I N S T I T U T D E S T A T I S T I Q U E
B I O S T A T I S T I Q U E E T
S C I E N C E S A C T U A R I E L L E S
(I S B A)

UNIVERSITÉ CATHOLIQUE DE LOUVAIN

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

2013/16

MULTIVARIATE HIGHER-DEGREE STOCHASTIC INCREASING CONVEXITY

DENUIT, M. and M. MESFIOUI

MULTIVARIATE HIGHER-DEGREE STOCHASTIC INCREASING CONVEXITY

Michel M. Denuit
Institut de statistique, biostatistique et sciences actuarielles (ISBA)
Université Catholique de Louvain
B-1348 Louvain-la-Neuve, Belgium

Mhamed Mesfioui
Département de mathématiques et d'informatique
Université du Québec à Trois-Rivières
Trois-Rivières (Québec)
Canada G9A 5H7

April 24, 2013

Abstract

Building on the seminal work by Shaked and Shanthikumar (1988a,b), Denuit et al. (1999, 2000, 2001) studied the stochastic s-increasing convexity properties of standard parametric families of distributions. However, the analysis is restricted there to a single parameter. As many standard families of distributions involve several parameters, multivariate higher-order stochastic convexity properties also deserve consideration for applications. This is precisely the topic of the present paper, devoted to stochastic $(s_1, s_2, \ldots, s_d)$-increasing convexity of distribution families indexed by a vector $(\theta_1, \theta_2, \ldots, \theta_d)$ of parameters. This approach accounts for possible correlation in multivariate mixture models.

Keywords: Multivariate higher-degree increasing convex order, ordering of mixtures, parametric families of distributions, concordance order, upper orthant order, orthant convex order.

Subject classification: 60E15

1 Introduction and motivation

Let us consider a family of random variables $\{X_\theta, \ \theta \in \Theta\}$ valued in a subset $\mathcal{S}$ of the real line $\mathbb{R}$, indexed by a single parameter $\theta \in \Theta \subseteq \mathbb{R}$. A classical problem in applied probability can be described as follows: given a function $g : \mathcal{S} \to \mathbb{R}$, does the associated function $g^\star$ defined as

$$g^\star \ : \ \Theta \to \mathbb{R} \ \ ; \ \ \theta \mapsto \mathbb{E}[g(X_\theta)] \tag{1.1}$$

inherit some properties from g? After Shaked and Shanthikumar (1988a,b) who considered non-decreasingness and convexity properties, Denuit et al. (1999, 2000, 2001) investigated the possible transmission of higher-order convexity properties from a function g to the function $g^\star$ given in (1.1).

In this paper, we extend this approach to the multivariate case, considering parametric families of random variables indexed with several parameters. Specifically, we consider a family of random variables $\{X_{\boldsymbol{\theta}}, \ \boldsymbol{\theta} \in \Theta\}$ valued in a subset $\mathcal{S}$ of the real line $\mathbb{R}$, indexed by a vector of parameters $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_d) \in \Theta \subseteq \mathbb{R}^d$. Given a function $g : \mathcal{S} \to \mathbb{R}$, define the associated function $g^\star$ as

$$g^\star \ : \ \Theta \to \mathbb{R} \ \ ; \ \ \boldsymbol{\theta} \mapsto \mathbb{E}[g(X_{\boldsymbol{\theta}})]. \tag{1.2}$$

For $d = 1$, we recover the univariate case (1.1) studied in Denuit et al. (1999, 2000, 2001). In this paper, we investigate the possible transmission of multivariate higher-order convexity properties from g to the function $g^\star$ given in (1.2).

The analysis conducted in the present paper has interesting applications for mixture models. Letting the vector parameter $\boldsymbol{\theta}$ become a random vector $\boldsymbol{T} = (T_1, \ldots, T_d)$, we see that $\mathbb{E}[g(X_{\boldsymbol{T}})] = \mathbb{E}[g^\star(\boldsymbol{T})]$. Hence, replacing $\boldsymbol{T}$ with another random vector dominating $\boldsymbol{T}$ in some appropriate stochastic order relation, the corresponding mixtures may be ordered accordingly. This is particularly interesting when there is some uncertainty about the parameter $\boldsymbol{\theta}$ which can be represented by a probability distribution, making the vector parameter random. In empirical illustrations, where $\boldsymbol{\theta}$ is estimated from the available data, $\boldsymbol{T}$ may represent a suitable estimator for $\boldsymbol{\theta}$. In risk analysis, $\boldsymbol{T}$ often accounts for some common factors influencing the outcome $X_{\boldsymbol{T}}$. Stengthening the dependence between the components of $\boldsymbol{T}$ often increases the risk associated to $X_{\boldsymbol{T}}$.

Some notions of multivariate stochastic convexity have been proposed in the literature, based on directional convexity. See, e.g., Shaked and Shanthikumar (1990), Chang et al. (1991), Meester and Shanthikumar (1993) and Fernandez-Ponce et al. (2008). Typically, stochastic directional convexity of parametric families allows one to study how the dependence and variability among multivariate random environments influences the variability of the mixed distribution by means of the increasing directionally convex order. As explained in Denuit and Mesfioui (2010), the multivariate higher-degree increasing convex orders considered in the present paper, comprising the upper orthant and the orthant convex orders as special cases, differ from the directionally convex orders and their higher-degree extensions.

The remainder of the paper is organized as follows. Section 2 recalls the definition of univariate and multivariate higher-degree increasing convex functions and of the integral stochastic order relations defined by means of these classes of functions. The concept studied in the present paper is then introduced in Section 3. Section 4 is devoted to the study of location-scale families and of some of their extensions. In Section 5, we consider sums of independent and identically distributed random variables, first with a fixed number of terms and then with a random number of terms. We pay special attention to the particular cases involving low-degree orders, including the univariate convex order and the multivariate upper orthant order and orthant convex order, as well as the intermediate cases considered in Denuit (2010). An appendix collects some technical results for the Binomial distribution. Before proceeding to the next section, let us introduce some notation used in the present paper. Henceforth, the half positive real line $(0, \infty)$ is denoted as $\mathbb{R}^+$. The set $\{1, 2, \ldots\}$ of the positive integers is denoted as $\mathbb{N}$.

2 Univariate and multivariate stochastic higher-degree increasing convex orders

Let s be a positive integer. The univariate s-increasing convex orders have been introduced by Denuit et al. (1998) to compare random variables valued in a subset $\mathcal{S}$ of the real line $\mathbb{R}$. Assume first that $\mathcal{S}$ is an interval and let $\mathcal{U}^{\mathcal{S}}_{s-\text{icx}}$ be the class of all the (regular) s-increasing convex functions $g : \mathcal{S} \to \mathbb{R}$, i.e. those functions g such that $\frac{d^k}{dx^k} g(x) \geq 0$ for all $x \in \mathcal{S}$ and $k = 1, 2, \ldots, s$. Given two random variables X and Y valued in $\mathcal{S}$, X is said to be smaller than Y in the s-increasing convex sense, denoted as $X \preceq^{\mathcal{S}}_{s-\text{icx}} Y$, if

$$\mathbb{E}[g(X)] \leq \mathbb{E}[g(Y)] \text{ for all } g \in \mathcal{U}^{\mathcal{S}}_{s-\text{icx}}, \tag{2.1}$$

provided that the expectations involved in (2.1) exist. For $s = 1$, we recover the usual stochastic order defined as

$$X \preceq^{\mathcal{S}}_{1-\text{icx}} Y \Leftrightarrow \Pr[X > t] \leq \Pr[Y > t] \text{ for all } t.$$

In words, $X \preceq^{\mathcal{S}}_{1-\text{icx}} Y$ means that Y tends to be larger than X as the probability that Y exceeds a given threshold t is larger than the corresponding probability for X, whatever t. The case $s = 2$ corresponds to the increasing convex order defined as

$$X \preceq^{\mathcal{S}}_{2-\text{icx}} Y \Leftrightarrow \mathbb{E}[(X - t)_+] \leq \mathbb{E}[(Y - t)_+] \text{ for all } t$$

where $(\cdot)_+$ denotes the positive part of its argument. The intuitive meaning of a ranking in the increasing convex order is as follows: $X \preceq^{\mathcal{S}}_{2-\text{icx}} Y$ means that Y tends to be larger and more variable than X. Notice that the function $x \mapsto (x - t)_+$ involved in the definition of $X \preceq^{\mathcal{S}}_{2-\text{icx}} Y$ has many interpretations in applied probability. It represents the payoff of a call option with exercise price t in finance or the amount paid by a reinsurer in a stop-loss treaty with priority t in actuarial science, for instance. We refer the reader to Shaked and Shanthikumar (2007) or to Denuit et al. (2005) for a detailed description of these stochastic order relations.

For $s \geq 3$ and $\mathcal{S}$ a subset of $\mathbb{N}$, more accurate comparisons can be obtained by considering test functions g involved in (2.1) defined on $\mathbb{N}$, replacing derivatives with finite differences as explained in Denuit and Lefevre (1997) and Denuit et al. (1999c). Specifically, let Δ be the forward difference operator, which is defined for a function $g : \mathbb{N} \to \mathbb{R}$ by $\Delta g(i) =$

$g(i+1)-g(i)$, $i \in \mathbb{N}$. Let Δ^k be the k-th order forward difference operator defined recursively by $\Delta^0 g = g$ and $\Delta^k g(i) = \Delta^{k-1}g(i+1) - \Delta^{k-1}g(i)$, $i \in \mathbb{N}$. Given two integer-valued random variables X and Y, X is smaller than Y in the s-increasing convex sense, written as $X \preceq^{\mathbb{N}}_{s-\mathrm{icx}} Y$, if (2.1) holds with

$$\mathcal{U}^{\mathbb{N}}_{s-\mathrm{icx}} = \{g : \mathbb{N} \to \mathbb{R} | \Delta^k g(i) \geq 0 \text{ for all } i \in \mathbb{N},\ \ k = 1, 2, \ldots, s\}. \tag{2.2}$$

The univariate s-increasing convex orders have been extended to the bivariate case by Denuit et al. (1999a) and to the multivariate case by Denuit and Mesfioui (2010). Specifically, let $\boldsymbol{s} = (s_1, \ldots, s_n)$ be a vector of positive integers and let $\mathcal{S}$ be a subset of the n-dimensional real space $\mathbb{R}^n$. Assume that $\mathcal{S}$ is the cartesian product of intervals $[a_1, b_1] \times \ldots \times [a_n, b_n]$ and let $\mathcal{U}^{\mathcal{S}}_{\boldsymbol{s}-\mathrm{icx}}$ be the class of all the functions $g : \mathcal{S} \to \mathbb{R}$ such that

$$\frac{\partial^{k_1+k_2+\cdots+k_n}}{\partial x_1^{k_1} \partial x_2^{k_2} \ldots \partial x_n^{k_n}} g(x_1, x_2, \ldots, x_n) \geq 0 \text{ for all } x_1, x_2, \ldots, x_n \in \mathcal{S},$$

$$k_i = 0, 1, \ldots, s_i,\ \ i = 1, 2, \ldots, n,\ \ k_1 + k_2 + \cdots + k_n \geq 1. \tag{2.3}$$

Consider two n-dimensional random vectors $(X_1, \ldots, X_n)$ and $(Y_1, \ldots, Y_n)$ valued in $\mathcal{S}$. Following Denuit and Mesfioui (2010), we say that $(X_1, \ldots, X_n)$ is smaller than $(Y_1, \ldots, Y_n)$ in the $\boldsymbol{s}$-increasing convex order, which is denoted by $(X_1, \ldots, X_n) \preceq^{\mathcal{S}}_{\boldsymbol{s}-\mathrm{icx}} (Y_1, \ldots, Y_n)$, if

$$\mathbb{E}[g(X_1, \ldots, X_n)] \leq \mathbb{E}[g(Y_1, \ldots, Y_n)] \text{ for all } g \in \mathcal{U}^{\mathcal{S}}_{\boldsymbol{s}-\mathrm{icx}}, \tag{2.4}$$

provided that the expectations involved in (2.4) exist. Recall from Denuit and Mesfioui (2010) that the stochastic inequality $(X_1, X_2, \ldots, X_n) \preceq^{\mathcal{S}}_{\boldsymbol{s}-\mathrm{icx}} (Y_1, Y_2, \ldots, Y_n)$ holds if, and only if, the inequality

$$\mathbb{E}\left[\prod_{i=1}^n g_i(X_i)\right] \leq \mathbb{E}\left[\prod_{i=1}^n g_i(Y_i)\right]$$

is fulfilled for all the non-negative s_i-increasing convex functions g_i, $i = 1, \ldots, n$. When $s_i = 1$ for all i, we recover the upper orthant order denoted by $(X_1, \ldots, X_n) \preceq_{\mathrm{uo}} (Y_1, \ldots, Y_n)$ corresponding to the the case where the inequality

$$\Pr[X_1 > x_1, X_2 > x_2, \ldots, X_n > x_n] \leq \Pr[Y_1 > x_1, Y_2 > x_2, \ldots, Y_n > x_n] \tag{2.5}$$

is valid for all $x_1, x_2, \ldots, x_n$. When $s_i = 2$ for all i, we recover the orthant convex order denoted by $(X_1, \ldots, X_n) \preceq_{\text{uo-cx}} (Y_1, \ldots, Y_n)$ for which the non-negative test functions $g_1, \ldots, g_n$ are non-decreasing and convex. We refer the reader to Shaked and Shanthikumar (2007) or to Denuit et al. (2005) for more details about these stochastic order relations.

Order relations for integer-valued random vectors are discussed in Denuit et al. (1999b). In the context of the present study, besides parameters valued in (an interval of) the real line, several standard parametric families also involve integer-valued parameters, such as the number of trials in the Binomial case. In this case, finite differences Δ are combined with partial derivatives. The applications developed in the present paper only require a single integer-valued parameter so that we only discuss this particular case. Considering a function g defined on $\mathbb{N} \times \mathcal{S}$, where $\mathcal{S}$ is the cartesian product of intervals, $(s_0, s_1, \ldots, s_n)$-increasing convexity means that $\Delta^j g$ belongs to $\mathcal{U}^{\mathcal{S}}_{(s_1,\ldots,s_n)-\text{icx}}$ for $j = 0, 1, \ldots, s_0$ where the function $\Delta^j g$ is defined recursively from

$$\Delta g(k) = g(k+1, x_1, \ldots, x_n) - g(k, x_1, \ldots, x_n) \text{ for } k \in \mathbb{N}$$

and $\Delta^j g = \Delta \circ \Delta^{j-1} g$ for $j \geq 2$; by convention, $\Delta^0 g = g$. Specifically, the partial derivatives of the functions $\Delta^j g$ have to fulfill the conditions stated in (2.3).

As most classical parametric families of distributions involve two parameters, the bivariate case is of particular interest. For $s_1 = s_2 = 1$, $\preceq^{\mathcal{S}}_{(1,1)-\text{icx}}$ coincides with the concordance order. Concordance conveys the idea of clustering of large and small events. Large and small values tend to be more often associated under the distribution that dominates the other one in the concordance order. The concordance order compares the strength of the positive quadrant dependence between the two components of the random couples as well as their relative sizes. The $\preceq^{\mathcal{S}}_{(1,1)-\text{icx}}$ order can be characterized by

$$\begin{aligned} & (X_1, X_2) \preceq^{\mathcal{S}}_{(1,1)-\text{icx}} (Y_1, Y_2) \\ \Leftrightarrow\quad & \mathbb{E}[g(X_1, X_2)] \leq \mathbb{E}[g(Y_1, Y_2)], \text{ for all non-decreasing } g \text{ such that } g^{(1,1)} \geq 0 \\ \Leftrightarrow\quad & \mathbb{E}[g_1(X_1)g_2(X_2)] \leq \mathbb{E}[g_1(Y_1)g_2(Y_2)], \text{ for all non-decreasing } g_1, g_2 \geq 0 \\ \Leftrightarrow\quad & \Pr[X_1 > t_1, X_2 > t_2] \leq \Pr[Y_1 > t_1, Y_2 > t_2], \text{ for all } t_1, t_2. \end{aligned}$$

The last inequality shows that $\preceq^{\mathcal{S}}_{(1,1)-\text{icx}}$ compares the respective probabilities assigned to quadrants $(t_1, +\infty) \times (t_2, +\infty)$ of the real plane. Let us provide the intuition behind a ranking in the $\preceq^{\mathcal{S}}_{(1,1)-\text{icx}}$-sense. Consider an individual faced with (X_1, X_2) and (Y_1, Y_2). This individual could select a target t_1 for the first component and a target t_2 for the second component and then rank (X_1, X_2) and (Y_1, Y_2) according to the probability that they meet the specified targets. As $(X_1, X_2) \preceq^{\mathcal{S}}_{(1,1)-\text{icx}} (Y_1, Y_2) \Rightarrow \Pr[X_1 > t_1] \leq \Pr[Y_1 > t_1]$ for all t_1 and $\Pr[X_2 > t_2] \leq \Pr[Y_2 > t_2]$ for all t_2, the components of (X_1, X_2) tend to be smaller compared to those of (Y_1, Y_2), the latter being also more positively related. In words, $(X_1, X_2) \preceq_{(1,1)-\text{icx}} (Y_1, Y_2)$ thus means that Y_1 and Y_2 tend to be larger and more positively related compared to X_1 and X_2.

If (X_1, X_2) and (Y_1, Y_2) have identical univariate marginals, that is,

$$\Pr[X_1 \leq t] = \Pr[Y_1 \leq t] \text{ and } \Pr[X_2 \leq t] = \Pr[Y_2 \leq t] \text{ hold for all } t, \tag{2.6}$$

then comparing (X_1, X_2) and (Y_1, Y_2) is only based on their dependence structure, i.e. on the way both components interact. In this case, we also have

$$\begin{aligned} (X_1, X_2) \preceq^{\mathcal{S}}_{(1,1)-\text{icx}} (Y_1, Y_2) \quad \Leftrightarrow\quad & \mathbb{C}\text{ov}[h_1(X_1), h_2(X_2)] \leq \mathbb{C}\text{ov}[h_1(X_1), h_2(Y_2)], \\ & \text{for all non-decreasing } h_1, h_2 \geq 0 \\ \Leftrightarrow\quad & \Pr[X_1 > t_1 | X_2 > t_2] \leq \Pr[Y_1 > t_1 | Y_2 > t_2] \text{ for all } t_1 \text{ and } t_2 \\ & \text{provided } \Pr[X_2 > t_2] > 0 \text{ and } \Pr[Y_2 > t_2] > 0. \end{aligned}$$

This intuitively means that the knowledge that Y_2 is large (that is, $Y_2 > t_2$) increases the probability that Y_1 is also large (that is, $Y_1 > t_1$) compared to (X_1, X_2). In terms of targets, the knowledge that Y_2 meets the target t_2 increases the probability that Y_1 also meets the target t_1 compared to (X_1, X_2).

Now, for $s_1 = 2$ and $s_2 = 1$, the stochastic order relation $\preceq^{\mathcal{S}}_{(2,1)-\text{icx}}$ is weaker than the concordance order $\preceq^{\mathcal{S}}_{(1,1)-\text{icx}}$. Denoting as $\mathbb{I}[A]$ the indicator function of the event A (equal to 1 if A is realized and to 0 otherwise), we have

$$
\begin{aligned}
& (X_1, X_2) \preceq^{\mathcal{S}}_{(2,1)} (Y_1, Y_2) \\
\Leftrightarrow\ & \mathbb{E}[g(X_1, X_2)] \leq \mathbb{E}[g(Y_1, Y_2)], \text{ for all non-decreasing } g \text{ such that } g^{(2,0)} \geq 0 \\
& \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad g^{(1,1)} \geq 0 \text{ and } g^{(2,1)} \geq 0 \\
\Leftrightarrow\ & \mathbb{E}[g_1(X_1)g_2(X_2)] \leq \mathbb{E}[g_1(Y_1)g_2(Y_2)], \text{ for all non-decreasing } g_1, g_2 \geq 0 \\
& \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{with } g_1 \text{ convex} \\
\Leftrightarrow\ & \left\{ \begin{array}{l} \mathbb{E}[(X_1 - t_1)_+] \leq \mathbb{E}[(Y_1 - t_1)_+] \text{ for all } t_1 \\ \Pr[X_2 > t_2] \leq \Pr[Y_2 > t_2] \text{ for all } t_2 \\ \mathbb{E}\Big[(X_1 - t_1)_+ \mathbb{I}[X_2 > t_2]\Big] \leq \mathbb{E}\Big[(Y_1 - t_1)_+ \mathbb{I}[Y_2 > t_2]\Big] \text{ for all } t_1 \text{ and } t_2. \end{array} \right.
\end{aligned}
$$

Note that the two components do not play a symmetric role in this comparison. The second component can be considered as a signal as the ordering of (X_1, X_2) and (Y_1, Y_2) in the (2,1)-increasing convex sense ensures that the random vectors $(X_1, h(X_2))$ and $(Y_1, h(Y_2))$ are ordered accordingly for any non-decreasing transformation h. In that respect, the second component can be normalized into any reference distribution, such as the unit uniform one, for instance. This also means that the scale on which the second attribute is measured is not relevant. Under (2.6), we have

$$
\begin{aligned}
& (X_1, X_2) \preceq^{\mathcal{S}}_{(2,1)-\text{icx}} (Y_1, Y_2) \\
\Leftrightarrow\ & \mathbb{E}\big[(X_1 - t_1)_+ \mathbb{I}[X_2 > t_2]\big] \leq \mathbb{E}\big[(Y_1 - t_1)_+ \mathbb{I}[Y_2 > t_2]\big], \text{ for all } t_1, t_2 \\
\Leftrightarrow\ & \mathbb{C}\text{ov}[g_1(X_1), g_2(Y_2)] \leq \mathbb{C}\text{ov}[g_1(Y_1), g_2(Y_2)], \text{ for all non-decreasing } g_1, g_2 \geq 0 \\
& \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{with } g_1 \text{ convex} \\
\Leftrightarrow\ & \mathbb{E}\Big[(X_1 - t_1)_+ \Big| X_2 > t_2\Big] \leq \mathbb{E}\Big[(Y_1 - t_1)_+ \Big| Y_2 > t_2\Big] \text{for all } t_1 \text{ and } t_2 \qquad (2.7) \\
& \qquad\qquad\qquad \text{provided } \Pr[X_2 > t_2] > 0 \text{ and } \Pr[Y_2 > t_2] > 0.
\end{aligned}
$$

Inequality (2.7) shows that $(X_1, X_2) \preceq^{\mathcal{S}}_{(2,1)-\text{icx}} (Y_1, Y_2)$ means that the knowledge that Y_2 is large (that is, $Y_2 > t_2$) increases the average part of Y_1 above any threshold t_1 compared to (X_1, X_2). This also means that the conditional distribution of Y_1 given $Y_2 > t_2$ dominates the conditional distribution of X_1 given $X_2 > t_2$ in the univariate increasing convex order. The order relation $\preceq^{\mathcal{S}}_{(2,1)-\text{icx}}$ has been used by Denuit (2010) to rank conditional expectations in the convex sense[1].

[1]Note, however, that in the main results of Denuit (2010), non-decreasingness must be replaced with increasingness to ensure that the transformation corresponding to the conditional expectation is one-to-one.

For $s_1 = s_2 = 2$, we get the bivariate orthant convex order

$$
\begin{aligned}
& (X_1, X_2) \preceq^{\mathcal{S}}_{(2,2)-\text{icx}} (Y_1, Y_2) \\
\Leftrightarrow\ & \mathbb{E}[g(X_1, X_2)] \leq \mathbb{E}[g(Y_1, Y_2)], \text{ for all non-decreasing } g \text{ such that } g^{(2,0)} \geq 0, g^{(0,2)} \geq 0 \\
& \qquad\qquad g^{(1,1)} \geq 0, g^{(2,1)} \geq 0, g^{(1,2)} \geq 0 \text{ and } g^{(2,2)} \geq 0 \\
\Leftrightarrow\ & \mathbb{E}[g_1(X_1)g_2(X_2)] \leq \mathbb{E}[g_1(Y_1)g_2(Y_2)], \text{ for all non-decreasing convex } g_1, g_2 \geq 0 \\
\Leftrightarrow\ & \begin{cases} \mathbb{E}[(X_1 - t_1)_+] \leq \mathbb{E}[(Y_1 - t_1)_+] \text{ for all } t_1 \\ \mathbb{E}[(X_2 - t_2)_+] \leq \mathbb{E}[(Y_2 - t_2)_+] \text{ for all } t_2 \\ \mathbb{E}\Big[(X_1 - t_1)_+(X_2 - t_2)_+\Big] \geq \mathbb{E}\Big[(Y_1 - t_1)_+(Y_2 - t_2)_+\Big] \text{ for all } t_1 \text{ and } t_2. \end{cases}
\end{aligned}
$$

The intuitive meaning of this characterization becomes clear when we restrict ourselves to random vectors with the same univariate marginals. If (2.6) is fulfilled, we have that $(X_1, X_2) \preceq^{\mathcal{S}}_{(2,2)-\text{icx}} (Y_1, Y_2)$ holds true if, and only if,

$$
\mathbb{C}\text{ov}\Big[(X_1 - t_1)_+, (X_2 - t_2)_+\Big] \leq \mathbb{C}\text{ov}\Big[(Y_1 - t_1)_+, (Y_2 - t_2)_+\Big] \text{ for all } t_1 \text{ and } t_2. \tag{2.8}
$$

This means that the exceedances $(X_1 - t_1)_+$ and $(X_2 - t_2)_+$ of X_1 and X_2 over the thresholds t_1 and t_2 are less correlated than the corresponding exceedances $(Y_1 - t_1)_+$ and $(Y_2 - t_2)_+$ of the dominating (Y_1, Y_2). In financial terms, (2.8) considers the correlation of the payoffs of call options written on assets X_1 and X_2 with respective exercise prices t_1 and t_2. In actuarial science, $\preceq^{\mathcal{S}}_{(2,2)-\text{icx}}$ compares the correlation between the payoffs of stop-loss reinsurance treaties.

3 Multivariate stochastic higher-degree increasing convexity and orderings of mixtures

Let $X_{\boldsymbol{\theta}}$ be a family of random variables valued in a subset $\mathcal{S} \subseteq \mathbb{R}$, and indexed by a vector of parameters $\boldsymbol{\theta} \in \Theta \subset \mathbb{R}^d$. To any function $g : \mathcal{S} \to \mathbb{R}$, let us associate the function $g^\star$ defined in (1.2). Our aim is study whether some higher-order increasing convexity properties of the function g can be transmitted to the function $g^\star$. This is formally stated next.

Definition 3.1. Let $\boldsymbol{s} = (s_1, \ldots, s_d)$ be a vector of positive integers and let $t = t(\boldsymbol{s})$ be a positive integer, depending on $\boldsymbol{s}$. If

$$
g \in \mathcal{U}^{\mathcal{S}}_{t-\text{icx}} \Rightarrow g^\star \in \mathcal{U}^{\Theta}_{\boldsymbol{s}-\text{icx}}
$$

provided that the expectation defining $g^\star$ exists, then the family $\{X_{\boldsymbol{\theta}}, \boldsymbol{\theta} \in \Theta\}$ is said to be stochastically $(t, \boldsymbol{s})$-increasing convex.

As explained in the introduction, this property is especially useful when mixtures are considered. Let $X_{\boldsymbol{T}}$ denote a random variable obtained from a mixture of $X_{\boldsymbol{\theta}}$, i.e. replacing the fixed parameter vector $\boldsymbol{\theta}$ with a random vector $\boldsymbol{T} = (T_1, \ldots, t_d)$ valued in the parameter space $\Theta \subseteq \mathbb{R}^d$. The distribution function of $X_{\boldsymbol{T}}$ is then given by

$$
\Pr[X_{\boldsymbol{T}} \leq x] = \int_{\theta_1} \ldots \int_{\theta_d} \Pr[X_{\boldsymbol{\theta}} \leq x] d \Pr[\boldsymbol{T} \leq \boldsymbol{\theta}], \;\; x \in \mathcal{S}.
$$

The next result shows how a stochastic order relation for $\boldsymbol{T}$ translates in a ranking of the corresponding $X_{\boldsymbol{T}}$.

Proposition 3.2. *If the family $\{X_{\boldsymbol{\theta}}, \boldsymbol{\theta} \in \Theta\}$ is $(t, \boldsymbol{s})$-increasing convex then*

$$\boldsymbol{T}_1 \preceq^{\Theta}_{\boldsymbol{s}-icx} \boldsymbol{T}_2 \quad \Rightarrow \quad X_{\boldsymbol{T}_1} \preceq^{\mathcal{S}}_{t-icx} X_{\boldsymbol{T}_2}.$$

Proof. This is a direct consequence of the representation $\mathbb{E}[g(X_{\boldsymbol{T}_i})] = \mathbb{E}[g^\star(\boldsymbol{T}_i)]$ for $i = 1, 2$ which ensures that for any t-increasing convex function g,

$$\begin{aligned} \mathbb{E}[g(X_{\boldsymbol{T}_1})] &= \mathbb{E}[g^\star(\boldsymbol{T}_1)] \\ &\leq \mathbb{E}[g^\star(\boldsymbol{T}_2)] \text{ as } g \in \mathcal{U}^{\mathcal{S}}_{t-\text{icx}} \Rightarrow g^\star \in \mathcal{U}^{\Theta}_{\boldsymbol{s}-\text{icx}} \text{ and } \boldsymbol{T}_1 \preceq^{\Theta}_{\boldsymbol{s}-\text{icx}} \boldsymbol{T}_2 \\ &= \mathbb{E}[g(X_{\boldsymbol{T}_1})], \end{aligned}$$

which ends the proof. □

4 Location-scale families and extensions

Location-scale families of distributions constitute a particular case of interest for many applications. Starting from a reference distribution, the whole family is obtained by scaling and shift. The bivariate stochastic increasing convex properties of these families are established in the next result.

Property 4.1. *Consider $\boldsymbol{\theta} = (\mu, \sigma)$ and define $X_{(\mu,\sigma)} = \mu + \sigma Z$ where Z is a real-valued random variable with probability density function f_Z symmetric about 0, i.e. $f_Z(z) = f_Z(-z)$ for all $z > 0$. Then,*

(i) the family $\{X_{\boldsymbol{\theta}}, \ \boldsymbol{\theta} \in \mathbb{R} \times \mathbb{R}^+\}$ is stochastically $(s_1 + s_2 + 1, (s_1, s_2))$-increasing convex for any integers $s_1 \geq 1$ and $s_2 \geq 1$ with s_2 odd. We thus have

$$(M_1, S_1) \preceq^{\mathbb{R}\times\mathbb{R}^+}_{(s_1,s_2)-icx} (M_2, S_2) \quad \Rightarrow \quad X_{(M_1,S_1)} \preceq^{\mathbb{R}}_{(s_1+s_2+1)-icx} X_{(M_2,S_2)}.$$

(ii) the family $\{X_{\boldsymbol{\theta}}, \ \boldsymbol{\theta} \in \mathbb{R} \times \mathbb{R}^+\}$ is stochastically $(s_1 + s_2, (s_1, s_2))$-increasing convex for any integers $s_1 \geq 1$ and $s_2 \geq 1$ with s_2 even. We thus have

$$(M_1, S_1) \preceq^{\mathbb{R}\times\mathbb{R}^+}_{(s_1,s_2)-icx} (M_2, S_2) \quad \Rightarrow \quad X_{(M_1,S_1)} \preceq^{\mathbb{R}}_{(s_1+s_2)-icx} X_{(M_2,S_2)}.$$

Proof. Consider $g \in \mathcal{U}^{\mathbb{R}}_{(s_1+s_2+1)-\text{icx}}$ with s_2 odd and define $g^\star(\mu, \sigma) = \mathbb{E}\left[g(X_{(\mu,\sigma)})\right]$. Clearly, we have for $k_1 = 0, 1, \ldots, s_1$ and $k_2 = 0, 1, \ldots, s_2$ such that $k_1 + k_2 \geq 1$,

$$\frac{\partial^{k_1+k_2}}{\partial\mu^{k_1}\partial\sigma^{k_2}} g^\star(\theta, \sigma) = \mathbb{E}\left[Z^{k_2} g^{(k_1+k_2)}(X_{(\mu,\sigma)})\right]$$

which is obviously non-negative if k_2 is even. For k_2 odd, we have

$$\begin{aligned} \frac{\partial^{k_1+k_2}}{\partial\mu^{k_1}\partial\sigma^{k_2}} g^\star(\theta, \sigma) &= \frac{\partial^{k_1+k_2}}{\partial\mu^{k_1}\partial\sigma^{k_2}} \left(\int_0^\infty \Big(g(\mu + \sigma z) - g(\mu - \sigma z) \Big) f_Z(z) dz \right) \\ &= \int_0^\infty z^{k_2} \Big(g^{(k_1+k_2)}(\mu + \sigma z) - g^{(k_1+k_2)}(\mu - \sigma z) \Big) f_Z(z) dz \end{aligned}$$

which is also positive since $g \in \mathcal{U}^{\mathbb{R}}_{(s_1+s_2+1)-\mathrm{icx}}$ so that $g^{(k_1+k_2)}$ is non-decreasing for every $k_1 = 0, 1, \ldots, s_1$ and $k_2 = 0, 1, \ldots, s_2$ such that $k_1 + k_2 \geq 1$ and the inequality $g^{(k_1+k_2)}(\mu + \sigma z) \geq g^{(k_1+k_2)}(\mu - \sigma z)$ thus holds for all $z \geq 0$. It follows that $g^\star \in \mathcal{U}^{\mathbb{R}\times\mathbb{R}^+}_{(s_1,s_2)-\mathrm{icx}}$ provided $g \in \mathcal{U}^{\mathbb{R}}_{(s_1+s_2+1)-\mathrm{icx}}$, that is, the family $X_{(\mu,\sigma)}$ is $(s_1 + s_2 + 1, (s_1, s_2))$-increasing convex, as announced. The proof for s_2 even is similar □

Let us now consider Normally distributed random variables.

Example 4.2. If Z is Normally distributed with zero mean and unit variance then $X_{(\mu,\sigma)} = \mu + \sigma Z$ corresponds to the family of Normally distributed random variables with standard deviation $\sigma > 0$ and mean $\mu \in \mathbb{R}$. The result stated in Property 4.1 has interesting consequences for Normal mixtures, letting σ and μ become random variables S and M, respectively, possibly correlated. Consider (M_1, S_1) and (M_2, S_2) with common marginals such that (M_2, S_2) is more concordant, i.e. more positively quadrant dependent, than (M_1, S_1). This means that the stochastic inequality $(M_1, S_1) \preceq^{\mathbb{R}\times\mathbb{R}^+}_{(1,1)-\mathrm{icx}} (M_2, S_2)$ holds. Property 4.1 then gives

$$X_{(M_1,S_1)} \preceq^{\mathbb{R}}_{3-\mathrm{icx}} X_{(M_2,S_2)}. \tag{4.1}$$

Increasing the quadrant dependence between the random mean M and standard deviation S makes the resulting Normal mixture $X_{(M,S)}$ larger in the 3-increasing convex sense.

Let us now assume that we increase the mean, i.e. $M_1 \preceq^{\mathbb{R}}_{1-\mathrm{icx}} M_2$ and make the standard deviation bigger and more variable, i.e. $S_1 \preceq^{\mathbb{R}^+}_{2-\mathrm{icx}} S_2$, with joint distribution such that $(M_1, S_1) \preceq^{\mathbb{R}\times\mathbb{R}^+}_{(1,2)-\mathrm{icx}} (M_2, S_2)$ holds (which is the case if the random variables are independent, for instance). Then, Property 4.1 shows that (4.1) also holds. Making the mean more variable and the standard deviation bigger, i.e. $(M_1, S_1) \preceq^{\mathbb{R}\times\mathbb{R}^+}_{(2,1)-\mathrm{icx}} (M_2, S_2)$, gives

$$X_{(M_1,S_1)} \preceq^{\mathbb{R}}_{4-\mathrm{icx}} X_{(M_2,S_2)} \tag{4.2}$$

Finally, if both the mean and standard deviation get bigger and more variable, i.e. $(M_1, S_1) \preceq^{\mathbb{R}\times\mathbb{R}^+}_{(2,2)-\mathrm{icx}} (M_2, S_2)$, we also have (4.2).

Let us now generalize the location-scale case as follows. Let $\boldsymbol{Z}$ be a random vector valued in a subset $\mathcal{T} \subseteq \mathbb{R}^k$. Here, we consider families of random variables $X_{\boldsymbol{\theta}}$ defined as

$$X_{\boldsymbol{\theta}} = h(\boldsymbol{Z}, \boldsymbol{\theta}), \quad \theta \in \Theta \subseteq \mathbb{R}^d \tag{4.3}$$

for some function $h : \mathcal{T} \times \Theta \to \mathcal{S}$, where $\mathcal{S} \subseteq \mathbb{R}$. The following technical result, inspired from Denuit and Mesfioui (2013), appears to be useful to deal with the family $X_{\boldsymbol{\theta}}$ built from model (4.3).

Lemma 4.3. *If $h : \Theta \to \mathcal{S}$ belongs to $\mathcal{U}^{\Theta}_{(s_1,\ldots,s_d)-icx}$ and $g : \mathcal{S} \to \mathbb{R}$ belongs to $\mathcal{U}^{\mathcal{S}}_{(s_1+\cdots+s_d)-icx}$, then the composition $g \circ h$ belongs to $\mathcal{U}^{\Theta}_{(s_1,\ldots,s_d)-icx}$.*

Proof. Notice that the $(k_1, \ldots, k_d)$th cross derivative of $g\circ h$ involves lower order derivatives of h together with derivatives of g up to the $(k_1+\cdots+k_{d-1})$th order of g. The conditions imposed on h and g ensure that these cross-derivatives are non-negative so that $g \circ h \in \mathcal{U}^{\Theta}_{(s_1,\ldots,s_d)-\mathrm{icx}}$, as announced. □

We are now in a position to state the following result about the multivariate stochastic increasing convexity of the family (4.3).

Proposition 4.4. *If* $h(\boldsymbol{z},\cdot)\in\mathcal{U}^{\Theta}_{(s_1,\ldots,s_d)-icx}$ *for any fixed* $\boldsymbol{z}\in\mathcal{T}$*, then the family* $\{X_{\boldsymbol{\theta}},\ \boldsymbol{\theta}\in\Theta\}$ *defined in (4.3) is stochastically* $(s_1+\cdots+s_d,(s_1,\ldots,s_d))$*-increasing convex, that is, if* $g\in\mathcal{U}^{\mathcal{S}}_{(s_1+\cdots+s_d)-icx}$*, then* $g^\star\in\mathcal{U}^{\Theta}_{(s_1,\ldots,s_d)-icx}$*. Moreover, one has*

$$\boldsymbol{T}_1\preceq^{\Theta}_{(s_1,\ldots,s_d)-icx}\boldsymbol{T}_2\quad\Rightarrow\quad X_{\boldsymbol{T}_1}\preceq^{\mathcal{S}}_{(s_1+\cdots+s_d)-icx}X_{\boldsymbol{T}_2}.$$

Proof. By Lemma 4.3, since $h(\boldsymbol{z},\cdot)\in\mathcal{U}^{\Theta}_{(s_1,\ldots,s_d)-\mathrm{icx}}$ for any $\boldsymbol{z}\in\mathcal{T}$, then for any $g\in\mathcal{U}^{\mathcal{S}}_{(s_1+\cdots+s_d)-\mathrm{icx}}$, we have $g\circ h(\boldsymbol{z},\cdot)\in\mathcal{U}^{\Theta}_{(s_1,\ldots,s_d)-\mathrm{icx}}$. The function $g^\star$ defined as

$$g^\star(\boldsymbol{\theta})=\mathbb{E}[g\circ h(\boldsymbol{Z},\boldsymbol{\theta})]$$

thus belongs to $\mathcal{U}^{\Theta}_{(s_1,\ldots,s_d)-\mathrm{icx}}$, which proves the result. □

Example 4.5. Let Z be a random variable uniformly distributed over the unit interval $[0,1]$ and let $\boldsymbol{\theta}=(a,d)\in\mathbb{R}\times\mathbb{R}^+$. Consider the family $X_{(a,d)}=a+dZ$ of random variables uniformly distributed over the interval $[a,a+d]$. This model is a special case of (4.3) with $h(z,a,d)=a+dz$. Clearly, for any positive integers s_1 and s_2, $h(z,\cdot,\cdot)\in\mathcal{U}^{\mathbb{R}\times\mathbb{R}^+}_{(s_1,s_2)-\mathrm{icx}}$ for all $z\in[0,1]$. Proposition 4.4 then shows that the family $X_{(a,d)}$ is $(s_1+s_2,(s_1,s_2))$-increasing convex. Consequently,

$$(A_1,D_1)\preceq^{\mathbb{R}\times\mathbb{R}^+}_{(s_1,s_2)-\mathrm{icx}}(A_2,D_2)\quad\Rightarrow\quad X_{(A_1,D_1)}\preceq^{\mathbb{R}}_{(s_1+s_2)-\mathrm{icx}}X_{(A_2,D_2)}.$$

For $s_1=s_2=1$, we see that increasing the quadrant dependence in the pair (A,D) makes the mixture $X_{(A,D)}$ larger in the 2-increasing convex sense (i.e., in the usual increasing convex order). Similarly, for $s_1=1$ and $s_2=2$ (or $s_1=2$ and $s_2=1$), increasing the pair (A,D) in the $\preceq^{\mathbb{R}\times\mathbb{R}^+}_{(1,2)-\mathrm{icx}}$-sense (or $\preceq^{\mathbb{R}\times\mathbb{R}^+}_{(2,1)-\mathrm{icx}}$-sense) makes the resulting mixture $X_{(A,D)}$ larger in the 3-increasing convex order. Finally, for $s_1=s_2=2$, increasing the pair (A,D) in the orthant convex order makes the mixture $X_{(A,D)}$ larger in the 4-increasing convex sense

Example 4.6. Consider $X_{(\alpha,\beta)}=\beta\big(\exp(\alpha Z)-1\big)$ where the random variable Z obeys the Negative Exponential distribution with unit mean and $\boldsymbol{\theta}=(\alpha,\beta)\in\mathbb{R}^+\times\mathbb{R}^+$. Then,

$$\Pr[X_{(\alpha,\beta)}>x]=P\left[\alpha Z>\ln\left(1+\frac{x}{\beta}\right)\right]=\left(1+\frac{x}{\beta}\right)^{-1/\alpha}$$

is Pareto distributed. We are in a position to apply Proposition 4.4 with $h(z,\alpha,\beta)=\beta\big(\exp(\alpha z)-1\big)$. Clearly, for any positive integers s_1 and s_2, $h(z,\cdot,\cdot)\in\mathcal{U}^{\mathbb{R}^+\times\mathbb{R}^+}_{(s_1,s_2)-\mathrm{icx}}$ for any fixed $z\geq 0$ so that

$$(A_1,B_1)\preceq^{\mathbb{R}^+\times\mathbb{R}^+}_{(s_1,s_2)-\mathrm{icx}}(A_2,B_2)\quad\Rightarrow\quad X_{(A_1,B_1)}\preceq^{\mathbb{R}^+}_{(s_1+s_2)-\mathrm{icx}}X_{(A_2,B_2)}.$$

For $s_1=s_2=1$, we see that increasing the quadrant dependence in the pair (A,B) makes the Generalized Pareto mixture $X_{(A,B)}$ larger in the 2-increasing convex sense (i.e., in the usual increasing convex order). Similarly, for $s_1=1$ and $s_2=2$ (or $s_1=2$ and $s_2=1$), increasing the pair (A,B) in the $\preceq^{\mathbb{R}^+\times\mathbb{R}^+}_{(1,2)-\mathrm{icx}}$-sense (or $\preceq^{\mathbb{R}^+\times\mathbb{R}^+}_{(2,1)-\mathrm{icx}}$-sense) makes the resulting mixture $X_{(A,B)}$ larger in the 3-increasing convex order. Finally, for $s_1=s_2=2$, increasing the pair (A,B) in the upper orthant convex order makes the mixture $X_{(A,D)}$ larger in the 4-increasing convex sense.

We also give the following result, which is closely related to Proposition 4.4 . The proof directly follows from Lemma 4.3 and is therefore omitted.

Property 4.7. *Let* $\{X_{\boldsymbol{\theta}},\ \theta\in\Theta\}$ *be a stochastically* $(t,\boldsymbol{s})$*-increasing convex family of random variables. Then,* $\{h(X_{\boldsymbol{\theta}}),\ \boldsymbol{\theta}\in\Theta\}$ *is also a stochastically* $(t,\boldsymbol{s})$*-increasing convex family of random variables, for any* $(s_1+s_2+\ldots+s_d)$*-increasing convex function* h.

Example 4.8. Let Z be Normally distributed with zero mean and unit variance. Define $X_{(\mu,\sigma)}=\exp(\sigma Z+\mu)$ with $\boldsymbol{\theta}=(\mu,\sigma)\in\mathbb{R}\times\mathbb{R}^+$. From Example 4.2, we know that the family $\sigma Z+\mu$ is stochastically $(s_1+s_2+1,(s_1,s_2))$-increasing convex for s_2 odd. Now, we are in a position to apply Property 4.7 to $h(x)=\exp(x)$ which is indeed (s_1+s_2)-increasing convex whatever s_1 and s_2. This shows that the Lognormal family of distributions is stochastically $(s_1+s_2+1,(s_1,s_2))$-increasing convex in its parameters μ and σ for s_2 odd. The same reasoning shows that the Lognormal family of distributions is stochastically $(s_1+s_2,(s_1,s_2))$-increasing convex in its parameters μ and σ for s_2 even.

5 Partial and compound sums

In this section, we consider sums of independent random variables, where the number of terms plays the role of one of the parameters. In addition, we let the summands depend on another parameter controlling their size and we study the stochastic increasing convexity of the resulting family of distributions.

Proposition 5.1. *Consider* $\boldsymbol{\theta}=(n,\phi)\in\mathbb{N}\times\mathbb{R}^+$ *and define the family*

$$X_{(n,\phi)}=\sum_{i=1}^{n}Y_{i,\phi},\quad n\in\{1,2,\ldots\}\tag{5.1}$$

where $Y_{i,\phi}$ *are independent and identically distributed non-negative random variables indexed with a single parameter* $\phi\geq 0$*. If* $Y_{i,\phi_1}\preceq^{\mathcal{S}}_{1-icx}Y_{i,\phi_2}$ *holds when* $\phi_1\leq\phi_2$ *then the family* $\{X_{\boldsymbol{\theta}},\ \boldsymbol{\theta}\in\Theta\}$ *defined in* (5.1) *is stochastically* $(s+1,(s,1))$*-increasing convex for any positive integer* s*, that is,*

$$g\in\mathcal{U}^{\mathbb{R}^+}_{(s+1)-icx}\Rightarrow g^\star\in\mathcal{U}^{\mathbb{N}\times\mathbb{R}^+}_{(s,1)-icx}.$$

We then have

$$(N_1,F_1)\preceq^{\mathbb{N}\times\mathbb{R}^+}_{(s,1)-icx}(N_2,F_2)\quad\Rightarrow\sum_{i=1}^{N_1}Y_{i,F_1}\preceq^{\mathbb{R}^+}_{(s+1)-icx}\sum_{i=1}^{N_2}Y_{i,F_2}.$$

Proof. Let us first establish the validity of the announced result for $s=1$. It well known that the stochastic ordering $\preceq^{\mathcal{S}}_{1-\text{icx}}$ is closed under convolution so that the stochastic inequality $X_{n,\phi_1}\preceq^{\mathbb{R}^+}_{1-\text{icx}}X_{n,\phi_2}$ holds for any integer n when $\phi_1\leq\phi_2$. Thus, if $g\in\mathcal{U}^{\mathbb{R}^+}_{2-\text{icx}}$ then

$$g^\star(n,\phi_1)=\mathbb{E}[g(X_{n,\phi_1})]\leq\mathbb{E}[g(X_{n,\phi_2})]=g^\star(n,\phi_2)$$

so that $g^\star$ is non-decreasing in ϕ. Also, it is easy to see that $n\mapsto g^\star(n,\phi)$ is non-decreasing, because the random variables $Y_{i,\phi}$ are non-negative. Now, let us show that $(n,\phi)\mapsto g^\star(n,\phi)$

is supermodular. To this end, it suffices to show that $\phi \mapsto g^\star(n+1,\phi) - g^\star(n,\phi)$ is non-decreasing. Noting that $g \in \mathcal{U}^{\mathbb{R}^+}_{2-\text{icx}}$ ensures that $x \mapsto g(x+s) - g(x) \in \mathcal{U}^{\mathbb{R}^+}_{1-\text{icx}}$ for all $s \in \mathbb{R}^+$ and $x \mapsto g(x+t) - g(t) \in \mathcal{U}^{\mathbb{R}^+}_{2-\text{icx}}$ for all $t \in \mathbb{R}^+$, we have for all $\phi_1 \leq \phi_2$ that

$$\begin{aligned}
\Delta g^\star(n,\phi_1) &= g^\star(n+1,\phi_1) - g^\star(n,\phi_1) \\
&= \int_{\mathcal{S}} \mathbb{E}\left[g(X_{n,\phi_1}+s) - g(X_{n,\phi_1})\right] d\,\Pr[Y_{n+1,\phi_1} \leq s] \\
&\leq \int_{\mathcal{S}} \mathbb{E}\left[g(X_{n,\phi_2}+s) - g(X_{n,\phi_2})\right] d\,\Pr[Y_{n+1,\phi_1} \leq s] \\
&= \int_{\mathcal{T}} \mathbb{E}\left[g(Y_{n+1,\phi_1}+t) - g(t)\right] d\,\Pr[X_{n,\phi_2} \leq t] \\
&\leq \int_{\mathcal{T}} \mathbb{E}\left[g(Y_{n+1,\phi_2}+t) - g(t)\right] d\,\Pr[X_{n,\phi_2} \leq t] \\
&= \Delta g^\star(n,\phi_2).
\end{aligned}$$

The announced result thus holds for $s = 1$. Now, let us show by induction that the result also holds for all $s \geq 2$. To this end, assume that the result is valid for some positive integer s_0 and let us then show that it also holds for $s_0 + 1$. Consider $g \in \mathcal{U}^{\mathbb{R}^+}_{(s_0+2)-\text{icx}}$. Then one observes that

$$\Delta g^\star(n,\phi) = \mathbb{E}\left[g(X_{n+1,\phi})\right] - \mathbb{E}\left[g(X_{n,\phi})\right] = \mathbb{E}[\tilde{g}(X_{n,\phi})]$$

where $\tilde{g}(t) = \mathbb{E}[g(Y_{n+1,\phi}+t) - g(t)]$. It is clear that $\tilde{g} \in \mathcal{U}^{\mathbb{R}^+}_{(s_0+1)-\text{icx}}$ because $g \in \mathcal{U}^{\mathbb{R}^+}_{(s_0+2)-\text{icx}}$. By induction, the function $(n,\phi) \mapsto \mathbb{E}[\tilde{g}(X_{n,\phi})]$ belongs to $\mathcal{U}^{\mathbb{N}\times\mathbb{R}^+}_{(s_0,1)-\text{icx}}$. Thus $\Delta g^\star \in \mathcal{U}^{\mathbb{N}\times\mathbb{R}^+}_{(s_0,1)-\text{icx}}$, that is, $g^\star \in \mathcal{U}^{\mathbb{N}\times\mathbb{R}^+}_{(s_0+1,1)-\text{icx}}$, which ends the proof. □

Hereafter, we present some illustrations of Proposition 5.1.

Example 5.2. Let $X_{(n,\phi)}$ be the Erlang family of distributions with parameters $n \in \mathbb{N}$ and $\phi > 0$. This model can be written in the form (5.1) with $Y_{i,\phi}$ independent and Exponentially distributed with mean ϕ. Clearly, $Y_{i,\phi_1} \preceq^{\mathbb{N}}_{1-\text{icx}} Y_{i,\phi_2}$ holds for all $\phi_1 \leq \phi_2$. Thus, from Proposition 5.1, we have

$$(N_1, F_1) \preceq^{\mathbb{N}\times\mathbb{R}^+}_{(s,1)-\text{icx}} (N_2, F_2) \qquad \Rightarrow \qquad X_{(N_1,F_1)} \preceq^{\mathbb{R}^+}_{(s+1)-\text{icx}} X_{(N_2,F_2)}$$

for any integer $s \geq 1$. In particular, increasing the quadrant dependence between K and F (in the case $s = 1$) makes the mixture of Erlang distributions larger in the 2-convex sense (i.e. in the usual convex order).

Example 5.3. Let $X_{(n,\phi)}$ obey the Negative Binomial distribution with integer parameter n and failure probability ϕ. This model is also a particular case of (5.1), with $Y_{i,\phi}$ Geometrically distributed with failure probability ϕ. Clearly, $Y_{i,\phi_1} \preceq^{\mathbb{N}}_{1-\text{icx}} Y_{i,\phi_2}$ for all $\phi_1 \leq \phi_2$. Consequently, Proposition 5.1 leads to

$$(N_1, F_1) \preceq^{\mathbb{N}\times[0,1]}_{(s,1)-\text{icx}} (N_2, F_2) \quad \Rightarrow \sum_{i=1}^{N_1} Y_{i,F_1} \preceq^{\mathbb{N}}_{(s+1)-\text{icx}} \sum_{i=1}^{N_2} Y_{i,F_2}$$

for any integer $s \geq 1$. Considering $s = 1$, this shows that increasing the quadrant dependence between N and F makes the Negative Binomial distribution smaller in the 2-convex sense (i.e. in the usual convex order).

In some cases, more general order relations can be established, as shown in the next example for the Binomial distribution.

Example 5.4. Let $X_{(n,p)}$ be distributed according to the Binomial distribution with exponent n and success probability p. In this case, $X_{(n,p)}$ can be represented as a sum of n independent Bernoulli random variables with common mean p so that Proposition 5.1 applies. It is nevertheless possible to get a stronger result in this particular situation, as shown next. Let $\boldsymbol{\theta} = (n, p)$, $n \in \{1, 2, \ldots\}$ and $p \in [0, 1]$. With $g^\star(n, p) = \mathbb{E}[g(X_{n,p})]$, formula (A.3) in appendix shows that

$$g \in \mathcal{U}^{\mathbb{N}}_{2-\mathrm{icx}} \Rightarrow g^\star \in \mathcal{U}^{\mathbb{N}\times[0,1]}_{(1,1)-\mathrm{icx}}.$$

It follows that

$$(N_1, P_1) \preceq^{\mathbb{N}\times[0,1]}_{(1,1)-\mathrm{icx}} (N_2, P_2) \quad \Rightarrow \quad X_{(N_1,P_1)} \preceq^{\mathbb{N}}_{2-\mathrm{icx}} X_{(N_2,P_2)}.$$

Now, we will show by induction that the implication

$$g \in \mathcal{U}^{\mathbb{N}}_{(s_1+s_2)-\mathrm{icx}} \Rightarrow g^\star \in \mathcal{U}^{\mathbb{N}\times[0,1]}_{(s_1,s_2)-\mathrm{icx}}$$

holds true for any positive integers s_1 and s_2. Assume that the result is valid for some positive integers (s_1, s_2) and let us establish that it then also holds true for $(s_1 + 1, s_2)$ and $(s_1, s_2 + 1)$. Consider $g \in \mathcal{U}^{\mathbb{N}}_{(s_1+s_2+1)-\mathrm{icx}}$. Then, $\Delta g \in \mathcal{U}^{\mathbb{N}}_{(s_1+s_2)-\mathrm{icx}}$. By induction, we see that the functions $(n, p) \mapsto \mathbb{E}[\Delta g(X_{n,p})]$ and $(n, p) \mapsto \mathbb{E}[\Delta g(X_{n-1,p})]$ both belong to the set $\mathcal{U}^{\mathbb{N}\times[0,1]}_{(s_1,s_2)-\mathrm{icx}}$. Therefore, from formulas (A.1) and (A.2) in appendix, we have that $\partial g^\star/\partial p$ and $\Delta g^\star$ defined as $\Delta g^\star(n, p) = g^\star(n + 1, p) - g^\star(n, p)$ both belong to $\mathcal{U}^{\mathbb{N}\times[0,1]}_{(s_1,s_2)-\mathrm{icx}}$ so that $g^\star \in \mathcal{U}^{\mathbb{N}\times[0,1]}_{(s_1,s_2+1)-\mathrm{icx}} \cap \mathcal{U}^{\mathbb{N}\times[0,1]}_{(s_1+1,s_2)-\mathrm{icx}}$. Hence, the result is true for $(s_1 + 1, s_2)$ and $(s_1, s_2 + 1)$, which completes the proof. As a consequence we have

$$(N_1, P_1) \preceq^{\mathbb{N}\times[0,1]}_{(s_1,s_2)-\mathrm{icx}} (N_2, P_2) \quad \Rightarrow \quad X_{(N_1,P_1)} \preceq^{\mathbb{N}}_{(s_1+s_2)-\mathrm{icx}} X_{(N_2,P_2)}.$$

where N_1 and N_2 are integer-valued random variables such that $N_i \geq s_i$ holds for $i = 1, 2$ and where P_1 and P_2 are random variables valued in the unit interval $[0, 1]$.

In particular, for $s_1 = s_2 = 1$, we see that strengthening the positive quadrant dependence between the number of trials N and the success probability P increases the Binomial mixture in the increasing convex order. Increasing (N, P) in the $\preceq^{\mathbb{N}\times[0,1]}_{(1,2)-\mathrm{icx}}$-sense or in the $\preceq^{\mathbb{N}\times[0,1]}_{(2,1)-\mathrm{icx}}$-sense makes the Binomial mixture larger in the $\preceq^{\mathbb{N}}_{3-\mathrm{icx}}$-sense. Finally, increasing (N, P) in the orthant convex order $\preceq^{\mathbb{N}\times[0,1]}_{(2,2)-\mathrm{icx}}$ increases the Binomial mixture in the $\preceq^{\mathbb{N}}_{4-\mathrm{icx}}$-sense.

The examples discussed above suggest to study compound sum families where the number of terms is no more fixed but becomes a random variable. As in Belzunce et al. (2006), we assume that (i) the common distribution of the independent summands is indexed by a single parameter (denoted as ϕ below) and (ii) the distribution for the number of terms is indexed by a single parameter (denoted as λ below).

Property 5.5. *Let $Y_{1,\phi}, Y_{2,\phi}, \ldots$ be a sequence of non-negative independent and identically distributed random variables indexed with a single parameter ϕ such that $Y_{i,\phi_1} \preceq^{\mathcal{S}}_{1-icx} Y_{i,\phi_2}$ when $\phi_1 \leq \phi_2$. Let $\{N_\lambda,\ \lambda \in \Lambda\}$ be a family of non-negative integer-valued random variables* *indexed with a parameter* λ*, independent of the* $Y_{i,\phi}$*'s. Let* $\{F_\mu,\ \ \mu \in \Gamma\}$ *be a family of random variables valued in a subset* $\mathcal{T} \subseteq \mathbb{R}$ *and independent of the* $Y_{i,\phi}$*'s. Let us now consider the family of random variables*

$$X_{\boldsymbol{\theta}} = \sum_{i=1}^{N_\lambda} Y_{i,F_\mu}$$

where $\boldsymbol{\theta} = (\lambda, \mu) \in \Theta = \Lambda \times \Gamma$*. If* $\{(N_\lambda, F_\mu),\ \ (\lambda,\mu) \in \Theta\}$ *is stochastically* $(s,1)$*-increasing convex, in the sense that* $h^\star(\lambda,\mu) = \mathbb{E}[h(N_\lambda, F_\mu)]$ *belongs to* $\mathcal{U}^{\Theta}_{(s,1)-icx}$ *whenever* $h \in \mathcal{U}^{\mathbb{N}\times\mathcal{T}}_{(s,1)-icx}$*. Then,*

$$g \in \mathcal{U}^{\mathbb{R}^+}_{(s+1)-icx} \Rightarrow g^\star \in \mathcal{U}^{\Theta}_{(s,1)-icx},$$

that is, $\{X_{\boldsymbol{\theta}},\ \ \boldsymbol{\theta} \in \Theta\}$ *is stochastically* $(s+1,(s,1))$*-increasing convex.*

Proof. Let $g \in \mathcal{U}^{\mathbb{R}^+}_{(s+1)-\text{icx}}$ and set

$$\widetilde{g}^\star(n,\phi) = \mathbb{E}\left[g\left(\sum_{i=1}^{n} Y_{i,\phi}\right)\right].$$

From Proposition 5.1, we see that $\widetilde{g}^\star \in \mathcal{U}^{\mathbb{N}\times\mathcal{T}}_{(s,1)-\text{icx}}$. Moreover, one has

$$\begin{aligned} g^\star(\lambda,\mu) &= \mathbb{E}[g(X_{\boldsymbol{\theta}})] \\ &= \mathbb{E}\Big[\mathbb{E}[g(X_{\boldsymbol{\theta}})|N_\lambda, F_\mu]\Big] \\ &= \mathbb{E}[\widetilde{g}^\star(N_\lambda, F_\mu)] \end{aligned}$$

which implies that $g^\star \in \mathcal{U}^{\Theta}_{(s,1)-\text{icx}}$. □

Property 5.5 covers the case where F_μ is degenerated at μ, i.e. $\Pr[F_\mu = \mu] = 1$, and N_λ is stochastically s-increasing convex in λ.

Example 5.6. Property 5.5 applies to compound Poisson distributions, i.e. when N_λ is Poisson distributed with mean λ. Assume that F_μ is degenerated at μ and consider $X_{(\lambda,\mu)}$ equal to the sum of N_λ independent random variables, distributed as $Y_{1,\mu}$ such that $Y_{1,\mu_1} \preceq^{\mathbb{R}^+}_{1-\text{icx}} Y_{1,\mu_2}$ when $\mu_1 \leq \mu_2$. In case of partial knowledge about the expected number of terms λ and about their expected value μ, it is common to let them become random variables L and M, say, possibly correlated (as the average size of each term may depend on their expected number). For instance, M and L may represent the effect of a common environment influencing both the number of terms and the size of each of them. As the Poisson distribution is known to be stochastically s-increasing convex after Denuit et al. (2000), we have

$$(L_1, M_1) \preceq^{\mathbb{R}^+\times\mathbb{R}^+}_{(s,1)-\text{icx}} (L_2, M_2) \Rightarrow X_{(L_1,M_1)} \preceq^{\mathbb{R}^+}_{(s+1)-\text{icx}} X_{(L_2,M_2)}.$$

For $s = 1$, we see that increasing the intensity of positive quadrant dependence between L and M makes the resulting mixture $X_{(L,M)}$ larger in the increasing convex order. Increasing s makes the resulting mixture $X_{(L,M)}$ larger in higher-degree increasing convex orders.

Acknowledgements

Michel Denuit acknowledges the financial support from the contract "Projet d'Actions de Recherche Concertées" No 12/17-045 of the "Communauté française de Belgique", granted by the "Académie universitaire Louvain". Mhamed Mesfioui acknowledges the financial support of the Natural Sciences and Engineering Research Council of Canada.

References

Belzunce, F., Ortega, E.-M., Pellerey, F., Ruiz, J.M. (2006). Variability of total claim amounts under dependence between claims severity and number of events. Insurance: Mathematics and Economics 38, 460-468.

Chang, C-S., Chao, X.L., Pinedo, M., Shanthikumar, J. G. (1991). Stochastic convexity for multidimensional processes and applications. IEEE Transactions on Automated Control 36, 1347-1355.

Denuit, M. (2010). Positive dependence of signals. Journal of Applied Probability 47, 893-897.

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

Denuit, M., Lefevre, 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., Lefevre, Cl. (2001). Stochastic s-(increasing) convexity. In Generalized Convexity and Generalized Monotonicity, pp.167-182, N. Hadjisavvas, J.E. Martinez-Legaz & J.-P. Penot Editors, Lecture Notes in Economics and Mathematical Systems 502, Springer, New York. (Proceedings of the 6th International Symposium on Generalized Convexity / Monotonicity, Karlovassi, Samos, Greece, August 30 - September 3 1999)

Denuit, M., Lefevre, Cl., Mesfioui, M. (1999a). A class of bivariate stochastic orderings with applications in actuarial sciences. Insurance: Mathematics and Economics 24, 31-50.

Denuit, M., Lefevre, Cl., Mesfioui, M. (1999b). Stochastic orderings of convex-type for discrete bivariate risks. Scandinavian Actuarial Journal, 32-51.

Denuit, M., Lefevre, Cl., Mesfioui, M. (1999c). On s-convex stochastic extrema for arithmetic risks. Insurance: Mathematics and Economics 25, 143-155.

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

Denuit, M., Lefevre, Cl., Shaked, M. (2000). Stochastic convexity of the Poisson mixture model, with applications in actuarial sciences. Methodology and Computing in Applied Probability 2, 231-254.

Denuit, M., Lefevre, Cl., Utev, S. (1999). Generalized stochastic convexity and stochastic ordering of mixtures. Probability in the Engineering and Informational Sciences 13, 275-291.

Denuit, M., Mesfioui, M. (2010). Generalized increasing convex and directionally convex orders. Journal of Applied Probability 47, 264-276.

Denuit, M., Mesfioui, M. (2013). Ordering functions of random vectors, with applications to partial sums. Journal of Theoretical Probability, in press.

Fernandez-Ponce, J.M., Ortega, E.M., Pellerey, F. (2008). Convex comparisons for random sums in random environments and applications. Probability in the Engineering and Informational Sciences 22, 389-413.

Meester, L.E., Shanthikumar, J.G. (1993). Regularity of stochastic processes. A theory based on directional convexity. Probability in the Engineering and Informational Sciences 7, 343-360.

Shaked, M., Shanthikumar, J.G. (1988a). Stochastic convexity and its applications. Advances in Applied Probability 20, 427-446.

Shaked, M., Shanthikumar, J.G. (1988b). Temporal stochastic convexity and concavity. Stochastic Processes and Their Applications 27, 1-20.

Shaked, M., Shanthikumar, J.G. (1990). Parametric stochastic convexity and concavity of stochastic processes. Annals of the Institute of Statistical Mathematics 42, 509-531.

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

A Some results for the Binomial distribution

Property A.1. *Let $\{X_{n,p}, n \in \mathbb{N}, p \in (0,1)\}$ be the family of Binomial distribution with mean np and variance $np(1-p)$. With $g^\star(n,p) = \mathbb{E}[g(X_{n,p})]$, define*

$$\Delta g^\star(n,p) = g^\star(n+1,p) - g^\star(n,p).$$

Then, we have

$$\begin{aligned}
\frac{\partial g^\star}{\partial p}(n,p) &= n\mathbb{E}\left[\Delta g(X_{n-1,p})\right] && \text{(A.1)}\\
\Delta g^\star(n,p) &= p\mathbb{E}[\Delta g(X_{n,p})] && \text{(A.2)}\\
\Delta \frac{\partial g^\star}{\partial p}(n,p) &= (n+1)p\mathbb{E}[\Delta^2 g(X_{n-1,p})] + \mathbb{E}[\Delta g(X_{n-1,p})]. && \text{(A.3)}
\end{aligned}$$

Proof. We have

$$\begin{aligned}
\frac{\partial g^\star}{\partial p}(n,p) &= \sum_{i=1}^{n-1} g(i) i \binom{n}{i} p^{i-1}(1-p)^{n-i} - \sum_{i=1}^{n-1} g(i)(n-i)\binom{n}{i} p^i (1-p)^{n-i-1}\\
&\quad + ng(n)p^{n-1} - ng(0)(1-p)^{n-1}
\end{aligned}$$

Using the well-known identities

$$i\binom{n}{i} = n\binom{n-1}{i-1} \quad \text{and} \quad (n-i)\binom{n}{i} = n\binom{n-1}{i}$$

we can write

$$\begin{aligned}
\frac{\partial g^\star}{\partial p}(n,p) &= \sum_{i=1}^{n} g(i) n \binom{n-1}{i-1} p^{i-1}(1-p)^{n-i} - \sum_{i=0}^{n-1} g(i) n \binom{n-1}{i} p^i(1-p)^{n-i-1}\\
&= \sum_{i=0}^{n-1} g(i+1) n \binom{n-1}{i} p^i (1-p)^{n-i-1} - \sum_{i=0}^{n-1} g(i) n \binom{n-1}{i} p^i (1-p)^{n-i-1}\\
&= n \sum_{i=0}^{n-1} \Delta g(i) \binom{n-1}{i} p^i (1-p)^{n-i-1}\\
&= n\mathbb{E}\left[\Delta g(X_{n-1,p})\right]
\end{aligned}$$

which ends the proof of (A.1). Now, let us show that the second equality is valid. Clearly, one has

$$\Delta g^\star(n,p) = \sum_{i=0}^{n+1} g(i)\binom{n+1}{i} p^i (1-p)^{n+1-i} - \sum_{i=0}^{n} g(i)\binom{n}{i} p^i (1-p)^{n-i}.$$

Since

$$\binom{n+1}{i} = \binom{n}{i} + \binom{n}{i-1}, \quad 1 \le i \le n,$$

we can write

$$
\begin{aligned}
\Delta g^\star(n,p) &= g(0)(1-p)^{n+1} - g(0)(1-p)^n + g(n+1)p^{n+1} \\
&\quad + \sum_{i=1}^{n} g(i)\left(\binom{n}{i} + \binom{n}{i-1}\right) p^i(1-p)^{n+1-i} - \sum_{i=1}^{n} g(i)\binom{n}{i}p^i(1-p)^{n-i} \\
&= g(0)(1-p)^{n+1} - g(0)(1-p)^n + g(n+1)p^{n+1} + \sum_{i=1}^{n} g(i)\binom{n}{i-1}p^i(1-p)^{n+1-i} \\
&\quad - \sum_{i=1}^{n} g(i)\binom{n}{i}p^{i+1}(1-p)^{n-i} \\
&= \sum_{i=1}^{n+1} g(i)\binom{n}{i-1}p^i(1-p)^{n+1-i} - \sum_{i=0}^{n} g(i)\binom{n}{i}p^{i+1}(1-p)^{n-i} \\
&= \sum_{i=0}^{n} g(i+1)\binom{n}{i}p^{i+1}(1-p)^{n-i} - \sum_{i=0}^{n} g(i)\binom{n}{i}p^{i+1}(1-p)^{n-i} \\
&= p\sum_{i=0}^{n} \big(g(i+1) - g(i)\big)\binom{n}{i}p^i(1-p)^{n-i} \\
&= p\sum_{i=0}^{n} \Delta g(i)\binom{n}{i}p^i(1-p)^{n-i} \\
&= p\mathbb{E}[\Delta g(X_{n,p})].
\end{aligned}
$$

Finally, let us establish the last formula (A.3). First, recall that for any functions h_1 and $h_2 : \mathbb{N} \to \mathbb{R}$, one has

$$\Delta\big(h_1(n)h_2(n)\big) = h_1(n+1)\Delta h_2(n) + h_2(n)\Delta h_1(n).$$

Considering $h_1(n) = n$ and $h_2(n) = \mathbb{E}\left[\Delta g(X_{n-1,p})\right] = (\Delta g)^\star(n-1,p)$, we get

$$
\begin{aligned}
\Delta\frac{\partial g^\star}{\partial p}(n,p) &= \Delta\big(h_1(n)h_2(n)\big) \\
&= (n+1)\Delta(\Delta g)^\star(n-1,p) + \mathbb{E}\left[\Delta g(X_{n-1,p})\right] \\
&= (n+1)p\mathbb{E}[\Delta^2 g(X_{n-1,p})] + \mathbb{E}\left[\Delta g(X_{n-1,p})\right]
\end{aligned}
$$

so that (A.3) is indeed valid. □