

ON MULTIPLY MONOTONE DISTRIBUTIONS, CONTINUOUS OR DISCRETE, WITH APPLICATIONS

CLAUDE LEFÈVRE,* *Université Libre de Bruxelles*

STÉPHANE LOISEL,** *Université de Lyon*

Abstract

This paper is concerned with the class of distributions, continuous or discrete, whose shape is monotone of finite integer order t . A characterization is presented as a mixture of a minimum of t independent uniform distributions. Then, a comparison of t -monotone distributions is made using the s -convex stochastic orders. A link is also pointed out with an alternative approach to monotonicity based on a stationary-excess operator. Finally, the monotonicity property is exploited to reinforce the classical Markov and Lyapunov inequalities. The results are illustrated by several applications to insurance.

Keywords: t -monotone function; s -convex stochastic order; stationary-excess operator; Markov and Lyapunov inequalities; insurance risk theory

2010 Mathematics Subject Classification: Primary 62E10; 60E15

Secondary 62P05; 60E10

1. Introduction

A frequent difficulty met in stochastic modeling is that only incomplete statistical knowledge is available or trustworthy on certain model components. This is especially true in insurance and finance where risks and losses are generally difficult to estimate. For instance, the partial information at disposal for a claim distribution could be its range, the first few moments, and some shape constraints.

The present work is concerned with a class of distributions whose shape is known to be monotone of finite integer order t . We will consider both absolutely continuous distributions with a t -monotone density function on $\mathbb{R}_+$, and discrete distributions with a t -monotone probability mass function (PMF) on $\mathbb{N}_0 = \{0, 1, \dots\}$. Let us recall the definition of t -monotonicity.

A function f on $\mathbb{R}_+$ is completely monotone if f is infinitely differentiable and $(-1)^k f^{(k)} \geq 0$ for all $k \geq 1$. By Bernstein's theorem, such a function can be represented as a scale mixture of exponentials (see, e.g. Feller (1971)). Functions f on $\mathbb{R}_+$ that satisfy a property of this kind up to a finite degree t have been introduced and studied in Williamson (1956), Lévy (1962), and Gneiting (1999). More precisely, we have the following definition.

Definition 1.1. A function $f(y)$, $y > 0$, is 1-monotone if it is nonnegative and nonincreasing. It is t -monotone, $t \geq 2$, if

$$(-1)^k f^{(k)}(y) \text{ is nonnegative, nonincreasing, and convex for } k = 0, \dots, t-2.$$

Received 24 February 2012; revision received 29 November 2012.

* Postal address: Département de Mathématique, Université Libre de Bruxelles, Campus de la Plaine CP 210, B-1050 Bruxelles, Belgium. Email address: clefevre@ulb.ac.be

** Postal address: Université de Lyon, Université Claude Bernard Lyon 1, I.S.F.A., 50 Avenue Tony Garnier, F-69007 Lyon, France. Email address: stephane.loisel@univ-lyon1.fr

In other words, $(-1)^k f^{(k)}(y) \geq 0$ for $k = 0, \dots, t-3$ and $(-1)^{t-2} f^{(t-2)}(y)$ is nonnegative, nonincreasing, and convex. A t -monotone function on $\mathbb{N}_0$ is defined similarly using the forward difference operator Δ (i.e. $\Delta f(j) = f(j+1) - f(j)$).

Definition 1.2. A function $f(j)$, $j \geq 0$, is t -monotone, $t \geq 1$, if

$$(-1)^k \Delta^k f(j) \geq 0 \quad \text{for } k = 1, \dots, t.$$

Multiple monotonicity of continuous distributions has been considered for different purposes in probability and statistics. So, t -monotonicity corresponds to the concept of beta(1, t)-unimodality with a mode at 0 (see the book by Bertin *et al.* (1997)). It is a special type of scaling relation discussed in, e.g. Pakes (1997) and Pakes and Navarro (2007). A link with Archimedean copulas and L_1 Dirichlet distributions is pointed out in, e.g. Constantinescu *et al.* (2011). The estimation of a t -monotone density is studied in, e.g. Balabdaoui and Wellner (2007). To our knowledge, multiple monotonicity of discrete distributions has not been investigated so far.

Our interest in t -monotonicity arises in part from insurance where such a property can realistically be imposed on certain risk distributions. The importance of the case $t = 1$, i.e. nonincreasing distributions, is well recognized in this area (see, e.g. Gerber (1972), Kaas and Goovaerts (1987), and Denuit *et al.* (2000)). Introducing an order t allows us to cover a hierarchical class of risk distributions with reinforced shape constraints. Of course, there are many other application fields where a monotonicity property could be relevant.

The paper is organized as follows. In Section 2 we present a representation of t -monotone distributions as a mixture of a minimum of t independent uniform distributions. In Section 3 we use the s -convex stochastic orders to compare t -monotone distributions and derive extremal distributions. In Section 4 we connect the present approach to t -monotonicity with an alternative approach that uses the t -fold iterate of a stationary-excess operator. In Section 5 we show how the classical Markov and Lyapunov inequalities can be strengthened under the additional assumption of t -monotonicity. In Section 6 we apply some of the bounds obtained to different risk measures in insurance.

2. t -monotone distributions

The purpose of this section is to provide a representation for random variables, continuous or discrete, that have a t -monotone density or PMF. In the continuous case, such a characterization was obtained in, e.g. Lévy (1962). In the discrete case, the result seems to be new.

2.1. Continuous case

For $t \geq 1$, let $\mathcal{M}_t(z)$ denote the random variable

$$\mathcal{M}_t(z) \equiv \min\{U_l(z), 1 \leq l \leq t\}, \quad (2.1)$$

where z is a positive real number and the $U_l(z)$ are t independent random variables uniform on the interval $[0, z]$. Clearly, a distributional representation for $\mathcal{M}_t(z)$ is

$$\mathcal{M}_t(z) \stackrel{D}{=} (1 - U^{1/t})z, \quad (2.2)$$

where U is a uniform $[0, 1]$ random variable. So, the density of $\mathcal{M}_t(z)$ is

$$\frac{d\mathbb{P}[\mathcal{M}_t(z) < x]}{dx} = \frac{t}{z} \left(1 - \frac{x}{z}\right)_+^{t-1}, \quad x > 0. \quad (2.3)$$

We easily obtain the iterated right-tail distribution functions (DFs) of $\mathcal{M}_t(z)$ (see (A.6) in Appendix A).

Property 2.1. For $i \geq 0$,

$$\bar{F}_{i+1}[\mathcal{M}_t(z), x] = \frac{t!}{(i+t)!} z^i \left(1 - \frac{x}{z}\right)_+^{t+i}, \quad x > 0. \quad (2.4)$$

Let us now consider a randomized random variable $\mathcal{M}_t(Z)$ obtained by substituting for z in (2.1) some exogenous random variable Z valued in $\mathbb{R}_+$. By (2.2),

$$\mathcal{M}_t(Z) \stackrel{D}{=} (1 - U^{1/t})Z, \quad (2.5)$$

where Z and U are independent variables. The random variable $\mathcal{M}_t(Z)$ can be viewed as a randomly scaled version of Z studied in, e.g. Pakes (1997) and Pakes and Navarro (2007). Note that the scaling factor used here, $1 - U^{1/t}$, has a beta(1, t) distribution. In the theory of unimodality, $\mathcal{M}_t(Z)$ is said to have a beta(1, t)-unimodal distribution (see, e.g. Bertin *et al.* (1997, p. 72)). We then have

$$\mathbb{E}([M_t(Z)]^i) = \frac{\mathbb{E}(Z^i)}{\binom{i+t}{t}}. \quad (2.6)$$

The introduction of the variable Z leads to a characterization of a t -monotone density. By (2.3), the density of $\mathcal{M}_t(Z)$ is

$$q_t(x) = \int_0^\infty \frac{t}{z} \left(1 - \frac{x}{z}\right)_+^{t-1} dF_Z(z), \quad x > 0, \quad (2.7)$$

where F_Z is the DF of Z . From Theorem 5 (with Equations (8.4) and (8.7)) of Lévy (1962), we then have the following (known) proposition.

Proposition 2.1. *The density of an $\mathbb{R}_+$ -valued random variable X is t -monotone if and only if $X \stackrel{D}{=} (1 - U^{1/t})Z$ for some $\mathbb{R}_+$ -valued variable Z . If X has a t -monotone density $q(x)$, $x > 0$, then the density of Z is given by*

$$\frac{dF_Z(z)}{dz} = (-1)^t \frac{z^t}{t!} [q(z)]^{(t)}, \quad z > 0.$$

For $t = 1$, this result corresponds to a classical Khintchine theorem for unimodal distributions with a mode at 0. As $t \rightarrow \infty$, we see that $t(1 - U^{1/t}) \xrightarrow{D} -\ln(U) \stackrel{D}{=} \text{Exp}(1)$, an exponential variable with mean 1. Thus, the limiting form of Proposition 2.1 and (2.5) show that a completely monotone density admits the representation (Bernstein's theorem)

$$q_\infty(x) = \int_0^\infty \frac{1}{z} e^{-x/z} dF_Z(z), \quad x > 0.$$

2.2. Discrete case

For $t \geq 1$, in this case $\mathcal{M}_t(z)$ denotes the random variable

$$\mathcal{M}_t(z) \equiv \min\{U_l(z+l), 1 \leq l \leq t\}, \quad (2.8)$$

where z is a nonnegative integer and the $U_l(z+l)$ are t independent discrete random variables uniform on the sets $\{0, \dots, z+l-1\}$, respectively.

Obviously, $\mathcal{M}_t(z) \leq U_1(z+1) \leq z$. Let us determine the iterated right-tail DFs of $\mathcal{M}_t(z)$. For the sequel, we define $\binom{k}{j} = 0$ if $k < 0$ and $j \geq 0$.

Property 2.2. For $i \geq 0$,

$$\bar{F}_{i+1}[\mathcal{M}_t(z), j] = \frac{\binom{z-j+t+i}{t+i}}{\binom{z+t}{t}}, \quad 0 \leq j \leq z. \quad (2.9)$$

In particular,

$$\mathbb{P}[\mathcal{M}_t(z) = j] = \frac{\binom{z-j+t-1}{t-1}}{\binom{z+t}{t}}, \quad 0 \leq j \leq z, \quad (2.10)$$

and

$$\mathbb{E} \binom{\mathcal{M}_t(z)}{i} = \frac{\binom{z}{i}}{\binom{z+t}{t}}. \quad (2.11)$$

Proof. We proceed by recurrence. From (2.8),

$$\begin{aligned} \bar{F}_1[\mathcal{M}_t(z), j] &= \mathbb{P}[\mathcal{M}_t(z) \geq j] \\ &= \mathbb{P}[U_1(z+1) \geq j] \cdots \mathbb{P}[U_t(z+t) \geq j] \\ &= \frac{z+1-j}{z+1} \cdots \frac{z+t-j}{z+t} \\ &= \frac{\binom{z-j+t}{t}}{\binom{z+t}{t}}, \quad 0 \leq j \leq z, \end{aligned}$$

i.e. (2.9) for $i = 0$. So, the PMF of $\mathcal{M}_t(z)$, given by

$$\mathbb{P}[\mathcal{M}_t(z) = j] = \mathbb{P}[\mathcal{M}_t(z) \geq j] - \mathbb{P}[\mathcal{M}_t(z) \geq j+1] = -\Delta \bar{F}_1[\mathcal{M}_t(z), j],$$

becomes (2.10) by (A.3) below. For $i \geq 1$, (A.10) below and induction yield

$$\begin{aligned} \bar{F}_{i+1}[\mathcal{M}_t(z), j] &= \frac{1}{\binom{z+t}{t}} \sum_{k=j}^{\infty} \binom{z-k+t+i-1}{t+i-1} \\ &= \frac{1}{\binom{z+t}{t}} \sum_{k=t+i-1}^{z+t+i-j-1} \binom{k}{t+i-1}, \quad 0 \leq j \leq z, \end{aligned}$$

which reduces to (2.9) using (A.1) below. By (A.11) below, the binomial moments are given by

$$\mathbb{E} \binom{\mathcal{M}_t(z)}{i} = \bar{F}_{i+1}[\mathcal{M}_t(z), i] = \frac{\binom{z+t}{t+i}}{\binom{z+t}{t}},$$

and (2.11) follows.

An interesting observation made by the referee is that $\mathcal{M}_t(z)$ admits a distributional representation analogous to (2.2).

Proposition 2.2. It holds that

$$\mathcal{M}_t(z) \stackrel{D}{=} \text{MBin}(z, 1 - U^{1/t}), \quad (2.12)$$

where U is a uniform $[0, 1]$ random variable and $\text{MBin}(\cdot, \cdot)$ denotes a mixed binomial random variable.

Proof. The PMF of $\text{MBin}(z, 1 - U^{1/t})$ is given by

$$\begin{aligned} m_j &\equiv \binom{z}{j} \mathbb{E}((1 - U^{1/t})^j U^{(z-j)/t}) \\ &= \binom{z}{j} \int_0^1 v^{z-j+t-1} (1 - v)^j \, dv \\ &= \binom{z}{j} B(z - j + t, j + 1), \quad 0 \leq j \leq z, \end{aligned}$$

after making the substitution $u = v^t$ and writing $B(\cdot, \cdot)$ for the usual beta function. This reduces to PMF (2.10) as announced.

Let us now replace z in (2.8) with some exogenous random variable Z valued in $\mathbb{N}_0$. By (2.10), the PMF of $\mathcal{M}_t(Z)$ is given by

$$p_t(j) \equiv \mathbb{P}[\mathcal{M}_t(Z) = j] = \mathbb{E}\left(\frac{\binom{Z-j+t-1}{t-1}}{\binom{Z+t}{t}}\right), \quad j \geq 0, \quad (2.13)$$

which is the discrete analogue of (2.7). By (2.11),

$$\mathbb{E}\left(\binom{\mathcal{M}_t(Z)}{i}\right) = \frac{\mathbb{E}\left(\binom{Z}{i}\right)}{\binom{i+t}{t}}, \quad i \geq 0. \quad (2.14)$$

Continuing with (2.12) leads us to a nice representation of $\mathcal{M}_t(Z)$ through a binomial thinning operator, ' $\odot$ ', due to Steutel and van Harn (1979). Recall that, with $a \in (0, 1)$ and $\mathbb{N}_0$ -valued Z ,

$$a \odot Z = \sum_{i=1}^Z I_i,$$

where the I_i are independent Bernoulli random variables with parameter a and independent of Z . Thus, we see that

$$\mathcal{M}_t(Z) \stackrel{D}{=} (1 - U^{1/t}) \odot Z, \quad (2.15)$$

where U and Z are independent. This was also pointed out to us by the referee.

The introduction of a variable Z allows us to characterize the t -monotonicity of a PMF.

Proposition 2.3. *The PMF of an $\mathbb{N}_0$ -valued random variable X is t -monotone if and only if $X \stackrel{D}{=} (1 - U^{1/t}) \odot Z$ for some $\mathbb{N}_0$ -valued Z . If X has a t -monotone PMF $p_X(j)$, $j \geq 0$, then the PMF of Z is given by*

$$\mathbb{P}[Z = z] = (-1)^t \binom{z+t}{t} \Delta^t p_X(z), \quad z \geq 0. \quad (2.16)$$

Proof. From (A.3) below, PMF (2.13) of $\mathcal{M}_t(Z)$ is such that, for $1 \leq k \leq t$,

$$\begin{aligned} \Delta^k p_t(j) &= \mathbb{E}\left(\frac{1}{\binom{Z+t}{t}} \Delta^k \binom{Z-j+t-1}{t-1}\right) \\ &= (-1)^k \mathbb{E}\left(\frac{\binom{Z-j+t-1-k}{t-1-k}}{\binom{Z+t}{t}}\right), \quad j \geq 0, \end{aligned}$$

so that (2.13) is t -monotone as desired.

Reciprocally, let X be a random variable whose PMF $p_X(j)$, $j \geq 0$, is t -monotone. Denote by $p(z)$ the right-hand side of (2.16) constructed using this PMF. The sequence $p(z)$, $z \geq 0$,

constitutes a PMF. Indeed, by the t -monotonicity of X , we have $p(z) \geq 0$ for all $z \geq 0$. Moreover, from identity (A.5) below, we obtain $\sum_{z \geq 0} p(z) = 1$. Now, let Z be a random variable which has precisely this PMF. By (2.13), the corresponding random variable $\mathcal{M}_t(Z)$ has the PMF

$$p_t(j) = (-1)^t \sum_{z=j}^{\infty} \binom{z-j+t-1}{t-1} \Delta^t p_X(z), \quad j \geq 0.$$

From identity (A.4) below, we know that the right-hand side reduces to $p_X(j)$. This means that, as announced, $\mathcal{M}_t(Z)$ and X have the same distribution.

3. s -convex orderings

In this section we compare t -monotone distributions and derive extremal distributions by using the s -convex stochastic orders, denoted by ' $\leq_{s\text{-cx}}$ ', where s is an integer greater than or equal to 1. These orders are briefly presented in Appendix A; let us recall that their definition is similar, but not identical, in the continuous and discrete cases. For $t = 1$, such a comparison problem has been discussed in Denuit *et al.* (1998), (1999b).

3.1. Continuous case

The proposition below states that an s -convex ordering on Z implies an s -convex ordering on $\mathcal{M}_t(Z)$. This is true for any t .

Proposition 3.1. *If $Z_1 \leq_{s\text{-cx}} Z_2$ then $\mathcal{M}_t(Z_1) \leq_{s\text{-cx}} \mathcal{M}_t(Z_2)$.*

Proof. Let u be an arbitrary value of U and consider the variable $(1 - u^{1/t})Z$. The (continuous) s -convex ordering is preserved by multiplication by a positive constant and by mixture (see Denuit *et al.* (1998)). Thus, applying these properties yields the announced assertion.

3.1.1. Convex extrema. Proposition 3.1 provides a simple way to construct s -convex extrema for t -monotone densities using standard s -convex extrema, i.e. when there is no monotonicity restriction on the densities. With the same goal, Lefèvre and Loisel (2010) proposed a different approach based on the use of the t -fold iterate of a stationary-excess operator. The present method has the advantage of being more easily applicable.

Specifically, let $\mathcal{B}_s([0, b]; \mu_1, \dots, \mu_{s-1})$ denote the class of all the random variables Z whose distributions have support in $[0, b]$ and which have the first $s - 1$ moments $\mathbb{E}(Z^i) = \mu_i$, $1 \leq i \leq s - 1$. Let $Z_{\min}^{(s)}$ and $Z_{\max}^{(s)}$ be the extrema in this class with respect to the order ' $\leq_{s\text{-cx}}$ '. A method to determine these extrema is described in, e.g. Shaked and Shanthikumar (2007, pp. 145–146).

Now, consider the class $\mathcal{B}_s([0, b]; \nu_1, \dots, \nu_{s-1}; t\text{-monotone})$ of all the random variables X with support in $[0, b]$, first $s - 1$ moments $\mathbb{E}(X^i) = \nu_i$, $1 \leq i \leq s - 1$, and which have a t -monotone density. By Proposition 2.1, there exists some random variable Z on $[0, b]$ such that $X \stackrel{D}{=} \mathcal{M}_t(Z)$. Remember that the moments μ_i of Z are obtained from ν_i through (2.6). As $Z_{\min}^{(s)} \leq_{s\text{-cx}} Z \leq_{s\text{-cx}} Z_{\max}^{(s)}$, Proposition 3.1 yields $\mathcal{M}_t(Z_{\min}^{(s)}) \leq_{s\text{-cx}} X \leq_{s\text{-cx}} \mathcal{M}_t(Z_{\max}^{(s)})$.

For instance, let $s = 2$. It is well known that, inside $\mathcal{B}_2([0, b]; \mu_1)$,

$$\begin{aligned} Z_{\min}^{(2)} &= \mu_1 \quad \text{almost surely,} \\ Z_{\max}^{(2)} &= \begin{cases} 0 & \text{with probability } 1 - \mu_1/b, \\ b & \text{with probability } \mu_1/b. \end{cases} \end{aligned}$$

Let $t = 1$, and consider $\mathcal{B}_2([0, b]; \nu_1; 1\text{-monotone})$. In this case, (2.6) implies that $\mu_1 = 2\nu_1$. As $\mathcal{M}_1(Z) = UZ$, we obtain

$$\mathcal{M}_1(Z_{\min}^{(2)}) = 2\nu_1 U, \quad (3.1)$$

$$\mathcal{M}_1(Z_{\max}^{(2)}) = \begin{cases} 0 & \text{with probability } 1 - 2\nu_1/b, \\ bU & \text{with probability } 2\nu_1/b. \end{cases} \quad (3.2)$$

Let $t = 2$, and consider $\mathcal{B}_2([0, b]; \nu_1; 2\text{-monotone})$. From (2.6), we have $\mu_1 = 3\nu_1$. Using (2.5), we then obtain

$$\mathcal{M}_2(Z_{\min}^{(2)}) = 3\nu_1(1 - U^{1/2}), \quad (3.3)$$

$$\mathcal{M}_2(Z_{\max}^{(2)}) = \begin{cases} 0 & \text{with probability } 1 - 3\nu_1/b, \\ b(1 - U^{1/2}) & \text{with probability } 3\nu_1/b. \end{cases} \quad (3.4)$$

Note that the density of $(1 - U^{1/2})$ is $2(1 - x)$, $0 \leq x \leq 1$.

The 1-convex minimum (3.1) has a density that is uniform on $[0, 2\nu_1]$ and takes the value 0 otherwise. Obviously, this density is nonincreasing, but it is not convex. This allows us to correct an erroneous assertion in Section 5 of Lefèvre and Loisel (2010): (3.1) is not, as claimed there, the 2-convex minimum in the class of the random variables whose density is nonincreasing convex. In fact, the true 2-convex minimum in that class is given by (3.3). The upper bound (3.2) has a nonincreasing convex density on $[0, b]$, with a probability mass at 0 (i.e. an infinite value of the density at 0 and a finite limit at 0_+). Thus, (3.2) is also the 2-convex maximum among the variables whose density is required to be 2-monotone on $[0, b]$ only; this is precisely what is stated in Lefèvre and Loisel (2010). The 2-convex maximum given by (3.4) is not that obtained in Lefèvre and Loisel (2010) because the 2-monotonicity required there is on $[0, b]$ only.

3.2. Discrete case

As in the continuous case, an s -convex ordering on Z is transferred to $\mathcal{M}_t(Z)$. This is a direct consequence of representation (2.15).

Proposition 3.2. *The assertion of Proposition 3.1 holds in the discrete case also.*

Proof. For any value u of U , consider the variable $(1 - u^{1/t}) \odot Z \stackrel{D}{=} \text{MBin}(Z, 1 - u^{1/t})$. If $Z_1 \leq_{s\text{-cx}} Z_2$ then $\text{MBin}(Z_1, 1 - u^{1/t}) \leq_{s\text{-cx}} \text{MBin}(Z_2, 1 - u^{1/t})$ (see Denuit *et al.* (1999c, Property 5.7)). As the (discrete) s -convex ordering is preserved by mixing (see Denuit and Lefèvre (1997)), the announced result follows.

Remark. Let f be a function on $\mathbb{N}_0$, and define the associated function g by

$$g(z) \equiv \mathbb{E}(f(\mathcal{M}_t(z))) = \sum_{j=0}^z \frac{\binom{z-j+t-1}{t-1}}{\binom{z+t}{t}} f(j), \quad z \geq 0. \quad (3.5)$$

We can show that

$$\binom{t+s}{s} \Delta^s g(z) = \sum_{j=0}^z \frac{\binom{j+s}{s} \binom{z-j+t-1}{t-1}}{\binom{z+s+t}{s+t}} \Delta^s f(j), \quad s \geq 0. \quad (3.6)$$

Thus, if the function f is s -convex (in the sense of (A.12) below) then the function g is also s -convex. By the definition of the s -convex order given in Definition A.2 below, this gives another proof of Proposition 3.2. For $t = 1$, identity (3.6) has been derived in Denuit *et al.* (1999) to prove the result in that case. The argument followed above is simpler and more enlightening than the use of (3.6). Note that, by a known combinatorial identity, the coefficients of $\Delta^s f(j)$ sum to 1. Thus, they constitute a PMF and the right-hand side corresponds to an expectation as in (3.5).

3.2.1. Convex extrema. The framework is similar. Let $\mathcal{B}_s(\{0, \dots, n\}; \mu_1, \dots, \mu_{s-1})$ be the class of all the random variables Z which are valued in a set $\{0, \dots, n\}$ and have the first $s - 1$ binomial moments $\mathbb{E}\binom{Z}{i} = \mu_i$, $1 \leq i \leq s - 1$. Denote by $Z_{\min}^{(s)}$ and $Z_{\max}^{(s)}$ the extrema in that class with respect to ' $\leq_{s\text{-cx}}$ '. A method for deriving these extrema is presented in Denuit and Lefèvre (1997).

Discrete extrema have received less attention in the literature. It is worth mentioning, however, that optimal bounds of this type have been investigated in branching theory for approximating the extinction probability and other functionals (see, e.g. Pakes (2003, pp. 705–706) for references to these bounds); see also, e.g. Lefèvre and Utev (1996) in epidemic theory.

Now, let $\mathcal{B}_s(\{0, \dots, n\}; \nu_1, \dots, \nu_{s-1}; t\text{-monotone})$ be the set of all the random variables X valued in $\{0, \dots, n\}$, with first $s - 1$ binomial moments $\mathbb{E}\binom{X}{i} = \nu_i$, $1 \leq i \leq s - 1$, and which have a t -monotone PMF. By Proposition 2.3, $X \stackrel{D}{=} \mathcal{M}_t(Z)$ for some random variable Z valued in $\{0, \dots, n\}$. The moments of Z and X are connected by (2.14). Applying Proposition 3.2 then yields $\mathcal{M}_t(Z_{\min}^{(s)}) \leq_{s\text{-cx}} X \leq_{s\text{-cx}} \mathcal{M}_t(Z_{\max}^{(s)})$.

For illustration, let $s = 2$. Inside $\mathcal{B}_2(\{0, \dots, n\}; \mu_1)$,

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

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

where ξ is the integer in $[0, n - 1]$ such that $\xi < \mu_1 \leq \xi + 1$ (see Denuit and Lefèvre (1997)). Let $t = 1$, and consider $\mathcal{B}_2(\{0, \dots, n\}; \nu_1; 1\text{-monotone})$. Following (2.14), we have $\mu_1 = 2\nu_1$; let ξ_1 be the corresponding value of ξ . By (2.8), $\mathcal{M}_1(Z) = U_1(Z + 1)$, so that

$$\mathcal{M}_1(Z_{\min}^{(2)}) = \begin{cases} 0, \dots, \xi_1 & \text{with equal probabilities } 2(\xi_1 + 1 - \nu_1)/(\xi_1 + 1)(\xi_1 + 2), \\ \xi_1 + 1 & \text{with probability } (2\nu_1 - \xi_1)/(\xi_1 + 2), \end{cases} \quad (3.7)$$

$$\mathcal{M}_1(Z_{\max}^{(2)}) = \begin{cases} 0 & \text{with probability } 1 - 2\nu_1/(n + 1), \\ 1, \dots, n & \text{with equal probabilities } 2\nu_1/n(n + 1). \end{cases} \quad (3.8)$$

Let $t = 2$, and consider $\mathcal{B}_2(\{0, \dots, n\}; \nu_1; 2\text{-monotone})$. From (2.14), we have $\mu_1 = 3\nu_1$; let ξ_2 be the corresponding value of ξ . Using (2.8), we obtain

$$\mathcal{M}_2(Z_{\min}^{(2)}) = j \in \{0, \dots, \xi_2 + 1\} \quad \text{with probability } (\xi_2 - j + 1)\pi_1 + (\xi_2 - j + 2)\pi_2, \quad (3.9)$$

where $\pi_1 = 2(\xi_2 + 1 - 3\nu_1)/(\xi_2 + 2)(\xi_2 + 1)$ and $\pi_2 = 2(3\nu_1 - \xi_2)/(\xi_2 + 3)(\xi_2 + 2)$, while

$$\mathcal{M}_2(Z_{\max}^{(2)}) = \begin{cases} 0 & \text{with probability } 1 - 3\nu_1/(n + 2), \\ j \in \{1, \dots, n\} & \text{with probability } 6\nu_1(n - j + 1)/n(n + 1)(n + 2). \end{cases} \quad (3.10)$$

The 1-convex maximum (3.8) has a PMF that is nonincreasing convex on $\{0, \dots, n\}$, but it is not convex on $\mathbb{N}_0$. Thus, (3.8) is the 2-convex maximum among the random variables whose PMF is required to be 2-monotone on $\{0, \dots, n\}$ only; this is what is asserted in Section 5 of Lefèvre and Loisel (2010). Of course, (3.8) is not the 2-convex maximum here because the 2-monotonicity is on $\mathbb{N}_0$. The true 2-convex maximum is given by (3.10). The 2-convex minimum (3.9) has a PMF that is nonincreasing convex on $\{0, \dots, n\}$. Thus, it is also optimal among the variables whose PMF is required to be 2-monotone on $\{0, \dots, n\}$ only; this result has been obtained in Lefèvre and Loisel (2010). Note that, by comparison, the PMF of the 1-convex minimum (3.7) is nonincreasing but nonconvex.

Other values of t may be considered without real difficulties. One can also deal with $s \geq 3$, provided the s -convex extrema for Z are available. This is possible, for instance, with $s = 3$. Then (2.14) leads to $\mu_1 = 2v_1$ and $\mu_2 = 3v_2$ when $t = 1$, and $\mu_1 = 3v_1$ and $\mu_2 = 6v_2$ when $t = 2$.

4. t -stationary-excess distributions

A different method to generate t -monotone distributions consists in using the t -fold iterate of a stationary-excess operator (see Lefèvre and Loisel (2010)). Our purpose in this section is to point out a link between that method and the present approach to t -monotonicity. This topic is partly related to the characterization of distributions through length-biasing, stationary-excess, and random scaling operations (see, e.g. Pakes (1996), (1997) and Pakes and Navarro (2007)).

4.1. Continuous case

4.1.1. *The length-bias transform.* Let Y be an $\mathbb{R}_+$ -valued random variable with finite mean and density q_Y . The length-bias transform $\mathcal{L}$ of Y is an $\mathbb{R}_+$ -valued random variable $\mathcal{L}(Y)$ with density

$$q_{\mathcal{L}(Y)}(z) = \frac{z q_Y(z)}{\mathbb{E}(Y)}, \quad z > 0.$$

For instance, if Y is gamma (α, n) then $\mathcal{L}(Y)$ is gamma $(\alpha, n + 1)$; if Y is Pareto (α, γ) (respectively Lognormal (μ, σ^2)), $\mathcal{L}(Y)$ is Pareto $(\alpha, \gamma - 1)$ (respectively Lognormal $(\mu + \sigma^2, \sigma^2)$). Length-biased distributions arise in many situations where the probability of selection is proportional to a size dimension (see, e.g. Patil and Rao (1978)). Note that the operator $\mathcal{L}$ yields a one-to-one correspondence.

Let us apply t times the operator $\mathcal{L}$ to Y , under the assumption that $\mathbb{E}(Y^t) < \infty$. We easily see that the resulting random variable $\mathcal{L}_t(Y)$ has the density

$$q_{\mathcal{L}_t(Y)}(z) = \frac{z^t q_Y(z)}{\mathbb{E}(Y^t)}, \quad z > 0. \quad (4.1)$$

Moreover, we get

$$\mathbb{E}([\mathcal{L}_t(Y)]^i) = \frac{\mathbb{E}(Y^{i+t})}{\mathbb{E}(Y^t)}, \quad i \geq 0. \quad (4.2)$$

In an actuarial context, the right-hand side for $i = 1$ corresponds to the size-biased pricing functional of an insurance risk or loss Y (see, e.g. Furman and Zitikis (2009)).

4.1.2. *The stationary-excess transform.* Let us first examine what becomes of the random variable $\mathcal{M}_t(Z) \stackrel{D}{=} (1 - U^{1/t})Z$ in the case where $Z \stackrel{D}{=} \mathcal{L}_t(Y)$ with density (4.1).

Property 4.1. *The density of $\mathcal{M}_t(\mathcal{L}_t(Y))$ is*

$$q_t(x) = \frac{t\mathbb{E}(Y-x)_+^{t-1}}{\mathbb{E}(Y^t)}, \quad x > 0. \quad (4.3)$$

Proof. Substituting (4.1) into (2.7) gives

$$q_t(x) = \frac{t}{\mathbb{E}(Y^t)} \int_0^t \frac{1}{z} \left(1 - \frac{x}{z}\right)_+^{t-1} z^t q_Y(z) dz, \quad x > 0,$$

and (4.3) follows.

Note that the density q_t reduces to a ratio between two expectations. From (2.6) and (4.2), the associated moments are

$$\mathbb{E}([\mathcal{M}_t(\mathcal{L}_t(Y))]^i) = \frac{\mathbb{E}(Y^{i+t})}{\binom{i+t}{t}\mathbb{E}(Y^t)}, \quad i \geq 0.$$

Now, let us consider the standard stationary-excess operator $\mathcal{J}$ (see, e.g. Cox (1962)). The operator $\mathcal{J}$ transforms the variable Y into a random variable $\mathcal{J}(Y)$ with density

$$q_{\mathcal{J}(Y)}(z) = \frac{\mathbb{P}[Y > y]}{\mathbb{E}(Y)}, \quad z > 0. \quad (4.4)$$

Obviously, density (4.4) is the same as (4.3) with $t = 1$. Let us apply t times the operator $\mathcal{J}$, which yields a random variable $\mathcal{J}_t(Y)$. As shown in Lefèvre and Loisel (2010, Equation (4.11)), the density of $\mathcal{J}_t(Y)$ is still given by (4.3), which gives the following result.

Proposition 4.1. *It holds that*

$$\mathcal{J}_t(Y) \stackrel{D}{=} (1 - U^{1/t})\mathcal{L}_t(Y). \quad (4.5)$$

Relation (4.5) represents the stationary-excess operator as a random contraction of the length-bias operator. It occurs as Lemma 4.1 of Pakes (1996) (with a different proof).

Note that, by virtue of Proposition 2.1 and (4.5), a stationary-excess density of order t is a t -monotone function. Moreover, from (4.5) and using Proposition 4.4 of Lefèvre and Loisel (2010), we obtain the following convex comparison result:

$$\text{if } Y_1 \leq_{(s+t)\text{-cx}} Y_2 \text{ then } \mathcal{M}_t(\mathcal{L}_t(Y_1)) \leq_{s\text{-cx}} \mathcal{M}_t(\mathcal{L}_t(Y_2)). \quad (4.6)$$

4.2. Discrete case

4.2.1. A length-bias-type transform. Let Y be an $\mathbb{N}_0$ -valued random variable with finite mean. We define an operator $\mathcal{L}$ that transforms Y into a random variable $\mathcal{L}(Y)$, also $\mathbb{N}_0$ -valued, whose PMF is defined by

$$\mathbb{P}[\mathcal{L}(Y) = z] = \frac{(z+1)\mathbb{P}[Y = z+1]}{\mathbb{E}(Y)}, \quad z \geq 0.$$

For instance, if Y is Poisson (λ) then $\mathcal{L}(Y)$ is also Poisson (λ); if Y is binomial (n, p) (respectively negative binomial (n, p)), $\mathcal{L}(Y)$ is binomial ($n-1, p$) (respectively negative binomial ($n+1, p$)). The operator $\mathcal{L}$ yields a one-to-one correspondence when $\mathbb{E}(Y)$ is fixed.

This is not true otherwise: for example, if v is a Bernoulli variable independent of Y then $\mathcal{L}(Y) \stackrel{D}{=} \mathcal{L}(vY)$.

Let us note that $\mathcal{L}$ differs slightly from the length-bias operator usually considered for discrete random variables (as in, e.g. Patil and Rao (1978)). This operator, $\tilde{\mathcal{L}}$ say, transforms Y into a variable $\tilde{\mathcal{L}}(Y)$ with PMF given by

$$\mathbb{P}[\tilde{\mathcal{L}}(Y) = z] = \frac{z\mathbb{P}[Y = z]}{\mathbb{E}(Y)}, \quad z \geq 1.$$

Thus, $\mathcal{L}(Y) \stackrel{D}{=} \tilde{\mathcal{L}}(Y) - 1$, i.e. $\mathcal{L}(Y)$ is a (-1) -translated length-biased version of Y .

Let us operate t times $\mathcal{L}$ to Y , provided that $\mathbb{E}(Y^t) < \infty$. For $t = 2$, this gives a variable $\mathcal{L}_2(Y)$ whose PMF is $\mathbb{P}[\mathcal{L}_2(Y) = z] = (z+1)\mathbb{P}[\mathcal{L}(Y) = z+1]/\mathbb{E}(\mathcal{L}(Y))$, $z \geq 0$, where $\mathbb{E}(\mathcal{L}(Y)) = 2\mathbb{E}\binom{Y}{2}/\mathbb{E}(Y)$. After t iterations, we get for $\mathcal{L}_t(Y)$ the following PMF:

$$\mathbb{P}[\mathcal{L}_t(Y) = z] = \frac{\binom{z+t}{t}\mathbb{P}[Y = z+t]}{\mathbb{E}\binom{Y}{t}}, \quad z \geq 0. \quad (4.7)$$

Its binomial moments are

$$\mathbb{E}\binom{\mathcal{L}_t(Y)}{i} = \frac{\binom{i+t}{t}\mathbb{E}\binom{Y}{t+i}}{\mathbb{E}\binom{Y}{t}}, \quad i \geq 0. \quad (4.8)$$

These formulae can also be obtained using the probability generating function of $\mathcal{L}_t(Y)$.

4.2.2. The stationary-excess transform. Let us consider the random variable $\mathcal{M}_t(z) \stackrel{D}{=} (1 - U^{1/t}) \odot Z$ in the case where $Z \stackrel{D}{=} \mathcal{L}_t(Y)$ with PMF (4.7).

Property 4.2. *The PMF of $\mathcal{M}_t(\mathcal{L}_t(Y))$ is*

$$p_t(j) = \frac{\mathbb{E}\binom{Y-j-1}{t-1}}{\mathbb{E}\binom{Y}{t}}, \quad j \geq 0. \quad (4.9)$$

Proof. From (2.13) and (4.7),

$$\begin{aligned} p_t(j) &= \sum_{z=j}^{\infty} \frac{\binom{z-j+t-1}{t-1}}{\binom{z+t}{t}} \frac{\binom{z+t}{t}\mathbb{P}[Y = z+t]}{\mathbb{E}\binom{Y}{t}} \\ &= \frac{1}{\mathbb{E}\binom{Y}{t}} \sum_{z=j+t}^{\infty} \binom{z-j-1}{t-1} \mathbb{P}[Y = z], \quad j \geq 0, \end{aligned}$$

which gives (4.9).

As in (4.3), the PMF p_t is a ratio between two expectations. From (2.14) and (4.8), its binomial moments are

$$\mathbb{E}\binom{\mathcal{M}_t(\mathcal{L}_t(Y))}{i} = \frac{\mathbb{E}\binom{Y}{i+t}}{\mathbb{E}\binom{Y}{t}}, \quad i \geq 0.$$

Now, let us consider the discrete stationary-excess operator $\mathcal{S}$ introduced in Lefèvre and Loisel (2010). The operator $\mathcal{S}$ transforms Y into an $\mathbb{N}$ -valued random variable $\mathcal{S}(Y)$ with PMF

$$\mathbb{P}[\mathcal{S}(Y) = z] = \frac{\mathbb{P}[Y \geq z+1]}{\mathbb{E}(Y)}, \quad z \geq 0. \quad (4.10)$$

Observe that PMFs (4.10) and (4.9) with $t = 1$ are identical. Applying t times $\mathcal{S}$ to Y yields a random variable $\mathcal{S}_t(Y)$ whose PMF is still given by (4.9) (see Lefèvre and Loisel (2010, Equation (4.2))). This yields the following result.

Proposition 4.2. *It holds that*

$$\mathcal{S}_t(Y) \stackrel{D}{=} (1 - U^{1/t}) \odot \mathcal{L}_t(Y). \quad (4.11)$$

By Proposition 2.3 and (4.11), a stationary-excess PMF of order t is a t -monotone function. We can show that the comparison result (4.6) also holds in the discrete case.

5. Refining simple standard inequalities

The goal here is to strengthen Markov and Lyapunov inequalities for distributions that are known to be t -monotone. Our study is mainly focused on the continuous case, which is more tractable. We thank S. Utev for fruitful discussions on this topic; see also Lefèvre and Utev (2012) for further results. As indicated below, such inequality refinements exist in the literature on unimodal distributions. For clarity, we write $X_t \equiv \mathcal{M}_t(Z)$ in this section.

5.1. Continuous case

First, note from (2.5) that (2.6) can be extended to any moment of order $r > 0$ by

$$\mathbb{E}(X_t^r) = \frac{\Gamma(r+1)\Gamma(t+1)}{\Gamma(r+t+1)}\mathbb{E}(Z^r). \quad (5.1)$$

5.1.1. Markov-type inequality. Let X be an $\mathbb{R}_+$ -valued random variable. The classical Markov inequality states that, for any $r \geq 0$,

$$\mathbb{P}[X \geq x] \leq \frac{\mathbb{E}(X^r)}{x^r}, \quad x > 0. \quad (5.2)$$

The bound may be refined when additional information on X is available. We derive below a tighter bound for a random variable X_t whose density is t -monotone.

Proposition 5.1. *For $r \geq 0$,*

$$\mathbb{P}[X_t \geq x] \leq c(r, t) \frac{\mathbb{E}(X_t^r)}{x^r}, \quad x > 0, \quad (5.3)$$

where $c(r, t)$ is a reducing factor given by

$$c(r, t) = \frac{\Gamma(r+t+1)}{\Gamma(r+1)\Gamma(t+1)} \left(\frac{r}{r+t}\right)^r \left(\frac{t}{r+t}\right)^t. \quad (5.4)$$

Proof. By (2.4),

$$x^r \mathbb{P}[X_t \geq x] = x^r \mathbb{E} \left(1 - \frac{x}{Z} \right)_+^t \equiv \mathbb{E}(\theta_Z(x)), \quad x > 0.$$

For Z fixed, the function $\theta_Z(x)$ has a maximum at $x = x_M \equiv [r/(r+t)]Z$ ($\in (0, Z)$). Thus,

$$x^r \mathbb{P}[X \geq x] \leq \mathbb{E}(\theta_Z(x_M)) = \left(\frac{r}{r+t}\right)^r \left(\frac{t}{r+t}\right)^t \mathbb{E}(Z^r), \quad x > 0. \quad (5.5)$$

From (5.1),

$$\mathbb{E}(Z^r) = \frac{\Gamma(r+t+1)}{\Gamma(r+1)\Gamma(t+1)} \mathbb{E}(X_t^r), \quad (5.6)$$

which, inserted into (5.5), gives (5.3) and (5.4).

From (5.4), we observe that $c(r, t) = c(t, r)$. The next property asserts that, as expected, the bound in (5.3) becomes tighter for larger values of t ; a proof is given in Appendix A.

Property 5.1. *The factor $c(r, t)$, $r > 0$, is decreasing with t , and*

$$\lim_{t \rightarrow \infty} c(r, t) = \frac{r^r e^{-r}}{\Gamma(r+1)}.$$

When $t = 0$, (5.3) and (5.4) reduce to (5.2). When $t = 1$ (nonincreasing densities), (5.3) and (5.4) give, for $r = 1$,

$$\mathbb{P}[X_1 \geq x] \leq \frac{\mathbb{E}(X_1)}{2x}, \quad (5.7)$$

and, for $r = 2$,

$$\mathbb{P}[X_1 \geq x] \leq \frac{4\mathbb{E}(X_1^2)}{9x^2}.$$

This refinement of Chebyshev's inequality in fact is the (earlier) Gauss inequality for unimodal distributions. A study of the case $t = 1$ is made in, e.g. the book of Dharmadhikari and Joag-Dev (1988, Section 1.5). Note that our proof of Proposition 5.1 for $t = 1$ is neater than the proof presented there. When $t = 2$ (nonincreasing convex densities), (5.3) and (5.4) yield, for $r = 1$,

$$\mathbb{P}[X_2 \geq x] \leq \frac{4\mathbb{E}(X_2)}{9x}, \quad (5.8)$$

and, for $r = 2$,

$$\mathbb{P}[X_2 \geq x] \leq \frac{3\mathbb{E}(X_2^2)}{8x^2}.$$

5.1.2. Lyapunov-type inequality. Considering $X \geq 0$ again, Jensen's inequality implies that, for any $r \geq 1$,

$$[\mathbb{E}(X)]^r \leq \mathbb{E}(X^r). \quad (5.9)$$

In particular, this yields the standard Lyapunov inequality: for $0 < r \leq s$,

$$[\mathbb{E}(X^r)]^{1/r} \leq [\mathbb{E}(X^s)]^{1/s}. \quad (5.10)$$

Let us show how to refine these two inequalities for a t -monotone random variable X_t .

Proposition 5.2. *For $r \geq 1$,*

$$[\mathbb{E}(X_t^r)]^r \leq \frac{\Gamma(r+t+1)}{(t+1)^r \Gamma(r+1) \Gamma(t+1)} \mathbb{E}(X_t^r), \quad (5.11)$$

and, for $0 < r \leq s$,

$$\left[\mathbb{E}(X_t^r) \frac{\Gamma(r+t+1)}{\Gamma(r+1) \Gamma(t+1)} \right]^{1/r} \leq \left[\mathbb{E}(X_t^s) \frac{\Gamma(s+t+1)}{\Gamma(s+1) \Gamma(t+1)} \right]^{1/s}. \quad (5.12)$$

Proof. From (5.1), we have, for $r \geq 0$,

$$[\mathbb{E}(X_t)]^r = \frac{1}{(t+1)^r} [\mathbb{E}(Z)]^r, \quad (5.13)$$

as well as the relation (5.6). By (5.9), $[\mathbb{E}(Z)]^r \leq \mathbb{E}(Z^r)$ for $r \geq 1$. Inserting this inequality into (5.13) and using (5.6) then yields (5.11). The proof of (5.12) is similar: it suffices to combine (5.6) with inequality (5.10) applied to Z , i.e. $[\mathbb{E}(Z^r)]^{1/r} \leq [\mathbb{E}(Z^s)]^{1/s}$.

For $t = 0$, (5.11) and (5.12) reduce to (5.9) and (5.10). For $t = 1$ and $r = 2$, (5.11) gives

$$[\mathbb{E}(X_1)]^2 \leq \frac{3\mathbb{E}(X_1^2)}{4},$$

or, equivalently, $[\mathbb{E}(X_1^2)] \leq 3 \text{var}(X_1)$, which is a known result (see, e.g. Dharmadhikari and Joag-Dev (1988, p. 9)). If $t = r = 2$ for instance,

$$[\mathbb{E}(X_2)]^2 \leq \frac{2\mathbb{E}(X_2^2)}{3}.$$

An inequality similar to (5.11) can be found in, e.g. Pečarić *et al.* (1992, p. 222), see also page 218 when $t = 1$. By comparison, our proof is especially simple.

5.2. Discrete case

This topic in the discrete case remains widely open. The difficulty comes from the less tractable representation (2.15) for $\mathcal{M}_t(Z)$. We just examine here how Markov's inequality in its simplest form, i.e. when $r = 1$, can be refined for a PMF that is nonincreasing. First, we note that if X is an $\mathbb{N}_0$ -valued random variable, we see that

$$\mathbb{P}[X \geq j+1] \leq \frac{\mathbb{E}(X)}{j+1}, \quad j \geq 0,$$

which is, of course, almost (5.2) for $r = 1$. The following inequality is similar to (5.7), but it is derived by a different argument.

Proposition 5.3. For $i \geq 0$,

$$\mathbb{P}[X_1 \geq j+1] \leq \frac{\mathbb{E}(X_1)}{2j+1}, \quad j \geq 0. \quad (5.14)$$

Proof. By definition,

$$\begin{aligned} \mathbb{E}(X_1) &= \sum_{i=1}^{\infty} \bar{F}_1(X_1, i) \\ &\geq \sum_{i=1}^{2j+1} \bar{F}_1(X_1, i) \\ &= \sum_{i=1}^j [\bar{F}_1(X_1, i) + \bar{F}_1(X_1, 2j+2-i)] + \bar{F}_1(X_1, j+1). \end{aligned} \quad (5.15)$$

As X_1 has a nonincreasing PMF, the function $\bar{F}_1(X_1, i)$ is convex. Thus,

$$\bar{F}_1(X_1, i) + \bar{F}_1(X_1, 2j+2-i) \geq 2\bar{F}_1(X_1, j+1)$$

for $1 \leq i \leq j$. Inserting this inequality into (5.15) then gives (5.14).

6. Illustrations in insurance

We present four applications in insurance of the bounds obtained in Sections 3 and 5. Many other examples in life insurance and risk theory could be considered (see, e.g. the books by Goovaerts *et al.* (1990), Kaas *et al.* (1994), (2008), and Asmussen and Albrecher (2010)).

6.1. Solvency capital requirement

In the context of solvency II, let us examine the problem of estimating the SCR (solvency capital requirement) for a given risk. A standard approximation formula is $\text{SCR} = q\sigma$, where $q > 0$ is a quantile factor and σ is the standard deviation of the random loss. Using a standard collective risk model, we have

$$\sigma^2 = \text{var}(W)\mathbb{E}(N) + [\mathbb{E}(W)]^2 \text{var}(N), \quad (6.1)$$

where N is the number of claims and W is an arbitrary claim amount independent of N .

Now, following Lefèvre and Loisel (2010), we consider the business line C_{27} , ‘drought and earthquake’, inside some French data. For heavy-tailed risks, an admissible value for the quantile factor is $q = 6$. Concerning the claims, we take $\mathbb{E}(W(C_{27})) = 1000$, $\text{var}(W(C_{27})) = 2500^2$, and N is assumed to have a bounded support $\{0, \dots, n\}$ with $\mathbb{E}(N) = 0.37$.

A reasonable assumption might be that the PMF of N is a nonincreasing convex function. In that case, $N_{\min}^{(t=2, s=2)} \leq_{2-\text{cx}} N \leq_{2-\text{cx}} N_{\max}^{(t=2, s=2)}$, where the lower and upper bounds have PMFs given by (3.9) and (3.10), respectively. Let us recall that the extrema in Lefèvre and Loisel (2010) are obtained for a PMF which is nonincreasing convex on $\{0, \dots, n\}$ only; they are denoted by $N_{\min}^{(2,2)}(n)$ and $N_{\max}^{(2,2)}(n)$. As indicated in Section 3.2, $N_{\min}^{(2,2)} = N_{\min}^{(2,2)}(n)$, but $N_{\max}^{(2,2)} \neq N_{\max}^{(2,2)}(n)$.

For the upper bound $N_{\max}^{(2,2)}$, we get, from (3.10),

$$\text{var}(N_{\max}^{(2,2)}) = \frac{6v_1}{n(n+1)(n+2)} \sum_{j=1}^n j^2(n-j+1) - v_1^2 = \frac{v_1(n+1)}{2} - v_1^2. \quad (6.2)$$

In Table 1 we provide the variance (6.2) and the SCR estimated using (6.1) for the business line C_{27} as a function of n . Of course, these bounds increase with n . They are also sharper than the corresponding bounds calculated with $N_{\max}^{(2,2)}(n)$. If $n = 10$ for instance, $\text{var}(N_{\max}^{(2,2)}(10)) = 2.453$ and $\text{SCR}(N_{\max}^{(2,2)}(10)) = 13098.2$.

6.2. Total risk of pension fund

Let us consider the example, discussed in Kaas and Goovaerts (1987), of a pension fund that covers the risk of an active married participant dying. The authors use a lifetable for Dutch government employees to estimate the first two moments of the total risk X of the pension

TABLE 1: Upper bounds on $\text{var}(N)$ and SCR for C_{27} when $t = s = 2$.

$n(C_{27})$	$\text{var}(N_{\max}^{(2,2)})$	$\text{SCR}(N_{\max}^{(2,2)})$
5	0.9731	10 875.73
10	1.8981	12 311.85
20	3.7481	14 770.97
30	5.5981	16 875.47
40	7.4481	18 745.18

fund; this gives $\mathbb{E}(X) = 27.63$ and $\mathbb{E}(X^2) = 1893$. Then, bounds for the stop-loss premium with a retention of 53.6982 are computed on the basis of these moments and with $n = 1000$ as the largest possible value of X . The restriction to unimodal distributions with a mode at 0 is also examined. The lower and upper bounds obtained are (0.41, 8.23) in the general case and (1.26, 7.16) under the unimodality assumption.

Now, let us suppose that the mean of X is known but not the variance, and that the PMF of X is not only nonincreasing but also convex. The absence of information on $\mathbb{E}(X^2)$ is a big drawback, of course. The 2-convex extrema for X , $[X_{\min}^{(t=2,s=2)}, X_{\max}^{(t=2,s=2)}]$, have PMFs given by (3.9) and (3.10). The bounds for the associated stop-loss premiums then follow easily and are equal to (1.29, 23.42). Observe that the lower bound is slightly higher, and thus better, than 1.26, while the upper bound is very large.

6.3. Percentile risk aversion

Following, e.g. De Jong and Madan (2011), under certain assumptions, the capital m required for a risk X with DF F_X is of the form

$$m = - \int_{-\infty}^{\infty} x \, d\Psi[F_X(x)] = -\mathbb{E}(X\psi(U_X))$$

for some concave distortion function Ψ on $[0, 1]$, or some nonincreasing percentile aversion function $\psi = d\Psi/dx$, where $U_X \equiv F_X(X)$ is uniformly distributed on $[0, 1]$. The risk margin is defined as $m + \mathbb{E}(X)$. If $\mathbb{E}(\psi(U_X)) = 1$, it is then given by

$$m + \mathbb{E}(X) = -\text{cov}(X, \psi(U_X)) = \sigma_\psi \sigma_X \rho_{-X},$$

where σ_ψ is the standard deviation of $\psi(U_X)$, σ_X is the standard deviation of X , and ρ_{-X} is negative of the correlation coefficient between X and $\psi(U_X)$. Note that σ_ψ does not depend on X , and $\rho_{-X} \geq 0$ as ψ is a nonincreasing function. The factor σ_ψ is a conservatism factor reflecting risk aversion, and ρ_X is the portion of σ_X taken into account in the risk margin.

De Jong and Madan (2011) considered a flexible class of risk aversion functions depending on a stress parameter γ . A plot of the density $\psi(u)$, $0 \leq u \leq 1$, is displayed in Figure 1 of their paper for different values of γ . As long as $\gamma > 0$, $\psi(u) \rightarrow \infty$ as $u \rightarrow 0$. The higher the level of γ , the higher the conservatism factor σ_ψ . The most cautious case in that figure is $\gamma = 0.75$, giving $\sigma_\psi = 2.06$. This case also seems to yield the smallest value of γ for which ψ is convex.

One could wonder whether it is possible to get high values for σ_ψ using other functions ψ that are nonincreasing convex on a bounded support $[0, b]$ with $\mathbb{E}(\psi(U_X)) = 1$. To answer this question, we evaluate the maximal value of σ_ψ for such a function ψ . Note that the DF of $\psi(U_X)$ is $1 - \psi^{-1}$, which is a concave function, so the density of $\psi(U_X)$ is nonincreasing. Thus, the highest level of σ_ψ is given by the standard deviation of the density (3.2) with $v_1 = 1$; it is equal to $\sqrt{2b/3 - 1}$. For instance, if $b = 100$, this bound gives 8.103.

Let us now add the constraint that the density of $\psi(U_X)$ is also convex. In that case, the maximal value for σ_ψ is given by the standard deviation of the density (3.4) with $v_1 = 1$; it is equal to $\sqrt{b/2 - 1}$. For $b = 100$, this bound is equal to 7, which is reasonable in comparison to the cautious value 2.06 considered in De Jong and Madan (2011).

6.4. Exponential premium principle

Goovaerts *et al.* (2003) showed that many risk measures and premium principles can be derived by minimizing a Markov bound. A typical example is the classical exponential

premium principle. For a continuous risk variable X , applying Markov's inequality to $\exp(\beta X)$, where $\beta > 0$, gives $\mathbb{P}[X \geq x] \leq \mathbb{E}(\exp(\beta X)) / \exp(\beta x)$, $x \geq 0$. This bound is nontrivial if it is at most 1. It equals 1 when $x \equiv \pi$ given by

$$\pi = \frac{1}{\beta} \ln \mathbb{E}(e^{\beta X}), \quad (6.3)$$

i.e. π is the exponential premium of parameter β . Under this choice, we similarly have

$$\mathbb{P}[X \geq \pi + y] \leq \frac{\mathbb{E}(e^{\beta X})}{e^{\beta(\pi+y)}} = e^{-\beta y}, \quad y \geq 0. \quad (6.4)$$

Suppose that the density of X is nonincreasing. Then $\exp(\beta X)$ has a nonincreasing density and applying (5.7) allows us to refine inequality (6.4) to

$$\mathbb{P}[X > \pi + y] \leq \frac{1}{2} e^{-\beta y}, \quad y \geq 0.$$

If the density of X is also convex, it follows similarly from (5.8) that the factor $\frac{1}{2}$ is replaced with $\frac{4}{9}$. Bound (6.4) and these improvements could be used to estimate, for instance, the probability that a stop-loss reinsurance treaty is activated.

In general, the value of β is defined by the market. Nevertheless, within an enterprise risk management process, insurance companies start to define so-called risk limits. One possibility, among others, could consist in setting a lower limit on the premium to guarantee that the probability of losing more than a fixed amount K (in excess of π) is smaller than a level ε . Using the Markov bound (6.3), the trigger level $\beta \equiv \beta(K, \varepsilon)$ satisfies $\exp(-\beta K) = \varepsilon$, so that $\beta = -(1/K) \ln(\varepsilon)$. If the density of X is nonincreasing, this level decreases to $\beta_1 = -(1/K) \ln(2\varepsilon)$. If the density is also convex, we obtain $\beta_2 = -(1/K) \ln(9\varepsilon/4)$. For example, if $K = 100$ and $\varepsilon = 0.05$, then $\beta = 0.0299$, $\beta_1 = 0.0230$ (a gain of 23%), and $\beta_2 = 0.0218$ (a gain of 27%).

As a consequence, more competitive premium levels above the risk limit are possible. Suppose, for instance, that X has a gamma distribution with parameters (θ, α) , i.e. $\mathbb{E}(\exp(\beta X)) = [\theta/(\theta - \beta)]^\alpha$. By (6.3), the lower risk limit on π is given by $(\alpha/\beta) \ln[\theta/(\theta - \beta)]$. Let us choose $\alpha = 0.2$ and $\theta = 0.1$, so that the density of X is decreasing convex. With β , β_1 , and β_2 above, the risk limits are then equal to 2.376, 2.273, and 2.256, respectively.

Appendix A

In this section we collect some notions and technical results used in Sections 2, 3, and 5.

A.1. Combinatorial identities

The following relations are straightforward. We have

$$\sum_{j=i}^m \binom{j}{i} = \binom{m+1}{i+1}, \quad 0 \leq i \leq m, \quad (A.1)$$

and if the operator Δ operates on j then, for $i, k \geq 0$,

$$\Delta^k \binom{j}{i+k} = \binom{j}{i}, \quad i+k \leq j, \quad (A.2)$$

$$\Delta^k \binom{m-j}{i+k} = (-1)^k \binom{m-j-k}{i}, \quad i+k \leq m-j. \quad (A.3)$$

The next relations are less standard for nonspecialists.

Lemma A.1. Let $f: \mathbb{N} \rightarrow \mathbb{R}$ be an arbitrary function. For $t \geq 1$,

$$(-1)^t \sum_{j=k}^{\infty} \binom{j-k+t-1}{t-1} \Delta^t f(j) = f(k), \quad k \geq 0, \quad (\text{A.4})$$

$$(-1)^t \sum_{j=0}^{\infty} \binom{j+t}{t} \Delta^t f(j) = \sum_{j=0}^{\infty} f(j). \quad (\text{A.5})$$

Proof. Let $g(k)$, $k \geq 0$, be the left-hand side of (A.4). Expanding Δ^t , we obtain

$$\begin{aligned} g(k) &= (-1)^t \sum_{j=k}^{\infty} \binom{j-k+t-1}{t-1} \sum_{l=0}^t \binom{t}{l} (-1)^{t-l} f(j+l) \\ &= \sum_{l=0}^t \binom{t}{l} (-1)^l \sum_{j=k+l}^{\infty} f(j) \binom{j-k-l+t-1}{t-1} \\ &= \sum_{j=k}^{\infty} f(j) \binom{j-k+t-1}{t-1} + \sum_{j=k+1}^{\infty} f(j) \sum_{l=1}^t \binom{t}{l} (-1)^l \binom{j-k-l+t-1}{t-1} \\ &= f(k) + \sum_{j=k+1}^{\infty} f(j) \Delta^t \binom{j-k-1}{t-1} \\ &= f(k), \end{aligned}$$

after using (A.2), and (A.4) follows. Now, summing (A.4) over k yields

$$(-1)^t \sum_{k=0}^{\infty} \sum_{j=k}^{\infty} \binom{j-k+t-1}{t-1} \Delta^t f(j) = \sum_{j=0}^{\infty} f(j),$$

which becomes (A.5) after permuting the two sums and using (A.1).

A.2. s -convex stochastic orders

These orders have been mostly studied in Denuit *et al.* (1998), (1999a) for continuous distributions, and in Lefèvre and Utev (1996) and Denuit and Lefèvre (1997) for discrete distributions. Basic points of the theory are recalled below.

Continuous case. With $\mathbb{R}_+$ -valued X , define the iterated right-tail DFs of X by $\bar{F}_1(X, x) = \mathbb{P}[X > x]$ and

$$\bar{F}_{i+1}(X, x) = \int_x^{\infty} \bar{F}_i(X, y) dy, \quad x \geq 0, i \geq 1. \quad (\text{A.6})$$

An equivalent expression is

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

Now, let s be some integer greater than or equal to 1. Denote by $\mathcal{F}_s$ the set of s -convex functions on $\mathbb{R}_+$, i.e.

$$\mathcal{F}_s = \{f: f^{(s)}(x) \geq 0, x \geq 0\}. \quad (\text{A.7})$$

A more general concept of s -convex functions is used in interpolation theory (see, e.g. Karlin and Studden (1966)). For the definition below, however, it is not restrictive to consider the class of functions (A.7).

Definition A.1. If X and Y are continuous random variables on $\mathbb{R}_+$, then X is smaller than Y in the s -convex stochastic sense, written $X \preceq_s Y$, when

$$\mathbb{E}f(X) \leq \mathbb{E}f(Y) \quad \text{for all functions } f \in \mathcal{F}_s, \quad (\text{A.8})$$

provided the expectations exist.

Note that $X \preceq_s Y$ implies that X and Y have necessarily the same first $s - 1$ moments. In fact, we can prove that a condition equivalent to (A.8) is

$$\mathbb{E}(X^i) = \mathbb{E}(Y^i), \quad i = 0, 1, \dots, s - 1, \quad \text{and} \quad \bar{F}_s(X, j) \leq \bar{F}_s(Y, j), \quad j \geq s. \quad (\text{A.9})$$

Discrete case. With $\mathbb{N}_0$ -valued X , the iterated right-tail DFs of X are defined by $\bar{F}_0(X, j) = \mathbb{P}[X = j]$, $j \geq 0$, and

$$\bar{F}_{i+1}(X, j) = \sum_{k=j}^{\infty} \bar{F}_i(X, k), \quad j \geq 0, i \geq 0. \quad (\text{A.10})$$

They can be also expressed as

$$\bar{F}_{i+1}(X, j) = \mathbb{E} \binom{X - j + i}{i}, \quad j \geq 0, i \geq 0. \quad (\text{A.11})$$

For s a positive integer, let $\mathcal{F}_s$ denote the set of functions that are s -convex on $\mathbb{N}_0$, i.e.

$$\mathcal{F}_s = \{f: \Delta^s f(j) \geq 0, j \geq 0\}. \quad (\text{A.12})$$

Definition A.2. If X and Y are random variables on $\mathbb{N}_0$, then $X \preceq_s Y$ when condition (A.8) is satisfied with respect to the class (A.12).

Here too, the conditions in (A.9) are equivalent to (A.8).

Proof of Property 5.1. Let us consider the function $c(r, t)$ for $t \in \mathbb{R}_+$. To show that this function is decreasing, we first obtain

$$\frac{d \log c(r, t)}{dt} = u(r) - u(0),$$

where

$$u(x) = \psi(x + t + 1) - \log(x + t), \quad x > 0,$$

with $\psi(x) = d \log \Gamma(x)/dx$. Thus, it suffices to prove that $u'(x) < 0$, i.e. $\psi'(x + 1) < 1/x$. Using a known expansion for $\psi'(x)$ (see, e.g. Abramowitz and Stegun (1972, Formula 6.4.10)), we obtain

$$\psi'(x + 1) = \sum_{i=1}^{\infty} \frac{1}{(x + i)^2} < \sum_{i=1}^{\infty} \left(\frac{1}{x + i - 1} - \frac{1}{x + i} \right) = \frac{1}{x},$$

as desired. For the limit as $t \rightarrow \infty$, we write

$$c(r, t) = \frac{r^r}{\Gamma(r + 1)} \left(1 - \frac{r}{r + t} \right)^t \left(\frac{t}{r + t} \right)^r \frac{\Gamma(r + t + 1)}{t^r \Gamma(t + 1)}.$$

By Formula 6.1.46 of Abramowitz and Stegun (1972), the last fraction on the right-hand side tends to 1, so the announced limit follows.

Acknowledgement

This research was conducted while C. Lefèvre was visiting the I.S.F.A., Université Lyon 1. It was partly funded by the research chair Actuariat Durable sponsored by Milliman Paris, the research chair Management de la Modélisation sponsored by BNP Paribas Cardif, and a grant from the Belgian FNRS and the Banque Nationale de Belgique. We are grateful to S. Utev for very helpful discussions. We also thank the anonymous referee for many useful remarks, improvements, and suggestions.

References

- ABRAMOWITZ, M. AND STEGUN, I. A. (eds) (1972). *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. Dover, New York.
- ASMUSSEN, S. AND ALBRECHER, H. (2010). *Ruin Probabilities*, 2nd edn. World Scientific, Hackensack, NJ.
- BALABDAOUI, F. AND WELLNER, J. A. (2007). Estimation of a k -monotone density: limit distribution theory and the spline connection. *Ann. Statist.* **35**, 2536–2564.
- BERTIN, E. M. J., CUCULESCU, I. AND THEODORESCU, R. (1997). *Unimodality of Probability Measures*. Kluwer, Dordrecht.
- CONSTANTINESCU, C., HASHORVA, E. AND JI, L. (2011). Archimedian copulas in finite and infinite dimensions—with applications to ruin problems. *Insurance Math. Econom.* **49**, 487–495.
- COX, D. R. (1962). *Renewal Theory*. John Wiley, New York.
- DE JONG, P. AND MADAN, D. B. (2011). Capital adequacy of financial enterprises. Working paper. Available at <http://ssrn.com/abstract=1761107>.
- DENUIT, M. AND LEFÈVRE, C. (1997). Some new classes of stochastic order relations among arithmetic random variables, with applications in actuarial sciences. *Insurance Math. Econom.* **20**, 197–213.
- DENUIT, M., DE VYLDER, E. AND LEFÈVRE, C. (1999a). Extremal generators and extremal distributions for the continuous s -convex stochastic orderings. *Insurance Math. Econom.* **24**, 201–217.
- DENUIT, M., LEFÈVRE, C. AND MESFIQUI, M. (1999b). On s -convex stochastic extrema for arithmetic risks. *Insurance Math. Econom.* **25**, 143–155.
- DENUIT, M., LEFÈVRE, C. AND SHAKED, M. (1998). The s -convex orders among real random variables, with applications. *Math. Inequal. Appl.* **1**, 585–613.
- DENUIT, M., LEFÈVRE, C. AND SHAKED, M. (2000). Stochastic convexity of the Poisson mixture model. *Methodology Comput. Appl. Prob.* **2**, 231–254.
- DENUIT, M., LEFÈVRE, C. AND UTEV, S. (1999c). Generalized stochastic convexity and stochastic orderings of mixtures. *Prob. Eng. Inf. Sci.* **13**, 275–291.
- DHARMADHIKARI, S. AND JOAG-DEV, K. (1988). *Unimodality, Convexity, and Applications*. Academic Press, Boston, MA.
- FELLER, W. (1971). *An Introduction to Probability Theory and Its Applications*, Vol. II, 2nd edn. John Wiley, New York.
- FURMAN, E. AND ZITIKIS, R. (2009). Weighted pricing functionals with applications to insurance: an overview. *N. Amer. Actuarial J.* **13**, 483–496.
- GERBER, H. U. (1972). Ein satz von Khintchin und die varianz von unimodalen. *Bull. Swiss Assoc. Actuaries*, 225–231.
- GNEITING, T. (1999). Radial positive definite functions generated by Euclid's hat. *J. Multivariate Anal.* **69**, 88–119.
- GOOVAERTS, M. J., KAAS, R., DHAENE, J. AND TANG, Q. (2003). A unified approach to generate risk measures. *ASTIN Bull.* **33**, 173–192.
- GOOVAERTS, M. J., KAAS, R., VAN HEERWAARDEN, A. E. AND BAUWELINCKX, T. (1990). *Effective Actuarial Methods*. North-Holland, Amsterdam.
- KAAS, R. AND GOOVAERTS, M. J. (1987). Unimodal distributions in insurance. *Bull. Assoc. R. Actuaire Belges* **81**, 61–66.
- KAAS, R., VAN HEERWAARDEN, A. E. AND GOOVAERTS, M. J. (1994). *Ordering of Actuarial Risks*. CAIRE, Brussels.
- KAAS, R., GOOVAERTS, M. J., DHAENE, J. AND DENUIT, M. (2008). *Modern Actuarial Risk Theory: Using R*. Springer, Heidelberg.
- KARLIN, S. AND STUDDEN, W. J. (1966). *Tchebycheff Systems: With Applications in Analysis and Statistics*. John Wiley, New York.
- LEFÈVRE, C. AND LOISEL, S. (2010). Stationary-excess operator and convex stochastic orders. *Insurance Math. Econom.* **47**, 64–75.
- LEFÈVRE, C. AND UTEV, S. (1996). Comparing sums of exchangeable Bernoulli random variables. *J. Appl. Prob.* **33**, 285–310.
- LEFÈVRE, C. AND UTEV, S. (2013). Convolution property and exponential bounds for symmetric monotone densities. *ESAIM Prob. Statist.* **17**, 605–613.

- LÉVY, P. (1962). Extensions d'un théorème de D. Dugué et M. Girault. *Z. Wahrscheinlichkeitsth.* **1**, 159–173.
- PAKES, A. G. (1996). Length biasing and laws equivalent to the log-normal. *J. Math. Anal. Appl.* **197**, 825–854.
- PAKES, A. G. (1997). Characterization by invariance under length-biasing and random scaling. *J. Statist. Planning Infer.* **63**, 285–310.
- PAKES, A. G. (2003). Biological applications of branching processes. In *Stochastic Processes: Modelling and Simulation* (Handbook Statist. **21**), eds D. N. Shanbhag and C. R. Rao, North-Holland, Amsterdam, pp. 693–773.
- PAKES, A. G. AND NAVARRO, J. (2007). Distributional characterizations through scaling relations. *Austral. N. Ze. J. Statist.* **49**, 115–135.
- PATIL, G. P. AND RAO, C. R. (1978). Weighted distributions and size-biased sampling with applications to wildlife populations and human families. *Biometrics* **34**, 179–189.
- PEČARIĆ, J. E., PROSCHAN, F. AND TONG, Y. L. (1992). *Convex Functions, Partial Orderings, and Statistical Applications*. Academic Press, Boston, MA.
- SHAKED, M. AND SHANTHIKUMAR, J. G. (2007). *Stochastic Orders*. Springer, New York.
- STEUTEL, F. W. AND VAN HARN, K. (1979). Discrete analogues of self-decomposability and stability. *Ann. Prob.* **7**, 893–899.
- WILLIAMSON, R. E. (1956). Multiply monotone functions and their Laplace transforms. *Duke Math. J.* **23**, 189–207.