



Stationary-excess operator and convex stochastic orders

Claude Lefèvre

*Département de Mathématique, Université Libre de Bruxelles, Boulevard du Triomphe
C.P.210, B-1050 Bruxelles, Belgique*

Stéphane Loisel

*Université de Lyon, Université Claude Bernard Lyon 1, Institut de Science Financière et
d'Assurances, 50 Avenue Tony Garnier, F-69007 Lyon, France*

Abstract

The present paper aims to point out how the stationary-excess operator and its iterates transform the s -convex stochastic orders and the associated moment spaces. This allows us to propose a new unified method on constructing s -convex extrema for distributions that are known to be t -monotone. Both discrete and continuous cases are investigated. Several extremal distributions under monotonicity conditions are derived. They are illustrated with some applications in insurance.

MSC: primary 60E15, 62P05; secondary 60E10

Keywords: Insurance risks; Stochastic orders; Monotone distributions; Conjugate operator; Stochastic extrema; Discrete and continuous versions.

1 Introduction

Concepts of stochastic orderings are useful in a number of applied probability models. This is especially true in insurance and finance when different risk scenarios are possible and have to be compared. Such questions arise, for instance, in life insurance, ruin theory and portfolio analysis. A theory of stochastic orders with various applications can be found in the books by Goovaerts et al. (1990), Kaas et al. (1994), Ross (1996), Müller and Stoyan (2002) and Shaked and Shanthikumar (1994), (2007).

This paper is concerned with the class of s -convex stochastic orders, where $s \in \mathbb{N}_0 = \{1, 2, \dots\}$. For $s = 1$, this order is the classical stochastic dominance. For $s = 2$, it is the well-known convex order and corresponds, in actuarial sciences, to the stop-loss order with fixed mean. For an arbitrary s , the s -convex order compares the s -th right-tail distribution functions of random variables that have the same first $s - 1$ moments. In the case of discrete distributions, this class of orders was studied in Lefèvre and Picard (1993), Fishburn and Lavalley (1995), Lefèvre and Utev (1996), Denuit and Lefèvre (1997) and Denuit et al. (1999a), (1999c), among others. The case, more traditional, of real-valued random variables was investigated by Rolski (1976), Levy (1992), Denuit et al. (1998), (1999b) and in many other works.

A directly related question is the construction of stochastic extrema with respect to the s -convex orders. In an actuarial context, the extremal risks represent the less and more dangerous risks, and their knowledge can yield accurate lower and upper bounds on various risk quantities

of interest (the premium, for instance). The problem of s -convex optimization corresponds to a traditional moment problem; see, e.g., Hoeffding (1955), Karlin and Studden (1966), Kemperman (1968), Utev (1985), Prékopa (1990) and Hürlimann (1999). In the discrete case, explicit extremal distributions are known for $s = 1, 2, 3, 4$ and were derived by Denuit and Lefèvre (1997), Denuit et al. (1999c) and Courtois et al. (2006). In the real case, explicit extrema were obtained by Jansen et al. (1986), De Vylder (1996) and Denuit et al. (1998), (1999b), inter alia.

The *stationary-excess* operator is a standard mathematical tool that plays an important role in renewal theory and survival analysis (Cox (1972)). Our purpose in the present work is to investigate how this operator and its iterates transform the s -convex stochastic orders and the associated moment spaces. Both discrete and continuous cases are discussed, with a special emphasis to the less traditional discrete case.

As a corollary, the followed approach allows us to propose a method on solving the s -convex optimization problem within the subset of distributions that are known to be t -monotone, $t \in \mathbb{N}_0$. For $t = 1$, this property is the nonincreasingness of the distributions, and for $t = 2$, it corresponds to their nonincreasingness and convexity. Monotonicity of order t in the discrete (resp. continuous) case means that the first t differences of the probability mass function (resp. derivatives of the density function) are assumed to be of alternating signs with a negative sign at the beginning. Our key theorem states that a problem of s -convex optimization among t -monotone distributions is (almost) equivalent, through the stationary-excess operator, to a problem of $s + t$ -convex optimization without monotonicity constraints.

It was already pointed out in several papers that the s -convex extrema can be improved when the distributions of interest are nonincreasing (and, more generally, unimodal); see, e.g., Denuit et al. (1998), (1999b), (1999c). Our work can be viewed as an extension of these results that is carried out by using the stationary-excess operator and its iterates.

The paper is organized as follows. In Section 2, we present the stationary-excess operator, under its usual definition and in a non-standard version specific to discrete distributions. In Section 3, we prove that this operator essentially transforms any $s + 1$ -convex stochastic order to an s -convex stochastic order for nonincreasing distributions. The result is extended in Section 4 by showing that the t -th iterate of the operator transforms any $s + t$ -convex order to an s -convex order for t -monotone distributions. In Section 5, we use this property, together with well-known convex extrema (without monotonicity conditions), to derive several explicit convex extrema for nonincreasing, possibly convex, distributions. Finally, the interest of these extrema is illustrated in Section 6 with some applications to insurance.

This work has been presented at the "2b) or not 2b) Conference" organized in June 2009 at the Université of Lausanne, in honor to Professor Hans U. Gerber. It gives us an opportunity to point out a nice paper by J. Keilson and H. U. Gerber (1971) on a related notion of discrete unimodality. We also thank Professor F. Dufresne for the excellent organization of the Conference.

2 Stationary-excess operator

In its classical version, the stationary-excess operator, H say, is built for any non-negative random variable X with distribution function F_X and mean $E(X) > 0$ (see Definition 2.3 below).

It is worth recalling that in a renewal process, if an interval between points is of distribution function F_X , then the associated stationary-excess mapping $H(F_X)$ gives the distribution function of the interval to the next point from an arbitrary time in equilibrium. To begin with, we are going to introduce a similar operator H that is specific to discrete distributions.

2.1 Discrete version

Let us assume that X is a *discrete non-negative* random variable with probability mass function $P_X = \{P(X = j), j \in \mathbb{N} = \{0, 1, \dots\}\}$ and mean $E(X) > 0$. Obviously, the classical stationary-excess operator may be applied here too. As explained by Whitt (1985), however, it is more appropriate to work with a discrete version that is directly applicable to discrete-time renewal processes. We propose to adopt the following definition.

Definition 2.1 *A discrete stationary-excess operator H maps any such random variable X to an associated discrete non-negative random variable X_H whose probability mass function $H(P_X)$ is defined by*

$$H(P_X)(j) \equiv P(X_H = j) = \frac{P(X \geq j + 1)}{E(X)}, \quad j \in \mathbb{N}. \quad (2.1)$$

Let us notice that a slightly different operator was investigated by Whitt (1985) for a discrete positive random variable X . Specifically, the associated random variable X_H is of probability mass function defined by

$$P(X_H = j) = \frac{P(X \geq j)}{E(X)}, \quad j \in \mathbb{N}_0.$$

Contrary to that operator, H defined by (2.1) does not yield a one-to-one correspondence on the set of probability measures on \mathbb{N} . Indeed, it is directly checked that X_H and $(\nu X)_H$ are equidistributed if ν is an indicator independent of X . Nevertheless, when $E(X)$ is fixed, H gives a one-to-one correspondence since

$$P(X = 0) = 1 - E(X) P(X_H = 0), \quad \text{and} \quad (2.2)$$

$$P(X = j) = E(X) [P(X_H = j - 1) - P(X_H = j)], \quad j = 1, 2, \dots \quad (2.3)$$

The definition (2.1) has the advantage to lead to a simple relationship between the binomial moments of X_H and X . We make the convention $\binom{x}{y} = 0$ if $x < y$.

Lemma 2.2

$$E \binom{X_H}{i} = \frac{1}{E(X)} E \binom{X}{i+1}, \quad i \in \mathbb{N}. \quad (2.4)$$

Proof. Let us consider the iterated right-tail distribution functions of X , i.e. $\bar{F}_0(X, j) = P(X = j)$ and

$$\bar{F}_{i+1}(X, j) = \sum_{k=j}^{\infty} \bar{F}_i(X, k), \quad i, j \in \mathbb{N}.$$

As proved in Lefèvre and Utev (1996), an equivalent expression is

$$\bar{F}_{i+1}(X, j) = \mathbb{E} \binom{X - j + i}{i}, \quad i, j \in \mathbb{N}. \quad (2.5)$$

Let us turn to the iterated right-tail distribution functions of X_H . By (2.1), we have

$$\bar{F}_0(X_H, j) = P(X_H = j) = \frac{\bar{F}_1(X, j + 1)}{\mathbb{E}(X)}, \quad j \in \mathbb{N}.$$

Arguing by induction, we then find

$$\bar{F}_{i+1}(X_H, j) = \frac{\bar{F}_{i+2}(X, j + 1)}{\mathbb{E}(X)}, \quad i, j \in \mathbb{N}. \quad (2.6)$$

Finally, taking $j = i$ in (2.5) and (2.6), we obtain the identity (2.4). \diamond

2.2 Continuous version

Let X be any *non-negative* real random variable with distribution function F_X and mean $\mathbb{E}(X) > 0$. The usual stationary-exces operator is defined as follows (Cox (1972)). For easiness, it will be named continuous subsequently.

Definition 2.3 *A continuous stationary-excess operator H maps any such random variable X to an associated non-negative random variable X_H whose distribution function $H(F_X)$ is defined by*

$$H(F_X)(x) \equiv P(X_H \leq x) = \frac{1}{\mathbb{E}(X)} \int_0^x [1 - F_X(y)] dy, \quad x \geq 0.$$

An equivalent expression is

$$P(X_H > x) = \frac{1}{\mathbb{E}(X)} \int_x^\infty P(X > y) dy, \quad x \geq 0. \quad (2.7)$$

This definition of H guarantees a one-to-one correspondence. Indeed, denoting the density function of X_H by q_{X_H} , one has

$$P(X > x) = \mathbb{E}(X) q_{X_H}(x), \quad x \geq 0, \quad (2.8)$$

and $\mathbb{E}(X)$ then follows as a consequence.

The moments of X_H and X are linked through a simple identity, known but rederived below for comparison with Lemma 2.2.

Lemma 2.4

$$\mathbb{E}(X_H^i) = \frac{\mathbb{E}(X^{i+1})}{(i+1)\mathbb{E}(X)}, \quad i \in \mathbb{N}. \quad (2.9)$$

Proof. The result is related to the i -th stop-loss transform of X , defined as

$$\Pi^{(i)}(X, x) = \mathbb{E}[(X - x)_+^i], \quad x \geq 0; \quad (2.10)$$

see, e.g., Cheng and Pai (2003). Indeed, considering the iterated right-tail distribution functions of X , i.e. $\bar{F}_1(X, x) = P(X > x)$ and

$$\bar{F}_{i+1}(X, x) = \int_x^\infty \bar{F}_i(X, y) dy, \quad x \geq 0, \quad i = 1, 2, \dots,$$

we easily see that

$$\bar{F}_{i+1}(X, x) = \frac{\Pi^{(i)}(X, x)}{i!}, \quad x \geq 0, \quad i \in \mathbb{N}. \quad (2.11)$$

For the transformed random variable X_H , one finds by induction that

$$\bar{F}_{i+1}(X_H, x) = \frac{\bar{F}_{i+2}(X, x)}{\mathbb{E}(X)}, \quad x \geq 0, \quad i \in \mathbb{N}. \quad (2.12)$$

From (2.12) and (2.11), (2.10), we then deduce that

$$\mathbb{E}[(X_H - x)_+^i] = \frac{\mathbb{E}[(X - x)_+^{i+1}]}{(i+1)\mathbb{E}(X)}, \quad x \geq 0,$$

which yields (2.9) when $x = 0$. \diamond

3 H-transform and convex orders

We start by giving a brief presentation of the s -convex stochastic orders and the associated moment spaces. Applying the stationary-excess operator, we then identify the image of these moment spaces and the corresponding stochastic orders. Both discrete and continuous cases will be examined.

3.1 Discrete problem

For any $s \in \mathbb{N}_0$, let \mathcal{F}_s denote the set of s -convex real functions on \mathbb{N} , i.e.

$$\mathcal{F}_s = \{f : \Delta^s f(j) \geq 0, \quad j \in \mathbb{N}\},$$

Δ being the usual forward difference operator [$\Delta f(j) = f(j+1) - f(j)$, $j \in \mathbb{N}$], and Δ^s the s -th iterate of Δ .

Definition 3.1 Consider any two probability mass functions P_1 and P_2 on \mathbb{N} . One says that P_1 is smaller than P_2 in the s -convex stochastic sense, written $P_1 \preceq_s P_2$, when

$$\langle f, P_1 \rangle \equiv \mathbb{E}_{P_1}(f) \leq \mathbb{E}_{P_2}(f) \equiv \langle f, P_2 \rangle \quad \text{for all functions } f \in \mathcal{F}_s. \quad (3.1)$$

As the functions $f(j) = j^i$ and $f(j) = -j^i$, $j \in \mathbb{N}$, belong to \mathcal{F}_s for all $i = 1, \dots, s-1$, the ordering $P_1 \preceq_s P_2$ necessarily implies that P_1 and P_2 have the same first $s-1$ moments. In fact, if X_1 and X_2 are two random variables with probability mass functions P_1 and P_2 , then an equivalent characterization of (3.1), written as $X_1 \preceq_s X_2$, is that

$$\begin{cases} \mathbb{E}(X_1^i) = \mathbb{E}(X_2^i), & i = 1, \dots, s-1, \text{ and} \\ \bar{F}_s(X_1, j) \leq \bar{F}_s(X_2, j), & j \geq s. \end{cases}$$

Now, let us consider the set of probability mass functions on \mathbb{N} with prescribed first s moments, where $s \in \mathbb{N}_0$. For the sequence, it will be more convenient to work on the basis of the combinatorial moments. So, we introduce a set

$$\mathcal{B}_{s+1} = \left\{ P_X : \mathbb{E} \binom{X}{i} = c_i, i = 1, \dots, s \right\}, \quad (3.2)$$

where c_1, \dots, c_s represent s given (admissible) constants. Clearly, the distributions in the set \mathcal{B}_{s+1} are susceptible to be ordered with respect to \preceq_{s+1} , the $(s+1)$ -convex stochastic order.

Let us take the discrete stationary-excess operator H defined by (2.1). We have seen before that H yields a one-to-one correspondence on the set of distributions with fixed mean. Thus, this property holds true on the previous set \mathcal{B}_{s+1} . Let us now apply the operator H to \mathcal{B}_{s+1} . On the image set $H(\mathcal{B}_{s+1})$, it is then natural to define a transformed order \preceq_H as follows.

Definition 3.2 Consider any two probability mass functions Q_1 and Q_2 in $H(\mathcal{B}_{s+1})$. One says $Q_1 \preceq_H Q_2$ when

$$H^{-1}(Q_1) \equiv P_1 \preceq_{s+1} P_2 \equiv H^{-1}(Q_2). \quad (3.3)$$

We are in a position to identify the set $H(\mathcal{B}_{s+1})$ and the transformed order \preceq_H . Let $Q_Y = \{Q_Y(j), j \in \mathbb{N}\}$ denote the probability mass function of a discrete non-negative random variable Y .

Proposition 3.3

$$H(\mathcal{B}_{s+1}) = \left\{ Q_Y : \mathbb{E} \binom{Y}{i} = \frac{c_{i+1}}{c_1}, i = 1, \dots, s-1, \text{ and} \right. \quad (3.4)$$

$$\left. \Delta Q_Y(j) \leq 0, j \in \mathbb{N}, \text{ with } c_1 Q_Y(0) \leq 1 \right\}, \quad (3.5)$$

and the order \preceq_H corresponds to the s -convex stochastic order \preceq_s .

Proof. From (2.4) and (3.2), we find that the $s-1$ first binomial moments of Y are provided by the formula (3.4). The remaining restrictions stipulated in (3.5) are straightforward from the relations (2.3) and (2.2).

Let us now identify \preceq_H . By virtue of (3.1) and (3.3), $Q_1 \preceq_H Q_2$ when

$$\langle f, H^{-1}(Q_1) \rangle \leq \langle f, H^{-1}(Q_2) \rangle \text{ for all } f \in \mathcal{F}_{s+1}. \quad (3.6)$$

Due to (2.2), (2.3), for any $Q = Q_Y$ in $H(\mathcal{B}_{s+1})$, the expansion of $\langle f, H^{-1}(Q) \rangle$ yields

$$\begin{aligned} \langle f, H^{-1}(Q) \rangle &= f(0) [1 - c_1 P(Y = 0)] + c_1 \sum_{j=1}^{\infty} f(j) [P(Y = j) - P(Y = j - 1)] \\ &= f(0) + c_1 \sum_{j=1}^{\infty} P(Y = j) [f(j + 1) - f(j)]. \end{aligned}$$

Therefore,

$$\langle f, H^{-1}(Q) \rangle = \langle (\delta_0 + c_1 \Delta)(f), Q \rangle, \quad (3.7)$$

where δ_0 is the operator such that $\delta_0(f) = f(0)$. In fact, the equivalence

$$\langle f, H^{-1}(Q) \rangle = \langle [(H^{-1})^*](f), Q \rangle \quad (3.8)$$

holds in general if $(H^{-1})^*$ is the conjugate operator of H^{-1} ; thus, in our case,

$$(H^{-1})^* = \delta_0 + c_1 \Delta. \quad (3.9)$$

Substituting (3.7) in (3.6), we have $Q_1 \preceq_H Q_2$ when

$$\langle \Delta(f), Q_1 \rangle \leq \langle \Delta(f), Q_2 \rangle \quad \text{for all } f \in \mathcal{F}_{s+1}.$$

As $\Delta^{s+1} = \Delta^s(\Delta)$, that condition means, by (3.1), $Q_1 \preceq_s Q_2$, hence the orders \preceq_H and \preceq_s are quite identical. \diamond

Note that the image set $H(\mathcal{B}_{s+1})$ involves $s - 1$ moment relations, the nonincreasingness and an additional technical constraint. Reciprocally, returning to \mathcal{B}_{s+1} through the operator H^{-1} has the effect of removing the nonincreasingness condition. To put it speaking, if for instance, all the combinatorial moments are equal to 1, Proposition 3.3 yields the following simple result.

Corollary 3.4 *The s -convex extremal distributions in the set*

$$\left\{ Q_Y : E \binom{Y}{i} = 1, i = 1, \dots, s - 1, \text{ and } \Delta Q_Y(j) \leq 0, j \in \mathbb{N} \right\}$$

correspond to the H -transform of the $s + 1$ -convex extremal ones in the set

$$\left\{ P_X : E \binom{X}{i} = 1, i = 1, \dots, s \right\}.$$

3.2 Continuous problem

A similar approach is applicable to random variables valued in \mathbb{R}^+ . Given any $s \in \mathbb{N}_0$, the set \mathcal{F}_s of s -convex real functions is defined here as

$$\mathcal{F}_s = \{f : \mathcal{D}^s f(x) \geq 0, x \geq 0\},$$

\mathcal{D} being the usual derivative operator. For two probability distributions P_1 and P_2 on \mathbb{R}^+ , $P_1 \preceq_s P_2$ when the condition (3.1) is again satisfied. Here too, an equivalent characterization is provided by the two conditions stated just after (3.1).

We still work on the set of distributions on \mathbb{R}^+ with prescribed first s moments, i.e.

$$\mathcal{B}_{s+1} = \{P_X : E(X^i) = \mu_i, i = 1, \dots, s\}. \quad (3.10)$$

Let us apply to \mathcal{B}_{s+1} the continuous stationary-excess operator H defined by (2.7). This yields a one-to-one correspondence with the image set $H(\mathcal{B}_{s+1})$. A transformed order \preceq_H may be defined on $H(\mathcal{B}_{s+1})$ exactly as by (3.3). Let Q_Y denote the probability distribution of a continuous non-negative random variable Y with density function $\{q_Y(y), y \geq 0\}$. The continuous analogue of Proposition 3.3 is then the following.

Proposition 3.5

$$H(\mathcal{B}_{s+1}) = \left\{ Q_Y : E(Y^i) = \frac{\mu_{i+1}}{(i+1)\mu_1}, i = 1, \dots, s-1, \text{ and} \right. \quad (3.11)$$

$$\left. \mathcal{D}q_Y(y) \leq 0, y \geq 0, \text{ with } \mu_1 q_Y(0) \leq 1 \right\}, \quad (3.12)$$

and the order \preceq_H corresponds to the continuous s -convex stochastic order \preceq_s .

Proof. The set $H(\mathcal{B}_{s+1})$ is straightforward from (2.8) and (2.9). Let us now identify the order \preceq_H . As with (3.7), consider any $Q = Q_Y$ in $H(\mathcal{B}_{s+1})$, with density function $q = q_Y$. From (2.8) and after integration, we get

$$\begin{aligned} \langle f, H^{-1}(Q) \rangle &= -\mu_1 \int_0^\infty f(y) \mathcal{D}q(y) dy \\ &= f(0) + \mu_1 \int_0^\infty \mathcal{D}f(y) dQ(y), \end{aligned}$$

i.e., the conjugate operator of H^{-1} is

$$(H^{-1})^* = \delta_0 + \mu_1 \mathcal{D}, \quad (3.13)$$

in place of (3.9). Thus, \preceq_H corresponds here also to the order \preceq_s . \diamond

4 Extension to H^t -transforms

At this point, a natural step is to apply successively the operator H and then investigate the effects on the s -convex orders and the associated moment spaces. So, given $t \in \mathbb{N}_0$, let H^t denote the t -th iterate of H ; its construction is simple, although not totally immediate.

4.1 Discrete problem

Given $s \in \mathbb{N}_0$, we start by introducing, similarly to (3.2), a set \mathcal{B}_{s+t} of all probability mass functions P_X on \mathbb{N} with prescribed first $s+t-1$ combinatorial moments (given by c_1, \dots, c_{s+t-1}). This set may be ordered with respect to the $(s+t)$ convex order.

Now, at step 1, consider the set $D_1 = D_1^{(c_1)}$ of all probability mass functions on \mathbb{N} with fixed mean c_1 . By construction, the operator $H \equiv H_1 = H_1^{(c_1)}$ is a one-to-one mapping from D_1 into

a new set $H_1(D_1) \equiv D_2$. By (3.4), the mean in D_2 is equal to c_2/c_1 ; put $D_2 = D_2^{(c_2/c_1)}$. At step 2, the operator $H_2 = H_2^{(c_2/c_1)}$ maps D_2 into $H_2(D_2) \equiv D_3$, and again by (3.4), $D_3 = D_3^{(c_3/c_2)}$. Finally, at step t , $H_t = H_t^{(c_t/c_{t-1})}$ maps D_t into $H_t(D_t) \equiv D_{t+1}$. The operator H^t is then the composition (denoted \circ) of the successive operators H_1, \dots, H_t , i.e.

$$H^t = H_t \circ \dots \circ H_2 \circ H_1 : D_1 \xrightarrow{H_1} D_2 \xrightarrow{H_2} \dots \xrightarrow{H_{t-1}} D_t \xrightarrow{H_t} D_{t+1}. \quad (4.1)$$

Let us apply H^t to the set to any random variable X in \mathcal{B}_{s+t} . Note that the moments c_1, \dots, c_t are well fixed within B_{s+t} . From (2.1) and (4.1), we see that X is mapped to an associated discrete non-negative random variable X_{H^t} defined as follows.

Definition 4.1 X_{H^t} has a probability mass function $H^t(P_X)$ given by

$$H^t(P_X)(j) \equiv P(X_{H^t} = j) = \frac{\bar{F}_t(X, j+t)}{c_t}, \quad j \in \mathbb{N}. \quad (4.2)$$

By (2.5), an equivalent expression is

$$P(X_{H^t} = j) = \frac{\mathbb{E} \binom{X-j-1}{t-1}}{c_t}, \quad j \in \mathbb{N}.$$

We underline that each H_k being one-to-one from D_k to D_{k+1} , H^t provides a one-to-one correspondence on \mathcal{B}_{s+t} . Furthermore, let us introduce, as with (3.3), an order \preceq_{H^t} on $H^t(\mathcal{B}_{s+t})$.

Definition 4.2 Two probability mass functions $Q_1, Q_2 \in H^t(\mathcal{B}_{s+t})$ satisfy $Q_1 \preceq_{H^t} Q_2$ when

$$(H^t)^{-1}(Q_1) \preceq_{s+t} (H^t)^{-1}(Q_2). \quad (4.3)$$

We are now able to generalize Proposition 3.3 to the transform by H^t . The previous notation Q_Y will be adopted here too, and it is convenient to put $c_0 = 1$.

Proposition 4.3

$$H^t(\mathcal{B}_{s+t}) = \left\{ Q_Y : \mathbb{E} \binom{Y}{i} = \frac{c_{i+t}}{c_t}, \quad i = 1, \dots, s-1, \quad \text{and} \right. \quad (4.4)$$

$$\left. (-1)^k \Delta^k Q_Y(j) \geq 0, \quad j \in \mathbb{N}, \quad k = 1, \dots, t, \quad \text{with} \right. \quad (4.5)$$

$$\left. (-1)^{t-k} c_t Q_Y(0) \leq \sum_{l=k-1}^{t-1} (-1)^{l-k+1} c_l, \quad k = 1, \dots, t \right\}, \quad (4.6)$$

and the order \preceq_{H^t} corresponds to the s -convex stochastic order \preceq_s .

Proof. The key tool is the conjugate operator of $(H^t)^{-1}$, built from the conjugate operators of $(H_1)^{-1}, \dots, (H_t)^{-1}$. By arguing as with (3.9), we see that

$$[(H_k)^{-1}]^* = \delta_0 + (c_k/c_{k-1})\Delta, \quad k = 1, \dots, t.$$

From (4.1), we then obtain

$$\begin{aligned} [(H^t)^{-1}]^* &= [(H_1)^{-1}]^* \circ [(H_2)^{-1}]^* \circ \dots \circ [(H_t)^{-1}]^* \\ &= [c_1\Delta + \delta_0] \circ [(c_2/c_1)\Delta + \delta_0] \circ \dots \circ [(c_t/c_{t-1})\Delta + \delta_0]. \end{aligned}$$

Note that, for $t = 2$ for example,

$$[a\Delta + \delta_0] \circ [b\Delta + \delta_0] = ab\Delta^2 + \delta_0 b\Delta + \delta_0,$$

as $\Delta\delta_0 = 0$ and $\delta_0^2 = \delta_0$. So, we get

$$\begin{aligned} [(\mathbf{H}^t)^{-1}]^* &= \sum_{k=0}^t \delta_0^k \prod_{l=k+1}^t [(c_l/c_{l-1})\Delta] \\ &= c_t \Delta^t + \delta_0 \sum_{k=1}^t (c_t/c_k) \Delta^{t-k}. \end{aligned} \quad (4.7)$$

Applying (4.7) to any function f on \mathbb{N} thus gives

$$[(\mathbf{H}^t)^{-1}]^*(f) = c_t \Delta^t f + \sum_{k=1}^t (c_t/c_k) (\Delta^{t-k} f)(0). \quad (4.8)$$

Let us first identify the order $\preceq_{\mathbf{H}^t}$. By (4.3) and (3.8), $Q_1 \preceq_{\mathbf{H}^t} Q_2$ when

$$\langle [(\mathbf{H}^{-1})^*](f), Q_1 \rangle \leq \langle [(\mathbf{H}^{-1})^*](f), Q_2 \rangle \quad \text{for all } f \in \mathcal{F}_{s+t}.$$

From (4.8), this means that

$$\langle \Delta^t f, Q_1 \rangle \leq \langle \Delta^t f, Q_2 \rangle \quad \text{for all } f \in \mathcal{F}_{s+t},$$

which shows the equivalence with the order \preceq_s .

We now determine the set $\mathbf{H}^t(\mathcal{B}_{s+t})$. Let us choose the function $f(j) = \binom{j}{i+t}$, $j \in \mathbb{N}$, for any fixed $i, t \in \mathbb{N}$. Since

$$\Delta^t \binom{j}{i+t} = \binom{j}{i}, \quad j \in \mathbb{N},$$

we have by (4.8) that

$$[(\mathbf{H}^t)^{-1}]^* \binom{j}{i+t} = c_t \binom{j}{i}, \quad j \in \mathbb{N}. \quad (4.9)$$

By (3.8), we then obtain from (4.9) the relations

$$\mathbf{E} \binom{X}{i+t} = c_t \mathbf{E} \binom{X_{\mathbf{H}^t}}{i}, \quad i \in \mathbb{N},$$

which generalize (2.4). The binomial moments in (4.4) then follow. Moreover, from (4.1) and using (3.5), we find that a probability mass function Q_Y necessarily satisfies the restrictions (4.5). This can also be seen from (4.2) as for $k = 1, \dots, t-1$,

$$\Delta^k \mathbf{E} \binom{X-j-1}{t-1} = (-1)^k \mathbf{E} \binom{X-j-1-k}{t-1-k}, \quad j \in \mathbb{N}.$$

So, Q_Y is a t -monotone function in the sense given in the Introduction. Finally, the t conditions (4.6) guarantee that the t successive transforms of P_X are true probability mass functions. Indeed, writing $X_{H^k} = H_k(X_{H^{k-1}})$, $k \geq 1$, with $X_{H^0} = X$, one gets (see (2.2))

$$P(X_{H^{k-1}} = 0) = 1 - (c_k/c_{k-1})P(X_{H^k} = 0),$$

that is

$$c_{k-1}P(X_{H^{k-1}} = 0) = c_{k-1} - c_k P(X_{H^k} = 0), \quad k = 1, \dots, t. \quad (4.10)$$

Note that for each k , $P(X_{H^{k-1}} = 0) \leq 1$ but the non-negativity is not guaranteed so far. By iterating (4.10) $t - k$ times, we then obtain

$$c_{k-1}P(X_{H^{k-1}} = 0) = \sum_{l=k-1}^{t-1} (-1)^{l-k+1} c_l + (-1)^{t-k+1} c_t P(X_{H^t} = 0), \quad k = 1, \dots, t.$$

Thus, the constraints $P(X_{H^{k-1}} = 0) \geq 0$ provide the restrictions (4.6). \diamond

Observe that the image set $H^t(\mathcal{B}_{s+t})$ involves $s - 1$ moment relations, the t -monotonicity and t other technical conditions. Going back to \mathcal{B}_{s+t} through the operator H^{-t} would allow us to remove the monotonicity condition.

4.2 Continuous problem

For $s, t \in \mathbb{N}_0$, consider now the set \mathcal{B}_{s+t} of distributions on \mathbb{R}^+ with prescribed first $s + t - 1$ moments (given by $\mu_1, \dots, \mu_{s+t-1}$).

We construct the operator H^t , the t -th iterate of the operator H defined by (2.7), by following a similar argument as for (4.1). Let $D_1 = D_1^{(\mu_1)}$ the set of all continuous distributions on \mathbb{R}^+ with fixed mean μ_1 . Then, H^t is defined as $H^t = H_t \circ \dots \circ H_2 \circ H_1$ where, for each $k = 1, \dots, t$, H_k is an operator mapping the set D_k into the set $H_k(D_k) = D_{k+1}$. From (2.9), one easily checks the mean in D_{k+1} is equal to $\mu_k/k\mu_{k-1}$. So, $H_k = H_k^{(\mu_k/k\mu_{k-1})}$ and it provides a one-to-one correspondence.

By definition of H^t and using (2.7), we see that any random variable X in \mathcal{B}_{s+t} is mapped to an associated non-negative random variable X_{H^t} whose survival function is given by

$$P(X_{H^t} > x) = \frac{t! \bar{F}_{t+1}(X, x)}{\mu_t}, \quad x \geq 0. \quad (4.11)$$

By (2.10) and (2.11), it can be rewritten as

$$P(X_{H^t} > x) = \frac{E[(X - x)_+^t]}{\mu_t}, \quad x \geq 0,$$

which generalizes (2.7). Here too, H^t is one-to-one on \mathcal{B}_{s+t} . A transformed order \preceq_{H^t} is then defined on $H^t(\mathcal{B}_{s+t})$ by stipulating again the condition (4.3). We adopt the same notation Q_Y and $q_Y(y)$ as before.

Proposition 4.4

$$H^t(\mathcal{B}_{s+t}) = \left\{ Q_Y : E(Y^i) = \frac{\mu_{i+t}}{\binom{i+t}{t} \mu_t}, \quad i = 1, \dots, s-1, \quad \text{and} \right. \quad (4.12)$$

$$\left. (-1)^k \mathcal{D}^{(k)} q_Y(y) \geq 0, \quad y \geq 0, \quad k = 1, \dots, t, \quad \text{with} \right. \quad (4.13)$$

$$\left. (-1)^{t-k} \mathcal{D}^{(t-k)} q_Y(0) \leq \frac{t! \mu_{k-1}}{(k-1)! \mu_t}, \quad k = 1, \dots, t \right\}, \quad (4.14)$$

and the order \preceq_{H^t} corresponds to the s -convex stochastic order \preceq_s .

Proof. Working with conjugate operators, one gets (see (3.13)) that

$$[(H_k)^{-1}]^* = \delta_0 + (\mu_k/k\mu_{k-1})\mathcal{D}, \quad k = 1, \dots, t.$$

Therefore,

$$\begin{aligned} [(H^t)^{-1}]^* &= [\mu_1\mathcal{D} + \delta_0] \circ [(\mu_2/2\mu_1)\mathcal{D} + \delta_0] \circ \dots \circ [(\mu_t/k\mu_{t-1})\mathcal{D} + \delta_0] \\ &= \frac{\mu_t}{t!} \mathcal{D}^t + \delta_0 \sum_{k=1}^t \frac{k! \mu_t}{t! \mu_k} \mathcal{D}^{t-k}. \end{aligned} \quad (4.15)$$

First, consider the function $f(x) = x^{i+t}$, $x \in \mathbb{R}^+$, for fixed $i, t \in \mathbb{N}$. Applying (4.15) to this function and using (3.8), we find that

$$E(X^{i+t}) = \binom{i+t}{t} \mu_t E(X_{H^t}^i), \quad i \in \mathbb{N},$$

hence (4.12). Now, as for (2.8), one has

$$P(X_{H^{k-1}} > x) = (\mu_k/k\mu_{k-1}) q_{H^k}(x), \quad k = 1, \dots, t, \quad (4.16)$$

which gives by differentiation

$$\mu_{k-1} q_{H^{k-1}}(x) = (-1/k) \mathcal{D}\{\mu_k q_{H^k}(x)\}, \quad k = 1, \dots, t. \quad (4.17)$$

By iterating (4.12) $t - k$ more times, we get

$$\mu_{k-1} q_{H^{k-1}}(x) = \frac{\mu_t}{k \dots t} (-1)^{t-k+1} \mathcal{D}^{(t-k+1)} q_{H^t}(x), \quad k = 1, \dots, t. \quad (4.18)$$

Thus, $q_{H^{k-1}} \geq 0$ yields the conditions (4.13). Moreover, (4.16) implies that

$$\mu_k q_{H^k}(0) \leq k\mu_{k-1}, \quad k = 1, \dots, t;$$

inserting (4.18) with $x = 0$ and k instead of $k - 1$ then leads to the constraints (4.14). \diamond

5 Convex extrema for monotone distributions

A further interest of Propositions 4.3 and 4.4 is to point out that the s -convex extrema for t -monotone distributions may be constructed from the simple $s + t$ -convex extrema (i.e. without monotonicity restriction).

Hereafter, we are going to derive the explicit expression of a few extrema for monotone distributions that are defined on a bounded non-negative support. In the discrete case, the random variables are valued in a set $\{0, \dots, n\}$, $n \in \mathbb{N}_0$; in the continuous case, the random variables are valued on an interval $[0, b]$, $b > 0$. We first obtain the $s = 1, 2, 3$ -convex extrema when $t = 1$ (nonincreasing distributions). These were found earlier by Denuit et al. (1999b) in the discrete case and Denuit et al. (1998) in the continuous case. The method of proof used there, however, is quite different as it relies on a Khinchine representation for unimodal distributions. Then, we examine the $s = 1, 2$ -convex extrema when $t = 2$ (convex nonincreasing distributions). These are seen to be the same as with $t = 1$ except for the minimum in the discrete case.

Let us introduce the class $\mathcal{B}_{s,t}(n)$ of all such discrete random variables Y that have prescribed first $s - 1$ moments $\nu_i = E(Y^i)$, $i = 1, \dots, s - 1$, and possess a t -monotone probability mass function Q_Y , i.e. for $k = 1, \dots, t$, $(-1)^k \Delta^k Q_Y(j) \geq 0$, $j = 0, \dots, n - k$. The s -convex extrema in this set are denoted by $Y_{min}^{(s,t)}(n)$ and $Y_{max}^{(s,t)}(n)$. Similarly, $\mathcal{B}_{s,t}(b)$ is the class of all continuous random variables Y with prescribed first $s - 1$ moments ν_i and a t -monotone density function q_Y , i.e. for $k = 1, \dots, t$, $(-1)^k \mathcal{D}^{(k)} q_Y(y) \geq 0$, $y \in [0, b]$. The s -convex extrema in the set are written $Y_{min}^{(s,t)}(b)$ and $Y_{max}^{(s,t)}(b)$.

For the discrete case, we observe that by the definition (2.1) of H , a random variable Y valued on $\{0, \dots, n\}$ is obtained as the transform of a random variable X that is valued on $\{0, \dots, n + 1\}$. This remark will be used in the sequel.

Let us recall that the usual s -convex extrema, i.e. when $t = 0$, inside $\mathcal{B}_s(n)$ or $\mathcal{B}_s(b)$, are explicitly known for $s = 1, 2, 3, 4$. They are recalled in the Appendix and denoted by $X_{min}^{(s)}(n)$, $X_{max}^{(s)}(n)$ and $X_{min}^{(s)}(b)$, $X_{max}^{(s)}(b)$.

1-convex extrema for nonincreasing distributions. This is the problem of extrema when $s = 1$ and $t = 1$. Let us write $\mathcal{U}(n)$ [$\mathcal{U}(b)$] for a uniform distribution on $\{0, \dots, n\}$ ($[0, b]$).

Corollary 5.1 *Inside $\mathcal{B}_{1,1}(n)$ [$\mathcal{B}_{1,1}(b)$],*

$$Y_{min}^{(1,1)}(n) [Y_{min}^{(1,1)}(b)] = 0 \quad a.s., \quad (5.1)$$

$$Y_{max}^{(1,1)}(n) [Y_{max}^{(1,1)}(b)] =_d \mathcal{U}(n) [\mathcal{U}(b)]. \quad (5.2)$$

Proof. For the discrete case, consider the set $H[\mathcal{B}_2(n + 1)]$ defined by (3.4), (3.5), with $n + 1$ substituted for n (see the remark above). We want to keep only the restriction of nonincreasing distributions (i.e. $\Delta Q_Y(j) \leq 0$, $j = 0, \dots, n - 1$). Note that as $s = 1$, there is no moment specification. Moreover, the condition $c_1 Q_Y(0) \leq 1$ will be satisfied by choosing $c_1 = 1$, for instance. Now, by Proposition 3.3, the associated 1-convex extrema are provided by the H -transform (2.1) of the 2-convex extrema in $\mathcal{B}_2(n + 1)$. Using (7.1) and (7.2) where $\xi = 0$, we then get

$$P[Y_{min}^{(1,1)}(n) = j] = P[X_{min}^{(2)}(n + 1) \geq j + 1] = \delta_{j,0}, \quad j = 0, \dots, n,$$

where $\delta_{j,0} = 1$ (resp. 0) if $j = 0$ (> 0), and

$$P[Y_{max}^{(1,1)}(n) = j] = P[X_{max}^{(2)}(n+1) \geq j+1] = 1/(n+1), \quad j = 0, \dots, n,$$

hence the extrema indicated above.

For the continuous case, $H[\mathcal{B}_2(b)]$ defined by (3.11), (3.12) contains now an undesired condition $\mu_1 q_Y(0) \leq 1$. It can be satisfied, however, by choosing $\mu_1 \rightarrow 0$. Then, applying the continuous operator H to the 2-convex minimum gives

$$P[Y_{min}^{(1,1)}(b) \leq x] = \frac{1}{\mu_1} \int_0^x [1 - F_{X_{min}^{(2)}(b)}(y)] dy = \frac{x}{\mu_1}, \quad 0 \leq x \leq \mu_1,$$

i.e. $Y_{min}^{(1,1)}(b)$ is uniform on $[0, \mu_1]$. As μ_1 tends to 0, this minimum degenerates to the single point 0. Using the 2-convex maximum (7.3), we have

$$P[Y_{max}^{(1,1)}(b) \leq x] = \frac{1}{\mu_1} \int_0^x [1 - F_{X_{max}^{(2)}(b)}(y)] dy = \frac{x}{b}, \quad 0 \leq x \leq b,$$

(independently of μ_1), i.e. $Y_{max}^{(1,1)}(b)$ is uniform on $[0, b]$ as announced. \diamond

2-convex extrema for nonincreasing distributions. This time, we work with $s = 2$ and $t = 1$. Denote by $\nu_1 = E(Y)$ the fixed mean in $\mathcal{B}_{2,1}(n)$ (and $H[\mathcal{B}_3(n+1)]$) or $\mathcal{B}_{2,1}(b)$ (and $H[\mathcal{B}_3(b)]$).

Corollary 5.2 *Inside $\mathcal{B}_{2,1}(n)$, putting $\tilde{\xi}$ for the integer in $[0, n-1]$ such that $\tilde{\xi} < 2\nu_1 \leq \tilde{\xi} + 1$,*

$$Y_{min}^{(2,1)}(n) = \begin{cases} 0, \dots, \tilde{\xi} & \text{with equal probabilities } 2(\tilde{\xi} + 1 - \nu_1)/(\tilde{\xi} + 1)(\tilde{\xi} + 2), \\ \tilde{\xi} + 1 & \text{with probability } (2\nu_1 - \tilde{\xi})/(\tilde{\xi} + 2), \end{cases} \quad (5.3)$$

$$Y_{max}^{(2,1)}(n) = \begin{cases} 0 & \text{with probability } 1 - 2\nu_1/(n+1), \\ 1, \dots, n & \text{with equal probabilities } 2\nu_1/n(n+1). \end{cases} \quad (5.4)$$

Inside $\mathcal{B}_{2,1}(b)$, putting \mathcal{I}_0 for a distribution degenerated in 0,

$$Y_{min}^{(2,1)}(b) =_d \mathcal{U}(2\nu_1), \quad (5.5)$$

$$Y_{max}^{(2,1)}(b) =_d (1 - 2\nu_1/b) \mathcal{I}_0 + (2\nu_1/b) \mathcal{U}(b). \quad (5.6)$$

Proof. As indicated in the Appendix, the unconstrained 3-convex extrema are given in terms of the moments $\mu_i = E(X^i)$, $i = 1, 2$. Let us start with the discrete case. By (3.4), $\nu_1 = c_2/c_1$ which implies that $\mu_2 = \mu_1(2\nu_1 + 1)$. For $X_{min}^{(3)}(n+1)$ given by (7.4), we have $\xi_1 - 1 < 2\nu_1 \leq \xi_1$, i.e. $\tilde{\xi} < 2\nu_1 \leq \tilde{\xi} + 1$ by putting $\tilde{\xi} = \xi_1 - 1$. Moreover,

$$X_{min}^{(3)}(n+1) = \begin{cases} 0 & \text{with probability } p_1 = 1 - p_2 - p_3, \\ \tilde{\xi} + 1 & \text{with probability } p_2 = \mu_1(\tilde{\xi} - 2\nu_1 + 1)/(\tilde{\xi} + 1), \\ \tilde{\xi} + 2 & \text{with probability } p_3 = \mu_1(2\nu_1 - \tilde{\xi})/(\tilde{\xi} + 2). \end{cases}$$

For $X_{max}^{(3)}(n+1)$ given by (7.5), we have $\xi_2 < (n-2\nu)/[(n+1)/\mu_1 - 1] \leq \xi_2 + 1$, so that $\xi_2 = 0$ as $\mu_1 \rightarrow 0$. Then,

$$X_{max}^{(3)}(n+1) = \begin{cases} 0 & \text{with probability } p_1 = 1 - p_2 - p_3, \\ 1 & \text{with probability } p_2 = \mu_1(n-2\nu_1)/n, \\ n+1 & \text{with probability } p_3 = \mu_1 2\nu_1/n(n+1). \end{cases}$$

Applying the discrete operator H then leads to the extrema (5.3) and (5.4).

Let us now discuss the continuous case. By (3.11), $\nu_1 = \mu_2/2\mu_1$. We may let $\mu_1 \rightarrow 0$ and choose $\mu_2 = 2\mu_1\nu_1$ accordingly. For the minimum, we get from (7.6) that

$$P[Y_{min}^{(2,1)}(b) \leq x] = (\mu_1/\mu_2)x, \quad 0 \leq x \leq \mu_2/\mu_1,$$

i.e. the result (5.5). For the maximum, (7.7) gives as $\mu_1 \rightarrow 0$

$$X_{max}^{(3)}(b) = \begin{cases} \mu_1(1-2\nu_1/b) & \text{with probability } 1-2\nu_1\mu_1/b^2, \\ b & \text{with probability } 2\nu_1\mu_1/b^2, \end{cases}$$

so that

$$P[Y_{max}^{(2,1)}(b) \leq x] = 1 - 2\nu_1/b + 2\nu_1x/b^2, \quad 0 \leq x \leq b,$$

i.e. the maximum (5.6). \diamond

3-convex extrema for nonincreasing distributions. We have now $s = 3$ and $t = 1$. Let $\nu_1 = E(Y)$ and $\nu_2 = E(Y^2)$ be the first two moments in $\mathcal{B}_{3,1}(n)$ (and $H[\mathcal{B}_4(n+1)]$) or $\mathcal{B}_{3,1}(b)$ (and $H[\mathcal{B}_4(b)]$).

Corollary 5.3 *Inside $\mathcal{B}_{3,1}(n)$, put $\tilde{\xi}_1$ and $\tilde{\xi}_2$ for the integers in $[0, n-1]$ such that*

$$\tilde{\xi}_1 < (3\nu_2 - \nu_1)/2\nu_1 \leq \tilde{\xi}_1 + 1, \quad \text{and} \quad \tilde{\xi}_2 < (2n\nu_1 - 3\nu_2 + \nu_1)/(n-2\nu_1) \leq \tilde{\xi}_2 + 1.$$

Then,

$$Y_{min}^{(3,1)}(n) = \begin{cases} 0 & \text{with probability } \frac{q_3}{\tilde{\xi}_1+2} + \frac{q_2}{\tilde{\xi}_1+1} + q_1, \\ 1, \dots, \tilde{\xi}_1 & \text{with equal probabilities } \frac{q_3}{\tilde{\xi}_1+2} + \frac{q_2}{\tilde{\xi}_1+1}, \\ \tilde{\xi}_1 + 1 & \text{with probability } \frac{q_3}{\tilde{\xi}_1+2}, \end{cases} \quad (5.7)$$

where $q_1 + q_2 + q_3 = 1$ and

$$q_2 = \frac{-3\nu_2 + 3\nu_1 + 2\tilde{\xi}_1\nu_1}{\tilde{\xi}_1}, \quad q_3 = \frac{3\nu_2 - \nu_1 - 2\tilde{\xi}_1\nu_1}{\tilde{\xi}_1 + 1},$$

while

$$Y_{max}^{(3,1)}(n) = \begin{cases} 0, \dots, \tilde{\xi}_2 & \text{with equal probabilities } \frac{X_3}{n+1} + \frac{X_2}{\tilde{\xi}_2+2} + \frac{X_1}{\tilde{\xi}_2+1}, \\ \tilde{\xi}_2 + 1 & \text{with probability } \frac{X_3}{n+1} + \frac{X_2}{\tilde{\xi}_2+2}, \\ \tilde{\xi}_2 + 2, \dots, n & \text{with equal probabilities } \frac{X_3}{n+1}, \end{cases} \quad (5.8)$$

where $\chi_1 + \chi_2 + \chi_3 = 1$ and

$$\chi_1 = \frac{(\tilde{\xi}_2 + 1)(n - 2\nu_1) + 3\nu_2 - \nu_1 - 2n\nu_1}{n - \tilde{\xi}_2}, \quad \chi_2 = \frac{2\nu_1(\tilde{\xi}_2 + n) - 3\nu_2 + \nu_1 - n\tilde{\xi}_2}{n - 1 - \tilde{\xi}_2}.$$

Inside $\mathcal{B}_{3,1}(b)$,

$$Y_{min}^{(3,1)}(b) =_d \frac{3\nu_2 - 4\nu_1^2}{3\nu_2} \mathcal{I}_0 + \frac{4\nu_1^2}{3\nu_2} \mathcal{U}\left(\frac{3\nu_2}{2\nu_1}\right), \quad (5.9)$$

$$Y_{max}^{(3,1)}(b) =_d \frac{3\nu_2 - 4\nu_1^2}{b^2 - 4b\nu_1 + 3\nu_2} \mathcal{U}(b) + \frac{b^2 - 4b\nu_1 + 4\nu_1^2}{b^2 - 4b\nu_1 + 3\nu_2} \mathcal{U}\left(\frac{2b\nu_1 - 3\nu_2}{b - 2\nu_1}\right). \quad (5.10)$$

Proof. In the discrete case, we have, by (3.4), $\nu_1 = c_2/c_1$, which yields $\mu_2 = \mu_1(2\nu_1 + 1)$, and $(\nu_2 - \nu_1)/2 = c_3/c_2$, which implies that $\mu_3 = \mu_1(3\nu_2 + 3\nu_1 + 1)$. Applying (2.1) to $X_{min}^{(4)}(n + 1)$ given by (7.8) yields

$$Y_{min}^{(3,1)}(n) = \begin{cases} 0, \dots, \theta - 1 & \text{with equal probabilities } 1/\mu_1, \\ \theta & \text{with probability } (p_2 + p_3 + p_4)/\mu_1, \\ \theta + 1, \dots, \eta - 1 & \text{with equal probabilities } (p_3 + p_4)/\mu_1, \\ \eta & \text{with probability } p_4/\mu_1. \end{cases}$$

Let us choose $\mu_1 \rightarrow 0$. To avoid absurdity, we must have $\theta = 0$. Moreover, we then find, after some simplifications, that

$$\begin{cases} p_2/\mu_1 = (3\nu_2 + \nu_1 - 4\nu_1\eta + \eta(\eta - 1))/\eta(\eta + 1), \\ p_3/\mu_1 = (-3\nu_2 + \nu_1 + 2\nu_1\eta)/\eta(\eta - 1), \\ p_4/\mu_1 = (3\nu_2 + \nu_1 - 2\nu_1\eta)/\eta(\eta - 1). \end{cases}$$

The condition $p_4 > 0$ implies that $\eta < 1 + (3\nu_2 - \nu_1)/2\nu_1$, and $p_3 \geq 0$ yields $\eta \geq (3\nu_2 - \nu_1)/2\nu_1$. Putting $\eta = 1 + \tilde{\xi}_1$, we then obtain for $\tilde{\xi}_1$ the value announced above.

Note that the condition $p_2 \geq 0$ is always verified. Indeed, this condition means that $3\nu_2 - 3\nu_1 - 4\tilde{\xi}_1\nu_1 + \tilde{\xi}_1(\tilde{\xi}_1 + 1) \geq 0$. It can be rewritten as $\tilde{\nu}_2 - (2\tilde{\xi}_1 + 1)\tilde{\nu}_1 + \tilde{\xi}_1(\tilde{\xi}_1 + 1) \geq 0$ where $\tilde{\nu}_1 = 2\nu_1$ and $\tilde{\nu}_2 = 3\nu_2 - \nu_1$. This is verified if $\tilde{\nu}_1$ and $\tilde{\nu}_2$ are the first two moments of some random variable on $\{0, \dots, n\}$, since $i^2 - (2\tilde{\xi}_1 + 1)i + \tilde{\xi}_1(\tilde{\xi}_1 + 1) \geq 0$ for all $0 \leq i \leq n$. In fact, the random variable Z in Denuit et al. (1999c) has precisely its first two moments equal to these $\tilde{\nu}_1$ and $\tilde{\nu}_2$; see formula (5.7) in that paper. So, we finally get the above result (5.7).

Consider $X_{max}^{(4)}(n + 1)$ given by (7.9). First, we observe that $(\mu_2n - \mu_3)/(\mu_1n - \mu_2) = 1 + (2n\nu_1 - 3\nu_2 + \nu_1)/(n - 2\nu_1)$; thus, $\zeta = \tilde{\xi}_2 + 1$ where $\tilde{\xi}_2$ is defined as above. Applying (2.1), we then get that

$$Y_{max}^{(3,1)}(n) = \begin{cases} 0, \dots, \tilde{\xi}_2 & \text{with equal probabilities } (p_2 + p_3 + p_4)/\mu_1, \\ \tilde{\xi}_2 + 1 & \text{with probability } (p_3 + p_4)/\mu_1, \\ \tilde{\xi}_2 + 2, \dots, n & \text{with equal probabilities } p_4/\mu_1. \end{cases}$$

Finally, one can check that $p_2/\mu_1 = \chi_1/(\tilde{\xi}_2 + 1)$, $p_3/\mu_1 = \chi_2/(\tilde{\xi}_2 + 2)$ and $p_4/\mu_1 = \chi_3/(n + 1)$, hence the result (5.8).

In the continuous case, (3.11) gives $\nu_1 = \mu_2/2\mu_1$ and $\nu_2 = \mu_3/3\mu_1$. So, we choose $\mu_2 = 2\mu_1\nu_1$ and $\mu_3 = 3\mu_1\nu_2$, with $\mu_1 \rightarrow 0$. For the minimum, we see from (7.10) that

$$X_{min}^{(4)}(b) = \begin{cases} r_- \rightarrow 0 & \text{with probability } (r_+ - \mu_1)/(r_+ - r_-) \rightarrow 1, \\ r_+ \rightarrow 3\nu_2/\nu_1 & \text{with probability } (\mu_1 - r_-)/(r_+ - r_-) \rightarrow 0. \end{cases}$$

Moreover, one can check that

$$r_-/\mu_1 \rightarrow (3\nu_2 - 4\nu_1^2)/3\nu_2,$$

implying also that $(\mu_1 - r_-)/(r_+ - r_-)\mu_1 \rightarrow (4\nu_1^2/3\nu_2)(2\nu_1/3\nu_2)$. All this then leads to the result (5.9) for $Y_{min}^{(3,1)}(b)$. For the maximum, we have by (7.11) that

$$X_{max}^{(4)}(b) = \begin{cases} 0 & \text{with probability } p_1 = 1 - p_2 - p_3, \\ \frac{2b\nu_1 - 3\nu_2}{b - 2\nu_1} & \text{with probability } p_2 = \mu_1 \frac{b - 2\nu_1}{2b\nu_1 - 3\nu_2} \frac{(b - 2\nu_1)^2}{b^2 - 4b\nu_1 + 3\nu_2}, \\ b & \text{with probability } p_3 = \mu_1 \frac{1}{b} \frac{3\nu_2 - 4\nu_1^2}{b^2 - 4b\nu_1 + 3\nu_2}. \end{cases}$$

The result (5.10) for $Y_{max}^{(3,1)}(b)$ then easily follows. \diamond

1-convex extrema for nonincreasing convex distributions. Here, $s = 1$ and $t = 2$. We observe that the 1-convex extrema obtained before for nonincreasing distributions are convex. As a consequence, they are extremal too for nonincreasing convex distributions.

2-convex extrema for nonincreasing convex distributions. This time, $s = 2$ and $t = 2$. The 2-convex extrema for nonincreasing distributions are again convex, except the minimum in the discrete case. Thus, apart this case which is discussed below, they remain extremal for nonincreasing convex distributions. Let $\nu_1 = E(Y)$ be the fixed mean in the set $\mathcal{B}_{2,2}(n)$ (and $H^2[\mathcal{B}_4(n+1)]$).

Corollary 5.4 *For the minimum inside $\mathcal{B}_{2,2}(n)$, putting $\tilde{\eta}$ for the integer in $[2, n]$ such that $\tilde{\eta} < 2 + 3\nu_1 \leq \tilde{\eta} + 1$,*

$$Y_{min}^{(2,2)}(n) = j \text{ with probability } (\tilde{\eta} - j - 1)\pi_1 + (\tilde{\eta} - j)\pi_2, \quad j = 0, \dots, \tilde{\eta} - 1, \quad (5.11)$$

where

$$\pi_1 = \frac{2(\tilde{\eta} - 1 - 3\nu_1)}{\tilde{\eta}(\tilde{\eta} - 1)}, \quad \text{and} \quad \pi_2 = \frac{2(2 + 3\nu_1 - \tilde{\eta})}{\tilde{\eta}(\tilde{\eta} + 1)}.$$

Proof. In (4.6), the two constraints (for $k = 1, 2$) yield the condition $0 \leq c_1 - c_2Q(0) \leq 1$. From now, we assume that $\mu_1 \rightarrow 0$. To satisfy the previous condition, one may choose, for instance, $E(X^2) = 2\mu_1$. Moreover, by (4.4), $\nu_1 = c_3/c_2$, i.e. $3(1 + \nu_1) = E(X^3 - X)/E(X^2 - X)$. Given our choice of $E(X^2)$, one then gets $E(X^3) = (4 + 3\nu_1)\mu_1$. Finally, as $\mu_1 \rightarrow 0$, we use the approximations $\mu_2 \simeq E(X^2)$ and $\mu_3 \simeq E(X^3)$.

Consider the 4-convex minimum given by (7.8). Under the first three moments above, one sees that $\theta = 0$ and the minimum then becomes

$$X_{min}^{(4)}(n+1) = \begin{cases} 0 & \text{with probability } p_1 = 1 - (\eta^2 - \eta + 3\nu_1 + 1)\mu_1/\eta(\eta + 1), \\ 1 & \text{with probability } p_2 = (\eta^2 - 3\eta + 2 + 3\nu_1)\mu_1/\eta(\eta - 1), \\ \eta & \text{with probability } p_3 = (\eta - 1 - 3\nu_1)\mu_1/\eta(\eta - 1), \\ \eta + 1 & \text{with probability } p_4 = (2 + 3\nu_1 - \eta)\mu_1/\eta(\eta + 1). \end{cases}$$

Furthermore, $p_1 \geq 0$ (as $\mu_1 \rightarrow 0$), $p_3 \geq 0$ means $1 + 3\nu_1 \leq \eta$, and $p_4 > 0$ requires $\eta < 2 + 3\nu_1$, which implies $p_2 > 0$ (since $\eta \geq 2$). All this leads to the announced value $\tilde{\eta}$.

By (4.2), the operator H^2 provides the distribution $P(Y = j) = [P(X = j + 2) + 2P(X = j + 3) + \dots + (n - j)P(X = n + 1)]/c_2$, $j = 0, \dots, n$. Applied to $X_{min}^{(4)}(n + 1)$, it gives, as $\mu_1 \rightarrow 0$, $P[Y_{min}^{(2,2)}(n) = \tilde{\eta} - 1] = p_4/c_2$, $P[Y_{min}^{(2,2)}(n) = \tilde{\eta} - 2] = (p_3 + 2p_4)/c_2$, $P[Y_{min}^{(2,2)}(n) = \tilde{\eta} - 3] = (2p_3 + 3p_4)/c_2$ and in general,

$$P[Y_{min}^{(2,2)}(n) = j] = (\tilde{\eta} - j - 1)p_3/c_2 + (\tilde{\eta} - j)p_4/c_2, \quad j = 0, \dots, \tilde{\eta} - 1.$$

By construction, $c_2 = \mu_1/2$, so that $p_3/c_2 = \pi_1$ and $p_4/c_2 = \pi_2$ as stated in (5.11). \diamond

6 Some numerical illustrations

We present hereafter three applications, somewhat non-standard, in insurance and biostatistics. More traditional illustrations in ruin theory and life insurance (as in Denuit and Lefèvre (1997) and Denuit et al. (1998), (1999b)) could be considered too.

(i) Solvency Capital Requirements for large claims. The new European regulation, Solvency II, to be in force in October 2012, imposes insurance companies to hold enough capital to deal with unfavorable events. The Directive was voted in April 2009 but many implementation measures are still being discussed currently. Some stakeholders recommend the creation of a so-called large loss module in order to cover claim amounts that are larger than attritional ones (but not of catastrophic type). First, for each risk, some buffer capital called Solvency Capital Requirement (SCR) has to be held. Then, at the company or group level, some aggregation formula takes diversification into account and defines the Basic Solvency Capital Requirement that the company must satisfy as

$$\sqrt{\sum_i \sum_j \rho_{ij} \text{SCR}_i \text{SCR}_j},$$

where SCR_i represents the Solvency Capital Requirement for the business line i and ρ_{ij} is a correlation-type parameter (without precise statistical meaning) between the lines i and j .

For each risk, the Solvency Capital Requirement is generally defined as the difference between a certain Value-at-Risk of the random loss and the expected loss. In most cases, a simplification is used and the SCR is defined by $\text{SCR} = q\sigma$ where σ is the standard error of the random loss and $q > 0$ is called a quantile factor. So, $q = 3$ is usually chosen for claim amounts with a moderate tail distribution (as it is in many Solvency II modules). For heavy-tailed risks, a more relevant value is $q = 5$ or 6 ; hereafter, we will take $q = 6$. A collective risk model is often adopted to describe the occurrence of large claims. Thus, the number of claims is a counting random variable N , and the large claim amounts are independent identically distributed random variables (distributed as W , say), independently of N . By the variance decomposition formula, the variance of the large loss aggregated claim amount is given by

$$\sigma^2 = \text{Var}(W)\text{E}(N) + [\text{E}(W)]^2 \text{Var}(N).$$

Various specialized softwares or internal models provide an estimation of the distribution of W and the mean $\text{E}(N)$. As only limited information is available, N is usually assumed to have

a Poisson distribution. For example, with respect to some French data, one could use a Poisson distribution with parameter $\lambda(C_{27}) = 0.37$ for the business line C_{27} (Drought and earthquake) and $\lambda(C_{35}) = 0.69$ for the business line C_{35} (Construction - damages to building). The Poisson assumption is partially satisfying as, in practice, N is usually over-dispersed ($\text{Var}(N)$ is significantly larger than $E(N)$) and only a bounded number of large claims are susceptible to occur (practitioners consider that observing more than a certain number n claims corresponds to a catastrophe). So, for the two previous business lines, one could set, for instance, $n(C_{27}) = 10$ and $n(C_{35}) = 20$.

A property of the Poisson distribution (with parameter λ) is that its probability mass function is nonincreasing convex if $\lambda \leq 2 - \sqrt{2} \simeq 0.5858$, and nonincreasing but not convex if $2 - \sqrt{2} < \lambda \leq 1$. Note that for our illustration, $\lambda(C_{27}) = 0.37 < 2 - \sqrt{2} < \lambda(C_{35}) = 0.69 < 1$. In fact, this shape for a counting distribution is quite frequent for standard (non-catastrophic) risks. Instead of a Poisson distribution, one could prefer to use a distribution with the same mean that is nonincreasing (convex) and has a bounded support $\{0, \dots, n\}$. The choice of such a distribution, however, is not an easy task. It is thus interesting to dispose of upper and lower bounds for the Solvency Capital Requirement.

The bounds for $\text{SCR} = q\sigma$ are simply given by

$$\text{SCR}(N_{\min}^{(2,2)}) = q\sqrt{\text{Var}(W)E(N) + [E(W)]^2 \text{Var}(N_{\min}^{(2,2)})},$$

and

$$\text{SCR}(N_{\max}^{(2,2)}) = q\sqrt{\text{Var}(W)E(N) + [E(W)]^2 \text{Var}(N_{\max}^{(2,2)})}.$$

These provide a prudent version of the SCR and allow us to situate any particular choice for the distribution of N (within the class of interest). Let us fix the mean $\nu_1 = E(N) = \lambda$. From (7.1), (7.2), one gets (writing $N^{(2,0)} \equiv N^{(2)}$)

$$\begin{aligned} \text{Var}(N_{\min}^{(2,0)}) &= \nu_1 - \xi(1 - 2\nu_1 + \xi) - \nu_1^2, \\ \text{Var}(N_{\max}^{(2,0)}) &= n\nu_1 - \nu_1^2. \end{aligned}$$

From (5.3), (5.4),

$$\begin{aligned} \text{Var}(N_{\min}^{(2,1)}) &= [-\tilde{\xi}^3 + \tilde{\xi}^2(-3 + 4\nu_1) + \tilde{\xi}(-2 + 11\nu_1 - 3\nu_1^2) + 6\nu_1(1 - \nu_1)]/3(\tilde{\xi} + 2), \\ \text{Var}(N_{\max}^{(2,1)}) &= \nu_1(2n + 1)/3 - \nu_1^2, \end{aligned}$$

and from (5.11) with the remark before Corollary 5.4,

$$\begin{aligned} \text{Var}(N_{\min}^{(2,2)}) &= (\tilde{\eta} - 1)(6\nu_1 + 2 - \tilde{\eta})/6 - \nu_1^2, \\ \text{Var}(N_{\max}^{(2,2)}) &= \text{Var}(N_{\max}^{(2,1)}), \end{aligned}$$

(as $1 + 2^2 + \dots + n^2 = n(n+1)(2n+1)/6$ and $1 + 2^3 + \dots + n^3 = n^2(n+1)^2/4$).

For $n(C_{27}) = 10$, $\lambda(C_{27}) = 0.37$, one has $\xi = 0$, $\tilde{\xi} = 0$, $\tilde{\eta} = 3$ and $\text{Var}(N_{\min}^{(2,1)}) = \text{Var}(N_{\min}^{(2,0)})$. For $n(C_{35}) = 20$, $\lambda(C_{35}) = 0.69$, then $\xi = 0$, $\tilde{\xi} = 1$, $\tilde{\eta} = 4$ and $\text{Var}(N_{\min}^{(2,1)}) > \text{Var}(N_{\min}^{(2,0)})$. Concerning the associated claim amounts, let us choose, for instance, $E[W(C_{27})] = 1000$, $\text{Var}[W(C_{27})] = 2500^2$ and $E[W(C_{35})] = 2000$, $\text{Var}[W(C_{35})] = 7000^2$ (in thousands of euros).

We present some numerical results in Tables 1 and 2 below. Table 1 shows that the Poisson case is much closer to the lower bounds than the upper bounds. This means that the Poisson assumption may lead to serious underestimations if one uses it by default. Besides, the upper bound is significantly improved when the distribution is known to be nonincreasing. The * indicates that the corresponding Poisson distribution is not convex; so, the bound given there is questionable if the Poisson distribution shape is taken for granted. From Table 2, one observes that, as expected, the choice of the maximal value n influences considerably the value of the variance of N ; its effect on the value of the SCR is, however, less important by comparison. Note that the lower bound does not depend on n .

	Line C_{27}	Line C_{35}
λ	$\lambda(C_{27}) = 0.37$	$\lambda(C_{35}) = 0.69$
n	$n(C_{27}) = 10$	$n(C_{35}) = 20$
$\text{Var}(N_{\min}^{(2,0)})$	0.2331	0.2139
$\text{Var}(N_{\min}^{(2,1)})$	0.2331	0.4672
$\text{Var}(N_{\min}^{(2,2)})$	0.2698	0.5939*
$\text{Var}(N)$ (Poisson case)	0.37	0.69
$\text{Var}(N_{\max}^{(2,2)})$	2.4531	8.9539*
$\text{Var}(N_{\max}^{(2,1)})$	2.4531	8.9539
$\text{Var}(N_{\max}^{(2,0)})$	3.5631	13.3239
$\text{SCR}(N_{\min}^{(2,0)})$	9573.0	35326.5
$\text{SCR}(N_{\min}^{(2,1)})$	9573.0	35839.0
$\text{SCR}(N_{\min}^{(2,2)})$	9641.7	36092.7*
SCR (Poisson case)	9827.0	36283.9
$\text{SCR}(N_{\max}^{(2,2)})$	13098.2	50065.2*
$\text{SCR}(N_{\max}^{(2,1)})$	13098.2	50065.2
$\text{SCR}(N_{\max}^{(2,0)})$	14543.8	55998.2

Table 1: Bounds on $\text{Var}(N)$ and SCR for C_{27} and C_{35} when $s = 2$ and $t = 0, 1, 2$.

$n(C_{27})$	$\text{Var}(N_{\max}^{(2,2)})$	$\text{SCR}(N_{\max}^{(2,2)})$
5	1.220	11276.6
10	2.453	13098.2
20	4.920	16135.7
30	7.386	18685.9
40	9.853	20927.5

Table 2: Bounds on $\text{Var}(N)$ and SCR for C_{27} in function of n when $s = t = 2$.

(ii) **Residual lifetime at high ages.** Yearly mortality rates at high ages (larger than 100) are difficult to estimate. Nowadays, mortality rates at age 100 in France are close to 0.36.

Until age 105, they slightly increase at a rate of around 0.015 per year; they tend to become stable after age 105. There is still some uncertainty about the pattern of mortality rates at high ages. Concerning, however, the residual lifetime floor at age 100, denoted by T_{100} , its distribution is observed to have a nonincreasing convex form whatever the pattern of mortality rates. This is shown in Figure 1 (resp. Figure 2) where the annual mortality rates increase of 0.015 -Hypothesis H1- (resp. remain constant -Hypothesis H2-); the life expectancy is then 1.68 (resp. 1.86) years. Note that at a lower age of 95, the mortality rate is of 0.20 with an increase of 0.03 per year up to age 105 say, after which the rate is taken constant -Hypothesis H3-; then, the residual lifetime floor T_{95} has a nonincreasing distribution which is no longer convex (see Figure 3).

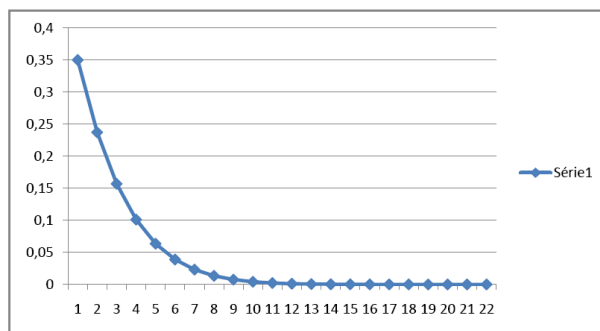


Figure 1: Probability mass function of T_{100} under H1.

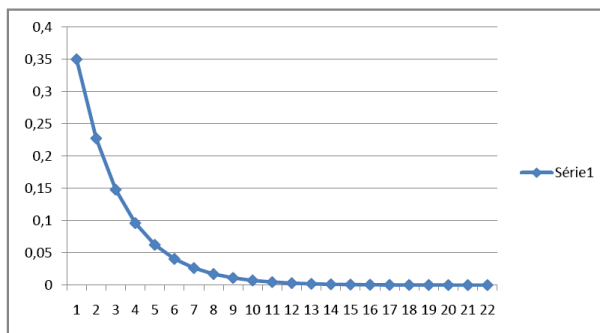


Figure 2: Probability mass function of T_{100} under H2.

Provided the residual lifetime floor has a nonincreasing (convex) distribution, its variance can be bounded by using again the previous extrema. Considering T_{100} , we choose $n = 15$ as the probability to live more than 115 years is extremely small (positive of course). Results are presented in Table 3 for both Hypotheses 1 and 2. We observe that the lower bounds are not very far from the empirical case, especially when information on the shape of the distribution is provided.

(iii) Number of asthma exacerbations. Asthma exacerbation is another term for an asthma attack in which the bronchial tubes through which air flows to the lungs suddenly tighten and become constricted. This makes it extremely difficult to breathe, resulting in an asthma exacerbation. There are many things that can trigger an asthma exacerbation and a

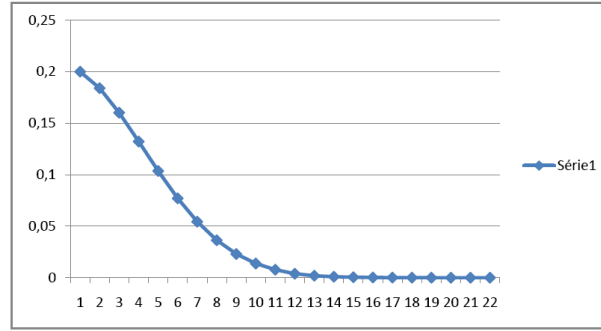


Figure 3: Probability mass function of T_{95} under $H3$.

Hypothesis	H1	H2
$E(T_{100})$	1.68	1.86
$\text{Var}(T_{100,min}^{(2,0)})$	0.22	0.22
$\text{Var}(T_{100,min}^{(2,1)})$	1.58	1.83
$\text{Var}(T_{100,min}^{(2,2)})$	2.26	2.69
$\text{Var}(T_{100})$ (empirical case)	3.77	5.27
$\text{Var}(T_{100,max}^{(2,2)})$	14.54	15.73
$\text{Var}(T_{100,max}^{(2,1)})$	14.54	15.73
$\text{Var}(T_{100,max}^{(2,0)})$	22.38	24.39

Table 3: Bounds on $\text{Var}(T_{100})$ when $s = 2$ and $t = 0, 1, 2$.

quick treatment is necessary to prevent a medical emergency. A study involving 202 patients during one year has given the following empirical distribution (Table 4) for the number N of asthma exacerbations. This distribution is nonincreasing but is easily seen to be non-convex. There is almost no overdispersion as $E(N) \simeq 1.1832$ and $\text{Var}(N) \simeq 1.1852$.

Given the observed mean of N , we computed the probability mass functions of the extrema for arbitrary or nonincreasing distributions valued on $\{0, \dots, n = 5\}$ (using (7.1),(7.2) and (5.3),(5.4), respectively). The results are presented in Table 5. Moreover, bounds for the variance of N are provided in Table 6: for $n = 5$, the bounds are rather narrow and for $n = 10$, the upper bound becomes much higher (as expected).

Values j of N	Frequencies	Proportions
0	67	0.33168
1	60	0.29703
2	53	0.26238
3	16	0.07921
4	5	0.02475
5	1	0.00495

Table 4: Empirical distribution for N .

j	$N_{min}^{(2,0)}$	$N_{max}^{(2,0)}$	$N_{min}^{(2,1)}$	$N_{max}^{(2,1)}$
0	0	0.7634	0.3028	0.60560
1	0.8168	0	0.3028	0.07888
2	0.1832	0	0.3028	0.07888
3	0	0	0.0916	0.07888
4	0	0	0	0.07888
5	0	0.2360	0	0.07888

Table 5: Probability mass functions of the extrema when $s = 2$ and $t = 0, 1$, with $n = 5$.

	$n = 5$	$n = 10$
$\text{Var}(N_{min}^{(2,0)})$	0.1496	0.1496
$\text{Var}(N_{min}^{(2,1)})$	0.6942	0.6942
$\text{Var}(N)$ (empirical case)	1.1852	1.1852
$\text{Var}(N_{max}^{(2,1)})$	2.9384	6.8823
$\text{Var}(N_{max}^{(2,0)})$	5.9259	11.8519

Table 6: Bounds on $\text{Var}(N)$ for $s = 2$ and $t = 0, 1$, with $n = 5$ or 10.

7 Appendix: convex extrema

The s -convex extrema are known under explicit form when $s = 1, 2, 3, 4$ (see the references given in the Introduction). They are recalled below, for discrete random variables valued in a set $\{0, \dots, n\}$, $n \in \mathbb{N}_0$, and for real-valued random variables with support in an interval $[0, b]$, $b > 0$. We denote by $\mathcal{B}_s(n)$ and $\mathcal{B}_s(b)$ the classes of all such discrete and real-valued random variables X that have prescribed first $s - 1$ moments $\mu_i = E(X^i)$, $i = 1, \dots, s - 1$.

1-convex extrema. Inside $\mathcal{B}_1(n)$ [$\mathcal{B}_1(b)$], $X_{min}^{(1)}(n) = 0$ [$X_{min}^{(1)}(b) = 0$] a.s., and $X_{max}^{(1)}(n) = n$ [$X_{max}^{(1)}(b) = b$] a.s..

2-convex extrema. Inside $\mathcal{B}_2(n)$, let ξ be the integer in $[0, n - 1]$ such that $\xi < \mu_1 \leq \xi + 1$.

$$X_{min}^{(2)}(n) = \begin{cases} \xi & \text{with probability } \xi + 1 - \mu_1, \\ \xi + 1 & \text{with probability } \mu_1 - \xi, \end{cases} \quad (7.1)$$

$$X_{max}^{(2)}(n) = \begin{cases} 0 & \text{with probability } 1 - \mu_1/n, \\ n & \text{with probability } \mu_1/n. \end{cases} \quad (7.2)$$

Inside $\mathcal{B}_2(b)$, $X_{min}^{(2)}(b) = \mu_1$ a.s., and

$$X_{max}^{(2)}(b) = \begin{cases} 0 & \text{with probability } 1 - \mu_1/b, \\ b & \text{with probability } \mu_1/b. \end{cases} \quad (7.3)$$

3-convex extrema. Inside $\mathcal{B}_3(n)$, let ξ_1 and ξ_2 be the integers in $[0, n - 1]$ such that $\xi_1 < \mu_2/\mu_1 \leq \xi_1 + 1$, and $\xi_2 < (n\mu_1 - \mu_2)/(n - \mu_1) \leq \xi_2 + 1$.

$$X_{min}^{(3)}(n) = \begin{cases} 0 & \text{with probability } p_1 = 1 - p_2 - p_3, \\ \xi_1 & \text{with probability } p_2 = [(\xi_1 + 1)\mu_1 - \mu_2]/\xi_1, \\ \xi_1 + 1 & \text{with probability } p_3 = (\mu_2 - \xi_1\mu_1)/(1 + \xi_1), \end{cases} \quad (7.4)$$

$$X_{max}^{(3)}(n) = \begin{cases} \xi_2 & \text{with probability } p_1 = 1 - p_2 - p_3, \\ \xi_2 + 1 & \text{with probability } p_2 = \frac{(n + \xi_2)\mu_1 - \mu_2 - n\xi_2}{n - 1 - \xi_2}, \\ n & \text{with probability } p_3 = \frac{(1 + \xi_2)(\xi_2 - \mu_1) + \mu_2 - \mu_1\xi_2}{(n - \xi_2)(n - 1 - \xi_2)}. \end{cases} \quad (7.5)$$

Inside $\mathcal{B}_3(b)$,

$$X_{min}^{(3)}(b) = \begin{cases} 0 & \text{with probability } (\mu_2 - \mu_1^2)/\mu_2, \\ \mu_2/\mu_1 & \text{with probability } \mu_1^2/\mu_2, \end{cases} \quad (7.6)$$

$$X_{max}^{(3)}(b) = \begin{cases} \frac{b\mu_1 - \mu_2}{b - \mu_1} & \text{with probability } \frac{(b - \mu_1)^2}{(b - \mu_1)^2 + \mu_2 - \mu_1^2}, \\ b & \text{with probability } \frac{\mu_2 - \mu_1^2}{(b - \mu_1)^2 + \mu_2 - \mu_1^2}. \end{cases} \quad (7.7)$$

4-convex extrema. Inside $\mathcal{B}_4(n)$, let θ and η be integers in $[0, n - 1]$ with $\theta + 1 < \eta$ such that $\alpha(\theta, \eta)$, $\alpha(\eta, \theta)$, $-\alpha(\theta - 1, \eta)$ and $-\alpha(\eta - 1, \theta) \geq 0$, where $\alpha(\theta, \eta) \equiv -\mu_3 + \mu_2(\theta + 2 + 2\eta) - \mu_1[(\theta + 1)\eta + (\theta + 1)(\eta + 1) + \eta(\eta + 1)] + (\theta + 1)\eta(\eta + 1)$. Let also ζ be the integer in $[1, n - 2]$

such that $\zeta < (\mu_2 n - \mu_3)/(\mu_1 n - \mu_2) \leq \zeta + 1$.

$$X_{min}^{(4)}(n) = \begin{cases} \theta & \text{with probability } p_1 = \alpha(\theta, \eta)/(\eta - \theta)(\eta - \theta + 1), \\ \theta + 1 & \text{with probability } p_2 = -\alpha(\theta - 1, \eta)/(\eta - \theta)(\eta - \theta - 1), \\ \eta & \text{with probability } p_3 = \alpha(\eta, \theta)/(\eta - \theta)(\eta - \theta - 1), \\ \eta + 1 & \text{with probability } p_4 = -\alpha(\eta - 1, \theta)/(\eta - \theta)(\eta - \theta + 1), \end{cases} \quad (7.8)$$

$$X_{max}^{(4)}(n) = \begin{cases} 0 & \text{with probability } p_1 = 1 - p_2 - p_3 - p_4, \\ \zeta & \text{with probability } p_2 = \frac{n\mu_1(\zeta+1) - \mu_2(\zeta+1+n) + \mu_3}{\zeta(n-\zeta)}, \\ \zeta + 1 & \text{with probability } p_3 = \frac{\mu_2(\zeta+n) - n\mu_1\zeta - \mu_3}{(\zeta+1)(n-\zeta-1)}, \\ n & \text{with probability } p_4 = \frac{\mu_3 - \mu_2(2\zeta+1) + \mu_1\zeta(\zeta+1)}{n(n-\zeta)(n-\zeta-1)}. \end{cases} \quad (7.9)$$

Inside $\mathcal{B}_4(b)$, put $r_-(r_+) = \{\mu_3 - \mu_1\mu_2 - (+) [(\mu_3 - \mu_1\mu_2)^2 - 4(\mu_2 - \mu_1^2)(\mu_1\mu_3 - \mu_2^2)]^{1/2}\}/2(\mu_2 - \mu_1^2)$.

$$X_{min}^{(4)}(b) = \begin{cases} r_- & \text{with probability } (r_+ - \mu_1)/(r_+ - r_-), \\ r_+ & \text{with probability } (\mu_1 - r_-)/(r_+ - r_-), \end{cases} \quad (7.10)$$

$$X_{max}^{(4)}(b) = \begin{cases} 0 & \text{with probability } p_1 = 1 - p_2 - p_3, \\ \frac{b\mu_2 - \mu_3}{b\mu_1 - \mu_2} & \text{with probability } p_2 = \frac{(b\mu_1 - \mu_2)^3}{(b\mu_2 - \mu_3)(b^2\mu_1 - 2b\mu_2 + \mu_3)}, \\ b & \text{with probability } p_3 = \frac{\mu_1\mu_3 - \mu_2^2}{b(b^2\mu_1 - 2b\mu_2 + \mu_3)}. \end{cases} \quad (7.11)$$

Acknowledgement

We are grateful to S. Utev for a number of helpful discussions and advice, and to M. Mesfioui for useful comments on a preliminary draft. We also thank the Referee for many remarks and suggestions. This work was conducted while C. Lefèvre was visiting the I.S.F.A., Université de Lyon. It was partly funded by an ANR research project (ANR-08-BLAN-0314-01).

References

- Cheng, Y., Pai, J.S., 2003. On the n th stop-loss transform order of ruin probability. *Insurance: Mathematics and Economics* 32, 51-60.
- Courtois, C., Denuit, M., Van Belleghem, S. (2006). Discrete s -convex extremal distributions: theory and applications. *Applied Mathematics Letters* 19, 1367-177.
- Cox, D.R., 1972. *Renewal Theory*. Methuen, London.
- Denuit, M., Lefèvre, C., 1997. Some new classes of stochastic order relations among arithmetic random variables, with applications in actuarial sciences. *Insurance: Mathematics and Economics* 20, 197-213.
- Denuit, M., Lefèvre, C., Shaked, M., 1998. The s -convex orders among real random variables, with applications. *Mathematical Inequalities and Applications* 1, 585-613.
- Denuit, M., Lefèvre, C., Utev, S., 1999a. Stochastic orderings of convex/concave type on an arbitrary grid. *Mathematics of Operations Research* 24, 585-613.
- Denuit, M., De Vylder, F.E., Lefèvre, C., 1999b. Extremal generators and extremal distributions for the continuous s -convex stochastic orderings. *Insurance: Mathematics and Economics* 24, 201-217.

Denuit, M., Lefèvre, C., Mesfioui, M., 1999c. On s -convex stochastic extrema for arithmetic risks. *Insurance: Mathematics and Economics* 25, 143-155.

De Vylder, F.E., 1996. *Advanced Risk Theory. A Self-Contained Introduction*. Editions de l'Université Libre de Bruxelles - Swiss Association of Actuaries, Bruxelles

Fishburn, P.C., Lavalley, I.H., 1995. Stochastic dominance on unidimensional grids. *Mathematics of Operation Research* 20, 513-525.

Goovaerts, M.J., Kaas, R., Van Heerwaarden, A.E., Bauwelinckx, T., 1990. *Effective Actuarial Methods*. North-Holland, Amsterdam.

Hoeffding, W., 1955. The extrema of the expected value of a function of independent random variables. *Annals of Mathematical Statistics* 26, 268-275.

Hürlimann, W., 1999. *Extremal Moment Methods and Stochastic Orders. Applications in Actuarial Sciences*. Monograph.

Jansen, K., Haezendonck, J., Goovaerts, M.J., 1986. Upper bounds on stop-loss premiums in case of known moments up to the fourth order. *Insurance: Mathematics and Economics* 5, 315-334.

Kaas, R., Van Heerwaarden, A.E., Goovaerts, M.J., 1994. *Ordering of Actuarial Risks*. CAIRE, Brussels.

Karlin, S., Studden, W.J., 1966. *Tchebycheff Systems: With Applications in Analysis and Statistics*. Wiley, New York.

Keilson J., Gerber, H.U., 1971. Some results for discrete unimodality. *Journal of the American Statistical Association* 66, 386-389.

Kemperman, J.H.B., 1968. The general moment problem, a geometric approach. *Annals of Mathematical Statistics* 39, 93-122.

Lefèvre, C., Picard, P., 1993. An unusual stochastic order relation with some applications in sampling and epidemic theory. *Advances in Applied Probability* 25, 63-81.

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

Levy, H., 1992. Stochastic dominance and expected utility: Survey and analysis. *Management Science* 38, 555-593.

Müller, A., Stoyan, D., 2002. *Comparison Methods for Stochastic Models and Risks*. Wiley, New York.

Prékopa, A., 1990. The discrete moment problem and linear programming. *Discrete Applied Mathematics* 27, 235-254.

Rolski, T., 1976. Order relations in the set of distribution functions and their applications in queueing theory. *Dissertations Math. CXXXII*, Polish Academy of Sciences, Warsaw.

Ross, S.M., 1996. *Stochastic Processes*. Wiley, New York.

Shaked, M., Shanthikumar, J.G., 1994. *Stochastic Orders and their Applications*. Academic, New York.

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

Utev, S.A., 1985. Extremal problems in moment inequalities. In *Limit Theorems of Probability Theory*, *Trudy Inst. Mat.* 5, "Nauka" Sib. Otdel., Novosibirsk, 56-75 (in Russian).

Whitt, W., 1985. The renewal-process stationary-excess operator. *Journal of Applied Probability* 22, 156-167.