On s-convex stochastic extrema for arithmetic risks

Michel Denuit*, Claude Lefèvre, M'hamed Mesfioui

Institut de Statistique et de Recherche Opérationnelle, Université Libre de Bruxelles, CP 210,
Boulevard du Triomphe, B-1050 Bruxelles, Belgium

Received 1 December 1998; received in revised form 1 May 1999

Abstract

Recently, Denuit and Lefèvre (Insurance: Mathematics and Economics 20 (1997) 197–213) have introduced a class of
discrete s-convex stochastic orderings for comparing arithmetic risks in actuarial sciences inter alia. The present paper is
concerned with the construction of the extremal distributions with respect to these orderings. Firstly, the general problem of
bounding such risks is studied in some details. Then, improved extrema are obtained for the case where the risks are known
to have a decreasing density function. For illustration, the results are applied to derive bounds for the probability of ruin in
the compound binomial risk model. ©1999 Elsevier Science B.V. All rights reserved.

Keywords: Discrete stochastic orderings; s-Convex orderings; Extremal distributions; Moment spaces; Decreasing density functions; Binomial risk model

1. Introduction

The theory of stochastic orderings generates considerable interest in actuarial sciences, for theoretical and practical
purposes. A number of actuarial applications and related topics are given in the comprehensive books by De Vylder
(1996), Goovaerts et al. (1990), Kaas et al. (1994) and Hürlimann (1998). We also refer the reader to the nice book
by Shaked and Shanthikumar (1994) for a multidisciplinary approach of the theory and its applications.

Quite recently, various discrete s-convex stochastic orderings have been introduced for comparing random variables (such
as actuarial risks) that are discrete by nature. A remarkable class investigated by Denuit and Lefèvre (1997a) is the
class of s-convex orderings among arithmetic random variables valued in some set \( D_n = \{0, 1, \ldots, n\} \), \( n \in \mathbb{R} \). Here
\( s \) is any positive integer less than or equal to \( n \). The present paper will precisely focus on this class of stochastic
orderings.

It is worth mentioning that these orderings have been generalized by Denuit et al. (1999e), using the concept of
divided difference operator, to compare any pair of discrete random variables. The continuous version of s-convex
orderings among continuous random variables has been studied in Denuit et al. (1998, 1999b). The bivariate

---

* Corresponding author. Present address: Institut de Statistique, Université Catholique de Louvain, Voie du Roman Pays 20, B-1348
Louvain-La-Neuve, Belgium. Tel.: +32-10-472835; fax: +32-10-473032
E-mail address: denuit@stat.ucl.ac.be (M. Denuit)

0167-6687/99/$ – see front matter ©1999 Elsevier Science B.V. All rights reserved.
PII: S0167-6687(99)00030-X
extension for comparing bivariate random vectors has been discussed by Denuit et al. (1999c) for the discrete case and by Denuit and Lefèvre (1997b) and Denuit et al. (1999d) for the continuous case.

Our purpose is to obtain the minimum and the maximum in the stochastic $s$-convex sense for random variables valued in $\mathcal{D}_n$. To begin with, we give in Section 2 some basic notions about the orderings. In particular, we recall that only random variables with identical first $s-1$ moments can be compared in the $s$-convex sense. This leads us to examine in Section 3 the associated moment space of all the random variables valued in $\mathcal{D}_n$ and with prescribed $s-1$ moments. Section 4 is devoted to the construction of extrema in this space. The cases $s = 1, 2, 3$ have been treated in Denuit and Lefèvre (1997a), and the general methodology outlined is used to derive the maximum for $s = 4$. In Section 5, we show how the above extrema can be improved when it is further assumed that the random variables have a decreasing density function. The key tool is a convenient representation theorem of such distributions provided by Khinchine’s theorem. Explicit expressions of the extrema are given for decreasing random variables have a density function. The key tool is a convenient representation theorem of such distributions provided by Khinchine’s theorem. Explicit expressions of the extrema are given for decreasing random variables having a density function.

It is worth emphasizing that a number of other applications can be developed in actuarial sciences. In particular, the discrete $s$-convex extrema allow us to construct stochastic bounds on a claim frequency distribution when only some of its moments are known. This leads us in a natural way to the important question of bounding the aggregate claims of an homogeneous portfolio. As shown in Denuit et al. (1999f), orderings on the claim frequency is transmitted to similar orderings on the aggregate claims when the claim amounts are independent and identically distributed.

The more realistic situation with dependent claim severities is rather delicate; when the number of claims is fixed, bounds were derived (Denuit et al., 1999a; Dhaene and Goovaerts, 1996). The question with a random number of claims, possibly correlated with the claim amounts, is a difficult task not yet studied, but is of great relevance for practical purposes.

2. Arithmetic $s$-convex stochastic orderings

Hereafter, we briefly recall some basic notions about the orderings under interest. A rather thorough study can be found in Denuit and Lefèvre (1997a).

As announced, random variables are assumed to take on values in the state space $\mathcal{D}_n = \{0, 1, \ldots, n\}$, for some $n \in \mathbb{N}$. Given a random variable $X$, we denote by $F_X$ its distribution function, by $S_X = 1 - F_X$ its survival function and by $f_X$ its (discrete) density function (i.e. $f_X(j) = P(X = j)$, $j \in \mathcal{D}_n$).

Now, let $s$ be any fixed positive integer in $\mathcal{D}_n$. We denote by $\Delta$ the usual forward difference operator, which is defined, for any function $\phi : \mathcal{D}_n \to \mathbb{R}$, by $\Delta \phi(i) = \phi(i + 1) - \phi(i)$ for all $i \in \mathcal{D}_n$ such that $i + 1 \in \mathcal{D}_n$ (i.e. $i \in \mathcal{D}_{n-1}$). Let $\Delta^k$, $k \in \mathcal{D}_n$, be the $k$th order forward difference operator defined recursively by $\Delta^0 \phi = \phi$ and, when $k \geq 1$, $\Delta^k \phi(i) = \Delta^{k-1} \phi(i + 1) - \Delta^{k-1} \phi(i)$ for all $i \in \mathcal{D}_n$ such that $i + k \in \mathcal{D}_n$ (i.e. $i \in \mathcal{D}_{n-k}$).

Definition 2.1. $X$ and $Y$ being two random variables valued in $\mathcal{D}_n$, $X$ is said to be smaller than $Y$ in the $s$-convex sense (written as $X \preceq_{s-convex} Y$) if $E \phi(X) \leq E \phi(Y)$ for all $s$-convex real functions $\phi$ on $\mathcal{D}_n$, that is within the class

$$\mathcal{U}_{s-convex}^{\mathcal{D}_n} = \{ \phi : \mathcal{D}_n \to \mathbb{R} | \Delta^k \phi(i) \geq 0, \text{ for all } i \in \mathcal{D}_{n-s} \}. \tag{2.1}$$

The $s$-convex ordering can be characterized in terms of the iterated right-tail distributions. Specifically, $X$ being a random variable valued in $\mathcal{D}_n$, let $F_0(X, j) = f_X(j)$, $j \in \mathcal{D}_n$, and define recursively, for $0 \leq k \leq n - 1$,

$$F_{k+1}(X, j) = \sum_{i=j}^{n} F_k(X, i), \quad j \in \mathcal{D}_n. \tag{2.2}$$

These iterated right tail distributions correspond to some factorial moments, namely...
\[ F_k(X, j) = E \left( \frac{X - j + k - 1}{k - 1} \right), \quad j \in \mathcal{D}_n, \ 1 \leq k \leq n, \]  

(2.3)

with the convention that \( \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \end{pmatrix} = 0 \) if \( \epsilon_1 < \epsilon_2 \).

**Characterization 2.2.** \( X \preceq \mathcal{D}_n Y \) if, and only if,

\[
\begin{align*}
EX^k &= EY^k, & k &= 1, \ldots, s - 1, \\
F_s(X, k) &\leq F_s(Y, k), & k &= s, \ldots, n.
\end{align*}
\]

(2.4)

An immediate consequence of (2.4) is that only random variables with identical first \( s - 1 \) moments can be compared in the \( s \)-convex sense. We denote by \( \mathcal{B}_s(\mathcal{D}_n; \mu_1, \mu_2, \ldots, \mu_{s-1}) \) the moment space of all the random variables valued in \( \mathcal{D}_n \) and with prescribed first \( s - 1 \) moments \( k \in \mathcal{E}X_k \), \( 1 \leq k \leq s - 1 \).

### 3. Geometry of moment spaces

In this section, we are going to enlighten the structure of the moment space \( \mathcal{B}_s(\mathcal{D}_n; \mu_1, \mu_2, \ldots, \mu_{s-1}) \), at least partly, with respect to the ordering \( \preceq \mathcal{D}_n \). This will be of interest for the construction of \( s \)-convex extrema.

Given a function \( \phi : \mathcal{D}_n \rightarrow \mathbb{R} \), the number of sign-changes of \( \phi \) on \( \mathcal{D}_n \) is defined as

\[ S^-(\phi) = \sup S^-[\phi(x_1), \phi(x_2), \ldots, \phi(x_k)], \]

(3.1)

where the supremum is extended over all sets \( \{x_1 < x_2 < \cdots < x_k\} \subseteq \mathcal{D}_n \), and \( S^-[y_1, y_2, \ldots, y_k] \) is the number of sign-changes of the sequence \( \{y_1, y_2, \ldots, y_k\} \), zero terms being discarded.

We start with a preliminary result stating that the difference of two density or distribution functions in \( \mathcal{B}_s(\mathcal{D}_n; \mu_1, \mu_2, \ldots, \mu_{s-1}) \) necessarily exhibits \( s \) or \( s - 1 \) changes of sign, respectively.

**Lemma 3.1.** Let \( X \) and \( Y \in \mathcal{B}_s(\mathcal{D}_n; \mu_1, \mu_2, \ldots, \mu_{s-1}) \), with \( F_X \neq F_Y \). Then, (i) \( S^-(f_X - f_Y) \geq s \) and (ii) \( S^-(F_X - F_Y) \geq s - 1 \).

**Proof.** The implication (i) has been established in Lemma 4.2 of Denuit and Lefèvre (1997a). Let us turn to the implication (ii). Here and later, we will use the classical *Abel transformation formula*, namely, given two sequences of real numbers \( \{a_0, a_1, \ldots, a_v\} \) and \( \{b_0, b_1, \ldots, b_v\} \) with \( v \geq 1 \), then

\[
\sum_{j=0}^{v-1} a_j b_j = a_v \sum_{j=0}^{v-1} b_j - \sum_{j=0}^{v-1} \left( a_{j+1} - a_j \right) \sum_{\ell=0}^{j} b_{\ell}. \quad (3.2)
\]

By (3.2), we see that for \( k = 1, 2, \ldots, n \),

\[
E \left( \frac{X}{k} \right) - E \left( \frac{Y}{k} \right) = - \sum_{j=0}^{n} \left( \frac{j}{k - 1} \right) (F_X - F_Y)(j). \quad (3.3)
\]

Since \( X \) and \( Y \) belong to the same moment space, the left-hand side member of (3.3) is equal to 0 for \( k = 1, 2, \ldots, s - 1 \). Therefore, we obtain that the right-hand side member of (3.3) reduces to 0 for \( k = 1, 2, \ldots, s - 1 \), which yields

\[
\sum_{j=0}^{n} j^k (F_X - F_Y)(j) = 0, \quad k = 0, 1, \ldots, s - 2,
\]
and thus
\[ \sum_{j=0}^{n} Q(j)(F_X - F_Y)(j) = 0 \quad (3.4) \]
for any polynomial \( Q \) of degree at most \( s - 2 \). Now, proceeding by absurd, let us assume that \( S^{-}(F_X - F_Y) \leq s - 2 \). Then, we may choose for \( Q \) in (3.4) the polynomial with single roots exactly at the roots of \( F_X - F_Y \). By construction, we have that
\[ Q(j)(F_X - F_Y)(j) \geq 0 \quad \text{for all} \quad j \in \mathcal{D}_n. \quad (3.5) \]
From (3.4) and (3.5), we thus deduce that \( Q(j)(F_X - F_Y)(j) = 0 \) for all \( j \in \mathcal{D}_n \), but this is in contradiction with the hypothesis \( F_X \neq F_Y \).

We now give a criterion on the distribution or density functions that allows us to compare two random variables in the \( s \)-convex sense. \( X \) and \( Y \) being two random variables valued in \( D_n \), we say that \( F_X \preceq F_Y \) (resp. \( f_X \preceq f_Y \)) near \( n \) if for some \( k \leq n - 1 \),
\[ F_X(j) = F_Y(j) \quad \text{for all} \quad j \leq k. \]

**Proposition 3.2.** Let \( X \) and \( Y \in B_s(D_n; \mu_1, \mu_2, \ldots, \mu_{s-1}) \). Then, the condition (i) \( S^{-}(F_X - F_Y) = s - 1 \) together with \( F_X \preceq F_Y \) near \( n \), or the condition (ii) \( S^{-}(f_X - f_Y) = s - 1 \) together with \( f_X \preceq f_Y \) near \( n \), imply that \( X \preceq_s Y \).

**Proof.** In Proposition 4.4 of Denuit and Lefèvre (1997a), it is established that within \( B_s(D_n; \mu_1, \mu_2, \ldots, \mu_{s-1}) \), \( S^{-}(F_X - F_Y) = s - 1 \) together with \( F_X \preceq F_Y \) near \( n \) implies \( X \preceq_s Y \). By Lemma 3.1, however, we see that the condition (i) can be substituted for (ii), hence (i). The condition (ii) then follows easily, as indicated in Corollary 4.5 of Denuit and Lefèvre (1997a).

The next result states that in \( B_s(D_n; \mu_1, \mu_2, \ldots, \mu_{s-1}) \), two random variables ordered in the \( s \)-convex sense and which have the same \( s \)th moment, are identically distributed. This shows \textit{inter alia} that \( \preceq_s \) constitutes a partial order relation in \( B_s(D_n; \mu_1, \mu_2, \ldots, \mu_{s-1}) \).

**Proposition 3.3.** Let \( X \) and \( Y \in B_s(D_n; \mu_1, \mu_2, \ldots, \mu_{s-1}) \). If \( X \preceq_s Y \) and \( E X^s = E Y^s \), then \( X \) and \( Y \) have the same distribution.

**Proof.** Let us first establish that
\[ f_X(j) = f_Y(j) \quad \text{for} \quad j \geq s. \quad (3.6) \]
From (2.3), we have
\[ \sum_{k=s}^{n} F_s(X, k) = \sum_{k=s}^{n} \sum_{j=k}^{n} \binom{j-k+s-1}{s-1} f_X(j) = \sum_{j=s}^{n} f_X(j) \sum_{k=0}^{j-s} \binom{k+s-1}{s-1} \]
and by the well-known formula
\[ \sum_{\ell=m_1}^{m_2} \binom{\ell}{m_1} = \binom{m_2+1}{m_1+1}, \]
we get
\[ \sum_{k=s}^{n} F_s(X, k) = E \binom{X}{s}. \quad (3.7) \]
Thus, for $X, Y \in B_s(D_n; \mu_1, \mu_2, \ldots, \mu_{s-1})$ with $EX^s = EY^s$, we obtain from (3.7) that
\[ \sum_{k=x}^{n} [F_s(Y, k) - F_s(X, k)] = E\left(\frac{Y}{s}\right) - E\left(\frac{X}{s}\right) = 0. \] (3.8)

But since $X \preceq_{s-cx} Y$, we know by the second part of (2.4) that the left-hand side member of (3.8) is non-negative. Therefore (3.8) yields
\[ F_s(Y, k) = F_s(X, k) \quad \text{for } k \geq s. \] (3.9)

By applying the operator $\Delta$ iteratively to the equality (3.9), we then deduce (3.6). It remains to show that
\[ f_X(j) = f_Y(j) \quad \text{for } j \leq s - 1. \] (3.10)

By the first part of (2.4), we have
\[ E\left(\frac{Y}{s-j}\right) = E\left(\frac{X}{s-j}\right) \quad \text{for } j \leq s - 1. \] (3.11)

Let us consider (3.11) successively for $j = s - 1; s - 2; \ldots, 1$. Using (3.6), we then obtain (3.10).

4. Extrema in $B_s(D_n; \mu_1, \mu_2, \ldots, \mu_{s-1})$

The construction of extremal distributions has a wide field of applications, especially when only partial informations on the distributions under study are available (see Sections 1 and 6). In our framework, the question is to derive random variables $X_{s}^{(\min)}$ and $X_{s}^{(\max)}$ belonging to $B_s(D_n; \mu_1, \mu_2, \ldots, \mu_{s-1})$ and such that
\[ X_{s}^{(\min)} \preceq_{s-cx} X \preceq_{s-cx} X_{s}^{(\max)} \quad \text{for all } X \in B_s(D_n; \mu_1, \mu_2, \ldots, \mu_{s-1}). \] (4.1)

The problem (4.1) has been discussed in Section 5 of Denuit and Lefèvre (1997a). Specifically, using the cut criterion on distribution functions (see Proposition 3.2), the extrema for $s = 1, 2, 3$ were obtained explicitly as follows.

**Property 4.1.** (a) $s = 1$. Then, $X_{s}^{(\min)} = 0$ and $X_{s}^{(\max)} = n$ almost surely.

(b) $s = 2$. Given $\mu_1$, let $\xi$ be the integer in $[0, n - 1]$ such that $\xi < \mu_1 \leq \xi + 1$. Then,
\[ X_{\min}^{(2)} = \begin{cases} \xi & \text{with the probability } r_1 = \frac{\xi + 1 - \mu_1}{n}, \\ \xi + 1 & \text{with the probability } r_2 = \mu_1 - \xi. \end{cases} \] (4.2)

and
\[ X_{\max}^{(2)} = \begin{cases} 0 & \text{with the probability } t_1 = 1 - \frac{\mu_1}{n}, \\ n & \text{with the probability } t_2 = \frac{\mu_1}{n}. \end{cases} \] (4.3)

(c) $s = 3$. Given $\mu_1$ and $\mu_2$, let $\xi_1$ and $\xi_2$ be the integers in $[0, n - 1]$ such that
\[ \frac{\mu_2}{\mu_1} \leq \xi_1 + 1 \quad \text{and} \quad \frac{\mu_2}{\mu_1} - \frac{\mu_1}{n} \leq \xi_2 \leq \frac{\mu_2}{\mu_1} \leq \xi_1 + 1. \]

Then,
\[ X_{\min}^{(3)} = \begin{cases} 0 & \text{with the probability } p_1 = 1 - p_2 - p_3, \\ \xi_1 & \text{with the probability } p_2 = \frac{(\xi_1 + 1)\mu_1 - \mu_2}{\xi_1}, \\ \xi_1 + 1 & \text{with the probability } p_3 = \frac{\mu_2 - \xi_1\mu_1}{1 + \xi_1}. \end{cases} \] (4.4)
and

\[
X_{\text{max}}^{(3)} = \begin{cases} 
\xi_2 & \text{with the probability } q_1 = \frac{(1 + \xi_2)(n - \mu_1) + \mu_2 - n\mu_1}{n - \xi_2}, \\
\xi_2 + 1 & \text{with the probability } q_2 = \frac{(n + \xi_2)\mu_1 - \mu_2 - n\xi_2}{n - 1 - \xi_2}, \\
n & \text{with the probability } q_3 = 1 - q_1 - q_2.
\end{cases}
\] (4.5)

Furthermore, it was outlined in Denuit and Lefèvre (1997a) that the theory of the discrete Tchebycheff systems provides a general methodology to find the \( s \)-convex extrema for larger values of \( s \). To be more precise, we are now going to apply this approach to derive the maximum for \( s = 4 \).

**Property 4.2.** (d) \( s = 4 \). Given \( \mu_1, \mu_2 \) and \( \mu_3 \), let \( \zeta \) be the integer in \([1, n - 2]\) such that

\[
\frac{\mu_2n - \mu_3}{\mu_1n - \mu_2} \leq \zeta + 1.
\] (4.6)

Then,

\[
X_{\text{max}}^{(4)} = \begin{cases} 
0 & \text{with the probability } v_1 = 1 - v_2 - v_3 - v_4, \\
\zeta & \text{with the probability } v_2 = \frac{n\mu_1(\zeta + 1) - \mu_2(\zeta + n + 1) + \mu_3}{\zeta(n - \zeta)}, \\
\zeta + 1 & \text{with the probability } v_3 = \frac{\mu_2(\zeta + n) - n\mu_1\xi - \mu_3}{(\zeta + 1)(n - \zeta - 1)}, \\
n & \text{with the probability } v_4 = \frac{\mu_3 - \mu_2(2\zeta + 1) + \mu_1\xi(\zeta + 1)}{n(n - \zeta)(n - \zeta - 1)}.
\end{cases}
\] (4.7)

**Proof.** Firstly, as explained in Denuit and Lefèvre (1997a), it can be deduced from the general theory that the support of \( X_{\text{max}}^{(4)} \) is of the form \([0, k, k + 1, n]\), for some integer \( k \) in \([1, n - 2]\). We will take for \( k \) the integer \( \zeta \) that satisfies (4.6). Let us check that, as announced, \( \zeta \) lies in the interval \([1, n - 2]\). This property can be rewritten as

\[
1 < \frac{\mu_2n - \mu_3}{\mu_1n - \mu_2} \leq n - 1.
\] (4.8)

The left-hand side inequality reduces to \( \mu_1n - \mu_2 < \mu_2n - \mu_3 \), which is obviously true since \( X(n - X) < X^2(n - X) \) on \([1, 2, \ldots, n - 1]\) and a biatomic random variable on \([0, n]\) cannot fulfill, in general, the moment conditions.

The right-hand side inequality of (4.8) is equivalent to \( \mu_3 - (2n - 1)\mu_2 + (n - 1)\mu_1 \geq 0 \), which is verified since \( X^2 - (2n - 1)X + (n - 1)n \geq 0 \) on \( D_n \) (the roots of the equation \( y^2 - (2n - 1)y + (n - 1)n \) being \( n - 1 \) and \( n \)). Now, let us determine the probability masses of \( X_{\text{max}}^{(4)} \), i.e. \( v_1, v_2, v_3, v_4 \). By definition, these have to satisfy the system of equations

\[
\begin{align*}
\xi v_2 + (\xi + 1)v_3 + nv_4 &= \mu_1, \\
\xi_2 v_2 + (\xi + 1)^2v_3 + n^2v_4 &= \mu_2, \\
\xi_2 v_2 + (\xi + 1)^3v_3 + n^3v_4 &= \mu_3,
\end{align*}
\] (4.9)

with the additional relation \( v_1 = 1 - v_2 - v_3 - v_4 \). The solutions correspond to the values indicated in (4.7). It remains to verify that these masses are non-negative. This is the case for \( v_2 \) and \( v_3 \) by the definition of \( \zeta \). To prove that \( v_4 \geq 0 \), it suffices to note that \( X^2 - (2\xi + 1)X + \xi(\xi + 1) \geq 0 \) on \( D_n \) (the roots of the associated equation being the integers \( \xi \) and \( \xi + 1 \)). Finally, from the observation that
with, in addition, a unique mode at 0. Our purpose is to derive the random variables $X_s$ such a mixture). Nevertheless, some discrete distributions are obtainable as mixtures of uniform distributions over uniform laws over $f_j$ cannot hold in general because the uniform distribution over $f_j$ is not a mixture of the uniform laws over $\{-j, -j + 1, \ldots, 0\}$ and $\{0, \ldots, k - 1, k\}$ (since the mass in zero gets inflated when we take such a mixture). Nevertheless, some discrete distributions are obtainable as mixtures of uniform distributions over

$$
\begin{vmatrix}
1 & 1 & 1 & 1 \\
X & \zeta & \zeta + 1 & n \\
X^2 & \zeta^2 & (\zeta + 1)^2 & n^2 \\
X^3 & \zeta^3 & (\zeta + 1)^3 & n^3
\end{vmatrix} \geq 0 \quad \text{on } \mathcal{D}_n,
$$

(since this is a Vandermonde's determinant), we easily get that $1 - v_2 - v_3 - v_4 \geq 0$ as required.

5. Improved extrema for decreasing densities

The extrema can be improved, of course, when supplementary informations on the distributions are provided. A situation of practical interest is when the distributions are known to have also a unique (given) mode $m$. In this section, we are going to investigate the special case where $m = 0$, that is when the density functions are decreasing. Thus, let us consider the subspace $B^s_1(\mathcal{D}_n; \mu_1, \mu_2, \ldots, \mu_{s-1})$ that contains all the distributions of $B_i(\mathcal{D}_n; \mu_1, \mu_2, \ldots, \mu_{s-1})$ with, in addition, a unique mode at 0. Our purpose is to derive the random variables $X^{(s)_m}$ and $X^{(s)_\max}$ belonging to $B^s_1(\mathcal{D}_n; \mu_1, \mu_2, \ldots, \mu_{s-1})$ and such that

$$X^{(s)_\min}_{\mathcal{D}_n} \leq X \leq X^{(s)_\max}_{\mathcal{D}_n} \text{ for all } X \in B^s_1(\mathcal{D}_n; \mu_1, \mu_2, \ldots, \mu_{s-1}). \quad (5.1)$$

We indicate that the case where the mode is interior (i.e. $m \geq 1$) seems to be mathematically much more difficult. We hope to be able to study it in a forthcoming work. The difficulty here is that for the moment, we do not have at our disposal a representation result such as Lemma 5.1 below. Rather surprisingly, the problem vanishes when the distributions are continuous instead of discrete.

Let us briefly expand on this unexpected phenomenon. For that, we first recall some classical notions from the continuous case. The definition of unimodality in terms of distribution functions was initially given by Khinchine who stated that a random variable $X$ (or, equivalently, its distribution function $F_X$) is unimodal about a mode $m$ if $F_X$ is convex on $]-\infty, m[$ and concave on $]m, +\infty[$. As a consequence, if $F_X$ is unimodal about $m$, then apart from a possible mass at $m$, $F_X$ is absolutely continuous. It is well known that the class of all distribution functions which are unimodal about a mode $m$ is convex under mixtures and closed under weak limits. Therefore, this class has extreme points and it is the closed convex hull of the set of these extreme points. Let us now take $m = 0$ (since this involves no loss of generality). Khinchine showed that any distribution function $F_X$ which is unimodal about 0 is a mixture of uniform distributions over $[-a, b]$, with $a, b \geq 0$. Since the uniform law over $[-a, b]$ is a mixture of the uniform distributions over $[0, b]$, $F_X$ is unimodal about 0 if, and only if, $F_X$ is a mixture of uniform distributions over $[0, b]$, with $b \in \mathbb{R}$. The latter convenient representation was used in Denuit et al. (1998, 1999b) in order to get improved extrema in the $s$-convex sense when unimodality is known to hold. Let us now turn to the discrete case. According to the definition of unimodality given above, the only discrete distributions which are unimodal are the degenerate ones. Therefore, we opt for the following classical alternative definition of unimodality: a distribution $\{f_X(n), n \in \pm\mathbb{N}\}$ (where $\pm\mathbb{N}$ stands for the set of positive and negative integers) is said unimodal about a mode $m$ if

$$f_X(n) \geq f_X(n - 1) \text{ for } n \leq m \quad \text{and} \quad f_X(n) \leq f_X(n - 1) \text{ for } n \geq m + 1.$$ 

In other words, a unimodal distribution $\{f_X(n), n \in \pm\mathbb{N}\}$ is such that $\Delta f_X(n)$ has only one change in sign, when the zero terms are ignored. As above, let us assume, without loss of generality, that $m = 0$. The extreme points of the set of all distributions on $\pm\mathbb{N}$ which are unimodal about 0 are the discrete uniform distributions over $\{-j, -j + 1, \ldots, k - 1, k\}$, where $j, k \in \mathbb{N}$. In contrast with the continuous case, however, we do not have that the uniform distributions over $\{\min(0,k), \ldots, \max(0,k)\}, k \in \pm\mathbb{N}$, are also the extreme points. Such a result cannot hold in general because the uniform distribution over $\{-j, -j + 1, \ldots, k - 1, k\}$ is not a mixture of the uniform laws over $\{-j, -j + 1, \ldots, 0\}$ and $\{0, \ldots, k - 1, k\}$ (since the mass in zero gets inflated when we take such a mixture). Nevertheless, some discrete distributions are obtainable as mixtures of uniform distributions over
\{ \min(0, k), \ldots, \max(0, k) \}, k \in \mathbb{N}. The condition is that the mass in 0 (i.e., in the mode) is high enough; see, e.g., Theorem 4.2 in Dharmadhikari and Joag-Dev (1988). But even in this particular case, technical problems arise from the extension of Lemma 5.2 below. Under the assumption of decreasing densities, however, the mathematics is more tractable.

To begin with, we give a convenient representation of any distribution on $\mathcal{D}_n$ with mode 0. This result corresponds to a particular case of Khinchine’s theorem (see, e.g., Theorem 4.2 in Dharmadhikari and Joag-Dev (1988)). For clarity, however, we will prove it. Let us denote by $U(0, z]$, $z \in \mathcal{D}_n$, the uniform distribution on $\{0, 1, \ldots, z\}$. Given some random variable $Z$ valued in $\mathcal{D}_n$, we write $MU(0, Z]$ for the mixed uniform distribution with random final point $Z$ as mixing parameter. Let “$\equiv_d$” denote the equality in distribution.

Lemma 5.1. A random variable $X$ valued in $\mathcal{D}_n$ has a unique mode at 0 if, and only if, $X \equiv_d MU(0, Z]$ where the random variable $Z$ is valued in $\mathcal{D}_n$ with density function $f_Z$ given by

$$
\begin{align*}
f_Z(j) &= -(j + 1) f_X(j), \quad j = 0, 1, \ldots, n - 1, \\
f_Z(n) &= (n + 1) f_X(n). 
\end{align*}
$$

(5.2)

Proof. Let us first check that $\{f_Z(j), j \in \mathcal{D}_n\}$ specified by (5.2) constitutes a density function. $f_X$ being decreasing, we see that $f_Z(j) \geq 0$ for all $j \in \mathcal{D}_n$. Moreover, from (5.2) and (3.2) we get

$$
\sum_{j=0}^{n} f_Z(j) = - \sum_{j=0}^{n-1} j \Delta f_X(j) - [f_X(n) - f_X(0)] + (n + 1) f_X(n)
$$

$$
= (n + 1) f_X(0) + \sum_{j=0}^{n-1} [f_X(j + 1) - f_X(0)] = 1.
$$

Now, the sufficient condition in the statement is directly verified. For the necessary condition, we have just to show that the solution $\{y_j, j \in \mathcal{D}_n\}$ to the system of linear equations

$$
f_X(j) = \sum_{k=j}^{n} \frac{y_k}{k+1}, \quad j \in \mathcal{D}_n, 
$$

(5.3)

corresponds precisely to $\{f_Z(j), j \in \mathcal{D}_n\}$. This follows from (5.3) by taking successively $j = n, n - 1, \ldots, 0$. \qed

By this lemma, any calculation involving $X$ can be expressed in terms of $Z$ and reciprocally. In particular, for any function $\phi : \mathcal{D}_n \to \mathbb{R}$,

$$
E\phi(X) = \sum_{j=0}^{n} f_Z(j) \frac{1}{j+1} \sum_{k=0}^{j} \phi(k),
$$

(5.4)

so that putting

$$
\tilde{\phi}(j) = \frac{1}{j+1} \sum_{k=0}^{j} \phi(k), \quad j \in \mathcal{D}_n,
$$

(5.5)

we get

$$
E\phi(X) = E\tilde{\phi}(Z).
$$

(5.6)
We also have

\[ \tilde{\mu}_i \equiv EZ^j = \sum_{j=0}^{n-1} j^i (j+1) f_X(j) - \sum_{j=0}^{n-1} j^i (j+1) f_X(j+1) + n^i (n+1) f_X(n), \]

which becomes

\[ \tilde{\mu}_i = E[X^i (X + 1)] - E[X (X - 1)^i] = (i + 1) \mu_i + \sum_{j=1}^{i-1} \binom{i}{j-1} (-1)^{i-j} \mu_j. \] (5.7)

Coming back to the problem of extrema, we would like to bound \( E\phi(X) \) within \( B_s(\mathcal{D}_n; \mu_1, \mu_2, \ldots, \mu_{s-1}) \), for any \( s \)-convex function \( \phi \) on \( \mathcal{D}_n \). By (5.6), this leads us to look at the properties of the associated function \( \phi \) defined in (5.5).

**Lemma 5.2.** Let \( \phi \in \mathcal{U}_{s-\text{cx}}^{\mathcal{D}_n} \), then \( \tilde{\phi} \in \mathcal{U}_{s-\text{cx}}^{\mathcal{D}_n} \).

**Proof.** We start by establishing that for \( 1 \leq t \leq n \),

\[ (j + t + 1) \Delta^t \phi(j) = \Delta^{t-1} \phi(j + 1) - t \Delta^{t-1} \phi(j), \quad j \in \mathcal{D}_{n-t}. \] (5.8)

We will proceed by recurrence. Take \( t = 1 \). Since for any functions \( \varphi \) and \( \psi : \mathcal{D}_n \to \mathbb{R} \),

\[ \Delta[\varphi \psi](j) = \varphi(j + 1) \Delta \psi(j) + \psi(j) \Delta \varphi(j), \] (5.9)

we have

\[ \Delta[(j + 1) \tilde{\phi}(j)] = (j + 2) \Delta \tilde{\phi}(j) + \tilde{\phi}(j), \quad j \in \mathcal{D}_{n-1}. \] (5.10)

On the other hand, from (5.5) we get

\[ \Delta[(j + 1) \tilde{\phi}(j)] = \Delta \sum_{k=0}^{j} \phi(k) = \phi(j + 1), \]

which, when combined with (5.10), yields (5.8). For \( t \geq 2 \), we have by (5.9) that

\[ \Delta[(j + t) \Delta^t \tilde{\phi}(j)] = (j + t + 1) \Delta^t \tilde{\phi}(j) + \Delta^{t-1} \tilde{\phi}(j), \]

while by the recurrence hypothesis,

\[ \Delta[(j + t) \Delta^{t-1} \tilde{\phi}(j)] = \Delta[\Delta^{t-2} \phi(j + 1) - (t - 1) \Delta^{t-2} \tilde{\phi}(j)] \]

\[ = \Delta^{t-1} \phi(j + 1) - (t - 1) \Delta^{t-1} \tilde{\phi}(j), \quad j \in \mathcal{D}_{n-1}, \]

hence (5.8) follows.

Now, using (5.8), we are going to prove, again by induction, that for \( 1 \leq s \leq n \),

\[ (j + s + 1) \binom{j + s}{s} \Delta^s \phi(j) = \sum_{k=0}^{j} \binom{k + s}{s} \Delta^s \phi(k), \quad j \in \mathcal{D}_{n-s}. \] (5.11)

which gives directly the assertion announced. Take \( s = 1 \). From (5.5) we obtain by (5.9) that

\[ (j + 2)(j + 1) \Delta \phi(j) = (j + 1) \phi(j + 1) - \sum_{k=0}^{j} \phi(k) = \sum_{k=0}^{j} (k + 1) \Delta \phi(k), \quad j \in \mathcal{D}_{n-1}, \]
that is (5.11). For \( s \geq 2 \), we get by (5.8) that the left-hand side member of (5.11) can be rewritten as
\[
\left( \frac{j + s}{s} \right) [\Delta^{s-1} \phi(j + 1) - s \Delta^{s-1} \phi(j)] = \left( \frac{j + s}{s} \right) \Delta^{s-1} \phi(j + 1) - (j + s) \left( \frac{j + s - 1}{s - 1} \right) \Delta^{s-1} \phi(j).
\]
The r.h.s. of the above relation becomes by the recurrence hypothesis
\[
\left( \frac{j + s}{s} \right) \Delta^{s-1} \phi(j + 1) - \sum_{k=0}^{j} \left( \frac{k + s - 1}{s - 1} \right) \Delta^{s-1} \phi(k),
\]
which reduces to
\[
\sum_{k=0}^{j} \left( \frac{k + s - 1}{s - 1} \right) [\Delta^{s-1} \phi(j + 1) - \Delta^{s-1} \phi(k)] = \sum_{k=0}^{j} \left( \frac{k + s - 1}{s - 1} \right) \sum_{i=k}^{j} \Delta^{s} \phi(i), \quad j \in \mathcal{D}_{n-s},
\]
hence the r.h.s. member of (5.11).

Now, by Lemmas 5.1 and 5.2 we see that the original extremal problem for \( X \) within \( \mathcal{B}^*_{s}(\mathcal{D}_n; \mu_1, \mu_2, \ldots, \mu_{s-1}) \) may be transferred, in an equivalent way, to the extremal problem for \( Z \) within \( \mathcal{B}^*_{s}(\mathcal{D}_n; \bar{\mu}_1, \bar{\mu}_2, \ldots, \bar{\mu}_{s-1}) \). The latter space is the moment space in which the \( \bar{\mu}_i \)'s are evaluated from the \( \mu_i \)'s using (5.7); there is no longer any constraint of unimodal type.

Therefore, we will proceed in two steps: (i) applying the methodology described in Section 4, we determine the extrema \( Z^{(s)\min} \) and \( Z^{(s)\max} \) within \( \mathcal{B}_s(\mathcal{D}_n; \bar{\mu}_1, \bar{\mu}_2, \ldots, \bar{\mu}_{s-1}) \);

(ii) using the representation Lemma 5.1, we then deduce that within \( \mathcal{B}^*_{s}(\mathcal{D}_n; \mu_1, \mu_2, \ldots, \mu_{s-1}) \), the extrema are
\[
X^{(s)\min} \equiv \mathcal{D} \mathcal{M} \mathcal{U}[0, Z^{(s)\min} \mathcal{D}] \quad \text{and} \quad X^{(s)\max} \equiv \mathcal{D} \mathcal{M} \mathcal{U}[0, Z^{(s)\max} \mathcal{D}].
\]

Notice that the distributions of \( X^{(s)\min} \) and \( X^{(s)\max} \) are thus mixtures of \( s \) (or fewer) discrete uniform laws. They are derived explicitly for \( s = 1, 2 \) and 3.

**Property 5.3.** (a) \( s = 1 \). Then, \( X^{(1)\min} \equiv 0 \) almost surely and \( X^{(1)\max} \) is uniformly distributed on \( \mathcal{D}_n \).

(b) \( s = 2 \). Given \( \mu_1 \), let \( \bar{\xi} \in \mathcal{D}_n \) be the integer in \([0, n-1] \) such that \( \bar{\xi} < 2\mu_1 \leq \bar{\xi} + 1 \). Then,
\[
f_{X^{(2)\min}}(j) = \begin{cases} 
2(\bar{\xi} + 1 - \mu_1) / (\bar{\xi} + 1)(\bar{\xi} + 2) & \text{for } j = 0, 1, \ldots, \bar{\xi}, \\
2\mu_1 - \bar{\xi} / \bar{\xi} + 2 & \text{for } j = \bar{\xi} + 1, \\
0 & \text{for } j \geq \bar{\xi} + 2,
\end{cases}
\]
and
\[
f_{X^{(2)\max}}(j) = \begin{cases} 
1 - 2\mu_1 / n + 1 & \text{for } j = 0, \\
2\mu_1 / n(n + 1) & \text{for } j = 1, 2, \ldots, n.
\end{cases}
\]

(c) \( s = 3 \). Given \( \mu_1 \) and \( \mu_2 \), let \( \bar{\xi}_1 \) and \( \bar{\xi}_2 \) be the integers in \([0, n-1] \) such that
\[
\bar{\xi}_1 < \frac{3\mu_2 - \mu_1}{2\mu_1} \leq \bar{\xi}_1 + 1 \quad \text{and} \quad \bar{\xi}_2 < \frac{2\mu_1 - 3\mu_2 + \mu_1}{n - 2\mu_1} \leq \bar{\xi}_2 + 1.
\]
Then,

\[ f_{X_{\min}^{(3)}}(j) = \begin{cases} 
\tilde{p}_1 + \frac{\tilde{p}_2}{\tilde{\xi}_1 + 1} + \frac{\tilde{p}_3}{\tilde{\xi}_1 + 2} & \text{for } j = 0, \\
\frac{\tilde{p}_2}{\tilde{\xi}_1 + 1} + \frac{\tilde{p}_3}{\tilde{\xi}_1 + 2} & \text{for } j = 1, \ldots, \tilde{\xi}_1, \\
\frac{\tilde{p}_3}{\tilde{\xi}_1 + 2} & \text{for } j = \tilde{\xi}_1 + 1, \\
0, & \text{for } j \geq \tilde{\xi}_1 + 2, 
\end{cases} \tag{5.15} \]

where

\[ \tilde{p}_1 = 1 - \tilde{p}_2 - \tilde{p}_3, \quad \tilde{p}_2 = \frac{2(\tilde{\xi}_1 + 1)\mu_1 - 3\mu_2 + \mu_1}{\tilde{\xi}_1}, \quad \tilde{p}_3 = \frac{3\mu_2 - \mu_1 - 2\tilde{\xi}_1\mu_1}{1 + \tilde{\xi}_1}, \]

and

\[ f_{X_{\max}^{(3)}}(j) = \begin{cases} 
\tilde{q}_1 + \frac{\tilde{q}_2}{\tilde{\xi}_2 + 1} + \frac{\tilde{q}_3}{n + 1} & \text{for } j = 0, 1, \ldots, \tilde{\xi}_2, \\
\frac{\tilde{q}_2}{\tilde{\xi}_2 + 1} + \frac{\tilde{q}_3}{n + 1} & \text{for } j = \tilde{\xi}_2 + 1, \\
\frac{\tilde{q}_3}{n + 1} & \text{for } j = \tilde{\xi}_2 + 2, \ldots, n, 
\end{cases} \tag{5.16} \]

where

\[ \tilde{q}_1 = \frac{(1 + \tilde{\xi}_2)(n - 2\mu_1) + 3\mu_2 - \mu_1 - 2n\mu_1}{n - \tilde{\xi}_2}, \quad \tilde{q}_2 = \frac{2(n + \tilde{\xi}_2)\mu_1 - 3\mu_2 + \mu_1 - n\tilde{\xi}_2}{n - 1 - \tilde{\xi}_2}, \quad \tilde{q}_3 = 1 - \tilde{q}_1 - \tilde{q}_2. \]

**Proof.** The result when \( s = 1 \) is a straightforward consequence from Proposition 3.2. Let us consider the cases \( s = 2 \) and \( s = 3 \). By (5.7), we get that \( \bar{\mu}_1 = 2\mu_1 \) and \( \bar{\mu}_2 = 3\mu_2 - \mu_1 \). Then, (4.2) and (4.3) yield \( Z_{\min}^{(2)} \) and \( Z_{\max}^{(2)} \), while (4.4) and (4.5) provide \( Z_{\min}^{(3)} \) and \( Z_{\max}^{(3)} \). By (5.12) and using formula (5.3), we easily deduce the density functions (5.13) and (5.14) for \( X_{\min}^{(2)*} \) and \( X_{\max}^{(2)*} \), and (5.15) and (5.16) for \( X_{\min}^{(3)*} \) and \( X_{\max}^{(3)*} \). \( \square \)

6. Bounding the eventual ruin probability

For illustration, we are going to derive bounds for the eventual ruin probability in the compound binomial risk process. We will thus continue the study made in Denuit and Lefèvre (1997a). In this model (see, e.g., De Vylder, 1996; Gerber, 1988), time is measured in discrete time units \( t \in \mathbb{N} \) and the number of claims is governed by a binomial process \( \{N(t), \ t \in \mathbb{N}\} \) with parameter \( q \), \( 0 < q < 1 \) (that is, in any time period, there occurs 1 or 0 claim with probabilities \( q \) and \( 1 - q \), respectively, and occurrences of claims in different time intervals are independent events). The claim amounts \( X_k, k \geq 1 \), are i.i.d. and distributed as an arithmetic random variable \( X \) valued in \( \{1, 2, 3, \ldots, n\} \), with mean \( \mu_X \); they are independent of the binomial process \( \{N(t), \ t \in \mathbb{N}\} \). The premium received in each period is equal to 1 and is assumed to be larger than the net premium, which means that \( 1 > q\mu_X \). The initial risk reserve is a non-negative integer amount \( z \).
Table 1
Bounds on $\psi_X(z)$ (when $q = 0.3$, $\mu_X = 2.7$, $n = 6$ or 12)

<table>
<thead>
<tr>
<th>$z$</th>
<th>$\psi_X^{\text{min}}$</th>
<th>$\psi_X^{\text{max}}$</th>
<th>$\psi_X^{\text{max}}, n = 6$</th>
<th>$\psi_X^{\text{max}}, n = 12$</th>
<th>$\psi_X^{\text{max}}, n = 12$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.728571</td>
<td>0.728571</td>
<td>0.728571</td>
<td>0.728571</td>
<td>0.728571</td>
</tr>
<tr>
<td>2</td>
<td>0.480962</td>
<td>0.546625</td>
<td>0.593910</td>
<td>0.643706</td>
<td>0.666147</td>
</tr>
<tr>
<td>4</td>
<td>0.311346</td>
<td>0.387966</td>
<td>0.458758</td>
<td>0.532307</td>
<td>0.591707</td>
</tr>
<tr>
<td>6</td>
<td>0.200707</td>
<td>0.278082</td>
<td>0.352556</td>
<td>0.425629</td>
<td>0.521182</td>
</tr>
<tr>
<td>8</td>
<td>0.129268</td>
<td>0.198956</td>
<td>0.272069</td>
<td>0.349879</td>
<td>0.453220</td>
</tr>
<tr>
<td>10</td>
<td>0.083240</td>
<td>0.142362</td>
<td>0.209674</td>
<td>0.282911</td>
<td>0.391735</td>
</tr>
<tr>
<td>12</td>
<td>0.053999</td>
<td>0.101868</td>
<td>0.161746</td>
<td>0.231856</td>
<td>0.341446</td>
</tr>
<tr>
<td>14</td>
<td>0.034512</td>
<td>0.072892</td>
<td>0.124723</td>
<td>0.188716</td>
<td>0.298159</td>
</tr>
<tr>
<td>16</td>
<td>0.022223</td>
<td>0.052158</td>
<td>0.096186</td>
<td>0.153970</td>
<td>0.259755</td>
</tr>
<tr>
<td>18</td>
<td>0.014309</td>
<td>0.037322</td>
<td>0.074176</td>
<td>0.125503</td>
<td>0.226139</td>
</tr>
<tr>
<td>20</td>
<td>0.009214</td>
<td>0.026705</td>
<td>0.057202</td>
<td>0.102317</td>
<td>0.196941</td>
</tr>
<tr>
<td>22</td>
<td>0.005933</td>
<td>0.019109</td>
<td>0.044113</td>
<td>0.083430</td>
<td>0.171596</td>
</tr>
<tr>
<td>24</td>
<td>0.003820</td>
<td>0.013674</td>
<td>0.034019</td>
<td>0.068017</td>
<td>0.149503</td>
</tr>
</tbody>
</table>

Let $U_X(t)$ denote the risk reserve at time $t$, i.e.

$$U_X(t) = z + t - \sum_{k=1}^{N(t)} X_k, \quad t \in \mathbb{N}. \quad (6.1)$$

The probability of eventual ruin is defined as $\psi_X(z) = P[U_X(t) < 0 \text{ for some } t | U_X(0) = z]$. We recall that it can be evaluated by a well-known formula (given, for instance, in De Vylder (1996), Theorem 11, p. 251).

Now, let us consider two such compound binomial risk processes, both with premium 1 per time unit and with $q$ as the binomial parameter, but with different discrete claim amounts, distributed as $X$ and $Y$, respectively. Let $\psi_X(.)$ and $\psi_Y(.)$ be the probabilities of eventual ruin for these models. Denuit and Lefèvre (1997a) proved that

$$X \preceq_{2-\text{(i)},x} Y \Rightarrow \psi_X(z) \leq \psi_Y(z) \quad \text{for all integer } z. \quad (6.2)$$

This result is applicable, for instance, to the following situation. Let us assume that the distribution of the claim amount $X$ is essentially unspecified, the only information available being that its mean is given by $\mu_X$ and, possibly, that $X$ has also a unique mode at 0. Then, the 2-convex extrema for $X$ are provided by (4.2) and (4.3), and (5.13) and (5.14), respectively. By (6.2), we thus deduce that the probability of eventual ruin $\psi_X(z)$ can be bounded by the corresponding probabilities $\psi_X^{\text{min}}(z)$, $\psi_X^{\text{max}}(z)$ and $\psi_Y^{\text{min}}(z)$, $\psi_Y^{\text{max}}(z)$, say, which are determined in the usual way.

Some numerical illustrations of these bounds are presented in Table 1 (for different values of the initial surplus level $z$ and of the largest claim amount $n$). The bounds are, of course, more accurate when the unimodality assumption is added. We observe that the improvement may be substantial, especially when the initial surplus $z$ is large.

Acknowledgements

The authors thank the referee for helpful remarks and suggestions on the work.

References


