I N S T I T U T   D E
S T A T I S T I Q U E

UNIVERSITÉ CATHOLIQUE DE LOUVAIN

D I S C U S S I O N
P A P E R

0122

MEASURING THE IMPACT OF DEPENDENCE BETWEEN CLAIMS OCCURRENCES

M. DENUIT, C. LEFÈVRE & S. UTEV

http://www.stat.ucl.ac.be

MEASURING THE IMPACT OF DEPENDENCE BETWEEN CLAIMS OCCURRENCES

By Michel Denuit [1] , Claude Lefèvre [2] and Sergey Utev [3]

July 4, 2001

Abstract

The purpose of this paper is to provide a quantitative measure of the impact of a possible dependence between the occurrences of claims in an individual risk model. Firstly, probabilistic distances, of stop-loss or total variation types, specific to arithmetic random variables are introduced and studied, especially in connection with related probabilistic orderings. Then, these results are applied to derive effective bounds for the distance between the total numbers of claims in the original model and under a standard independence assumption.

Classifications codes: IM11, IM12, IM30, IM54
Key words: Individual risk model; Dependent claim occurrences; Positive and negative association; s-convex probabilistic orderings and metrics
Submitted: November 2000

1 Introduction

The importance of dependencies among risks is well recognized in actuarial theory and practice. A standard independence assumption is usually made for convenience, but it does not reflect reality in a number of situations. Typical cases arise for policies covering natural disasters (hurricane, tornado, flood ...) or groups of individuals (household, staff of a large company ...). As illustrated numerically by Kaas (1993), dependencies can have, for instance, disastrous effects on stop-loss premiums. Recently, problems raised by a possible dependence between risks have received increasing attention in the literature; see, e.g., Dhaene and Goovaerts (1997) and the references therein, as well as the various examples given below.

The present paper is concerned with those situations, quite frequent in practice, where dependence affects primarily occurrences of claims (rather than claims severities). Our main objective is to measure quantitatively the impact of such a dependence in an individual risk model.

In Section 2, we illustrate the relevance of this problem by presenting a broad class of examples that describe various specific dependence structures, of global, local or mixed form.

[1]M. Denuit: Institut de Statistique, Université Catholique de Louvain, Voie du Roman Pays, 20, B-1348 Louvain-la-Neuve, Belgique. Email address: denuit@stat.ucl.ac.be

[2]Cl. Lefèvre: Institut de Statistique et de Recherche Opérationnelle, Université Libre de Bruxelles, Boulevard du Triomphe, C.P. 210, B-1050 Bruxelles, Belgique. Email address: clefevre@ulb.ac.be

[3]S. Utev: NCEPH, The Australian National University, Canberra, ACT 0200, Australia. Email address: sergey.utev@anu.edu.au

Most come from the actual literature. A new illustration (in the actuarial context) is Example 2.5 which treats of exchangeable individual risks within a community of finite size.

A mathematical tool for getting reasonable bounds on the impact is through probabilistic distances. As dependence affects occurrences of claims, it is natural to have recourse to metrics that are specific to arithmetic random variables. In Section 3, we introduce new classes of such metrics, which are of stop-loss or total variation types with variable degrees. We mention that the stop-loss metrics examined are the discrete analogues of those discussed by Rachev and Rüschendorf (1990). Then, in the spirit of Lefèvre and Utev (1998), we will point out and exploit a close connection of these distances with related discrete s-convex probabilistic orderings (see, e.g., Denuit et al. (1999)).

In Section 4, we return to the problem of risks dependency, with special attention to the different examples presented before. Applying results derived for metrics, we will be in a position to construct bounds for the distance between the total numbers of claims in the original model and under a standard independence assumption. These bounds are effective, that is easily computable and reasonably small when the dependence is weak. We mention that the form of results obtained here have been partly motivated by Boutsikas and Koutras (2000), but the technique we will follow is totally distinct. Moreover, our work is slightly overlapping with a paper by Genest et al. (2000) devoted to another question of similar nature.

2 Dependence between claims occurrences

Let us sketch briefly our framework. Consider a portfolio consisting of n policies over a given period. To each policy $i, 1 \leq i \leq n$, is associated a Bernoulli variate I_i, with mean q_i, that indicates claims occurrence in policy i. The statistic of interest is $I_\bullet = I_1 + I_2 + \ldots + I_n$, the total number of policies generating claims.

Usually it is assumed that the claims occurrence indicators $I_i, 1 \leq i \leq n$, are independent. As explained in the Introduction, however, this assumption is quite unrealistic in various situations. For the Examples below, the kind of dependence encountered is, in general, positive or negative association, possibly conditional. We recall that a family of r.v. $\{X_i, 1 \leq i \leq n\}$ is (positively) associated if for every pair of coordinatewise nondecreasing functions f, g,

$$\mathrm{Cov}[f(X_1, \ldots, X_n), g(X_1, \ldots, X_n)] \geq 0 \tag{2.1}$$

(Esary et al. (1967)). It is said to be negatively associated -NA- if for every pair of disjoint subsets $\mathcal{S}_1, \mathcal{S}_2$ of $\{1, \ldots, n\}$, and every pair of coordinatewise nondecreasing functions f, g,

$$\mathrm{Cov}[f(X_i, i \in \mathcal{S}_1), g(X_i, i \in \mathcal{S}_2)] \leq 0 \tag{2.2}$$

(Joag-Dev and Proschan (1983)). In the special case $n = 2$, association between X_1 and X_2 implies that $P(X_1 \geq x_1, X_2 \geq x_2) \geq P(X_1 \geq x_1)P(X_2 \geq x_2)$ for all x_1, x_2, which is the definition of positively quadrant dependence (Lehmann (1966)). NA between X_1 and X_2 is equivalent to negatively quadrant dependence.

Basically, one can distinguish three kinds of dependence structure between the I_i's: global, local and mixed. Global means that risks are influenced by a common environment, as in Examples 2.1 - 2.5 below.

Example 2.1 Consider the Fréchet space of all indicators random vectors $\mathbf{I} = (I_1, \ldots, I_n)$ with fixed marginal means $(q_1, \ldots, q_n)$. The extremal distributions in that space correspond to the Fréchet upper and lower bounds. In the actuarial literature, these bounds have been discussed (in a general setting) by, e.g., Dhaene and Goovaerts (1996, 1997), Müller (1997), Hu and Wu (1999) and Dhaene and Denuit (1999). The upper bound is the distribution of a vector $\mathbf{I}^{(c)} = (I_1^{(c)}, \ldots, I_n^{(c)})$, called comonotonic, and given by

$$\mathbf{I}^{(c)} = [I(U \geq 1 - q_1), \ldots, I(U \geq 1 - q_n)], \tag{2.3}$$

where $I(\ldots)$ denotes the indicator of $(\ldots)$, and U is a uniform variate on $[0, 1]$. As for the lower bound, under the condition that $\sum_{i=1}^{n} q_i \leq 1$ (which often holds true in practice), it is reachable and is the distribution of a vector $\mathbf{I}^{(m)} = (I_1^{(m)}, \ldots, I_n^{(m)})$, called mutually exclusive, and defined by

$$P(I_i^{(m)} = 1) = q_i, \text{ and } P(I_i^{(m)} = 1, I_j^{(m)} = 1) = 0 \text{ for all } i \neq j \in \{1, \ldots, n\}. \tag{2.4}$$

It is easily checked that $\mathbf{I}^{(c)}$ is associated, while $\mathbf{I}^{(m)}$ is NA.

Example 2.2 Marceau et al. (1999) proposed an individual model subjected to a possible global shock that can produce claims occurrences to each policy. This yields

$$I_i = \min(J_i + J_0, 1), \quad i = 1, \ldots, n, \tag{2.5}$$

where $J_0, J_1, \ldots, J_n$ are independent Bernoulli variates with means r_i. Thus, I_i, $1 \leq i \leq n$, are Bernoulli distributed with mean $q_i = r_0 + (1 - r_0) r_i$. Moreover, it is clear that $\mathbf{I}$ is associated.

Example 2.3 In the following model motivated by Genest et al. (2000), dependence is represented in $\mathbf{I}$ by writing its joint distribution function as

$$P(I_1 \leq i_1, \ldots, I_n \leq i_n) = C_\phi[F_1(i_1), \ldots, F_n(i_n)],$$

where $F_1, \ldots, F_n$ are the marginal d.f.'s of $I_1, \ldots, I_n$, and C_ϕ is an Archimedean copula, i.e. of the form

$$C_\phi(u_1, \ldots, u_n) = \phi^{-1}[\phi(u_1) + \ldots + \phi(u_n)],$$

with $\phi : (0, 1] \to \mathbb{R}^+$ such that $\phi(1) = 0$ and $(-1)^j [\phi(u)]^{(j)} \geq 0$, $1 \leq j \leq n$. If the latter condition is verified for all $n \geq 1$, ϕ is completely monotone so that ϕ^{-1} is the Laplace transform of a distribution, $K(\theta)$ say, on $\mathbb{R}^+$. Then, following Marshall and Olkin (1988), the joint d.f. of $\mathbf{I}$ can be represented as a mixture of powers, namely

$$P(I_1 \leq i_1, \ldots, I_n \leq i_n) = \int_{\theta=0}^{+\infty} \prod_{j=1}^{n} [F_j^*(i_j)]^\theta dK(\theta), \tag{2.6}$$

where $F_j^* = \exp[-\phi(F_j)]$ is the d.f. of a Bernoulli variate with mean $1 - \exp[-\phi(1 - q_j)]$, $1 \leq j \leq n$. Now, since the distributions $\{F_j^*\}^\theta$ are monotone in $\theta \in \mathbb{R}^+$, we have that $\mathbf{I}$ is associated (see, e.g., Jogdeo (1978)).

Example 2.4 In the same vein, $\mathbf{I}$ might be extracted from an infinite sequence of indicators $\{I_i, i \geq 1\}$ that are exchangeable. Then, by de Finetti's theorem, there exists some mixing r.v. Q such that conditionally on Q, the I_i's are independent Bernoulli variates. Such a situation is discussed, for instance, by Schlesinger (1999) to factor out a correlated component of loss distributions for catastrophic risks.

Example 2.5 Finite exchangeability can lead to qualitatively different situations. The following model is partly inspired from Lefèvre and Utev (1997) in a context of epidemics modelling. Let us consider a community of finite size n, and suppose that it undergoes disasters, in number $t \geq 1$ say, over a given period. Each disaster j, $1 \leq j \leq t$, is assumed to affect, independently, a random number D_j of the n individuals. During disaster j, any given individual escapes damage independently and with probability p_j. Of interest is the vector of indicators I_i, $1 \leq i \leq n$, that individual i does not succeed in escaping damage for the whole set of the t disasters. For convenience, let us introduce the complementary indicators $\bar{I}_i = 1 - I_i$, $1 \leq i \leq n$. Clearly, $\bar{\mathbf{I}}$ is a family of n exchangeable Bernoulli variates. Their joint distribution can be defined through the parameters $q(k) = P(\bar{I}_1 = 1, \ldots, \bar{I}_k = 1)$, $1 \leq k \leq n$. To determine these parameters, we first observe that for the successive disasters j, $1 \leq j \leq t$, the numbers of individuals that are effectively affected, are independent r. v.'s R_j with mixed binomial law $\text{Bin}(D_j, 1 - p_j)$. We then deduce that

$$q(k) = \prod_{j=1}^{t} E\left[\binom{n-k}{R_j} \Big/ \binom{n}{R_j}\right], \quad k = 1, \ldots, n. \tag{2.7}$$

We observe that conditionally on the R_j's, $\bar{\mathbf{I}}$ (and $\mathbf{I}$ too) has a permutation distribution and is thus NA (see, e.g., Joag-Dev and Proschan (1983)). This is not true unconditionally, however.

Local dependence means that risks are grouped into independent classes, as in Examples 2.6 - 2.8 below, while mixed dependence corresponds to risks which are submitted to both global and local dependencies, as in Example 2.9.

Example 2.6 Sets of local dependence have been examined, e.g., in Goovaerts and Dhaene (1996). In particular, they consider a portfolio of $c+s$ life insurances providing a death benefit, subdivided into c couples and s singles. Individual death occurrences are indicated by the Bernoulli variates $(I_{1,i}, I_{2,i})$ inside couple i, $1 \leq i \leq c$, and by I_i for single i, $1 \leq i \leq s$. Dependencies are only allowed within each couple i, between $I_{1,i}$ and $I_{2,i}$. It seems reasonable to assume that each of these c pairs of indicators is associated (see, e.g., Frees et al. (1996)).

Example 2.7 Going back to Example 2.2, shocks could affect only certain subgroups of the whole population. Then, for subgroup $\mathcal{C}_j$, $1 \leq j \leq k$ say,

$$I_{j,i} = \min(J_i + J_{0,j}, 1), \quad i \in \mathcal{C}_j, \tag{2.8}$$

$J_1, \ldots, J_n; J_{0,1}, \ldots, J_{0,k}$ being independent Bernoulli variates with means $r_1, \ldots, r_n; r_{0,1}, \ldots, r_{0,k}$, respectively.

Example 2.8 The epidemic-type model of Example 2.5 can be extended to take account of varying degrees of risks exposure amongst individuals. More precisely, the whole population is now partitioned into r homogeneous subgroups $\mathcal{C}_i$ with size n_i, $1 \leq i \leq r$. Claims occurrences in the different subgroups are assumed to be independent. Within subgroup $\mathcal{C}_i$, each disaster $j, 1 \leq j \leq t$, affects a random number $D_{j,i}$ of the n_i individuals, and any of these individuals escapes damage independently and with probability $p_{j,i}$.

Example 2.9 Mixing is possible, for instance, with cars accidents occurrences in a certain geographical region. These events are generally considered as independent (yielding the indicators $\mathbf{I}^{\perp}$), but under special weather conditions (fog, snow ...), they could become strongly associated (yielding the indicators $\mathbf{I}^{(c)}$ above). Let V be a Bernoulli variate with mean π to indicate such difficult driving conditions. Then,

$$\mathbf{I} = (1 - V)\mathbf{I}^{\perp} + V\mathbf{I}^{(c)}. \tag{2.9}$$

3 Tool: discrete probabilistic metrics

We purpose to evaluate distances between r.v.'s which are basically arithmetic. This leads us to introduce classes of discrete probabilistic metrics, especially in connection with related discrete probabilistic orderings. We will also discuss links with their continuous counterparts.

3.1 s-convex orderings

The so-called s-convex orderings, $s \in \mathbb{N}_0$, form a hierarchical family of probabilistic orderings, with various fields of applications in probability and statistics. For univariate r.v.'s, they were studied, e.g., by Lefèvre and Utev (1996) and Denuit and Lefèvre (1997) in the discrete case, and by Kaas and Hesselager (1995) and Denuit et al. (1998) in the continuous case.

Let us recall some standard definitions and results, firstly in the discrete case. Given a r.v. X valued in $\mathbb{N}$, put $\overline{F}_0^{(a)}(X, j) = P(X = j), j \in \mathbb{N}$, and define recursively the k-th iterated right tail of X, $k = 1, 2, \ldots$, as

$$\overline{F}_{k+1}^{(a)}(X, j) = \sum_{i=j}^{\infty} \overline{F}_k^{(a)}(X, i), \quad j \in \mathbb{N}. \tag{3.1}$$

These tails can be expressed as

$$\overline{F}_k^{(a)}(X, j) = E\binom{X - j + k - 1}{k - 1}, \quad j \in \mathbb{N}. \tag{3.2}$$

Definition 3.1 *X and Y being two r.v.'s valued in $\mathbb{N}$, X is smaller than Y in the s-convex sense, written as $X \preceq_s^{\mathbb{N}} Y$, if*

$$EX^k = EY^k \text{ for } k = 1, \ldots, s - 1, \text{ and } \overline{F}_s^{(a)}(X, j) \leq \overline{F}_s^{(a)}(Y, j), \quad j \in \mathbb{N}.$$

An equivalent characterization of the ordering $X \preceq_s^{\mathbb{N}} Y$ is that

$$E\phi(X) \leq E\phi(Y) \text{ for all discrete s-convex functions } \phi, \tag{3.3}$$

i.e. for any $\phi : \mathbb{N} \to \mathbb{R}$ that satisfies $\Delta^s \phi(j) \geq 0, j \in \mathbb{N}$, where Δ^s is the usual forward difference operator of degree s.

Passing to the continuous version, let X be a r.v. valued in $\mathbb{R}^+$, with d.f. F. Put $\overline{F}_0^{(c)}(X,x) = dF(x)/dx, x \in \mathbb{R}^+$, and define recursively the k-th iterated right tail of X, $k = 1, 2, \ldots$, as

$$\overline{F}_{k+1}^{(c)}(X,t) = \int_{x=t}^{\infty} \overline{F}_k^{(c)}(X,x)dx, \quad t \in \mathbb{R}^+. \tag{3.4}$$

These tails can be rewritten as

$$(k-1)!\overline{F}_k^{(c)}(X,t) = E(X-t)_+^{k-1}, \quad t \in \mathbb{R}^+. \tag{3.5}$$

Definition 3.2 *X and Y being two r.v.'s valued in $\mathbb{R}^+$, X is smaller than Y in the s-convex sense, written as $X \preceq_s^{\mathbb{R}^+} Y$, if*

$$EX^k = EY^k \text{ for } k = 1, \ldots, s-1, \text{ and } \overline{F}_s^{(c)}(X,t) \leq \overline{F}_s^{(c)}(Y,t), \quad t \in \mathbb{R}^+.$$

Here too, $\preceq_s^{\mathbb{R}^+}$ can be characterized by (3.3) for all continuous s-convex functions ϕ, i.e. for any $\phi : \mathbb{R}^+ \to \mathbb{R}$ with s-th derivative that satisfies $\phi^{(s)}(t) \geq 0, t \in \mathbb{R}^+$.

Now, obviously any arithmetic r.v. X may be viewed as a particular continuous r.v. (with zero probability outside $\mathbb{N}$). Thus, both discrete and continuous versions of the iterated right tails are computable for such a X. The Lemma below states that $\overline{F}_k^{(c)}(X,t)$ corresponds to some average of the $\overline{F}_k^{(a)}(X,j)$'s; its proof is given in Appendix.

Lemma 3.3 *Given a r.v. X valued in $\mathbb{N}$, the following identity holds true for any $s \geq 1$:*

$$\overline{F}_s^{(c)}(X,t) = \sum_{j=0}^{+\infty} \overline{F}_s^{(a)}(X,j)\alpha_s(j,t), \quad t \in \mathbb{R}^+, \tag{3.6}$$

where the coefficients $\alpha_s(j,t), j \geq 1$, are nonnegative and of sum equal to 1.

As a straightforward consequence of Lemma 3.3, we get the implication (3.7) below, which states that within the family of arithmetic r.v.'s, the ordering $\preceq_s^{\mathbb{N}}$ is stronger than $\preceq_s^{\mathbb{R}^+}$. This result, already known, was derived, e.g., in Denuit et al. (1999) using a different argument. It was also pointed out there that the reciprocal implication holds true only for $s = 1$ (i.e. the stochastic ordering) and $s = 2$ (i.e. the convex ordering).

Property 3.4 *X and Y being two r.v.'s valued in $\mathbb{N}$, then*

$$X \preceq_s^{\mathbb{N}} Y \Rightarrow X \preceq_s^{\mathbb{R}^+} Y, \tag{3.7}$$

and the converse is true when $s = 1$ or $s = 2$.

3.2 s-convex metrics

We examine here probabilistic metrics which are of stop-loss or total variation types with variable degrees. These will be defined for r.v.'s that are either arithmetic or continuous. By analogy with the s-convex orderings, the metrics are said to be s-convex, $s \in \mathbb{N}$.

Definition 3.5 *(Stop-loss metrics) X and Y being two r.v.'s valued in $\mathbb{N}$, the stop-loss distance of degree s between them is given as*

$$d_s^{(a)}(X,Y) = \sup_{j \in \mathbb{N}} |\overline{F}_s^{(a)}(X,j) - \overline{F}_s^{(a)}(Y,j)|.$$

If X and Y are valued in $\mathbb{R}^+$, the stop-loss distance is defined by

$$d_s^{(c)}(X,Y) = \sup_{t \in \mathbb{R}^+} |\overline{F}_s^{(c)}(X,t) - \overline{F}_s^{(c)}(Y,t)|.$$

In the actuarial literature, these stop-loss metrics have been discussed, for the continuous case, by Rachev and Rüschendorf (1990); see also, e.g., Gerber (1984) and De Pril and Dhaene (1992).

The discrete version has received very little attention. In this case, both discrete and continuous versions may be applied. It is clear that they are identical when $s = 1$ or $s = 2$: indeed, for $s = 1$,

$$d_1^{(c)}(X,Y) = d_1^{(a)}(X,Y) = \sup_{j \in \mathbb{N}} |P(X \geq j) - P(Y \geq j)| \equiv d_K(X,Y),$$

d_K being the classical Kolmogorov metric, and for $s = 2$,

$$d_2^{(c)}(X,Y) = d_2^{(a)}(X,Y) = \sup_{j \in \mathbb{N}} |E(X-j)_+ - E(Y-j)_+| \equiv d_{sl}(X,Y),$$

d_{sl} being the stop-loss metric of common use.

The distances are different, however, when $s \geq 3$. For illustration, let us choose $X = 3$ a.s. and $Y = 5$ a.s. Then, we get, for $s = 3$, that $d_3^{(c)}(X,Y) = 8$ (take $t = 0$), and $d_3^{(a)}(X,Y) = 11$ (take $j = 0$). Nevertheless, by Lemma 3.3 we directly obtain the implication (3.8) below, which states that within the family of arithmetic r.v.'s, the distance $d_s^{(c)}$ underestimates the true distance $d_s^{(a)}$.

Property 3.6 *X and Y being two r.v.'s valued in $\mathbb{N}$, then*

$$d_s^{(c)}(X,Y) \leq d_s^{(a)}(X,Y), \quad \textit{for any } s \in \mathbb{N}, \tag{3.8}$$

and the equality is true when $s = 1$ or $s = 2$.

In Theorem 3.10, we will point out a remarkable situation where $d_s^{(c)} = d_s^{(a)}$ even for $s \geq 3$.

Definition 3.7 *(Total variation metrics) X and Y being two r.v.'s valued in $\mathbb{N}$, the total variation distance of degree s between them is given as*

$$D_s^{(a)}(X,Y) = \sum_{j=0}^{\infty} |\overline{F}_s^{(a)}(X,j) - \overline{F}_s^{(a)}(Y,j)|.$$

If X and Y are valued in $\mathbb{R}^+$, the total variation distance is defined by

$$D_s^{(c)}(X,Y) = \int_{t=0}^{\infty} |\overline{F}_s^{(c)}(X,t) - \overline{F}_s^{(c)}(Y,t)|dt.$$

In the discrete case, both discrete and continuous versions are equivalent when $s = 0$ or $s = 1$. Indeed, for $s = 0$,

$$D_0^{(c)}(X,Y) = D_0^{(a)}(X,Y) = \sum_{j=0}^{\infty} |P(X=j) - P(Y=j)| \equiv 2d_{tv}(X,Y),$$

d_{tv} being the standard total variation distance, and for $s = 1$,

$$D_1^{(c)}(X,Y) = D_1^{(a)}(X,Y) = \sum_{j=0}^{\infty} |P(X>j) - P(Y>j)| \equiv d_W(X,Y),$$

d_W being the Wasserstein-Kantorovich distance. These metrics are classical, especially for the study of Poisson approximations (see, e.g., the book by Barbour et al. (1992), and Goovaerts and Dhaene (1996) in risk theory).

The two versions become distinct when $s \geq 2$. So, let us consider again the above example where $X = 3$ a.s. and $Y = 5$ a.s. By Theorem 3.10 below, we get that $D_2^{(c)}(X,Y) = d_3^{(c)}(X,Y) = 8 \neq 11 = d_3^{(a)}(X,Y) = D_2^{(a)}(X,Y)$.

3.3 Metrics and orderings

Hereafter, all the r.v.'s are assumed to be arithmetic. We start by deriving some general inequalities between metrics. Property 3.8 is immediate from Definitions 3.5, 3.7 and (3.1).

Property 3.8 *X and Y being two r.v.'s valued in $\mathbb{N}$, then*

$$d_s^{(a)}(X,Y) \leq D_s^{(a)}(X,Y), \quad \textit{and} \quad d_{s+1}^{(a)}(X,Y) \leq D_s^{(a)}(X,Y), \quad \textit{for any } s \in \mathbb{N}.$$

As a preliminary to the next Property, we recall the well-known inequality

$$d_{tv}(X,Y) \leq d_W(X,Y), \quad \text{i.e.} \quad D_0^{(a)}(X,Y) \leq 2D_1^{(a)}(X,Y).$$

This comes from the simple observation that $P(X \neq Y) = P(|X-Y| \geq 1) \leq E|X-Y|$. The result can be generalized as follows.

Property 3.9 *X and Y being two r.v.'s valued in $\mathbb{N}$, then*

$$D_s^{(a)}(X,Y) \leq 2D_{s+1}^{(a)}(X,Y), \quad \textit{and} \quad d_s^{(a)}(X,Y) \leq 2d_{s+1}^{(a)}(X,Y), \quad \textit{for any } s \in \mathbb{N}.$$

Proof. By definition,

$$\begin{aligned}\overline{F}_s^{(a)}(Y,j)-\overline{F}_s^{(a)}(X,j) &= [\overline{F}_{s+1}^{(a)}(Y,j)-\overline{F}_{s+1}^{(a)}(Y,j+1)]-[\overline{F}_{s+1}^{(a)}(X,j)-\overline{F}_{s+1}^{(a)}(X,j+1)]\\ &= [\overline{F}_{s+1}^{(a)}(Y,j)-\overline{F}_{s+1}^{(a)}(X,j)]-[\overline{F}_{s+1}^{(a)}(Y,j+1)-\overline{F}_{s+1}^{(a)}(X,j+1)].\end{aligned} \tag{3.9}$$

The announced inequalities then follow from Definitions 3.5 and 3.7. $\diamond$

A particular case of interest for the sequel is when $s=0$ or 1, yielding

$$\begin{aligned} & D_0^{(a)}(X,Y)/2 \le D_1^{(a)}(X,Y) \le 2D_2^{(a)}(X,Y),\\ \text{i.e.}\quad & d_{tv}(X,Y)\le d_W(X,Y)\le 2\sum_{j=0}^{\infty}|E(X-j)_+-E(Y-j)_+|.\end{aligned} \tag{3.10}$$

Now, since orderings and metrics are complementary notions, it is useful to enlight possible links between them. In Lefèvre and Utev (1998), the problem of preservation of metrics with respect to orders has been investigated. This time, we are going to examine how the knowledge that X is smaller or larger than Y in the s-convex sense can influence the measure of their s-convex distances.

For brevity, we will write that two random variables X and Y valued in $\mathbb{N}$ are *s-convex-ordered*, $s\in\mathbb{N}$, when either $X\preceq_s^{\mathbb{N}} Y$, or $Y\preceq_s^{\mathbb{N}} X$ holds.

Theorem 3.10 *X and Y being two r.v.'s valued in $\mathbb{N}$ which are s-convex ordered, then*

$$\begin{aligned} D_s^{(a)}(X,Y) &= D_s^{(c)}(X,Y)=d_{s+1}^{(a)}(X,Y)=d_{s+1}^{(c)}(X,Y)\\ &= \left|E\binom{X}{s}-E\binom{Y}{s}\right| = \frac{1}{s!}|EX^s-EY^s|.\end{aligned} \tag{3.11}$$

Proof. Suppose that $X\preceq_s^{\mathbb{N}} Y$, for instance. Then, by definition,

$$D_s^{(a)}(X,Y)=\sum_{j=0}^{\infty}[\overline{F}_s^{(a)}(Y,j)-\overline{F}_s^{(a)}(X,j)].$$

Moreover, since X and Y have the same first $s-1$ moments, we see from (3.2) that $\overline{F}_s^{(a)}(Y,j)=\overline{F}_s^{(a)}(X,j)$ for any $j\le s-1$. Thus, we obtain, using (3.1), that

$$D_s^{(a)}(X,Y)=\sum_{j=s}^{\infty}[\overline{F}_s^{(a)}(Y,j)-\overline{F}_s^{(a)}(X,j)]=\overline{F}_{s+1}^{(a)}(Y,s)-\overline{F}_{s+1}^{(a)}(X,s),$$

and from (3.2) and s-convex ordering,

$$D_s^{(a)}(X,Y)=E\binom{Y}{s}-E\binom{X}{s}=\frac{1}{s!}[EY^s-EX^s].$$

We also get

$$D_s^{(c)}(X,Y)=\int_{t=0}^{\infty}[\overline{F}_s^{(a)}(Y,t)-\overline{F}_s^{(a)}(X,t)]dt=\frac{1}{s!}(EY^s-EX^s),$$

using (3.5) and after direct integration, so that $D_s^{(c)}(X,Y) = D_s^{(a)}(X,Y)$. On the other hand, we have

$$\begin{aligned} d_{s+1}^{(a)}(X,Y) &= \sup_{j\in\mathbb{N}}[\overline{F}_{s+1}^{(a)}(Y,j) - \overline{F}_{s+1}^{(a)}(X,j)] \\ &\geq \overline{F}_{s+1}^{(a)}(Y,s) - \overline{F}_{s+1}^{(a)}(X,s) = D_s^{(a)}(X,Y), \end{aligned}$$

and since the inverse inequality holds too by Property 3.8, we deduce that $d_{s+1}^{(a)}(X,Y) = D_s^{(a)}(X,Y)$. Finally, we observe that

$$d_{s+1}^{(a)}(X,Y) = \frac{1}{s!}(EY^s - EX^s) \leq \frac{1}{s!}\sup_{t\in\mathbb{R}^+}[E(Y-t)_+^s - E(X-t)_+^s] = d_{s+1}^{(c)}(X,Y),$$

and by Property 3.6, we conclude that $d_{s+1}^{(a)}(X,Y) = d_{s+1}^{(c)}(X,Y)$. $\diamond$

Property 3.11 *X and Y being two r.v.'s valued in $\mathbb{N}$ which are s-convex ordered, then*

$$d_{s-1}^{(a)}(X,Y) \leq d_s^{(a)}(X,Y) \leq d_{s+1}^{(a)}(X,Y).$$

Proof. The inequality $d_s^{(a)} \leq d_{s+1}^{(a)}$ is a direct consequence of Property 3.8 and Theorem 3.10. Furthermore, as in (3.9), we can write that

$$\overline{F}_{s-1}^{(a)}(Y,j) - \overline{F}_{s-1}^{(a)}(X,j) = [\overline{F}_s^{(a)}(Y,j) - \overline{F}_s^{(a)}(X,j)] - [\overline{F}_s^{(a)}(Y,j+1) - \overline{F}_s^{(a)}(X,j+1)] \equiv a_j - b_j,$$

say. If $X \preceq_s^{\mathbb{N}} Y$, we have $0 \leq a_j, b_j \leq d_s^{(a)}(X,Y)$, which leads to

$$d_{s-1}^{(a)}(X,Y) = \sup_{j\in\mathbb{N}}|\overline{F}_{s-1}^{(a)}(Y,j) - \overline{F}_{s-1}^{(a)}(X,j)| \leq \sup_{j\in\mathbb{N}}\max(a_j,b_j) \leq d_s^{(a)}(X,Y).$$

The same inequality is obtained when $Y \preceq_s^{\mathbb{N}} X$. $\diamond$

Now, let us start by focusing on the special situation of 2-convex ordering. From (3.10), Theorem 3.10 and Property 3.11, we obtain the Corollary below (note that the result for $D_2^{(c)}$ is given in Kaas (1993)).

Corollary 3.12 *Let X and Y be two r.v.'s valued in $\mathbb{N}$ which are 2-convex ordered. Define*

$$\epsilon_{X,Y} = |\mathrm{Var}(Y) - \mathrm{Var}(X)|/2. \tag{3.12}$$

Then, the following inequalities hold true:

$$\begin{cases} D_2^{(a)}(X,Y) = D_2^{(c)}(X,Y) = d_3^{(a)}(X,Y) = d_3^{(c)}(X,Y) = \epsilon_{X,Y}, \\ d_K(X,Y) \leq d_{sl}(X,Y) \leq \epsilon_{X,Y}, \\ d_{tv}(X,Y) \leq d_W(X,Y) \leq 2\epsilon_{X,Y}. \end{cases} \tag{3.13}$$

Illustration 3.13 A classical subject in probability is the approximation of $I_\bullet^\perp$, the sum of n independent indicators I_i with mean q_i, by a Poisson variate N with mean $q_\bullet = q_1 + \ldots + q_n$ (see, e.g., Barbour et al. (1992)). First, we notice that $I_\bullet^\perp \preceq_2^{\mathbb{N}} N$ (this is obvious for $n = 1$, and then for any integer n by the preservation of the convex ordering under convolution). Therefore, Corollary 3.12 is applicable with $\epsilon_{N,I_\bullet^\perp} = (q_1^2 + \ldots + q_n^2)/2$, yielding the well-known upper bound

$$d_{tv}(N, I_\bullet^\perp) \leq d_W(N, I_\bullet^\perp) \leq \sum_{i=1}^n q_i^2, \tag{3.14}$$

but also the upper bound, which seems to be new,

$$d_K(N, I_\bullet^\perp) \leq d_{sl}(N, I_\bullet^\perp) \leq \frac{1}{2}\sum_{i=1}^n q_i^2. \tag{3.15}$$

Next, let us treat the general case where $s \geq 2$.

Corollary 3.14 *Let X and Y be two r.v.'s valued in $\mathbb{N}$ which are s-convex ordered. Define*

$$\epsilon_{X,Y}^{(s)} = |EX^s - EY^s|/s! \tag{3.16}$$

Then,

$$d_{tv}(X, Y) \leq d_W(X, Y) \leq 2^{s-1}\epsilon_{X,Y}^{(s)}, \tag{3.17}$$

and

$$d_k^{(a)}(X, Y) \leq 2^{(s-1-k)_+}\epsilon_{X,Y}^{(s)}, \quad \textit{for } k \leq s+1, \tag{3.18}$$

which yields, in particular,

$$d_K(X, Y) \leq 2^{s-2}\epsilon_{X,Y}^{(s)}, \quad \textit{and} \;\; d_{sl}(X, Y) \leq 2^{(s-3)_+}\epsilon_{X,Y}^{(s)}. \tag{3.19}$$

Proof. For (3.17), it suffices to apply $s-1$ times the former inequality of Property 3.9 and then to use Theorem 3.10. For (3.18), apply $(s-1-k)_+$ times the latter inequality of Property 3.9, then Property 3.11 for $d_{s-1}^{(a)}$ and finally Theorem 3.10. (3.19) follows from (3.18) with $k = 1$ and $k = 2$, respectively. $\diamond$

Illustration 3.15 Let $\{X_i, 1 \leq i \leq n\}$ and $\{Y_i, 1 \leq i \leq n\}$ be two families of independent r.v.'s that satisfy $X_i \preceq_s^{\mathbb{N}} Y_i$, $1 \leq i \leq n$. Put $S_n = X_1 + \ldots + X_n$ and $T_n = Y_1 + \ldots + Y_n$. Since convex ordering is preserved by convolution, Corollary 3.14 is applicable to the r.v.'s S_n and T_n, with

$$\epsilon_{S_n,T_n}^{(s)} = \frac{1}{s!}\sum_{i=1}^n (EY_i^s - EX_i^s),$$

which gives for d_{tv} an upper bound of the type encountered in limit theorems for sums of independent r.v.'s.

Remark 3.16 In some applications, the r.v.'s X and Y that are s-convex ordered, take values on some finite set $\{0, 1, \ldots, n\}$. Working with the complementary r.v.'s $n - X$ and $n - Y$, it is clear from the definitions that

$$\begin{aligned} \epsilon_{X,Y}^{(s)} &= \epsilon_{n-X,n-Y}^{(s)}, \\ d_k^{(a)}(X, Y) &= d_k^{(a)}(n - X, n - Y), \quad k = 1, \ldots, s+1, \\ D_k^{(a)}(X, Y) &= D_k^{(a)}(n - X, n - Y), \quad k = 1, \ldots, s. \end{aligned} \tag{3.20}$$

4 Dependence versus independence

Coming back to the individual risk model, our purpose is to evaluate the impact of dependence between claims occurrences $\mathbf{I} = (I_1, \ldots, I_n)$. As before, we denote by $I_\bullet = I_1 + \ldots + I_n$ the total number of policies with claims.

Firstly, we are going to derive general upper bounds for distances between $I_\bullet$ and the sum $I'_\bullet$ corresponding to another portfolio $\mathbf{I}' = (I'_1, \ldots, I'_n)$.

Theorem 4.1 *Assume that $I_\bullet$ and $I'_\bullet$ are s-convex ordered. Then, Corollary 3.14 is applicable to distances between $I_\bullet$ and $I'_\bullet$, with*

$$\epsilon^{(s)}_{I_\bullet, I'_\bullet} = \left| \sum_{j_1 < \ldots < j_s} [E(I_{j_1} \ldots I_{j_s}) - E(I'_{j_1} \ldots I'_{j_s})] \right|. \tag{4.1}$$

Proof. By Corollary 3.14, we have just to establish that $\epsilon^{(s)}_{I_\bullet, I'_\bullet} = |E(I_\bullet)^s - E(I'_\bullet)^s|/s!$ can be reexpressed as in (4.1). For that, given any real vector $\mathbf{a} = (a_1, \ldots, a_n)$ with $a_i = 0$ or 1 for all $1 \leq i \leq n$, let us introduce symmetric functions T_k defined as

$$T_k \equiv T_{k,\mathbf{a}} = \sum_{j_1 \neq \ldots \neq j_k} a_{j_1} \ldots a_{j_k}, \quad k = 1, \ldots, n. \tag{4.2}$$

We begin by deriving the following recurrence relation which will play a key role:

$$T_1^k = b_{1,k} T_k + \ldots + b_{k-1,k} T_{k-1} + T_k, \quad k = 1, \ldots, n, \tag{4.3}$$

where the coefficients $b_{i,k}$ are constants that do not depend on the vector $\mathbf{a}$. Indeed, we observe that for $1 \leq k \leq n-1$,

$$\begin{aligned} T_1 T_k &= \left(\sum_{j=1}^{n} a_j \right) \left(\sum_{j_1 \neq \ldots \neq j_k} a_{j_1} \ldots a_{j_k} \right) \\ &= \left(\sum_{j \neq j_1 \neq \ldots \neq j_k} a_j [a_{j_1} \ldots a_{j_k}] \right) + \left(\sum_{j_1 \neq \ldots \neq j_k \, ; j \in \{j_1, \ldots, j_k\}} a_j [a_{j_1} \ldots a_{j_k}] \right) \\ &= T_{k+1} + k T_k. \end{aligned}$$

Proceeding by induction, we then obtain (4.3).

Let us show how (4.1) follows from (4.3) and the assumptions of Theorem 4.1. Since $I_\bullet$ and $I'_\bullet$ are s-convex ordered, we know that

$$E(I_\bullet)^j = E(I'_\bullet)^j, \quad j = 1, \ldots, s-1,$$

which can be rewritten as

$$E(T_{1,\mathbf{I}})^j = E(T_{1,\mathbf{I}'})^j, \quad j = 1, \ldots, s-1. \tag{4.4}$$

By (4.3), we see that (4.4) is equivalent to

$$ET_{j,\mathbf{I}} = ET_{j,\mathbf{I}'}, \quad j = 1, \ldots, s-1. \tag{4.5}$$

Finally, considering (4.3) with $k = s$ and using (4.5), we deduce that

$$\epsilon^{(s)}_{I_\bullet,I'_\bullet} = |ET_{s,\mathbf{I}} - ET_{s,\mathbf{I'}}|/s!,$$

hence the identity (4.1) by (4.2). $\diamond$

Now, we may take for $I'_\bullet$ the sum $I^\perp_\bullet$ corresponding to the portfolio $\mathbf{I}^\perp = (I^\perp_1, \ldots, I^\perp_n)$ in which the indicators $I^\perp_i, 1 \leq i \leq n$ have the same parameters q_i as the original I_i, but are assumed this time to be independent.

¿From Theorem 4.1, we then get the first part of Corollary 4.2 that provides upper bounds for distances between $I_\bullet$ and $I^\perp_\bullet$. These bounds, quite effective, rely on the validity of the assumption that $I_\bullet$ and $I^\perp_\bullet$ are s-convex ordered. We recall in Lemma 4.3 below that this assumption is satisfied for $s = 2$ when, in particular, the I_i's are associated or NA. This leads to the second part of Corollary 4.2.

Corollary 4.2 *Assume that $I_\bullet$ and $I^\perp_\bullet$ are s-convex ordered. Then, Theorem 4.1 is applicable to distances between $I_\bullet$ and $I^\perp_\bullet$, with*

$$\epsilon^{(s)}_{I_\bullet,I^\perp_\bullet} = \left| \sum_{j_1<\ldots<j_s} [E(I_{j_1} \ldots I_{j_s}) - q_{j_1} \ldots q_{j_s}] \right|. \tag{4.6}$$

In particular, assume that $\mathbf{I}$ is associated or NA. Then, Corollary 3.12 is applicable to distances between $I_\bullet$ and $I^\perp_\bullet$, with

$$\epsilon^{(2)}_{I_\bullet,I^\perp_\bullet} = \left| \sum_{i<j} \mathrm{Cov}(I_i, I_j) \right|. \tag{4.7}$$

For the proof of Lemma 4.3, the reader is referred to Dhaene and Goovaerts (1996) and Denuit et al. (2000) for positive association, and to Shao (2000) for negative association. Note that association could be replaced by a slightly weaker form of dependence.

Lemma 4.3 *Assume that a family of r.v. $\{X_i, 1 \leq i \leq n\}$ is associated (resp. NA). Denote by $\{X^\perp_i, 1 \leq i \leq n\}$ an independent version of this family with the same marginal distributions. Then,*

$$X^\perp_1 + \ldots + X^\perp_n \preceq^{\mathbb{N}}_2 \ \textit{(resp. } \succeq^{\mathbb{N}}_2\textit{)}\ X_1 + \ldots + X_n. \tag{4.8}$$

At this stage, a further step to the approximation of the sum of dependent indicators by independent (less dependent) ones, is to apply to the independent case known results such as Poisson approximation, normal approximation, inequalities. For instance, let N be a Poisson variable with mean $q_\bullet = q_1 + \ldots + q_n$. From Corollary 4.2, the triangular inequality and the bounds (3.14), (3.15) we obtain the following corollary.

Corollary 4.4 *Assume that $I_\bullet$ and $I^\perp_\bullet$ are s-convex ordered. Then,*

$$d_{tv}(I_\bullet, N) \leq d_W(I_\bullet, N) \leq 2^{s-1}\epsilon^{(s)}_{I_\bullet,I^\perp_\bullet} + \sum_{i=1}^n q_i^2, \tag{4.9}$$

$$\max[d_K(I_\bullet, N), d_{sl}(I_\bullet, N)] \leq 2^{s-2}\epsilon^{(s)}_{I_\bullet,I^\perp_\bullet} + \frac{1}{2}\sum_{i=1}^n q_i^2. \tag{4.10}$$

We mention that for associated indicators I_i, Boutsikas and Koutras (2000) have derived bounds for d_K and d_W that are similar to those given in Corollaries 4.2 and 4.4 for $s = 2$. Their approach, however, is quite different and exploits directly the association property. Moreover, it is worth noticing that in the case of associated indicators, s-convex ordering between $I_\bullet$ and $I_\bullet^\perp$ for $s > 2$ implies that the indicators I_i are independent (since $\epsilon^{(2)}_{I_\bullet, I_\bullet^\perp} = 0$ by (4.7)), that is, for $s > 2$ the above results are essentially new and cannot be derived from known results for associated indicators.

Let us apply the bounds given by Theorem 4.1, Corollaries 4.2 and 4.4 to the different examples discussed in Section 2.

Example 4.5 (2.1 continued, when $\sum_{i=1}^n q_i \leq 1$). By Lemma 4.3, we know that $I_\bullet^{(m)} \preceq_2^{\mathbb{N}} I_\bullet^\perp \preceq_2^{\mathbb{N}} I_\bullet^{(c)}$. Thus, distances between $I_\bullet^{(m)}$ and $I_\bullet^\perp$, or $I_\bullet^{(c)}$ and $I_\bullet^\perp$ can be evaluated by Corollary 4.2. Note that for $i \neq j$,

$$\mathrm{Cov}(I_i^{(m)}, I_j^{(m)}) = -q_i q_j,$$

and putting $q_1 \leq \ldots \leq q_n$, we get, for $i < j$,

$$\mathrm{Cov}(I_i^{(c)}, I_j^{(c)}) = P(U \geq 1 - q_i) - q_i q_j = q_i(1 - q_j).$$

The values of $\epsilon^{(2)}_{I_\bullet^{(c)}, I_\bullet^\perp}$ and $\epsilon^{(2)}_{I_\bullet^{(m)}, I_\bullet^\perp}$ then follow from (4.7).

Example 4.6 (2.2 continued). We have here that $I_\bullet^\perp \preceq_2^{\mathbb{N}} I_\bullet$, and

$$\mathrm{Cov}(I_i, I_j) = r_0(1 - r_0) r_i r_j - q_i q_j = r_0(1 - r_0)(1 - r_i)(1 - r_j).$$

Example 4.7 (2.4 continued). The I_i's, $1 \leq i \leq n$, have a mixed Bernoulli distribution with random mean Q. Consider n indicators I_i' with similar laws but with random mean Q', and assume that Q and Q' are s-convex ordered. Then, $I_\bullet$ and $I_\bullet'$ have mixed binomial distributions that are s-convex ordered (see, e.g., Denuit et al. (1999)). Thus, Theorem 4.1 is applicable with

$$\epsilon^{(s)}_{I_\bullet, I_\bullet'} = \binom{n}{s} |E(Q)^s - E(Q')^s|.$$

In particular, since $EQ \preceq_2^{\mathbb{R}^+} Q$, we have that $I_\bullet^\perp \preceq_2^{\mathbb{N}} I_\bullet$, and Corollary 4.2 holds with

$$\epsilon^{(2)}_{I_\bullet, I_\bullet^\perp} = \binom{n}{2} \mathrm{Var}(Q).$$

Example 4.8 (2.5 continued). Here again, we want to compare $I_\bullet$ and $I_\bullet^\perp$ by applying Corollary 4.2. By Remark 3.16, it is equivalent to work with $\bar{I}_\bullet = \bar{I}_1 + \ldots + \bar{I}_n$ and $\bar{I}_\bullet^\perp = \bar{I}_1^\perp + \ldots + \bar{I}_n^\perp$. Now, let us construct a model for $\bar{I}_\bullet$. We observe that if D_j has a binomial law $\mathrm{Bin}(n, d_j)$, for some $d_j \in (0, 1)$, then R_j has a binomial law $\mathrm{Bin}(n, a_j)$ where $a_j = d_j(1 - p_j)$. Moreover, it is intuitive, and can be easily checked, that if this is the case for times $j = 1, \ldots, t$, the indicators $\bar{I}_i$ are i.i.d. Bernoulli variates with mean

$$q(1) = \prod_{j=1}^t E\left[\binom{n-1}{R_j} \Big/ \binom{n}{R_j}\right] = \prod_{j=1}^t (1 - a_j).$$

This suggests us to discuss a situation where

$$D_j \succeq_s^{\mathbb{N}} \operatorname{Bin}(n, d_j), \quad j = 1, \ldots, t. \tag{4.11}$$

We then have that $R_j \succeq_s^{\mathbb{N}} \operatorname{Bin}(n, a_j)$ (Denuit et al. (1999)), and therefore, by a result of Lefèvre and Utev (1996),

$$\bar{I}_\bullet \preceq_s^{\mathbb{N}} \bar{I}_\bullet^\perp = \operatorname{Bin}[n, q(1)]. \tag{4.12}$$

Let us compute the error term $\epsilon^{(s)}_{I_\bullet, I_\bullet^\perp}$ in terms of the deviations

$$\delta_j = E(D_j)^s - E[\operatorname{Bin}(n, d_j)]^s, \quad j = 1, \ldots, t. \tag{4.13}$$

By (4.11), we know that

$$E(D_j)^k = E[\operatorname{Bin}(n, d_j)]^k, \ \text{ and } \ E(R_j)^k = E[\operatorname{Bin}(n, a_j)]^k, \quad k = 1, \ldots, s-1. \tag{4.14}$$

Combining (4.14) and (4.13), we find that

$$E(R_j)^s - E[\operatorname{Bin}(n, a_j)]^s = (1 - p_j)^s \delta_j, \quad j = 1, \ldots, t. \tag{4.15}$$

Now, from Remark 3.16, (4.6) and (2.7), we see that

$$\epsilon^{(s)}_{I_\bullet, I_\bullet^\perp} = \binom{n}{s} \{q(s) - [q(1)]^s\} = \binom{n}{s} \left\{ \prod_{j=1}^t E\left[\binom{n-s}{R_j} \Big/ \binom{n}{R_j} \right] - \prod_{j=1}^t (1 - a_j)^s \right\}. \tag{4.16}$$

By (4.14), we get

$$\begin{aligned} E\left[\binom{n-s}{R_j} \Big/ \binom{n}{R_j} \right] &= E\left[\binom{n-s}{\operatorname{Bin}(n, a_j)} \Big/ \binom{n}{\operatorname{Bin}(n, a_j)} \right] + \frac{E(R_j)^s - E[\operatorname{Bin}(n, a_j)]^s}{n_{[s]}} \\ &= (1 - a_j)^s + \frac{E(R_j)^s - E[\operatorname{Bin}(n, a_j)]^s}{n_{[s]}}, \end{aligned} \tag{4.17}$$

where $n_{[s]} \equiv n(n-1)\ldots(n-s+1)$. Thus, inserting (4.17) with (4.15) in (4.16), we obtain that

$$\epsilon^{(s)}_{I_\bullet, I_\bullet^\perp} = \binom{n}{s} \prod_{j=1}^t (1 - a_j)^s \left\{ \prod_{j=1}^t \left[1 + \frac{\delta_j (1 - p_j)^s}{(1 - a_j)^s n_{[s]}} \right] - 1 \right\}.$$

In particular, putting

$$\gamma = \max_{1 \le j \le t} \{\delta_j (1 - p_j)^s / (1 - a_j)^s\},$$

we deduce that

$$\epsilon^{(s)}_{I_\bullet, I_\bullet^\perp} \le \binom{n}{s} \left[\left(1 + \frac{\gamma}{n_{[s]}} \right)^t - 1 \right] \le \frac{\gamma t}{s!} \exp\left(\frac{\gamma t}{n_{[s]}} \right). \tag{4.18}$$

Example 4.9 (2.6 continued). For this portfolio,

$$I_\bullet = \sum_{i=1}^{c}(I_{1,i} + I_{2,i}) + \sum_{i=1}^{s} I_i,$$

where the pairs of indicators $(I_{1,i}, I_{2,i})$ are assumed to be associated and with means $(q_{1,i}, q_{2,i})$. By Lemma 4.3 and (4.7),

$$\epsilon^{(2)}_{I_\bullet, I_\bullet^\perp} = \sum_{i=1}^{c} \mathrm{Cov}(I_{1,i}, I_{2,i}),$$

so that by Cauchy-Schwarz inequality,

$$\epsilon^{(2)}_{I_\bullet, I_\bullet^\perp} \leq \sum_{i=1}^{c} \sqrt{q_{1,i}(1-q_{1,i})q_{2,i}(1-q_{2,i})}.$$

The bound (4.9) that follows for $d_{tv}(I_\bullet, N)$, can be compared with the bound given in, e.g., formula (3.1) of Goovaerts and Dhaene (1996).

Example 4.10 (2.7 continued). As with Example 4.4, we have that

$$I^\perp_{j,\bullet} = \sum_{i \in \mathcal{C}_j} I^\perp_{j,i} \preceq_2^{\mathbb{N}} \sum_{i \in \mathcal{C}_j} I_{j,i} = I_{j,\bullet}, \quad j = 1, \ldots, k,$$

and convex ordering being preserved under convolution,

$$I_\bullet^\perp = \sum_{j=1}^{k} I^\perp_{j,\bullet} \preceq_2^{\mathbb{N}} \sum_{j=1}^{k} I_{j,\bullet} = I_\bullet.$$

Therefore, from (4.7) we get

$$\epsilon^{(2)}_{I_\bullet, I_\bullet^\perp} = \sum_{j=1}^{k} r_{0,j}(1-r_{0,j}) \sum_{i<l;\, i,l \in \mathcal{C}_j} (1-r_i)(1-r_l).$$

Example 4.11 (2.8 continued). Arguing as for (4.11), we assume that

$$D_{j,i} \succeq_s^{\mathbb{N}} \mathrm{Bin}(n, d_{j,i}), \quad j = 1, \ldots, t, \quad i = 1, \ldots, r. \tag{4.19}$$

Define

$$\gamma_i = \max_{1 \leq j \leq t} \{\delta_{j,i}(1-p_{j,i})^s/(1-a_{j,i})^s\}, \quad i = 1, \ldots, r,$$

where

$$a_{j,i} = d_{j,i}(1-p_{j,i}), \text{ and } \delta_{j,i} = E(D_{j,i})^s - E[\mathrm{Bin}(n, d_{j,i})]^s, \quad j = 1, \ldots, t, \quad i = 1, \ldots, r.$$

Since claims occurrences in different subgroups are independent, we may estimate the error term for each subgroup by a bound (4.18). Then, applying the result obtained in Illustration 3.15, we see that Corollary 4.2 is applicable to the whole population with

$$\epsilon^{(s)}_{I_\bullet, I_\bullet^\perp} \leq \sum_{i=1}^{r} \frac{\gamma_i t}{s!} \exp\left(\frac{\gamma_i t}{n_{i[s]}}\right).$$

Let us now consider two r.v.'s X and Y, valued in $\mathbb{N}$, that are mixtures of some random variable Θ. Denote $X_\theta \equiv [X|\Theta = \theta]$, and $Y_\theta \equiv [Y|\Theta = \theta]$. We say that X and Y are conditionally s-convex ordered, $s \in \mathbb{N}$, when either $X_\theta \preceq_s^{\mathbb{N}} Y_\theta$, or $Y_\theta \preceq_s^{\mathbb{N}} X_\theta$ for any θ within the support of Θ. We underline that the ordering is not required to be in the same direction for every θ (i.e., one could have $X_{\theta_1} \preceq_s^{\mathbb{N}} Y_{\theta_1}$ and $Y_{\theta_2} \preceq_s^{\mathbb{N}} X_{\theta_2}$ for some $\theta_1 \neq \theta_2 \in \Theta$).

Notice that since s-convex orderings are preserved under mixture, $X_\theta \preceq_s^{\mathbb{N}} Y_\theta$ for every $\theta \in \Theta$ implies that $X \preceq_s^{\mathbb{N}} Y$. Furthermore, the above metrics too being preserved under mixture, $d_s^{(a)}(X_\theta, Y_\theta) \leq \epsilon^{(s)}(\theta)$ for every $\theta \in \Theta$, for instance, implies that $d_s^{(a)}(X, Y) \leq E[\epsilon^{(s)}(\Theta)]$. Therefore, Corollary 3.14 yields the following result.

Corollary 4.12 *Let X and Y be two r.v.'s valued in $\mathbb{N}$ which are conditionally s-convex ordered. Define*

$$\epsilon^{(s)} = E|E(Y^s|\Theta) - E(X^s|\Theta)|/s! \tag{4.20}$$

Then, the inequalities (3.17)-(3.19) hold true.

At first view, it could be thought that since both orderings and metrics are preserved under mixture, the whole construction of conditioning is artificial and superfluous. This is not the case, however: we are going to present an example where X and Y are conditionally 2-convex ordered but not unconditionally (and thus, Corollary 4.12 is applicable but not Corollary 3.14).

Example 4.13 Let Θ and Λ be independent Bernoulli variates with respective means π and $1/n$ ($n \in \mathbb{N}$). Then, let us construct the two following r.v.'s X and Y that are mixtures of Θ:

$$X = \Theta + (1 - \Theta)n\Lambda, \tag{4.21}$$

and Y has a binomial law $\mathrm{Bin}(n, 1/n)$ (thus, Y_θ is distributed as Y for any $\theta \in \Theta$). Note that X, Y are valued in $\{0, \ldots, n\}$ and have the same mean 1.

We start by showing that X and Y are conditionally 2-convex ordered. For that, let us introduce n exchangeable Bernoulli variates $L_1, \ldots, L_n$, independent of Θ and of sum equal to 1. Obviously, X can be rewritten as

$$X = \sum_{i=1}^{n} [\Theta L_i + (1 - \Theta)\Lambda]. \tag{4.22}$$

Observe that in (4.22), the r.v.'s $[\Theta L_i + (1 - \Theta)\Lambda]$ are exchangeable Bernoulli variates with mean $1/n$. We then see that given $\Theta = 1$,

$$[X|\Theta = 1] = 1 = EY \preceq_2^{\mathbb{R}^+} Y.$$

Moreover, given $\Theta = 0$,

$$[X|\Theta = 0] = \Lambda = \begin{cases} 0 \text{ with probability } 1 - 1/n, \\ n \text{ with probability } 1/n, \end{cases}$$

which is known to be the 2-convex maximum for distributions with support in $\{0, 1, \ldots, n\}$ and mean 1, yielding

$$Y \preceq_2^{\mathbb{N}} [X|\Theta = 0].$$

Now, let us show that unconditionally, X and Y are not necessarily 2-convex ordered. Using (4.21), we find that

$$\mathrm{Var}(X) = (1-\pi)(n-1) > 1 - 1/n = \mathrm{Var}(Y) \quad \text{when} \quad 1-\pi > 1/n;$$

thus, in this case, $X \not\preceq_2^{\mathbb{N}} Y$. On the other hand, we get that

$$E(X-1)_+ = (1-\pi)(1-1/n) < (1-1/n)^n = E(Y-1)_+ \quad \text{when} \quad 1-\pi < (1-1/n)^{n-1} \sim e^{-1};$$

thus, in that case, $Y \not\preceq_2^{\mathbb{N}} X$. In conclusion, for any fixed $0 < 1-\pi < e^{-1}$, both orderings are not valid when n is sufficiently large.

Example 4.14 (2.9 continued). For the mixture model (2.9),

$$[I_\bullet | V = 0] = I_\bullet^\perp, \quad \text{and} \quad [I_\bullet | V = 1] = I_\bullet^{(c)},$$

which yields, for any previous distance δ,

$$\delta(I_\bullet, I_\bullet^\perp) = \pi\delta(I_\bullet^{(c)}, I_\bullet^\perp).$$

Since $I_\bullet$ and $I_\bullet^\perp$ are conditionally 2-convex ordered, it now suffices to bound $\delta(I_\bullet^{(c)}, I_\bullet^\perp)$ as in Example 4.5.

Acknowledgments

The research of S. Utev was supported by a grant of the Fonds National de la Recherche Scientifique Belge, for a visit at the Institut de Statistique et de Recherche Opérationnelle of the Université Libre de Bruxelles.

References

Barbour, A. D., Holst, L. and Janson, S. (1992). *Poisson Approximation.* Oxford University Press, Oxford.

Boutsikas, M. V. and Koutras, M. V. (1999). A bound for the distribution of the sum of discrete associated or NA random variables, with applications. Mimeo.

Denuit, M., Dhaene, J. and Ribas, C. (2000). Does positive dependence between individual risks increase stop-loss premiums? *Insurance: Mathematics and Economics*, to appear.

Denuit, M. and Lefèvre, Cl. (1997). Some new classes of stochastic order relations among arithmetic random variables, with applications in actuarial sciences. *Insurance: Mathematics and Economics* **20**, 197-214.

Denuit, M., Lefèvre, Cl. and Shaked, M. (1998). The s-convex orders among real random variables, with applications. *Mathematical Inequalities and their Applications* **1**, 585-613.

Denuit, M., Lefèvre, Cl. and Utev, S. (2000). Stochastic orderings of convex/concave - type on an arbitrary grid. *Mathematics of Operations Research* **24**, 835-846.

De Pril, N. and Dhaene, J. (1992). Error bounds for compound Poisson approximation of the individual risk model. *ASTIN Bulletin* **22**, 135-148.

Dhaene, J. and Denuit, M. (1999). The safest dependence structure among risks. *Insurance: Mathematics and Economics* **25**, 11-21.

Dhaene, J. and Goovaerts, M. J. (1996). Dependency of risks and stop-loss order. *ASTIN Bulletin* **26**, 201-212.

Dhaene, J. and Goovaerts, M. J. (1997). On the dependency of risks in the indidual life model. *Insurance: Mathematics and Economics* **19**, 243-253.

Esary, J. D., Proschan, F. and Walkup, D. (1967). Association of random variables with applications. *Annals of Mathematical Statistics* **38**, 1466-1474.

Frees, E. W., Carriere, J. F. and Valdez, E. (1996). Annuity valuation with dependent mortality. *The Journal of Risk and Insurance* **63**, 229-261.

Genest, Ch., Marceau, E. and Mesfioui, M. (2000). Compound Poisson approximations for individual models with dependent risks. Mimeo.

Gerber, H. U. (1984). Error bounds for the compound Poisson approximation. *Insurance: Mathematics & Economics* **3**, 191-194.

Goovaerts, M. J. and Dhaene, J. (1996). The compound Poisson approximation for a portfolio of dependent risks. *Insurance: Mathematics and Economics* **18**, 81-85.

Hu, T. and Wu, Z. (1999). On the dependence of risks and stop-loss premiums. *Insurance: Mathematics and Economics* **24**, 323-332.

Joag-Dev, K. and Proschan, F. (1983). Negative association of random variables with applications. *Annals of Statistics* **11**, 286-295.

Jogdeo, K. (1978). On a probability bound of Marshall and Olkin. *Annals of Statistics* **6**, 232-234.

Kaas, R. (1993). How to (and how not to) compute stop-loss premiums in practice. *Insurance: Mathematics and Economics* **13**, 241-254.

Kaas, R. and Hesselager, O. (1995). Ordering claim size distributions and mixed Poisson probabilities. *Insurance: Mathematics and Economics* **17**, 193-201.

Lefèvre, Cl. and Utev, S. (1996). Comparing sums of exchangeable Bernoulli random variables. *Journal of Applied Probability* **33**, 285-310.

Lefèvre, Cl. and Utev, S. (1998). On order-preserving properties of probability metrics. *Journal of Theoretical Probability* **11**, 907-920.

Lefèvre, Cl. and Utev, S. (1997). Mixed Poisson approximation in the collective epidemic models. *Stochastic Processes and their Applications* **69**, 217-246.

Lehmann, E. L. (1966). Some concepts of dependence. *Annals of Mathematical Statistics* **43**, 1137-1153.

Marceau, E., Cossette, H., Gaillardetz, P. and Rioux, J. (1999). Dependence in the individual risk model. Mimeo

Marshall, A. W. and Olkin, I. (1988). Families of multivariate distributions. *Journal of the American Statistical Association* **83**, 834-841.

Müller, A. (1997). Stop-loss order for portfolios of dependent risks. *Insurance: Mathematics and Economics* **21**, 219-223.

Rachev, S. T. and Rüschendorf, L. (1990). Approximation of sums by compound Poisson distributions with respect to stop-loss distances. *Advances in Applied Probability* **22**, 350-374.

Schlesinger, H. (1999). Decomposing catastrophic risk. *Insurance: Mathematics and Economics* **24**, 95-101.

Shao, Q. M. (2000). A comparison theorem on moment inequalities between negatively associated and independent random variables. *Journal of Theoretical Probability* **13**, 343-356.

5 Appendix

Proof of Lemma 3.3. Let $x \in \mathbb{N}$ and $t \in \mathbb{R}^+$. Write $t = k_0 + \delta$ for some $k_0 \in \mathbb{N}$ and $0 \leq \delta < 1$. We are going to establish that the following algebraic expansion holds true for any $m \in \mathbb{N}$:

$$(x-t)_+^m = \sum_{j=0}^{m} \alpha_{m+1}(j,t)\psi_m^+(x-k_0-1+j), \tag{5.1}$$

where the function $\psi_m^+ : \mathbf{Z} \to \mathbb{N}$ is an analogue to the mth descending factorial defined as

$$\psi_m^+(\xi) = \xi_+(\xi-1)_+ \ldots (\xi-m+1)_+,$$

and the coefficients $\alpha_{m+1}(j,t)$ are nonnegative. Note that the sum of the $\alpha_{m+1}(j,t)$'s is then equal to 1: indeed, dividing (5.1) by x^m, and then letting $x \to \infty$ gives 1 for the l.h.s and $\sum_{j=0}^{m} \alpha_{m+1}(j,t)$ for the r.h.s.

Now, let us choose $m = s-1 \in \mathbb{N}$ in (5.1). Putting $i = k_0 + s - j$ $(\in \mathbb{N})$ yields

$$\frac{(x-t)_+^{s-1}}{(s-1)!} = \sum_{i=k_0+1}^{k_0+s} \alpha_s(k_0+s-i,t)\binom{x-i+s-1}{s-1}.$$

Thus, substituting the r.v. X valued in $\mathbb{N}$ for x and taking the expectation, we deduce, by (3.2) and (3.5), the announced identity (3.6).

It remains to prove (5.1). This is obvious for $m = 0$ (remember that $x \leq k_0$ or $x \geq k_0 + 1$). Similarly, (5.1) holds too for $m = 1$ since $(x-t)_+$ can be decomposed as

$$(x-t)_+ = \delta(x-k_0-1)_+ + (1-\delta)(x-k_0)_+.$$

Proceeding by recurrence, let us show that if (5.1) is true for m, then it is still true for $m+1$. For that, we write that

$$\begin{aligned}(x-t)_+^{m+1} &= (x-t)_+(x-t)_+^m \\ &= [\delta(x-k_0-1)_+ + (1-\delta)(x-k_0)_+] \sum_{j=0}^{m} \alpha_{m+1}(j,t)\psi_m^+(x-k_0-1+j),\end{aligned}$$

and putting $y = x - k_0 - 1$,

$$(x-t)_+^{m+1} = \delta \sum_{j=0}^{m} \alpha_{m+1}(j,t)y_+\psi_m^+(y+j) + (1-\delta)\sum_{j=0}^{m} \alpha_{m+1}(j,t)(y+1)_+\psi_m^+(y+j). \tag{5.2}$$

Here, we will have recourse to the following relation, which is derived after: for any integers i, j satisfying

$$j - m \leq i \leq j + 1, \tag{5.3}$$

then

$$(y+i)_+ \psi_m^+(y+j) = a(m,j,i)\psi_{m+1}^+(y+j+1) + b(m,j,i)\psi_{m+1}^+(y+j), \tag{5.4}$$

where

$$0 \leq a(m,j,i) = 1 - b(m,j,i) = (m+i-j)/(m+1) \leq 1. \tag{5.5}$$

Inserting (5.4), (5.5) with $i = 0$ and 1 in (5.2), we then obtain

$$\begin{aligned}
(x-t)_+^{m+1} &= \delta \sum_{j=0}^{m} \alpha_{m+1}(j,t)[a(m,j,0)\psi_{m+1}^+(y+j+1) + b(m,j,0)\psi_{m+1}^+(y+j)] \\
&\quad + (1-\delta)\sum_{j=0}^{m} \alpha_{m+1}(j,t)[a(m,j,1)\psi_{m+1}^+(y+j+1) + b(m,j,1)\psi_{m+1}^+(y+j)] \\
&\equiv \sum_{j=0}^{m+1} \alpha_{m+2}(j,t)\psi_{m+1}^+(y+j), \quad \text{say,}
\end{aligned}$$

where the new coefficients $\alpha_{m+2}(j,t)$ are clearly nonnegative.

Let us derive (5.4), (5.5) when (5.3) is satisfied. This is obvious for (5.5). To get (5.4), we observe that it suffices to establish that

$$(y+i)\psi_m(y+j) = a(m,j,i)\psi_{m+1}(y+j+1) + b(m,j,i)\psi_{m+1}(y+j), \tag{5.6}$$

where $\psi_m : \mathbf{Z} \to \mathbf{Z}$ is defined as $\psi_m(\xi) = \xi(\xi-1)\ldots(\xi-m+1)$. Now, we directly check that

$$(y+i) = a(m,j,i)(y+j+1) + b(m,j,i)(y+j-m). \tag{5.7}$$

Moreover, we have

$$\begin{aligned}
(y+j+1)\psi_m(y+j) &= \psi_{m+1}(y+j+1), \qquad (5.8)\\
(y+j-m)\psi_m(y+j) &= \psi_{m+1}(y+j). \qquad (5.9)
\end{aligned}$$

Combining (5.8), (5.9) with (5.7) then leads to (5.6). $\diamond$