac{\int_{\mathbb{R}_+^n} \left(x_i/\left\|\boldsymbol{x}\right\|\right) \nu\left(\mathrm{d}\boldsymbol{x}\right)}{\int_{\mathbb{R}_+^n} \left(x_j/\left\|\boldsymbol{x}\right\|\right) \nu\left(\mathrm{d}\boldsymbol{x}\right)}$$

and

$$\mathrm{E}[e^{-\left\langle \boldsymbol{t},\boldsymbol{Z}^{[1]}\right\rangle}] = ... = \mathrm{E}[e^{-\left\langle \boldsymbol{t},\boldsymbol{Z}^{[n]}\right\rangle}]$$

as well as

$$\mathrm{E}[e^{-\left\langle \boldsymbol{t},\boldsymbol{Y}^{[1]}\right\rangle}] = ... = \mathrm{E}[e^{-\left\langle \boldsymbol{t},\boldsymbol{Y}^{[n]}\right\rangle}].$$

Note that

$$\mathrm{E}[e^{-t\left\langle \mathbf{1},\boldsymbol{Y}^{[i]}\right\rangle}] = \mathrm{E}[e^{-t\left\langle \mathbf{1},\boldsymbol{Z}^{[i]}\right\rangle}]\mathrm{E}[e^{-t\langle \mathbf{1},\boldsymbol{X}\rangle}] = \mathrm{E}[e^{-t\left\langle \mathbf{1},\boldsymbol{Z}^{[i]}\right\rangle}]\mathrm{E}[e^{-tS}]$$

and that

$$\begin{aligned}
\mathrm{E}[e^{-t\widetilde{S}}] &= \frac{\mathrm{E}[Se^{-tS}]}{\mathrm{E}[S]} = -\frac{1}{\mathrm{E}[S]}\frac{\partial}{\partial t}\mathrm{E}[e^{-tS}] = -\frac{1}{\mathrm{E}[S]}\frac{\partial}{\partial t}\mathrm{E}[e^{-t\langle \mathbf{1},\boldsymbol{X}\rangle}] \\
&= \frac{1}{\mathrm{E}[S]}\int_{\mathbb{R}_+^n} \frac{\langle \mathbf{1},\boldsymbol{x}\rangle}{\left\|\boldsymbol{x}\right\|} e^{-t\langle \mathbf{1},\boldsymbol{x}\rangle} \nu\left(\mathrm{d}\boldsymbol{x}\right) \exp\left(\int_{\mathbb{R}_+^n} \frac{e^{-t\langle \mathbf{1},\boldsymbol{x}\rangle}-1}{\left\|\boldsymbol{x}\right\|} \nu\left(\mathrm{d}\boldsymbol{x}\right)\right) \\
&= \sum_{i=1}^n \frac{\mathrm{E}[X_i]}{\mathrm{E}[S]}\mathrm{E}[e^{-t\left\langle \mathbf{1},\boldsymbol{Z}^{[i]}\right\rangle}]\mathrm{E}[e^{-t\langle \mathbf{1},\boldsymbol{X}\rangle}] \\
&= \sum_{i=1}^n \frac{\mathrm{E}[X_i]}{\mathrm{E}[S]}\mathrm{E}[e^{-t\left\langle \mathbf{1},\boldsymbol{Z}^{[i]}\right\rangle}]\mathrm{E}[e^{-tS}] \\
&= \mathrm{E}[e^{-t\left\langle \mathbf{1},\boldsymbol{Y}^{[i]}\right\rangle}] \text{ for } i = 1, ..., n,
\end{aligned}$$

in accordance with (3.4).

It finally follows that

$$\mathrm{E}[X_i|S=s] = \mathrm{E}[X_i]\frac{f_{\tilde{S}}\left(s\right)}{f_S\left(s\right)} = \frac{\mathrm{E}[X_i]}{\mathrm{E}[S]}s = \frac{\int_{\mathbb{R}_+^n} \left(x_i/\left\|\boldsymbol{x}\right\|\right) \nu\left(\mathrm{d}\boldsymbol{x}\right)}{\sum_{j=1}^n \int_{\mathbb{R}_+^n} \left(x_j/\left\|\boldsymbol{x}\right\|\right) \nu\left(\mathrm{d}\boldsymbol{x}\right)}s.$$

Again, this result can also be obtained with the help of (3.3).

4 Application to dependence by mixture

4.1 Size-biased transform of a mixture

Arratia et al. (2019, Lemma 2.6) studied the size-biased transform of mixtures. This section summarizes their findings, that appear to be useful for studying dependence induced by correlated latent factors in the next sections. We provide the reader with an elementary reasoning, for convenience and because this result will also be central to the extension to size-biasing conditionally independent risks.

Consider a family of non-negative random variables $\{X(\lambda),\ \lambda \geq 0\}$ indexed by a single, non-negative parameter λ. Let Λ be a mixing parameter with distribution function F_Λ. The corresponding mixture $X(\Lambda)$ has distribution function

$$\mathrm{P}[X(\Lambda) \leq x] = \int_0^\infty \mathrm{P}[X(\lambda) \leq x] \mathrm{d}F_\Lambda(\lambda).$$

Define $\widetilde{X}(\lambda)$ to be the size-biased version of $X(\lambda)$, with distribution function

$$\mathrm{P}[\widetilde{X}(\lambda) \leq z] = \frac{\mathrm{E}\big[X(\lambda)\mathrm{I}[X(\lambda) \leq z]\big]}{\mathrm{E}[X(\lambda)]}.$$

Then, the size-biased version $\widetilde{X(\Lambda)}$ of the mixture $X(\Lambda)$ corresponds to the mixture of the non-negative random variables $\{\widetilde{X}(\lambda),\ \lambda \geq 0\}$ with mixing parameter $\Lambda^\star$ distributed according to

$$\mathrm{d}F_{\Lambda^\star}(\lambda) = \frac{\mathrm{E}[X(\lambda)]}{\mathrm{E}[X(\Lambda)]} \mathrm{d}F_\Lambda(\lambda). \tag{4.1}$$

This result can be rewritten as

$$\widetilde{X(\Lambda)} \stackrel{\mathrm{d}}{=} \widetilde{X}(\Lambda^\star) \tag{4.2}$$

and is easily obtained as follows:

$$\begin{aligned}
\mathrm{P}[\widetilde{X(\Lambda)} \leq x] &= \frac{1}{\mathrm{E}[X(\Lambda)]} \int_0^x t \mathrm{d}F_{X(\Lambda)}(t) \\
&= \frac{1}{\mathrm{E}[X(\Lambda)]} \int_0^\infty \int_0^x t \mathrm{d}F_{X(\lambda)}(t) \mathrm{d}F_\Lambda(\lambda) \\
&= \frac{1}{\mathrm{E}[X(\Lambda)]} \int_0^\infty \mathrm{E}[X(\lambda)] F_{\widetilde{X}(\lambda)}(t) \mathrm{d}F_\Lambda(\lambda) \\
&= \int_0^\infty F_{\widetilde{X}(\lambda)}(t) \mathrm{d}F_{\Lambda^\star}(\lambda)
\end{aligned}$$

which shows that the announced distributional equality (4.2) is indeed valid.

In the particular case where $\mathrm{E}[X(\lambda)]$ is proportional to λ, that is, if the identity $\mathrm{E}[X(\lambda)] = a\lambda$ holds for some positive constant a, we obtain

$$\mathrm{d}F_{\Lambda^\star}(\lambda) = \frac{\lambda}{\mathrm{E}[\Lambda]} \mathrm{d}F_\Lambda(\lambda),$$

so that $\Lambda^\star \stackrel{\mathrm{d}}{=} \widetilde{\Lambda}$ and $\widetilde{X(\Lambda)} \stackrel{\mathrm{d}}{=} \widetilde{X}(\widetilde{\Lambda})$.

4.2 Common mixture model

4.2.1 Definition

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 $X_1, X_2, \ldots, X_n$ with $X_i \overset{\mathrm{d}}{=} X_i(\Lambda)$. Given the environmental parameter Λ, the individual risks are independent. Formally, the joint distribution function of the portfolio vector $\boldsymbol{X}$ can be written as

$$\begin{aligned} F_{\boldsymbol{X}}(t_1, \ldots, t_n) &= \mathrm{E}\big[\mathrm{P}[X_1 \leq t_1, \ldots, X_n \leq t_n | \Lambda]\big] \\ &= \int_0^\infty \left(\prod_{i=1}^n \mathrm{P}[X_i(\lambda) \leq t_i] \right) \mathrm{d}F_\Lambda(\lambda). \end{aligned} \tag{4.3}$$

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, with the second one common to all risks. This common shock then plays the role of Λ in the common mixture model (4.3). The multivariate Gamma distribution proposed in Furman and Landsman (2005) is built in this way, defining $X_i = \beta_i Y_0 + Y_i$ for some $\beta_i > 0$, where $Y_0, Y_1, \ldots, Y_n$ are independent, Gamma-distributed random variables. Here, Y_0 plays the role of Λ. This is also the approach followed by Furman and Landsman (2010) to derive a multivariate extension of the Tweedie distribution. The multivariate Pareto II distribution used by Asimit et al. (2013) is also obtained in this way, specifying $X_i = \mu_i + Y_i/Y_0$ where $Y_0, Y_1, \ldots, Y_n$ are independent random variables with $Y_1, \ldots, Y_n$ obeying the same Negative Exponential distribution and Y_0 following the Gamma distribution. The multivariate Pareto distribution proposed by Asimit et al. (2010) corresponds to $X_i = \min\{\sigma_i Y_0 + \mu_i, Y_i\}$ with $\sigma_i > 0$, where $Y_0, Y_1, \ldots, Y_n$ are independent, Pareto-distributed random variables. In credit risk modeling, time-to-defaults X_i are sometimes assumed to be subject to a competing risk mechanism with $X_i = \min\{Y_i, Y_0\}$ where $Y_0, Y_1, \ldots, Y_n$ are independent, positive random variables. The common factor Y_0 impacting all times-to-default accounts for a systematic shock. See, e.g., Giesecke (2003).

4.2.2 Association

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 $\boldsymbol{Y}^{[i]}$ with distribution function (2.6) larger so that size-biasing induces a safety margin in this case. This is formally stated in the next result.

Property 4.1. *Consider the risks $X_1, X_2, \ldots, X_n$ with joint distribution function (4.3). Assume that*

$$\mathrm{P}[X_i(\lambda) \leq t] \geq \mathrm{P}[X_i(\lambda + \delta) \leq t] \text{ for all } t,\ \lambda > 0,\ \delta > 0, \text{ and } i = 1, 2, \ldots, n.$$

Then, the random vector $\boldsymbol{Y}^{[i]}$ *with distribution function* (2.6) *is larger than* $\boldsymbol{X}$ *in the sense of multivariate first-order stochastic dominance, that is, the inequality* $\mathrm{E}[g(\boldsymbol{X})] \leq \mathrm{E}[g(\boldsymbol{Y}^{[i]})]$ *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 $X_1, \ldots, X_n$ are associated random variables, that is, the covariance $\mathrm{Cov}[g_1(\boldsymbol{X}), g_2(\boldsymbol{X})]$ is non-negative for all non-decreasing functions g_1 and g_2. The announced result then follows from Property 2.1. □

4.2.3 Representation formula

When the risks $X_1, X_2, \ldots, X_n$ are conditionally independent, given Λ, with joint distribution function (4.3), it turns out that each $\boldsymbol{Y}^{[i]}$ also obeys a common mixture model with a change in the ith conditional distribution and in the mixing distribution. This is formally stated next.

Proposition 4.2. *Consider the risks* $X_1, X_2, \ldots, X_n$ *with joint distribution function* (4.3). *Then, each* $\boldsymbol{Y}^{[i]}$ *also possesses a joint distribution function of the form* (4.3), *that is,*

$$F_{\boldsymbol{Y}^{[i]}}(t_1, \ldots, t_n) = \int_0^\infty \left(\prod_{j \neq i} \mathrm{P}[X_j(\lambda) \leq t_j] \right) \mathrm{P}[\widetilde{X}_i(\lambda) \leq t_i] \mathrm{d}F_{\Lambda_i^\star}(\lambda)$$

where $\Lambda_i^\star$ *is distributed according to* (4.1), *that is,*

$$\mathrm{d}F_{\Lambda_i^\star}(\lambda) = \frac{\mathrm{E}[X_i(\lambda)]}{\mathrm{E}[X_i(\Lambda)]} \mathrm{d}F_\Lambda(\lambda).$$

Proof. It suffices to write

$$\begin{aligned} F_{\boldsymbol{Y}^{[i]}}(y_1, \ldots, y_n) &= \frac{\mathrm{E}\big[X_i \mathrm{I}[X_1 \leq y_1, \ldots, X_n \leq y_n]\big]}{\mathrm{E}[X_i]} \\ &= \frac{\mathrm{E}\left[\mathrm{E}\left[X_i \prod_{j=1}^n \mathrm{I}[X_j \leq y_j] \middle| \Lambda\right]\right]}{\mathrm{E}[X_i]} \\ &= \mathrm{E}\left[\frac{\mathrm{E}\big[X_i \mathrm{I}[X_i \leq y_i] \big| \Lambda\big]}{\mathrm{E}[X_i]} \prod_{j \neq i} \mathrm{P}[X_j \leq y_j | \Lambda] \right]. \end{aligned}$$

This ends the proof. □

4.2.4 Individual model of risk theory

Assume that $\Lambda \in [0, 1]$ and that each risk X_i is of the form $X_i = I_i C_i$. Given $\Lambda = \lambda$, $I_i = I_i(\lambda)$ is Bernoulli(λ) distributed, that is,

$$\mathrm{P}[I_i = 1 | \lambda] = 1 - \mathrm{P}[I_i = 0 | \lambda] = \lambda.$$

We assume that the costs C_i are independent of Λ with respective means $\mathrm{E}[C_i] = \mu_i$. Then, $X_i = X_i(\lambda)$ is such that $\mathrm{E}[X_i(\lambda)] = \lambda\mu_i$ for every $i = 1, 2, \ldots, n$. Given Λ, all the random variables $I_1, \ldots, I_n$, $C_1, \ldots, C_n$ are assumed to be independent.

Each $\boldsymbol{Y}^{[i]}$ obeys a mixture of $\left(X_1(\lambda), \ldots, \widetilde{C}_i, \ldots, X_n(\lambda)\right)$ with mixing parameter $\Lambda_i^\star$ distributed according to

$$\mathrm{d}F_{\Lambda_i^\star}(\lambda) = \frac{\lambda\mu_i}{\mathrm{E}[\Lambda]\mu_i}\mathrm{d}F_\Lambda(\lambda) \Leftrightarrow \Lambda_i^\star \overset{\mathrm{d}}{=} \widetilde{\Lambda} \text{ for all } i.$$

Thus, the following representations hold true:

$$\begin{aligned}
\mathrm{E}[X_i|S > t] &= \mathrm{E}[\Lambda]\mu_i \frac{\mathrm{P}\left[\widetilde{C}_i + \sum_{j\neq i} X_j\left(\widetilde{\Lambda}\right) > t\right]}{\mathrm{P}\left[\sum_{j=1}^n X_j(\Lambda) > t\right]} \\
\mathrm{E}[X_i|S = t] &= \mathrm{E}[\Lambda]\mu_i \frac{f_{\widetilde{C}_i + \sum_{j\neq i} X_j\left(\widetilde{\Lambda}\right)}(t)}{f_{\sum_{j=1}^n X_j(\Lambda)}(t)}.
\end{aligned}$$

Example 4.3. *Assume that given $\Lambda \in (0,1)$, each I_i is Bernoulli distributed with the same mean Λ where Λ follows the Beta(η, β) distribution. The costs C_i are independent of Λ and obey Gamma(α_i, τ) distributions. Given Λ, all the random variables $I_1, ..., I_n, C_1, ..., C_n$ are assumed to be independent.*

It is easy to see that $\widetilde{C}_i$ follows the Gamma$(\alpha_i + 1, \tau)$ distribution. Moreover,

$$\boldsymbol{Y}^{[i]} \overset{\mathrm{d}}{=} \left(X_1\left(\widetilde{\Lambda}\right), ..., \widetilde{C}_i, ..., X_n\left(\widetilde{\Lambda}\right)\right)$$

where $\widetilde{\Lambda}$ follows the Beta$(\eta + 1, \beta)$ distribution. Since

$$\begin{aligned}
& f_{\widetilde{C}_i + \sum_{j\neq i} X_1(\widetilde{\Lambda})|\widetilde{\Lambda}=\lambda}(s) \\
= \; & \exp(-s\tau) \sum_{j=0}^{n-1} \lambda^j (1-\lambda)^{n-1-j} \sum_{\sum_{l\neq i} k_l = j : k_l \in \{0,1\}} \frac{\tau}{\Gamma\left(\alpha_i + \sum_{l\neq i} k_l\alpha_l + 1\right)} (\tau s)^{\alpha_i + \sum_{l\neq i} k_l\alpha_l}
\end{aligned}$$

we then have

$$\begin{aligned}
& f_{\widetilde{C}_i + \sum_{j\neq i} X_1(\widetilde{\Lambda})}(s) \\
= \; & \exp(-s\tau) \sum_{j=0}^{n-1} \frac{\mathrm{B}(\eta + 1 + j, \beta + n - 1 - j)}{\mathrm{B}(\eta + 1, \beta)} \sum_{\sum_{l\neq i} k_l = j : k_l \in \{0,1\}} \frac{\tau}{\Gamma\left(\alpha_i + \sum_{l\neq i} k_l\alpha_l + 1\right)} (\tau s)^{\alpha_i + \sum_{l\neq i} k_l\alpha_l}.
\end{aligned}$$

If $\alpha_1 = \ldots = \alpha_n = \alpha$, we thus obtain

$$\begin{aligned}
& f_{\widetilde{C}_i + \sum_{j\neq i} X_1(\widetilde{\Lambda})}(s) \\
= \; & \exp(-s\tau) \sum_{j=0}^{n-1} \frac{\mathrm{B}(\eta + 1 + j, \beta + n - 1 - j)}{\mathrm{B}(\eta + 1, \beta)\,\mathrm{B}(j + 1, n)} \frac{\Gamma(n + 1 + j)}{\Gamma(n - j)} \frac{\tau}{\Gamma(j\alpha + 1)} (\tau s)^{j\alpha + 1}
\end{aligned}$$

and $\mathrm{E}[X_i|S = s] = \frac{\alpha_i}{\alpha_\bullet} s$ holds true in this case, as expected.

4.2.5 Compound mixed Poisson model

Let us now assume that given $\Lambda = \lambda$ each X_i is a compound Poisson sum. Precisely, 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, $i = 1, 2, \ldots, n$, and independent of $(N_1, \ldots, N_n)$. We denote as $\mu_i = \mathrm{E}[C_i]$ the mean claim severity for X_i. Each N_i obeys the Poisson$(a_i\Lambda)$ distribution given Λ, for some positive constants $a_1, \ldots, a_n$. Claim severities thus remain independent but the random vector of claim frequencies $(N_1, \ldots, N_n)$ 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).

Here, we see that

$$\mathrm{d}F_{\Lambda_i^\star}(\lambda) = \frac{a_i\lambda\mu_i}{a_i\mathrm{E}[\Lambda]\mu_i}\mathrm{d}F_\Lambda(\lambda) = \frac{\lambda}{\mathrm{E}[\Lambda]}\mathrm{d}F_\Lambda(\lambda)$$

so that $\Lambda_1^\star, \ldots, \Lambda_n^\star$ are identically distributed and $\Lambda_i^\star \stackrel{\mathrm{d}}{=} \widetilde{\Lambda}$ for all i. Since $\widetilde{X}_i(\lambda) \stackrel{\mathrm{d}}{=} X_i(\lambda) + \widetilde{C}_i$ where the random variable $\widetilde{C}_i$ is independent of $X_i(\lambda)$, we get

$$F_{\boldsymbol{Y}^{[i]}}(t_1, \ldots, t_n) = \int_0^\infty \left(\prod_{j\neq i} \mathrm{P}[X_j(\lambda) \leq t_j]\right) \mathrm{P}[X_i(\lambda) + \widetilde{C}_i \leq t_i]\mathrm{d}F_{\widetilde{\Lambda}}(\lambda).$$

Thus, the following representations hold true:

$$\begin{aligned}
\mathrm{E}[X_i|S > t] &= a_i\mathrm{E}[\Lambda]\mu_i \frac{\mathrm{P}\left[\sum_{j=1}^n X_j\big(\widetilde{\Lambda}\big) + \widetilde{C}_i > t\right]}{\mathrm{P}\left[\sum_{j=1}^n X_j(\Lambda) > t\right]} \\
\mathrm{E}[X_i|S = t] &= a_i\mathrm{E}[\Lambda]\mu_i \frac{f_{\sum_{j=1}^n X_j\left(\widetilde{\Lambda}\right)+\widetilde{C}_i}(t)}{f_{\sum_{j=1}^n X_j(\Lambda)}(t)}.
\end{aligned}$$

Now, assume that severities also depend on Λ, that is, $C_{ik} = C_{ik}(\Lambda)$ with $\mathrm{E}[C_{ik}(\lambda)] = \lambda\mu_i$. Given Λ, all the random variables are supposed to be independent. Then, $\boldsymbol{Y}^{[i]}$ is a mixture of

$$\big(X_1(\lambda), \ldots, X_i(\lambda) + \widetilde{C}_i(\lambda), \ldots, X_n(\lambda)\big)$$

with mixing parameter $\Lambda_i^\star$ distributed as

$$\mathrm{d}F_{\Lambda_i^\star}(\lambda) = \frac{a_i\lambda^2\mu_i}{a_i\mathrm{E}[\Lambda^2]\mu_i}\mathrm{d}F_\Lambda(\lambda) = \frac{\lambda^2}{\mathrm{E}[\Lambda^2]}\mathrm{d}F_\Lambda(\lambda).$$

Thus, the mixing parameters $\Lambda_i^\star$ are identically distributed for all i, that is, $\Lambda_i^\star \stackrel{\mathrm{d}}{=} \Lambda^\star$ for all i. We have the following representations:

$$\begin{aligned}
\mathrm{E}[X_i|S > t] &= a_i\mathrm{E}[\Lambda^2]\mu_i \frac{\mathrm{P}\left[\sum_{j=1}^n X_j\big(\Lambda^\star\big) + \widetilde{C}_i\big(\Lambda^\star\big) > t\right]}{\mathrm{P}\left[\sum_{j=1}^n X_j(\Lambda) > t\right]} \\
\mathrm{E}[X_i|S = t] &= a_i\mathrm{E}[\Lambda^2]\mu_i \frac{f_{\sum_{j=1}^n X_j\left(\Lambda^\star\right)+\widetilde{C}_i\left(\Lambda^\star\right)}(t)}{f_{\sum_{j=1}^n X_j(\Lambda)}(t)}.
\end{aligned}$$

Some families of distributions are closed in the sense that if Λ obeys a distribution belonging to the family then this is also the case for $\Lambda^\star$. Gamma distributions satisfy this property.

4.3 General mixture models

4.3.1 Definition

Let us now extend the common mixture model by letting each risk X_i depend on its own latent variable Θ_i, that is,

$$F_{\boldsymbol{X}}(t_1,\ldots,t_n)=\int_0^\infty\cdots\int_0^\infty\left(\prod_{i=1}^n \mathrm{P}[X_i(\theta_i)\le t_i]\right)\mathrm{d}F_{\boldsymbol{\Theta}}(\theta_1,\ldots,\theta_n). \tag{4.4}$$

If $\Theta_1=\ldots=\Theta_n=\Lambda$ then we are back to the common mixture model (4.3) studied in the preceding section.

Mixture models (4.4) are also widely used in insurance studies. In multiperil insurance modeling for instance, policyholders may own several insurance products (up to n, say). Given $\boldsymbol{\Theta}$, the number of claims N_i recorded for product i is often assumed to be Poisson distributed with parameter $\lambda_i\Theta_i$, where λ_i is the a priori expected claim frequency for product i, based on policyholder's specific risk profile. Denuit and Lu (2020) assumed that $\boldsymbol{\Theta}$ obeys the Wishart distribution with dimension n. As another example, Denuit et al. (2015) introduced the Max-factor individual risk model for credit portfolios. Given $\boldsymbol{\Theta}\in[0,1]^n$, default indicators I_i are Bernoulli distributed with

$$\mathrm{P}[I_i=1|\Theta_i]=1-\mathrm{P}[I_i=0|\Theta_i]=\Theta_i,$$

with

$$\Theta_i=F_\Psi\Big(\max\{\nu_i+\sigma_i\Psi_i,\mu_0+\sigma_0\Psi_0\}\Big)$$

where F_Ψ is a suitable distribution function, $\nu_i\in\mathbb{R}$ and $\sigma_i\ge 0$. There is thus a competition between the risk-specific effect $\nu_i+\sigma_i\Psi_i$ and the global effect $\mu_0+\sigma_0\Psi_0$ and only the largest one impacts on the occurrences of losses.

4.3.2 Association

As it was the case for common mixture models, 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 θ_i and the components of $\boldsymbol{\Theta}$ are positively related. This makes the random vector $\boldsymbol{Y}^{[i]}$ with distribution function (2.6) larger so that size-biasing induces a safety margin in this case. This is formally stated in the next result.

Property 4.4. *Consider the risks $X_1,X_2,\ldots,X_n$ with joint distribution function (4.4). Assume that*

$$\mathrm{P}[X_i(\theta_i)\le t]\ge\mathrm{P}[X_i(\theta_i+\delta)\le t] \text{ for all } t,\ \theta_i>0,\ \delta>0, \text{ and } i=1,2,\ldots,n,$$

and $\boldsymbol{\Theta}$ is associated. Then, the random vector $\boldsymbol{Y}^{[i]}$ with distribution function (2.6) is larger than $\boldsymbol{X}$ in the sense of multivariate first-order stochastic dominance, that is, the inequality $\mathrm{E}[g(\boldsymbol{X})]\le\mathrm{E}[g(\boldsymbol{Y}^{[i]})]$ holds true for every non-decreasing function g such that the expectations exist.

Proof. We know from Property 7.2.17 in Denuit et al. (2005) that the risks $X_1, \ldots, X_n$ are associated random variables. The announced result then follows from Property 2.1. □

4.3.3 Representation formula

The next result shows that each $\boldsymbol{Y}^{[i]}$ obeys a common mixture model with a change in the ith conditional distribution and in the mixing distribution.

Proposition 4.5. *Consider the risks $X_1, X_2, \ldots, X_n$ with joint distribution function* (4.4)*. Then, each $\boldsymbol{Y}^{[i]}$ also possesses a joint distribution function of the form* (4.4)*, that is,*

$$F_{\boldsymbol{Y}^{[i]}}(t_1, \ldots, t_n) = \int_0^\infty \cdots \int_0^\infty \left(\prod_{j \neq i} \mathrm{P}[X_j(\theta_j) \leq t_j] \right) \mathrm{P}[\widetilde{X}_i(\theta_i) \leq t_i] \mathrm{d}F_{\boldsymbol{\Theta}^{\star[i]}}(\theta_1, \ldots, \theta_n)$$

where the mixing parameter $\boldsymbol{\Theta}^{\star[i]}$ has joint distribution

$$\mathrm{d}F_{\boldsymbol{\Theta}^{\star[i]}}(\theta_1, \ldots, \theta_n) = \frac{\mathrm{E}[X_i(\theta_i)]}{\mathrm{E}[X_i]} \mathrm{d}F_{\boldsymbol{\Theta}}(\theta_1, \ldots, \theta_n).$$

Proof. It suffices to write

$$\begin{aligned} F_{\boldsymbol{Y}^{[i]}}(y_1, \ldots, y_n) &= \frac{\mathrm{E}\big[X_i \mathrm{I}[X_1 \leq y_1, \ldots, X_n \leq y_n]\big]}{\mathrm{E}[X_i]} \\ &= \frac{\mathrm{E}\left[\mathrm{E}\left[X_i \prod_{j=1}^n \mathrm{I}[X_j \leq y_j] \middle| \boldsymbol{\Theta}\right]\right]}{\mathrm{E}[X_i]} \\ &= \mathrm{E}\left[\frac{\mathrm{E}\big[X_i \mathrm{I}[X_i \leq y_i] \big| \Theta_i\big]}{\mathrm{E}[X_i]} \prod_{j \neq i} \mathrm{P}[X_j \leq y_j | \Theta_j] \right]. \end{aligned}$$

This ends the proof. □

4.3.4 Individual model of risk theory

Assume that each risk X_i is of the form $X_i = I_i C_i$ and that $\boldsymbol{\Theta} \in [0,1]^n$. Given $\boldsymbol{\Theta} = \boldsymbol{\theta}$, $I_i = I_i(\theta_i)$ obeys the Bernoulli(θ_i) distribution, that is,

$$\mathrm{P}[I_i = 1 | \theta_i] = 1 - \mathrm{P}[I_i = 0 | \theta_i] = \theta_i$$

for every $i = 1, 2, \ldots, n$. Given $\boldsymbol{\Theta}$, all the random variables I_j and C_j are assumed to be independent. Then, $Y^{[i]}$ is a mixture of

$$\big(X_1(\theta_1), \ldots, \widetilde{C}_i, \ldots, X_n(\theta_n)\big)$$

with mixing parameter $\boldsymbol{\Theta}^{[i]}$ defined in Proposition 4.5.

4.3.5 Compound mixed Poisson model

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. Given $\mathbf{\Theta} = \boldsymbol{\theta}$, each N_i obeys the Poisson$(a_i\theta_i)$ distribution, for some positive constants $a_1, \ldots, a_n$, and severities are such that $\mathrm{E}[C_{ik}(\theta_i)] = \theta_i\mu_i$. Given $\mathbf{\Theta}$, all the random variables are assumed to be independent. Then, $\boldsymbol{Y}^{[i]}$ is a mixture of

$$\big(X_1(\theta_1), \ldots, X_i(\theta_i) + \widetilde{C}_i(\theta_i), \ldots, X_n(\theta_n)\big)$$

with mixing parameter $\mathbf{\Theta}^{\star[i]}$ distributed as

$$\mathrm{d}F_{\mathbf{\Theta}^{\star[i]}}(\theta_1, \ldots, \theta_n) = \frac{a_i\theta_i^2\mu_i}{a_i\mathrm{E}[\Theta_i^2]\mu_i}\mathrm{d}F_{\mathbf{\Theta}}(\theta_1, \ldots, \theta_n) = \frac{\theta_i^2}{\mathrm{E}[\Theta_i^2]}\mathrm{d}F_{\mathbf{\Theta}}(\theta_1, \ldots, \theta_n).$$

Sometimes, $\mathbf{\Theta}$ and $\mathbf{\Theta}^{\star[i]}$ obey distributions in the same family. This is for instance the case when $\mathbf{\Theta}$ follows a multivariate Liouville distribution.

If the severities C_{ik} are independent of $\mathbf{\Theta}$ then we see that $Y^{[i]}$ is a mixture of

$$\big(X_1(\theta_1), \ldots, X_i(\theta_i) + \widetilde{C}_i, \ldots, X_n(\theta_n)\big)$$

with mixing parameter $\mathbf{\Theta}^{\star[i]}$. Since

$$\mathrm{d}F_{\mathbf{\Theta}^{\star[i]}}(\theta_1, \ldots, \theta_n) = \frac{a_i\theta_i\mu_i}{a_i\mathrm{E}[\Theta_i]\mu_i}\mathrm{d}F_{\mathbf{\Theta}}(\theta_1, \ldots, \theta_n) = \frac{\theta_i}{\mathrm{E}[\Theta_i]}\mathrm{d}F_{\mathbf{\Theta}}(\theta_1, \ldots, \theta_n),$$

we see that $\mathbf{\Theta}^{\star[i]} \overset{\mathrm{d}}{=} \mathbf{\Theta}^{[i]}$ for all $i = 1, 2, \ldots, n$. Here also, if $\mathbf{\Theta}$ obeys a Liouville distribution then this is also the case for $\mathbf{\Theta}^{[i]}$.

References

- Arratia, R., Goldstein, L., Kochman, F. (2019). Size bias for one and all. Probability Surveys 16, 1-61.

- Asimit, A.V., Furman, E., Vernic, R. (2010). On a multivariate Pareto distribution. Insurance: Mathematics and Economics 46, 308-316.

- Asimit, A.V., Vernic, R., Zitikis, R. (2013). Evaluating risk measures and capital allocations based on multi-losses driven by a heavy-tailed background risk: The multivariate Pareto-II model. Risks 1, 14-33.

- Cohen, A., Sackrowitz, H.B. (1995). On stochastic ordering of random vectors. Journal of Applied Probability 32, 960-965.

- Denuit, M. (2019). Size-biased transform and conditional mean risk sharing, with application to P2P insurance and tontines. ASTIN Bulletin 49, 591-617.

- Denuit, M., Dhaene, J. (2012). Convex order and comonotonic conditional mean risk sharing. Insurance: Mathematics and Economics 51, 265-270.

- Denuit, M., Dhaene, J., Goovaerts, M.J., Kaas, R. (2005). Actuarial Theory for Dependent Risks: Measures, Orders and Models. Wiley, New York.

- Denuit, M., Kiriliouk, A., Segers, J. (2015). Max-factor individual risk models with application to credit portfolios. Insurance: Mathematics and Economics 62, 162-172.

- Denuit, M., Lu, Y. (2020). Wishart-Gamma random effects for multivariate experience ratemaking, frequency-severity experience ratemaking and micro loss-reserving. Submitted for publication.

- 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., Kuznetsov, A., Zitikis, R. (2018). Weighted risk capital allocations in the presence of systematic risk. Insurance: Mathematics and Economics 79, 75-81.

- 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. (2008). 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.

- Giesecke, K. (2003). A simple exponential model for dependent defaults. Journal of Fixed Income 13, 74-83.

- Guo, X., Li, J., Liu, D., Wang, J. (2016). Preserving the Rothschild–Stiglitz type of increasing risk with background risk. Insurance: Mathematics and Economics 70, 144-149.

- Gupta, R.D., Richards, D.S.P. (1987). Multivariate Liouville distributions. Journal of Multivariate Analysis 23, 233-256.

- Horn, R.A., Steutel, F.W. (1978). On multivariate infinitely divisible distributions. Stochastic Processes and Their Applications 6, 139-151.

- Jain, K., Nanda, A.K. (1995). On multivariate weighted distributions. Communications in Statistics-Theory and Methods 24, 2517-2539.

- 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 23, 82-97.

- Navarro, J., Ruiz, J.M., Del Aguila, Y. (2006). Multivariate weighted distributions: A review and some extensions. Statistics 40, 51-64.

- Shaked, M., Shanthikumar, J.G. (2007). Stochastic Orders. Springer, New York.

- Wang, S. (1998). Aggregation of correlated risk portfolios: Models and algorithms. Proceedings of the Casualty Actuarial Society, 848-939.

- Zabell, S. (1979). Continous versions of regular conditional distributions. The Annals of Probability 7, 159-165.