S-Convex Approximations

Michel Denuit
Institut de Statistique
Université Catholique de Louvain
Voie du Roman Pays, 20
B-1348 Louvain-la-Neuve, Belgium
denuit@stat.ucl.ac.be

Claude Lefèvre
Institut de Statistique et de Recherche Opérationnelle
Université Libre de Bruxelles
CP 210, Boulevard du Triomphe
B-1050 Bruxelles, Belgium
clefevre@ulb.ac.be

Moshe Shaked
Department of Mathematics
University of Arizona
Tucson, Arizona 85721
USA
shaked@math.arizona.edu

November 1999

Abstract

In this paper we examine the structure of the s-convex stochastic extrema in the continuous case. After having briefly recalled the key tools, we revisit the theory of moment spaces. Specifically, using representation results of moment matrices *à la* Lindsay (1989a), we provide conditions for the admissibility of moment sequences in terms of lower bounds on the number of support points of the corresponding cumulative distribution functions (cdf's, in short). We then point out two special distributions with a minimal number of support points; these distributions happen to be the s-convex extremal distributions. The supports of these extrema are investigated: it is shown that the support points correspond to the roots of some polynomials, and an efficient method for the complete determination of the cdf's of these extrema is described. A study of the goodness of fit, of the approximation of an arbitrary element in a moment space by one of the stochastic s-convex extrema, is then given. Using standard ideas from linear regression, we derive Tchebycheff-type inequalities which extend previous results of Lindsay (1989b), and we establish some limit theorems involving the moment matrices. Finally we detail some applications in insurance theory and in economics. Namely, we provide bounds on the Lundberg's coefficient in risk theory, on the actual interest rate relating to a life insurance policy, and on the certainty equivalence of an investment. These bounds are obtained with the aid of the s-convex extrema, and are determined only by the support and the first few moments of the underlying distribution.

AMS Subject Classification: 60E15

Key words and phrases: Moment spaces, s-convex orders, stochastic extrema, stochastic approximations, Tchebycheff-type inequalities, risk theory, Lundberg's coefficient, life insurance, certainty equivalence.

1 Introduction and Motivation

For any integer $s \geq 1$, the s-convex stochastic order is a partial order on a set of cumulative distribution functions (cdf's, in short) which share the same first $s-1$ moments. These orders generalize the well known stochastic dominance and convex orders (corresponding to $s = 1$ and $s = 2$, respectively). For arithmetic random variables these orders have been introduced in Denuit & Lefèvre (1997b), and for continuous random variables they have been introduced in Denuit, Lefèvre & Shaked (1998). Furthermore, Denuit, Lefèvre & Utev (1999) considered the s-convex order for discrete random variables valued in an arbitrary countable subset of the real line $\mathbb{R}$. Many results available for the stochastic dominance and convex orders remain valid for the entire class of the s-convex orders. A useful feature of the s-convex orders is that it is possible to derive stochastic extrema in moment spaces, that is, amongst all the cdf's with identical $s-1$ first moments, and whose support is contained in a bounded interval $[a, b]$, say. These extrema have been obtained by Denuit & Lefèvre (1997b) and Denuit, Lefèvre & Mesfioui (1999) in the discrete case, and by Denuit, Lefèvre & Shaked (1998) and Denuit, De Vylder & Lefèvre (1999) in the continuous case. It is worth mentioning that the question of extrema with respect to some stochastic ordering among cdf's with prescribed moments has been studied in many earlier works in probability and statistics; see, for example, Harris (1959, 1962), Karlin & Shapley (1953), Karlin & Studden (1966), Rolski (1976) and Stoyan (1983, Section 1.9).

In the present paper, we further examine the structure of the s-convex stochastic extrema in the continuous case. After having briefly recalled the key tools in Section 2, we revisit the theory of moment spaces in Section 3. To be specific, using representation results of moment matrices *à la* Lindsay (1989a), we provide conditions for the admissibility of moment sequences in terms of lower bounds on the number of support points of the corresponding cdf's. In Section 4, we point out two special distributions with a minimal number of support points; these distributions happen to be the s-convex extremal distributions. The supports of these extrema are investigated; it is shown *inter alia* that the support points correspond to the roots of some polynomials, and an efficient method for the complete determination of the cdf's of these extrema is described. Then, Section 5 is devoted to the study of the goodness of fit of the approximation of an arbitrary element in a moment space by one of the stochastic s-convex extrema. Using standard ideas from linear regression, we derive Tchebycheff-type inequalities which extend previous results of Lindsay (1989b), and we establish some limit theorems involving the moment matrices. Finally, in Section 6, we detail some applications in insurance theory and in economics. Namely, we provide bounds on the Lundberg's coefficient in risk theory, on the actual interest rate relating to a life insurance policy, and on the certainty equivalence of an investment. These bounds are obtained with the aid of the s-convex extrema, and are determined only by the support and the first few moments of the underlying distribution.

2 The s-Convex Orders

Let $\mathcal{S}$ be a subinterval (finite, half finite, or infinite) of the real line $\mathbb{R}$. Let ϕ be a real function defined on $\mathcal{S}$. For a positive integer s and for $x_0 < x_1 < \cdots < x_s \in \mathcal{S}$ define recursively the divided differences

$$[x_k]\phi = \phi(x_k), \qquad k = 0, 1, \ldots, s,$$

and

$$[x_0, x_1, \ldots, x_k]\phi = \frac{[x_1, x_2, \ldots, x_k]\phi - [x_0, x_1, \ldots, x_{k-1}]\phi}{x_k - x_0}, \qquad k = 1, 2, \ldots, s.$$

The function ϕ is called s-convex if

$$[x_0, x_1, \ldots, x_s]\phi \geq 0 \quad \text{for all} \quad x_0 < x_1 < \cdots < x_s \in \mathcal{S}.$$

Many properties of s-convex functions can be found, for instance, in Popoviciu (1933) and in Pečarić, Proschan & Tong (1992, Section 1.3). It is worth mentioning that the 1-convex functions are the non-decreasing functions, and that the 2-convex functions are the usual convex functions on $\mathcal{S}$. If the sth derivative $\phi^{(s)}$ of ϕ exists, then ϕ is s-convex if, and only if, $\phi^{(s)} \geq 0$ on $\mathcal{S}$. The class of the s-convex functions is the closure of the class

$$\{\phi : \mathcal{S} \to \mathbb{R} \big| \phi^{(s)} \geq 0\}$$

in the topology of the pointwise convergence.

Let X and Y be two random variables that take on values in $\mathcal{S}$. If

$$E[\phi(X)] \leq E[\phi(Y)] \qquad \text{for all } s\text{-convex functions } \phi,$$

provided the expectations exist, then X is said to be smaller than Y in the s-convex sense (denoted $X \preceq^{\mathcal{S}}_{s\text{-cx}} Y$). Note that the 1-convex order is the usual stochastic order (see, for example, Section 1.A in Shaked & Shanthikumar, 1994), and the 2-convex order is the usual convex order (see, for example, Section 2.A in Shaked & Shanthikumar, 1994).

Let us now give a sufficient (but not necessary) condition for the order $\preceq^{\mathcal{S}}_{s\text{-cx}}$. For a real function ϕ defined on $\mathcal{S}$, define the number of sign changes of ϕ on $\mathcal{S}$ by

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

where $S^-[\phi(x_1), \phi(x_2), \ldots, \phi(x_n)]$ is the number of sign changes of the indicated sequence, zero terms being discarded, and the supremum is extended over all $x_1 < x_2 < \cdots < x_n$, $x_i \in \mathcal{S}$, $n < \infty$. Let X and Y be two random variables, taking on values in $\mathcal{S}$, with the same $s-1$ moments. Denote by F_X and F_Y the cdf's of X and Y, respectively. In Denuit, Lefèvre & Shaked (1998) it is shown that

$$S^-(F_X - F_Y) = s - 1 \Rightarrow X \preceq^{\mathcal{S}}_{s\text{-cx}} Y, \tag{2.1}$$

provided the last sign of $F_X - F_Y$ is a $+$. Also, if $S^-(f_X - f_Y) = s$, and if the last sign of $f_Y - f_X$ is a $+$, then $X \preceq^{\mathcal{S}}_{s\text{-cx}} Y$, where f_X and f_Y denote the (discrete or continuous) probability density functions of X and Y, respectively.

3 Moment Spaces

Every polynomial of degree at most $s-1$ is an s-convex function. Thus $X \preceq_{s\text{-cx}}^{\mathcal{S}} Y$ implies that $E[X^k] = E[Y^k]$, $k = 1, 2, \ldots, s-1$; that is, only random variables with identical first $s-1$ moments can be compared in the s-convex sense. We denote by $\mathcal{B}_s(\mathcal{S}; \mu_1, \mu_2, \ldots, \mu_{s-1})$ the class of all (cdf's of) random variables with support in $\mathcal{S}$ having $\mu_1, \mu_2, \ldots, \mu_{s-1}$ as their first $s-1$ moments. In the remainder of the paper, we assume that $\mathcal{S}$ is compact; that is, $\mathcal{S} = [a, b]$ for some $-\infty < a < b < \infty$. Then $\mathcal{B}_s([a, b]; \mu_1, \mu_2, \ldots, \mu_{s-1})$ is a Krein-Milman space; that is, it is convex and compact and it is the closed convex hull of its extremal points. Each of the extremal points of $\mathcal{B}_s([a, b]; \mu_1, \mu_2, \ldots, \mu_{s-1})$ has at most s points of support; see, for example, De Vylder (1996).

Of course not every vector $(\mu_1, \mu_2, \ldots, \mu_{s-1}) \in \mathbb{R}^{s-1}$ corresponds to a moments sequence. We denote by $\mathcal{D}_s([a, b])$ the set of all the possible $(s-1)$-dimensional moments vectors; that is,

$$\mathcal{D}_s([a, b]) = \{(\mu_1, \mu_2, \ldots, \mu_{s-1}) \in \mathbb{R}^{s-1} \big| \mathcal{B}_s([a, b]; \mu_1, \mu_2, \ldots, \mu_{s-1}) \neq \emptyset\}.$$

Note that $\mathcal{D}_s([a, b])$ is the convex hull of $\{(\xi, \xi^2, \ldots, \xi^{s-1}) \big| \xi \in [a, b]\}$. Let $\mathcal{D}_s^{\circ}([a, b])$ denote the topological interior of $\mathcal{D}_s([a, b])$. For example, if $\xi \in [a, b]$ then $(\xi, \xi^2, \ldots, \xi^{s-1}) \in \mathcal{D}_s([a, b])$ because the random variable $X =_{\text{a.s.}} \xi$ has this moments sequence, but $(\xi, \xi^2, \ldots, \xi^{s-1}) \notin \mathcal{D}_s^{\circ}([a, b])$.

Put $\mu_0 \equiv 1$, and for any non-negative integer k, and for any sequence $\mu_0, \mu_1, \ldots, \mu_{2k+2}$ of real numbers, define the $(k+1) \times (k+1)$ matrices

$$\boldsymbol{P}_k = \begin{pmatrix} \mu_0 & \mu_1 & \mu_2 & \cdots & \mu_k \\ \mu_1 & \mu_2 & \mu_3 & \cdots & \mu_{k+1} \\ \mu_2 & \mu_3 & \mu_4 & \cdots & \mu_{k+2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mu_k & \mu_{k+1} & \mu_{k+2} & \cdots & \mu_{2k} \end{pmatrix}, \qquad \boldsymbol{Q}_k = \begin{pmatrix} \mu_1 & \mu_2 & \mu_3 & \cdots & \mu_{k+1} \\ \mu_2 & \mu_3 & \mu_4 & \cdots & \mu_{k+2} \\ \mu_3 & \mu_4 & \mu_5 & \cdots & \mu_{k+3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mu_{k+1} & \mu_{k+2} & \mu_{k+3} & \cdots & \mu_{2k+1} \end{pmatrix},$$

and

$$\boldsymbol{R}_k = \begin{pmatrix} \mu_2 & \mu_3 & \mu_4 & \cdots & \mu_{k+2} \\ \mu_3 & \mu_4 & \mu_5 & \cdots & \mu_{k+3} \\ \mu_4 & \mu_5 & \mu_6 & \cdots & \mu_{k+4} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mu_{k+2} & \mu_{k+3} & \mu_{k+4} & \cdots & \mu_{2k+2} \end{pmatrix}.$$

For $k = 0, 1, \ldots$, consider the condition

$$\boldsymbol{Q}_k - a\boldsymbol{P}_k \text{ and } b\boldsymbol{P}_k - \boldsymbol{Q}_k \text{ are non-negative definite,} \tag{3.1}$$

and for $k = 1, 2, \ldots$, consider the condition

$$\boldsymbol{P}_k \text{ and } -ab\boldsymbol{P}_{k-1} + (a+b)\boldsymbol{Q}_{k-1} - \boldsymbol{R}_{k-1} \text{ are non-negative definite.} \tag{3.2}$$

From results in De Vylder (1996, pages 362–366) it follows that

$$(\mu_1, \mu_2, \ldots, \mu_{2s+1}) \in \mathcal{D}_{2s+2}([a,b])$$
$$\text{if, and only if, (3.1) holds for } k = s \text{ and (3.2) holds for } k = s, \qquad s \geq 0,$$

and

$$(\mu_1, \mu_2, \ldots, \mu_{2s}) \in \mathcal{D}_{2s+1}([a,b])$$
$$\text{if, and only if, (3.1) holds for } k = s-1 \text{ and (3.2) holds for } k = s, \qquad s \geq 1.$$

The topological interior $\mathcal{D}_s^\circ([a,b])$ of $\mathcal{D}_s([a,b])$ can be characterized by strict inequalities as follows. For any square matrix $\boldsymbol{A}$, let $\|\boldsymbol{A}\|$ denote the determinant of $\boldsymbol{A}$. For $k = 0, 1, \ldots$, consider the condition

$$\|\boldsymbol{Q}_k - a\boldsymbol{P}_k\| > 0 \text{ and } \|b\boldsymbol{P}_k - \boldsymbol{Q}_k\| > 0, \tag{3.3}$$

and for $k = 1, 2, \ldots$, consider the condition

$$\|\boldsymbol{P}_k\| > 0 \text{ and } \|-ab\boldsymbol{P}_{k-1} + (a+b)\boldsymbol{Q}_{k-1} - \boldsymbol{R}_{k-1}\| > 0. \tag{3.4}$$

From results in De Vylder (1996, page 371) it follows that for $s \geq 0$ we have

$$(\mu_1, \mu_2, \ldots, \mu_{2s+1}) \in \mathcal{D}_{2s+2}^\circ([a,b])$$
$$\text{if, and only if, (3.3) holds for } k \leq s \text{ and (3.4) holds for } k \leq s, \tag{3.5}$$

and that for $s \geq 1$ we have

$$(\mu_1, \mu_2, \ldots, \mu_{2s}) \in \mathcal{D}_{2s+1}^\circ([a,b])$$
$$\text{if, and only if, (3.3) holds for } k \leq s-1 \text{ and (3.4) holds for } k \leq s. \tag{3.6}$$

Explicit expressions of $\mathcal{D}_s^\circ([a,b])$ for $s = 1$ to 5 can be found in De Vylder (1996).

For any cdf F, denote by $\mathrm{supp}(F)$ the support of F; and for any set A, let $\#(A)$ denote the number of points in A.

Proposition 3.1. (i) *Let $s \geq 0$ be an integer and let $(\mu_1, \mu_2, \ldots, \mu_{2s+1}) \in \mathcal{D}_{2s+2}([a,b])$. Then*

$$(\mu_1, \mu_2, \ldots, \mu_{2s+1}) \in \mathcal{D}_{2s+2}^\circ([a,b])$$
$$\iff \#(\mathrm{supp}(F) - \{a\}) \geq s+1 \textit{ and } \#(\mathrm{supp}(F) - \{b\}) \geq s+1$$
$$\textit{for all } F \in \mathcal{B}_{2s+2}([a,b]; \mu_1, \mu_2, \ldots, \mu_{2s+1}).$$

(ii) *Let $s \geq 1$ be an integer and let $(\mu_1, \mu_2, \ldots, \mu_{2s}) \in \mathcal{D}_{2s+1}([a,b])$. Then*

$$(\mu_1, \mu_2, \ldots, \mu_{2s}) \in \mathcal{D}_{2s+1}^\circ([a,b])$$
$$\iff \#(\mathrm{supp}(F)) \geq s+1 \textit{ and } \#(\mathrm{supp}(F) - \{a,b\}) \geq s$$
$$\textit{for all } F \in \mathcal{B}_{2s+1}([a,b]; \mu_1, \mu_2, \ldots, \mu_{2s}).$$

Proof. The result follows from Proposition A.1 in Appendix A. All four determinants in (A.1)–(A.4) are positive for all $k \leq s$ if, and only if, $\#(\mathrm{supp}(F) - \{a\}) \geq s+1$ and $\#(\mathrm{supp}(F) - \{b\}) \geq s+1$. The result stated in (i) now follows from (3.5).

The proof of (ii) is similar. The determinants in (A.1) and (A.2) are positive for all $k \leq s-1$, and the determinants in (A.3) and (A.4) are positive for all $k \leq s$, if, and only if, $\#(\mathrm{supp}(F)) \geq s+1$ and $\#(\mathrm{supp}(F) - \{a,b\}) \geq s$. The result stated in (ii) now follows from (3.6). $\square$

4 Two Special Extremal Distributions

Every moment space $\mathcal{B}_r([a,b];\mu_1,\mu_2,\ldots,\mu_{r-1})$ possesses two special extremal distributions when $(\mu_1,\mu_2,\ldots,\mu_{r-1}) \in \mathcal{D}_r^\circ([a,b])$. When r is odd, that is, $r = 2s+1$, these two extremal distributions in $\mathcal{B}_{2s+1}([a,b];\mu_1,\mu_2,\ldots,\mu_{2s})$ are $(s+1)$-atomic. The first puts positive masses $p_1,p_2,\ldots,p_{s+1}$ on the atoms $a,x_2,\ldots,x_{s+1}$, with $a < x_2 < \cdots < x_{s+1} < b$. Such a distribution is called 'initial' in Denuit, De Vylder & Lefèvre (1999) because it puts a positive mass on the left endpoint a. The $2s+1$ quantities $x_2,\ldots,x_{s+1},p_1,p_2,\ldots,p_{s+1}$ are uniquely determined by the $2s$ moment constraints, and the probability constraint $\sum_{i=1}^{s+1} p_i = 1$. Traditionally these quantities are found by solving the $2s+1$ equations $p_1 a^j + p_2 x_2^j + \cdots + p_{s+1} x_{s+1}^j = \mu_j$, $j = 0,1,\ldots,2s$. However, we describe below a far better alternative. Explicitly, first note that these variables satisfy

$$p_1 a^j + p_2 x_2^j + \cdots + p_{s+1} x_{s+1}^j = \mu_j, \qquad j = 0,1,\ldots,s. \tag{4.1}$$

Next, using (4.1) (denoting $\mu_{j;\theta} = \mu_j - \theta\mu_{j-1}$, $j = 1,2,\ldots$, $\theta \in \mathbb{R}$) we may write

$$\begin{pmatrix} 1 & x & x^2 & \cdots & x^s \\ \mu_{1;a} & \mu_{2;a} & \mu_{3;a} & \cdots & \mu_{s+1;a} \\ \mu_{2;a} & \mu_{3;a} & \mu_{4;a} & \cdots & \mu_{s+2;a} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mu_{s;a} & \mu_{s+1;a} & \mu_{s+2;a} & \cdots & \mu_{2s;a} \end{pmatrix}$$
$$= \begin{pmatrix} 0 & 0 & \cdots & 0 & 1 \\ p_2(x_2-a) & p_3(x_3-a) & \cdots & p_{s+1}(x_{s+1}-a) & 0 \\ p_2x_2(x_2-a) & p_3x_3(x_3-a) & \cdots & p_{s+1}x_{s+1}(x_{s+1}-a) & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ p_2x_2^{s-1}(x_2-a) & p_3x_3^{s-1}(x_3-a) & \cdots & p_{s+1}x_{s+1}^{s-1}(x_{s+1}-a) & 0 \end{pmatrix}$$
$$\times \begin{pmatrix} 1 & x_2 & x_2^2 & \cdots & x_2^s \\ 1 & x_3 & x_3^2 & \cdots & x_3^s \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{s+1} & x_{s+1}^2 & \cdots & x_{s+1}^s \\ 1 & x & x^2 & \cdots & x^s \end{pmatrix}.$$

Since the p_j's are positive and the x_j's are distinct it is seen that $x_2, \ldots, x_{s+1}$ must be the s distinct roots of the equation

$$\left\| \begin{matrix} 1 & x & x^2 & \cdots & x^s \\ \mu_{1;a} & \mu_{2;a} & \mu_{3;a} & \cdots & \mu_{s+1;a} \\ \mu_{2;a} & \mu_{3;a} & \mu_{4;a} & \cdots & \mu_{s+2;a} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mu_{s;a} & \mu_{s+1;a} & \mu_{s+2;a} & \cdots & \mu_{2s;a} \end{matrix} \right\| = 0. \tag{4.2}$$

Equation (4.2) determines $x_2, \ldots, x_{s+1}$, and, plugging these into (4.1), also $p_1, p_2, \ldots, p_{s+1}$ for the 'initial' extremal distribution are determined.

From (2.1) it is easily seen that the random variable that is distributed according to this 'initial' extremal distribution is the smallest, in the sense of the order $\preceq^{[a,b]}_{(2s+1)\text{-cx}}$, of all random variables with cdf in $\mathcal{B}_{2s+1}([a, b]; \mu_1, \mu_2, \ldots, \mu_{2s})$. It will be denoted below by $X^{(2s+1)}_{\min}$.

The other special extremal distribution in $\mathcal{B}_{2s+1}([a, b]; \mu_1, \mu_2, \ldots, \mu_{2s})$ is the one which puts positive masses $p_1, p_2, \ldots, p_{s+1}$ on the atoms $x_1, x_2, \ldots, x_s, b$, with $a < x_1 < x_2 < \cdots < x_s < b$. Such a distribution is called 'final' in Denuit, De Vylder & Lefèvre (1999) because it puts a positive mass on the right endpoint b. Proceeding as above, it can be shown that $x_1, x_2, \ldots, x_s$ are the s distinct roots of the equation

$$\left\| \begin{matrix} 1 & x & x^2 & \cdots & x^s \\ \mu_{1;b} & \mu_{2;b} & \mu_{3;b} & \cdots & \mu_{s+1;b} \\ \mu_{2;b} & \mu_{3;b} & \mu_{4;b} & \cdots & \mu_{s+2;b} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mu_{s;b} & \mu_{s+1;b} & \mu_{s+2;b} & \cdots & \mu_{2s;b} \end{matrix} \right\| = 0,$$

and the probabilities $p_1, p_2, \ldots, p_{s+1}$ are the solutions of the Vandermonde system

$$p_1 x_1^j + p_2 x_2^j + \cdots + p_{s+1} b^j = \mu_j, \qquad j = 0, 1, \ldots, s.$$

From (2.1) it is seen that the random variable that is distributed according to this 'final' extremal distribution is the largest, in the sense of the order $\preceq^{[a,b]}_{(2s+1)\text{-cx}}$, of all random variables with cdf in $\mathcal{B}_{2s+1}([a, b]; \mu_1, \mu_2, \ldots, \mu_{2s})$. It will be denoted below by $X^{(2s+1)}_{\max}$.

These 'initial' and 'final' $(s + 1)$-atomic distributions are precisely those achieving the bounds $\#(\text{supp}(F)) = s + 1$ and $\#(\text{supp}(F) - \{a, b\}) = s$ in Proposition 3.1(ii).

When r is even, that is, $r = 2s + 2$, the two extremal distributions in $\mathcal{B}_{2s+2}([a, b]; \mu_1, \mu_2, \ldots, \mu_{2s+1})$ that we want to point out are the following. The first is $(s + 1)$-atomic, and it puts positive masses $p_1, p_2, \ldots, p_{s+1}$ on the atoms $x_1, x_2, \ldots, x_{s+1}$, with $a < x_1 < x_2 < \cdots < x_{s+1} < b$. Such a distribution is called 'narrow' in Denuit, De Vylder & Lefèvre (1999) because it does not put any mass on a and b. The $2s+2$ quantities $x_1, x_2, \ldots, x_{s+1}, p_1, p_2, \ldots, p_{s+1}$ are uniquely determined by the $2s+1$ moment constraints, and the probability constraint $\sum_{i=1}^{s+1} p_i = 1$. Explicitly, again we first note that these variables satisfy

$$p_1 x_1^j + p_2 x_2^j + \cdots + p_{s+1} x_{s+1}^j = \mu_j, \qquad j = 0, 1, \ldots, s. \tag{4.3}$$

Next, using (4.3), we consider the matrix factorization

$$\begin{pmatrix} 1 & x & x^2 & \cdots & x^{s+1} \\ 1 & \mu_1 & \mu_2 & \cdots & \mu_{s+1} \\ \mu_1 & \mu_2 & \mu_3 & \cdots & \mu_{s+2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mu_s & \mu_{s+1} & \mu_{s+2} & \cdots & \mu_{2s+1} \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 0 & \cdots & 0 & 1 \\ p_1 & p_2 & \cdots & p_{s+1} & 0 \\ p_1 x_1 & p_2 x_2 & \cdots & p_{s+1} x_{s+1} & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ p_1 x_1^s & p_2 x_2^s & \cdots & p_{s+1} x_{s+1}^s & 0 \end{pmatrix} \begin{pmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{s+1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{s+1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{s+1} & x_{s+1}^2 & \cdots & x_{s+1}^{s+1} \\ 1 & x & x^2 & \cdots & x^{s+1} \end{pmatrix}$$

from which it is seen that $x_1, x_2, \ldots, x_{s+1}$ must be the $s+1$ distinct roots of the equation

$$\begin{Vmatrix} 1 & x & x^2 & \cdots & x^{s+1} \\ 1 & \mu_1 & \mu_2 & \cdots & \mu_{s+1} \\ \mu_1 & \mu_2 & \mu_3 & \cdots & \mu_{s+2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mu_s & \mu_{s+1} & \mu_{s+2} & \cdots & \mu_{2s+1} \end{Vmatrix} = 0. \qquad (4.4)$$

Equation (4.4) determines $x_2, \ldots, x_{s+1}$, and, plugging these into (4.3), also $p_1, p_2, \ldots, p_{s+1}$ for the 'narrow' extremal distribution are determined.

The random variable that is distributed according to this 'narrow' extremal distribution is the smallest, in the sense of the order $\preceq_{(2s+2)\text{-cx}}^{[a,b]}$, of all random variables with cdf in $\mathcal{B}_{2s+2}([a,b]; \mu_1, \mu_2, \ldots, \mu_{2s+1})$. It will be denoted below by $X_{\min}^{(2s+2)}$.

The other special extremal distribution in $\mathcal{B}_{2s+2}([a,b]; \mu_1, \mu_2, \ldots, \mu_{2s+1})$ is $(s+2)$-atomic which puts positive masses $p_1, p_2, \ldots, p_{s+2}$ on the atoms $a, x_2, \ldots, x_{s+1}, b$, with $a < x_2 < \cdots < x_{s+1} < b$. Such a distribution is called 'broad' in Denuit, De Vylder & Lefèvre (1999) because it puts positive masses on both a and b. Proceeding as above, it can be shown that $x_2, \ldots, x_{s+1}$ are the s distinct roots of the equation

$$\begin{Vmatrix} 1 & x & x^2 & \cdots & x^s \\ \mu_{2;a,b} & \mu_{3;a,b} & \mu_{4;a,b} & \cdots & \mu_{s+2;a,b} \\ \mu_{3;a,b} & \mu_{4;a,b} & \mu_{5;a,b} & \cdots & \mu_{s+3;a,b} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mu_{s+1;a,b} & \mu_{s+2;a,b} & \mu_{s+3;a,b} & \cdots & \mu_{2s+1;a,b} \end{Vmatrix} = 0,$$

with $\mu_{j;a,b} = \mu_j - (a+b)\mu_{j-1} + ab\mu_{j-2}$, and the probabilities $p_1, p_2, \ldots, p_{s+2}$ are the solutions of the Vandermonde system

$$p_1 a^j + p_2 x_2^j + \cdots + p_{s+1} x_{s+1}^j + p_{s+2} b^j = \mu_j, \qquad j = 0, 1, \ldots, s+1.$$

The random variable that is distributed according to this 'broad' extremal distribution is the largest, in the sense of the order $\preceq_{(2s+2)\text{-cx}}^{[a,b]}$, of all random variables with cdf in $\mathcal{B}_{2s+2}([a,b]; \mu_1, \mu_2, \ldots, \mu_{2s+1})$. It will be denoted below by $X_{\max}^{(2s+2)}$.

These 'narrow' $(s+1)$-atomic distribution and 'broad' $(s+2)$-atomic distribution achieve the bounds $\#(\text{supp}(F)-\{a\}) = s+1$ and $\#(\text{supp}(F)-\{b\}) = s+1$ in Proposition 3.1(i).

Improvements of the stochastic bounds $X_{\min}^{(s)}$ and $X_{\max}^{(s)}$ can be obtained using Khinchine's theorem when unimodality is known to hold. To be specific, let $\mathcal{B}_s^*([a,b];\mu_1,\mu_2,\ldots,\mu_{s-1};m)$ denote the restriction of $\mathcal{B}_s([a,b];\mu_1,\mu_2,\ldots,\mu_{s-1})$ to its elements which are unimodal about $m \in [a,b]$. We know from Denuit, De Vylder & Lefèvre (1999, Proposition 5.3) that there exists a random variable $X_{\min}^{(s)*}$, with a unimodal density with mode m and the same first $s-1$ moments as X, such that

$$X_{\min}^{(s)} \preceq_{s\text{-cx}}^{[a,b]} X_{\min}^{(s)*} \preceq_{s\text{-cx}}^{[a,b]} X \quad \text{for all } X \in \mathcal{B}_s^*([a,b];\mu_1,\mu_2,\ldots,\mu_{s-1};m).$$

Similarly, there exists a random variable $X_{\max}^{(s)*}$, with a unimodal density with mode m and the same first $s-1$ moments as X, such that

$$X \preceq_{s\text{-cx}}^{[a,b]} X_{\max}^{(s)*} \preceq_{s\text{-cx}}^{[a,b]} X_{\max}^{(s)} \quad \text{for all } X \in \mathcal{B}_s^*([a,b];\mu_1,\mu_2,\ldots,\mu_{s-1};m).$$

In order to describe explicitly the cdf's of $X_{\min}^{(s)*}$ and $X_{\max}^{(s)*}$, let us denote by $\text{Unif}[a,b]$ the uniform cdf over the interval $[\min(a,b),\max(a,b)]$ (when $a=b$ then $\text{Unif}[a,b]$ denotes the degenerate distribution at a). If the extremum (that is, $X_{\min}^{(s)}$ or $X_{\max}^{(s)}$) in $\mathcal{B}_s([a,b];\mu_1,\mu_2,\ldots,\mu_{s-1})$ has the support points $x_i(\mu_1,\mu_2,\ldots,\mu_{s-1})$, $i=1,2,\ldots,l$, with the respective probabilities $p_i(\mu_1,\mu_2,\ldots,\mu_{s-1})$, $i=1,2,\ldots,l$, then the cdf F of the improved extremum (that is, of $X_{\min}^{(s)*}$ or $X_{\max}^{(s)*}$) is a mixture of l uniform cdf's given by

$$F = \sum_{i=1}^{l} p_i(\tilde{\mu}_1,\tilde{\mu}_2,\ldots,\tilde{\mu}_{s-1})\,\text{Unif}[m, x_i(\tilde{\mu}_1,\tilde{\mu}_2,\ldots,\tilde{\mu}_{s-1})],$$

where

$$\tilde{\mu}_j = (j+1)\mu_j - jm\mu_{j-1}, \qquad j=1,2,\ldots. \tag{4.5}$$

The s-convex stochastic bounds allow for many applications in different disciplines. They often appear in the literature under a variety of names (especially $X_{\min}^{(2s)}$, following Mammana's (1954) results). Their study in the context of s-convex stochastic orders often throws a new light on existing methodologies. For instance, in Denuit, Lefèvre & Shaked (1999) it is shown that the s-convex stochastic extrema play a central role in the theory of canonical moments. For an $X \in \mathcal{B}_s([a,b];\mu_1,\mu_2,\ldots,\mu_{s-1})$, the sth canonical moment of X, denoted by c_s, can be cast as

$$c_s = \frac{EX^s - E\big[X_{\min}^{(s)}\big]^s}{E\big[X_{\max}^{(s)}\big]^s - E\big[X_{\min}^{(s)}\big]^s}.$$

Thus, c_s is simply the position of the sth moment of X relative to its possible range (since the function $x \mapsto x^s$ is s-convex, so that the extrema $X_{\min}^{(s)}$ and $X_{\max}^{(s)}$ achieve the bounds on EX^s). Canonical moments have been used *inter alia* by Lau (1996) (using results of Studden (1982); see also Dette and Studden (1997)) to approximate an unknown mixing distribution in a discrete mixture model. His analysis is based on the identification of the minimal integer ζ such that $c_\zeta = 0$ or 1. From Proposition 3.8 in Denuit, Lefèvre & Shaked (1998), we then immediately deduce that

$$c_\zeta = 0 \Longrightarrow X =_{\text{st}} X_{\text{min}}^{(\zeta)} \quad \text{and} \quad c_\zeta = 1 \Longrightarrow X =_{\text{st}} X_{\text{max}}^{(\zeta)}.$$

Analogous methods have been developed by Lindsay (1989b) and Lindsay & Roeder (1997). These authors, however, only used $X_{\text{min}}^{(2s)}$ (because it is the only extremum which puts no probability mass on the extremities a and b, and it remains thus valid for random variables with infinite support). Some other extrema could be useful in their framework, and improve some technical points in their method.

5 Approximations by Extremal Distributions

Reformulated in terms of moment matrices and s-convex stochastic extrema, the classical Tchebycheff inequality states that, given any $X \in \mathcal{B}_2([a,b];\mu_1)$, we have

$$P\big[|X - X_{\text{min}}^{(2)}| > h\big] \leq \frac{\|\boldsymbol{P}_1\|}{h^2} \qquad \text{for any } h > 0.$$

In this section we aim to evaluate the closeness between elements of $\mathcal{B}_s([a,b];\mu_1,\mu_2,\ldots,\mu_{s-1})$ and the special extremal distributions described in Section 4 by generalizing the Tchebycheff inequality. We will use standard ideas from the theory of linear regression briefly recalled in Appendix B, where the moment matrices, defined in Section 3, play naturally an important role. Part (i) of the following result is Theorem 2C of Lindsay (1989a).

Theorem 5.1. (i) *Let $X \in \mathcal{B}_{2s+2}([a,b];\mu_1,\mu_2,\ldots,\mu_{2s+1})$, and let $x_1, x_2, \ldots, x_{s+1}$ be the $s+1$ support points of the 'narrow' distribution in $\mathcal{B}_{2s+2}([a,b];\mu_1,\mu_2,\ldots,\mu_{2s+1})$. Then, if μ_{2s+2} denotes the $(2s+2)$nd moment of X, we have*

$$P\Big(\inf_{i\in\{1,2,\ldots,s+1\}} |X - x_i| > h\Big) \leq \frac{\|\boldsymbol{P}_{s+1}\|}{\|\boldsymbol{P}_s\| h^{2s+2}} \qquad \textit{for any } h > 0.$$

(ii) *Let $X \in \mathcal{B}_{2s+2}([a,b];\mu_1,\mu_2,\ldots,\mu_{2s+1})$, and let $x_1 = a, x_2, \ldots, x_{s+1}, x_{s+2} = b$ be the $s+2$ support points of the 'broad' distribution in $\mathcal{B}_{2s+2}([a,b];\mu_1,\mu_2,\ldots,\mu_{2s+1})$. Then, if μ_{2s+2} denotes the $(2s+2)$nd moment of X, we have*

$$P\Big(\inf_{i\in\{1,2,\ldots,s+2\}} |X - x_i| > h\Big) \leq \frac{\| -ab\boldsymbol{P}_s + (a+b)\boldsymbol{Q}_s - \boldsymbol{R}_s\|}{\| -ab\boldsymbol{P}_{s-1} + (a+b)\boldsymbol{Q}_{s-1} - \boldsymbol{R}_{s-1}\| h^{2s+2}} \qquad \textit{for any } h > 0.$$

(iii) *Let $X \in \mathcal{B}_{2s+1}([a,b];\mu_1,\mu_2,\ldots,\mu_{2s})$, and let $x_1 = a, x_2, \ldots, x_{s+1}$ be the $s+1$ support points of the 'initial' distribution in $\mathcal{B}_{2s+1}([a,b];\mu_1,\mu_2,\ldots,\mu_{2s})$. Then, if μ_{2s+1} denotes the $(2s+1)$st moment of X, we have*

$$P\Big(\inf_{i\in\{1,2,\ldots,s+1\}} |X - x_i| > h\Big) \leq \frac{\|\boldsymbol{Q}_s - a\boldsymbol{P}_s\|}{\|\boldsymbol{Q}_{s-1} - a\boldsymbol{P}_{s-1}\| h^{2s+1}} \qquad \textit{for any } h > 0. \tag{5.1}$$

(iv) *Let* $X \in \mathcal{B}_{2s+1}([a,b];\mu_1,\mu_2,\ldots,\mu_{2s})$*, and let* $x_1, x_2, \ldots, x_{s+1} = b$ *be the* $s+1$ *support points of the 'final' distribution in* $\mathcal{B}_{2s+1}([a,b];\mu_1,\mu_2,\ldots,\mu_{2s})$*. Then, if* μ_{2s+1} *denotes the* $(2s+1)$*st moment of* X*, we have*

$$P\Big(\inf_{i\in\{1,2,\ldots,s+1\}} |X-x_i| > h\Big) \le \frac{\|b\boldsymbol{P}_s - \boldsymbol{Q}_s\|}{\|b\boldsymbol{P}_{s-1} - \boldsymbol{Q}_{s-1}\| h^{2s+1}} \qquad \text{for any } h>0. \tag{5.2}$$

Proof. (i) Since the minimization problem in Proposition B.1(i) (with $n=s$) is completely determined by the first $2s+1$ moments of X, we can solve for this infimum by using the 'narrow' distribution in $\mathcal{B}_{2s+2}([a,b];\mu_1,\mu_2,\ldots,\mu_{2s+1})$ for X. It is then possible to make the expectation zero if, and only if, the polynomial $t^{s+1}-\sum_{i=0}^{s}\beta_i^* t^i$ is the polynomial $\prod_{i=1}^{s+1}(t-x_i)$. Therefore,

$$\frac{\|\boldsymbol{P}_{s+1}\|}{\|\boldsymbol{P}_s\|} = E\Big[\prod_{i=1}^{s+1}(X-x_i)^2\Big]. \tag{5.3}$$

Now,

$$\begin{aligned} P(\inf|X-x_i|>h) &= P\big(\inf(X-x_i)^2 > h^2\big) \\ &\le P\Big(\prod_{i=1}^{s+1}(X-x_i)^2 > h^{2s+2}\Big) \\ &\le \frac{E\left(\prod_{i=1}^{s+1}(X-x_i)^2\right)}{h^{2s+2}}, \end{aligned} \tag{5.4}$$

where the last inequality follows from the Markov Inequality. Combination of (5.3) and (5.4) gives the stated result.
(ii) We proceed as in (i). The minimization problem in Proposition B.1(ii) (with $n=s-1$) is completely determined by the first $2s+1$ moments of X. Therefore, we may assume that X follows the 'broad' distribution in $\mathcal{B}_{2s+2}([a,b];\mu_1,\mu_2,\ldots,\mu_{2s+1})$. It is then possible to make the expectation zero if, and only if, the polynomial $(t-a)(b-t)\big(t^s-\sum_{i=0}^{s-1}\beta_i t^i\big)^2$ is the polynomial $(t-a)(b-t)\prod_{i=2}^{s+1}(t-x_i)^2$. Hence

$$\frac{\|-ab\boldsymbol{P}_s+(a+b)\boldsymbol{Q}_s-\boldsymbol{R}_s\|}{\|-ab\boldsymbol{P}_{s-1}+(a+b)\boldsymbol{Q}_{s-1}-\boldsymbol{R}_{s-1}\|} = E\Big[(X-a)(b-X)\prod_{i=2}^{s+1}(X-x_i)^2\Big]. \tag{5.5}$$

Now,

$$\begin{aligned} P(\inf|X-x_i|>h) &\le P\Big((X-a)(b-X)\prod_{i=2}^{s+1}(X-x_i)^2 > h^{2s+2}\Big) \\ &\le \frac{E\big((X-a)(b-X)\prod_{i=2}^{s+1}(X-x_i)^2\big)}{h^{2s+2}}. \end{aligned} \tag{5.6}$$

Combination of (5.5) and (5.6) gives the stated result. The proof of (iii)-(iv) is similar and is omitted. □

The upper bounds in Theorem 5.1 are useful to evaluate the goodness of fit of the approximation of $X \in \mathcal{B}_s([a,b];\mu_1,\mu_2,\ldots,\mu_{s-1})$ by either $X_{\min}^{(s)}$ or $X_{\max}^{(s)}$: if the right-hand side of the inequalities in (i)–(iv) is small then the approximation can be regarded as satisfactory. Some numerical computations of these bounds are given later in Subsection 6.1.2; it is seen there that these bounds are sometimes not very tight, but this is not surprising because, like in the classical Tchebycheff inequality, very little is assumed to be known about the estimated cdf. These bounds, however, are certainly of theoretical interest.

It is well-known that given a sequence of random variables $\{X_n,\ n=1,2,\ldots\}$ such that $EX_n \to \mu_1$ and $\operatorname{Var}[X_n] \to 0$ as $n \to \infty$, then the sequence converges in probability to the random variable which is equal to μ_1 almost surely. In our context, this means that if $\|\boldsymbol{P}_1^{(n)}\| \to 0$ as $n \to \infty$, then $\lim_{n\to\infty} X_n = X_{\min}^{(2)}$. Lindsay (1989b, Lemma 3C) indicates how to extend this result to the case where $\|\boldsymbol{P}_k^{(n)}\| \to 0$ as $n \to \infty$. In the next result we extend these ideas to the other extrema; part (i1) of the next result is essentially the same as the above mentioned result of Lindsay (1989b). Recall from Proposition 3.1 the conditions for a moments sequence to be in the topological interior $\mathcal{D}_s^\circ([a,b])$ of $\mathcal{D}_s([a,b])$.

Theorem 5.2. *Let $\{X_n,\ n=1,2,\ldots\}$ be a sequence of random variables with support in $[a,b]$ and moments $\mu_k^{(n)} = EX_n^k$, $k \geq 1$, $n \geq 1$, and denote by $\boldsymbol{P}_k^{(n)}$, $\boldsymbol{Q}_k^{(n)}$ and $\boldsymbol{R}_k^{(n)}$ the moment matrices corresponding to X_n, $n \geq 1$. We then have the following results:*

(i) *Assume $\#\{\operatorname{supp}(X_n)-\{a\}\} \geq s+1$ and $\#\{\operatorname{supp}(X_n)-\{b\}\} \geq s+1$ for all n, and $\mu_k^{(n)} \to \mu_k$ for $k=1,2,\ldots,2s+1$, with $(\mu_1,\mu_2,\ldots,\mu_{2s+1}) \in \mathcal{D}_{2s+2}^0([a,b])$:*

   (i1) *If $\|\boldsymbol{P}_{s+1}^{(n)}\| \to 0$ as $n \to \infty$, then X_n converges weakly to $X_{\min}^{(2s+2)}$.*

   (i2) *If $\|-ab\boldsymbol{P}_s^{(n)} + (a+b)\boldsymbol{Q}_s^{(n)} - \boldsymbol{R}_s^{(n)}\| \to 0$ as $n \to \infty$, then X_n converges weakly to $X_{\max}^{(2s+2)}$.*

(ii) *Assume $\#\{\operatorname{supp}(X_n)\} \geq s+1$ and $\#\{\operatorname{supp}(X_n)-\{a,b\}\} \geq s$ for all n, and $\mu_k^{(n)} \to \mu_k$ for $k=1,2,\ldots,2s$, with $(\mu_1,\mu_2,\ldots,\mu_{2s}) \in \mathcal{D}_{2s+1}^0([a,b])$:*

   (ii1) *If $\|\boldsymbol{Q}_s^{(n)} - a\boldsymbol{P}_s^{(n)}\| \to 0$ as $n \to \infty$, then X_n converges weakly to $X_{\min}^{(2s+1)}$.*

   (ii2) *If $\|b\boldsymbol{P}_s^{(n)} - \boldsymbol{Q}_s^{(n)}\| \to 0$ as $n \to \infty$, then X_n converges weakly to $X_{\max}^{(2s+1)}$.*

The proof of Theorem 5.2 is similar to the proof of Lemma 3C of Lindsay (1989b) and is therefore omitted.

6 Some Applications

6.1 Insurance Theory

6.1.1 Theoretical background

Let $\varphi_X(\cdot)$ be the Laplace transform of some given non-negative random variable X; that is, $\varphi_X(t) = E\exp(-tX)$, $t \geq 0$. Many classical problems in risk theory involve the resolution of the equation

$$\varphi_X(t) = P_k(t), \tag{6.1}$$

in the unknown $t \geq 0$, where $P_k(\cdot)$ is a non-decreasing polynomial function on $[0,\infty)$ of degree k (usually, $k \leq 2$). When all that is known about X is that it belongs to $\mathcal{B}_s([0,b];\mu_1,\mu_2,\ldots,\mu_{s-1})$, then (6.1) cannot be solved explicitly. The aim of this Subsection is to show that the s-convex extrema described in Section 4 allow accurate approximations for the solution of (6.1). In this respect, we expand on previous works by Eckberg (1977), Whitt (1983) and Lefèvre, Hallin & Narayan (1986). There is another situation where this methodology is useful albeit the cdf F_X of X is fully known. This is the case when F_X is analytically complicated and then solving (6.1) may be practically impossible. Therefore, approximating F_X by the simpler cdf's of the s-convex extrema yields bounds on the unknown solutions (which are often accurate enough for practical purposes).

Since the function $x \mapsto (-1)^s \exp(-tx)$ is s-convex, we can construct two functions $\varphi^{(s)}_{\min}(\cdot)$ and $\varphi^{(s)}_{\max}(\cdot)$ (see details below) such that the inequalities

$$\varphi^{(s)}_{\min}(t) \leq \varphi_X(t) \leq \varphi^{(s)}_{\max}(t), \qquad t \geq 0, \tag{6.2}$$

hold for any X in $\mathcal{B}_s([0,b];\mu_1,\mu_2,\ldots,\mu_{s-1})$. Explicitly, when s is even, then $\varphi^{(s)}_{\min}(t) = \varphi_{X^{(s)}_{\min}}(t)$ and $\varphi^{(s)}_{\max}(t) = \varphi_{X^{(s)}_{\max}}(t)$, while when s is odd, then $\varphi^{(s)}_{\min}(t) = \varphi_{X^{(s)}_{\max}}(t)$ and $\varphi^{(s)}_{\max}(t) = \varphi_{X^{(s)}_{\min}}(t)$. Explicit expressions of these bounds are easily deduced from Tables 6.1 and 6.2 for s up to 5 (thus using the mean, variance, skewness and kurtosis of the underlying unknown distribution); in Table 6.2 we use the notation

$$\begin{aligned}\varrho = &\big((\mu_1-b)(\mu_4-b\mu_3)-(\mu_2-b\mu_1)(\mu_3-b\mu_2)\big)^2\\ &-4\big((\mu_1-b)(\mu_3-b\mu_2)-(\mu_2-b\mu_1)^2\big)\big((\mu_2-b\mu_1)(\mu_4-b\mu_3)-(\mu_3-b\mu_2)^2\big).\end{aligned}$$

The bounds $\varphi^{(s)}_{\min}$ and $\varphi^{(s)}_{\max}$ in (6.2) can be improved to $\varphi^{(s)*}_{\min}$ and $\varphi^{(s)*}_{\max}$ (see the definitions below) when X is further known to have a unimodal density with a mode m. When s is even we take $\varphi^{(s)*}_{\min}(t) = \varphi_{X^{(s)*}_{\min}}(t)$ and $\varphi^{(s)*}_{\max}(t) = \varphi_{X^{(s)*}_{\max}}(t)$ (where $X^{(s)*}_{\min}$ and $X^{(s)*}_{\max}$ are described in Section 4), while when s is odd we take $\varphi^{(s)*}_{\min}(t) = \varphi_{X^{(s)*}_{\max}}(t)$ and $\varphi^{(s)*}_{\max}(t) = \varphi_{X^{(s)*}_{\min}}(t)$. Explicit expressions of these bounds are easily deduced from Tables 6.3 and 6.4; in Table 6.3 the notations $\tilde{r}_\pm$, $\tilde{t}_\pm$ and $\tilde{q}_\pm$ are those from Table 6.1, with the $\tilde{\mu}_j$'s (defined in (4.5)) substituted for the μ_j's; in Table 6.4 the notations $\tilde{p}_1$, $\tilde{p}_2$, $\tilde{p}_\pm$ and $\tilde{z}_\pm$ are those from Table 6.2, again with the $\tilde{\mu}_j$'s substituted for the μ_j's.

Solving the equation $\varphi^{(s)}_{\min}(t) = P_k(t)$ yields the root $t^{(s)}_1$, say, and solving $\varphi^{(s)}_{\max}(t) = P_k(t)$ yields the root $t^{(s)}_2$, say. The solution $\tilde{t}$, say, of (6.1) then satisfies $t^{(s)}_2 \leq \tilde{t} \leq t^{(s)}_1$. In case unimodal distributions with known modes are concerned, it is preferable to substitute $\varphi^{(s)*}_{\min}$ and $\varphi^{(s)*}_{\max}$ for $\varphi^{(s)}_{\min}$ and $\varphi^{(s)}_{\max}$, respectively.

The same problem with $\phi_X(t) = E\exp(tX)$, the moment generating function of X, can be handled similarly. Since $x \mapsto \exp(tx)$ is s-convex for all s, we have that $\phi^{(s)}_{\min}(t) \leq \phi_X(t) \leq \phi^{(s)}_{\max}(t)$ with $\phi^{(s)}_{\min}(t) = \phi_{X^{(s)}_{\min}}(t)$ and $\phi^{(s)}_{\max}(t) = \phi_{X^{(s)}_{\max}}(t)$. As above, these provide bounds on the root of the equation $\phi_X(t) = P_k(t)$, where P_k is a monotone polynomial function.

s	Support points	Probability masses
1	0	1
2	μ_1	1
3	0	$\frac{\mu_2-\mu_1^2}{\mu_2}$
	$\frac{\mu_2}{\mu_1}$	$\frac{\mu_1^2}{\mu_2}$
4	$r_+ = \frac{\mu_3-\mu_1\mu_2+\sqrt{(\mu_3-\mu_1\mu_2)^2-4(\mu_2-\mu_1^2)(\mu_1\mu_3-\mu_2^2)}}{2(\mu_2-\mu_1^2)}$	$\frac{\mu_1-r_-}{r_+-r_-}$
	$r_- = \frac{\mu_3-\mu_1\mu_2-\sqrt{(\mu_3-\mu_1\mu_2)^2-4(\mu_2-\mu_1^2)(\mu_1\mu_3-\mu_2^2)}}{2(\mu_2-\mu_1^2)}$	$1-\frac{\mu_1-r_-}{r_+-r_-}$
5	0	$1-q_+-q_-$
	$t_+ = \frac{\mu_1\mu_4-\mu_2\mu_3+\sqrt{(\mu_1\mu_4-\mu_2\mu_3)^2-4(\mu_1\mu_3-\mu_2^2)(\mu_2\mu_4-\mu_3^2)}}{2(\mu_1\mu_3-\mu_2^2)}$	$q_+ = \frac{\mu_2-t_-\mu_1}{t_+(t_+-t_-)}$
	$t_- = \frac{\mu_1\mu_4-\mu_2\mu_3-\sqrt{(\mu_1\mu_4-\mu_2\mu_3)^2-4(\mu_1\mu_3-\mu_2^2)(\mu_2\mu_4-\mu_3^2)}}{2(\mu_1\mu_3-\mu_2^2)}$	$q_- = \frac{\mu_2-t_+\mu_1}{t_-(t_--t_+)}$

Table 6.1: *Probability distribution of* $X_{\min}^{(s)} \in \mathcal{B}_s([0,b];\mu_1,\mu_2,\ldots,\mu_{s-1})$.

6.1.2 Lundberg's coefficient

In the classical risk model, individual claim amounts $X_1, X_2, \ldots$ are assumed to be independent and identically distributed with a common cdf F having finite mean and second moment, such that $F(0)=0$. The number of claims in the time interval $[0,t]$ is assumed to be independent of the individual claim amounts and to form a Poisson process $\{N_t,\ t\geq 0\}$ with constant rate λ. Let the premium rate $c>0$ be such that the inequality $c>\lambda EX$ holds. Further, let $\psi(z)$ be the ultimate ruin probability with an initial capital z; that is, the probability that the process $Z_t = z+ct-X_1-X_2-\cdots-X_{N_t}$, $t\geq 0$, describing the wealth of the insurance company, ever falls below zero. If the moment generating function of X exists, Lundberg's inequality provides an exponential upper bound on ψ, namely $\psi(z)\leq\exp(-\kappa z)$, where κ is the Lundberg's adjustment coefficient satisfying the integral equation $E\exp(\kappa X) = 1+c\kappa/\lambda$.

As an example, consider claim amounts with mean $\mu_1=1$, second moment μ_2 and upper bound b. Solving for ω the equations (see line $s=3$ in Table 6.1)

$$\phi_{X_{\min}^{(3)}}(\omega) = \frac{\mu_2-1}{\mu_2}+\frac{1}{\mu_2}\exp(\omega\mu_2) = 1+\frac{c\omega}{\lambda}$$

and (see line $s=3$ in Table 6.2)

$$\phi_{X_{\max}^{(3)}}(\omega) = \frac{(b-1)^2}{(b-1)^2+\mu_2-1}\exp\left(\omega\frac{b-\mu_2}{b-1}\right)+\frac{\mu_2-1}{(b-1)^2+\mu_2-1}\exp(\omega b) = 1+\frac{c\omega}{\lambda}$$

yields $\kappa_{\max}$ and $\kappa_{\min}$, respectively, such that $\kappa_{\min}\leq\kappa\leq\kappa_{\max}$. As an illustration, let $b=5$, $c=12$, $\lambda=10$, and consider $\kappa_{\min}$ and $\kappa_{\max}$ as functions of μ_2 (ranging from $\mu_1^2=1$ to $b\mu_1=5$; these are the conditions defining the admissible values for μ_2). We then get the

s	Support points	Probability masses
1	b	1
2	0 b	$\frac{b-\mu_1}{b}$ $\frac{\mu_1}{b}$
3	$\frac{b\mu_1-\mu_2}{b-\mu_1}$ b	$\frac{(b-\mu_1)^2}{(b-\mu_1)^2+\mu_2-\mu_1^2}$ $\frac{\mu_2-\mu_1^2}{(b-\mu_1)^2+\mu_2-\mu_1^2}$
4	0 $\frac{\mu_3-b\mu_2}{\mu_2-b\mu_1}$ b	$1-p_1-p_2$ $p_1=\frac{(\mu_2-b\mu_1)^3}{(\mu_3-b\mu_2)(\mu_3-2b\mu_2+b^2\mu_1)}$ $p_2=\frac{\mu_1\mu_3-\mu_2^2}{b(\mu_3-2b\mu_2+b^2\mu_1)}$
5	$z_+=\frac{(\mu_1-b)(\mu_4-b\mu_3)-(\mu_2-b\mu_1)(\mu_3-b\mu_2)+\sqrt{\varrho}}{2((\mu_1-b)(\mu_3-b\mu_2)-(\mu_2-b\mu_1)^2)}$ $z_-=\frac{(\mu_1-b)(\mu_4-b\mu_3)-(\mu_2-b\mu_1)(\mu_3-b\mu_2)-\sqrt{\varrho}}{2((\mu_1-b)(\mu_3-b\mu_2)-(\mu_2-b\mu_1)^2)}$ b	$p_+=\frac{\mu_2-(b+z_-)\mu_1+bz_-}{(z_+-z_-)(z_+-b)}$ $p_-=\frac{\mu_2-(b+z_+)\mu_1+bz_+}{(z_--z_+)(z_--b)}$ $1-p_+-p_-$

Table 6.2: *Probability distribution of* $X^{(s)}_{\max} \in \mathcal{B}_s([0,b];\mu_1,\mu_2,\ldots,\mu_{s-1})$.

graph depicted in Figure 6.1; some numerical values are given in Table 6.5. It is seen that the bounds are quite accurate, and are particularly so when μ_2 is large.

For the case described here (that is, when $a=0$, $b=5$, $2s=2$, $\mu_1=1$ and $\mu_2=1.5$) some of the upper bounds of Theorem 5.1(iii) and (iv) (for various values of h and various allowable values of μ_3) are given in Table 6.6. In this table we replaced by 1 any value of the right hand side of (5.1) or of (5.2) which was greater than 1. As was mentioned after Theorem 5.1, we see that the bounds provided by this theorem are often not very tight, but this is not surprising because, like in the classical Tchebycheff inequality, very little (namely, just the first two moments) is assumed to be known about the cdf of X. These bounds, however, are of theoretical interest.

If the X_i's are further known to be unimodal about m, we get better bounds on κ by solving for ω the equations (see line $s=3$ in Table 6.3)

$$\phi_{X^{(3)*}_{\min}}(\omega)=\frac{\tilde\mu_2-\tilde\mu_1^2}{\tilde\mu_2 m\omega}\{\exp(\omega m)-1\}+\frac{\tilde\mu_1^2}{\tilde\mu_2\big(m-\frac{\tilde\mu_2}{\tilde\mu_1}\big)}\{\exp(\omega m)-\exp(\omega\tilde\mu_2/\tilde\mu_1)\}=1+\frac{c\omega}{\lambda}$$

and (see line $s=3$ in Table 6.4)

$$\begin{aligned}\phi_{X^{(3)*}_{\max}}(\omega)=&\frac{(b-\tilde\mu_1)^2}{(b-\tilde\mu_1)^2+\tilde\mu_2-\tilde\mu_1^2}\cdot\frac{b-\tilde\mu_1}{mb-m\tilde\mu_1-b\tilde\mu_1+\tilde\mu_2}\Big\{\exp(\omega m)-\exp\big(\omega\frac{b\tilde\mu_1-\tilde\mu_2}{b-\tilde\mu_1}\big)\Big\}\\&+\frac{\tilde\mu_2-\tilde\mu_1^2}{(b-\tilde\mu_1)^2+\tilde\mu_2-\tilde\mu_1^2}\cdot\frac{1}{b-m}\{\exp(\omega b)-\exp(\omega m)\}\\=&1+\frac{c\omega}{\lambda}.\end{aligned}$$

s	Distributions
1	$\mathrm{Unif}[0,m]$
2	$\mathrm{Unif}[m,\tilde\mu_1]$
3	$\frac{\tilde\mu_2-\tilde\mu_1^2}{\tilde\mu_2}\,\mathrm{Unif}[0,m]+\frac{\tilde\mu_1^2}{\tilde\mu_2}\,\mathrm{Unif}[m,\tilde\mu_2/\tilde\mu_1]$
4	$\frac{\tilde\mu_1-\tilde r_-}{\tilde r_+-\tilde r_-}\,\mathrm{Unif}[m,\tilde r_+]+\left(1-\frac{\tilde\mu_1-\tilde r_-}{\tilde r_+-\tilde r_-}\right)\mathrm{Unif}[m,\tilde r_-]$
5	$(1-\tilde q_+-\tilde q_-)\,\mathrm{Unif}[0,m]+\tilde q_+\,\mathrm{Unif}[m,\tilde t_+]+\tilde q_-\,\mathrm{Unif}[m,\tilde q_-]$

Table 6.3: *Probability distribution of* $X_{\min}^{(s)*}\in\mathcal{B}_s^*([0,b];\mu_1,\mu_2,\ldots,\mu_{s-1};m)$.

s	Distributions
1	$\mathrm{Unif}[m,b]$
2	$\frac{b-\tilde\mu_1}{b}\,\mathrm{Unif}[0,m]+\frac{\tilde\mu_1}{b}\,\mathrm{Unif}[m,b]$
3	$\frac{(b-\tilde\mu_1)^2}{(b-\tilde\mu_1)^2+\tilde\mu_2-\tilde\mu_1^2}\,\mathrm{Unif}\left[m,\frac{b\tilde\mu_1-\tilde\mu_2}{b-\tilde\mu_1}\right]+\frac{\tilde\mu_2-\tilde\mu_1^2}{(b-\tilde\mu_1)^2+\tilde\mu_2-\tilde\mu_1^2}\,\mathrm{Unif}[m,b]$
4	$(1-\tilde p_1-\tilde p_2)\,\mathrm{Unif}[0,m]+\tilde p_1\,\mathrm{Unif}\left[m,\frac{\tilde\mu_3-b\tilde\mu_2}{\tilde\mu_2-b\tilde\mu_1}\right]+\tilde p_2\,\mathrm{Unif}[m,b]$
5	$\tilde p_+\,\mathrm{Unif}[m,\tilde z_+]+\tilde p_-\,\mathrm{Unif}[m,\tilde z_-]+(1-\tilde p_+-\tilde p_-)\,\mathrm{Unif}[m,b]$

Table 6.4: *Probability distribution of* $X_{\max}^{(s)*}\in\mathcal{B}_s^*([0,b];\mu_1,\mu_2,\ldots,\mu_{s-1};m)$.

6.1.3 Determination of the interest rate relating to a life insurance policy

Assume that you are given an insurer's price list, and that you know that the amount of premium charged for a whole life insurance coverage is π. As in Belgium, assume further that the lifetable the insurer has to use is fixed by the regulatory authorities. Then, you would like to determine the actual interest rate δ used to set the premium of your contract; that is, the value of δ satisfying

$$\varphi_L(\delta)=E\exp(-\delta L)=\pi, \tag{6.3}$$

where L is the future random lifetime (and assuming, without real loss of generality, that the capital paid by the insurer is equal to one monetary unit). Solving (6.3) explicitly for δ is mostly impossible. However, the bounds described in Section 4, studied in Section 5 and illustrated in Subsection 6.1.1, often give an accurate approximation of δ if some moments of L are known.

6.2 Economics

In the framework of the expected utility theory (Von Neumann & Morgenstern (1947)), each decision-maker is assumed to possess a utility function that expresses his preferences. To be specific, a utility function u assigns a value $u(x)$ to a monetary amount of x; u describes

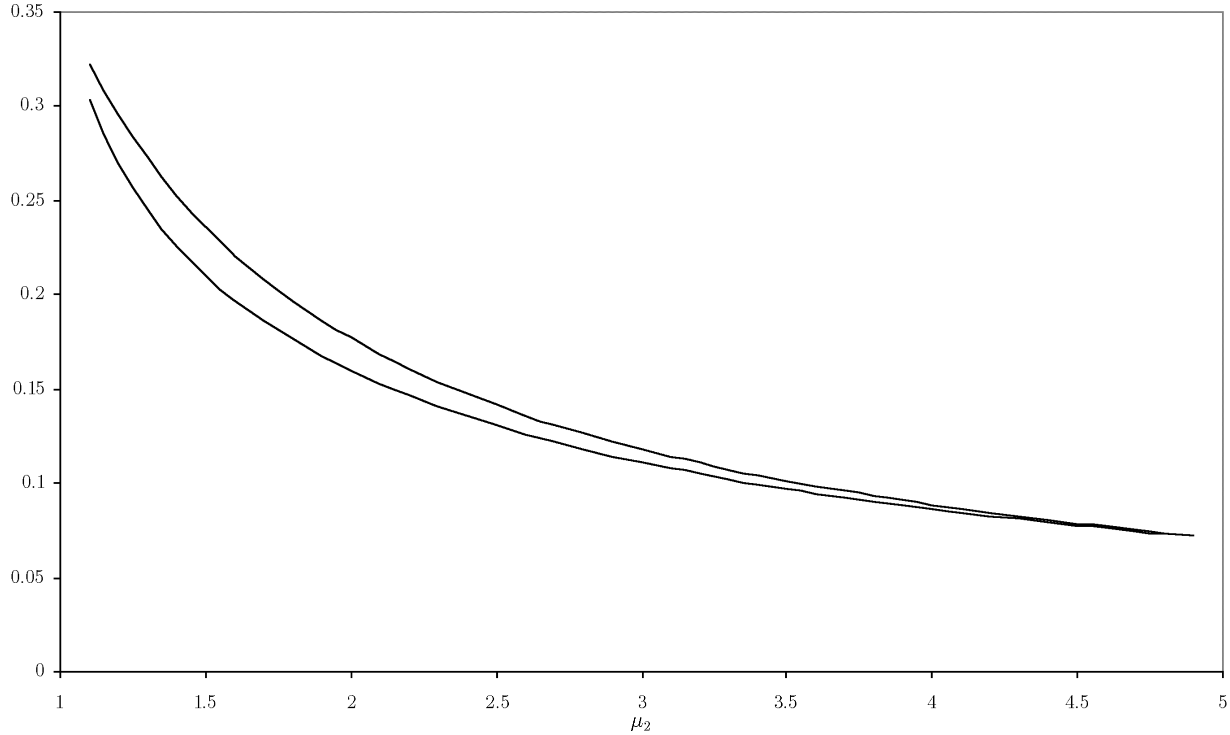

Figure 6.1: *Bounds on Lundberg's coefficient* κ *when* $\mu_1 = 1$, $b = 5$, $c = 12$ *and* $\lambda = 10$.

μ_2	1.5	2.0	2.5	3.0	3.5	4.0	4.5
$\kappa_{\min}$	.21041	.15965	.13057	.11114	.09705	.08630	.07777
$\kappa_{\max}$	.23613	.17710	.14168	.11807	.10120	.08855	.07871

Table 6.5: *Bounds on Lundberg's coefficient* κ *when* $\mu_1 = 1$, $b = 5$, $c = 12$ *and* $\lambda = 10$.

'how much the decision-maker appreciates a fortune of x.' The decision-maker is then said to base his preferences on the 'expected utility hypothesis' if he acts in order to maximize his expected utility. This means that for all uncertain future fortunes X and Y, the decision-maker with utility u will prefer Y over X if, and only if, $Eu(X) \leq Eu(Y)$, provided the expectations exist.

All reasonable utility functions share certain common properties. For instance, it seems natural that an individual always prefers more wealth to less wealth. Therefore, the utility function u is always assumed to be non-decreasing; if u is differentiable, the lure of gain is expressed by the condition $u^{(1)} \geq 0$. Another important concept is the notion of risk aversion: a decision-maker is said to be risk averse when his utility function is concave on its domain; if u is twice differentiable, this reduces to $u^{(2)} \leq 0$. One way to justify the concavity assumption is to note that for any concave function u we have, by Jensen inequality, that $E[u(X)] \leq u(E[X])$ for all random fortunes X. Therefore, a risk averse decision-maker always prefers a certain fortune to a random income with the same expected value. Further, it is natural to require that the risk aversion of the decision-maker decreases with his wealth; a necessary condition for this is that u has to be 3-convex (see Menezes, Geiss & Tressler

h	μ_3	(5.1)	(5.2)	h	μ_3	(5.1)	(5.2)
1	2.5	0.25	1	2	2.5	0.03125	0.24219
	3	0.75	1		3	0.09375	0.17969
	3.5	1	0.94		3.5	0.15625	0.11719
	4	1	0.44		4	0.21875	0.05469

Table 6.6: *The bounds* (5.1) *and* (5.2) *when* $a = 0$, $b = 5$, $2s = 2$, $\mu_1 = 1$ *and* $\mu_2 = 1.5$.

(1980)); when u is differentiable, this reduces to $u^{(3)} \geq 0$. A detailed discussion of risk aversion can be found, for example, in Scarsini (1994).

The certainty equivalence of a random fortune X, denoted as C_X, is defined as the root of the equation

$$u(C_X) = E[u(X)].$$

The certainty equivalence C_X of a random fortune X is therefore defined by the condition that the decision-maker is indifferent between receiving the random amount X or the fixed amount C_X. It is worth mentioning that the inequality $C_X \leq E[X]$ holds for any non-decreasing concave utility function u. Therefore, a profit-seeking risk averse decision-maker always prefers certainty to uncertainty, even if the certain income is (to a certain amount) less than the expected uncertain income. A closely related concept is that of the risk premium. The risk premium of X is defined as the difference between the expected fortune and the certainty equivalence; that is, $R_X = E[X] - C_X$. As a consequence, we have, by definition, that

$$u(E[X] - R_X) = E[u(X)].$$

This can be interpreted as follows: the risk premium R_X is equal to the amount of money the individual is willing to pay in order to get a certain fortune $E[X]$ rather than the random fortune X; that is, to replace uncertainty by certainty.

The determination of C_X is difficult, or even impossible, when the cdf of X is complicated or when it is not fully known. However, as it is described in this paper, tight bounds on C_X can be obtained with the help of the s-convex extrema. Since all reasonable utility functions are 1, 2 and 3-convex, the extrema corresponding to $s = 1$, 2 and 3 are useful.

APPENDICES

A Lindsay-Type Representation of Moment-Matrices Determinants

Proposition A.1. *For any cdf* $F \in \mathcal{B}_{2k+2}([a,b];\mu_1,\mu_2,\ldots,\mu_{2k+1})$, *the representations*

$$\|\boldsymbol{Q}_k - a\boldsymbol{P}_k\| = \frac{1}{(k+1)!}E\Big[\prod_{i=0}^{k}(X_i - a)\prod_{0\leq j<l\leq k}(X_j - X_l)^2\Big], \tag{A.1}$$

and

$$\left\| b\boldsymbol{P}_k - \boldsymbol{Q}_k \right\| = \frac{1}{(k+1)!} E\left[\prod_{i=0}^{k} (b - X_i) \prod_{0 \le j < l \le k} (X_j - X_l)^2 \right], \tag{A.2}$$

hold, where $X_0, X_1, \ldots, X_k$ are independent and identically distributed random variables with the common cdf F. Also, for any $F \in \mathcal{B}_{2k+1}([a,b]; \mu_1, \mu_2, \ldots, \mu_{2k})$, the representations

$$\left\| \boldsymbol{P}_k \right\| = \frac{1}{(k+1)!} E\left[\prod_{0 \le j < l \le k} (X_j - X_l)^2 \right], \tag{A.3}$$

and

$$\left\| -ab\boldsymbol{P}_{k-1} + (a+b)\boldsymbol{Q}_{k-1} - \boldsymbol{R}_{k-1} \right\| = \frac{1}{k!} E\left[\prod_{i=0}^{k-1} (X_i - a)(b - X_i) \prod_{0 \le j < l \le k-1} (X_j - X_l)^2 \right], \tag{A.4}$$

hold, where $X_0, X_1, \ldots, X_k$ are independent and identically distributed random variables with the common cdf F.

Proof. For the proof of (A.1)–(A.4) we will use the following result of Wilks (1960); see also Theorem 2A in Lindsay (1989a). For $k \ge 1$, let $\boldsymbol{Y}^t = (Y_0, Y_1, \ldots, Y_k)$ be a random vector with a $(k+1)$-dimensional cdf G. Assume that the second moment matrix $\boldsymbol{C} = E[\boldsymbol{Y}\boldsymbol{Y}^t]$ exists. Let $\boldsymbol{Y}_0, \boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_k$ be $k+1$ independent replicates from G and define $\boldsymbol{A}$ to be the $(k+1) \times (k+1)$ random matrix with jth column $\boldsymbol{Y}_j$. Then the random variable $\left\| \boldsymbol{A} \right\|$ has mean 0 and variance $(k+1)! \left\| \boldsymbol{C} \right\|$; that is

$$\left\| \boldsymbol{C} \right\| = \frac{E\left\| \boldsymbol{A} \right\|^2}{(k+1)!}.$$

Now let $X, X_0, X_1, \ldots, X_k$ be independent random variables with the common cdf F. If we choose $\boldsymbol{Y}^t = (\sqrt{X-a}, X\sqrt{X-a}, X^2\sqrt{X-a}, \ldots, X^k\sqrt{X-a})$ then we have $\boldsymbol{C} = \boldsymbol{Q}_k - a\boldsymbol{P}_k$ and

$$\left\| \boldsymbol{A} \right\| = \prod_{i=0}^{k} \sqrt{X_i - a} \prod_{0 \le j < l \le k} (X_j - X_l),$$

and (A.1) follows. If we argue similarly with $\boldsymbol{Y}^t = (\sqrt{b-X}, X\sqrt{b-X}, X^2\sqrt{b-X}, \ldots, X^k\sqrt{b-X})$ we get (A.2). Formula (A.3) is equation (1.2) in Lindsay (1989); it can also be obtained as (A.1) and (A.2) with $\boldsymbol{Y}^t = (1, X, X^2, \ldots, X^k)$. Finally, taking $\boldsymbol{Y}^t = (\sqrt{(X-a)(b-X)}, X\sqrt{(X-a)(b-X)}, X^2\sqrt{(X-a)(b-X)}, \ldots, X^{k-1}\sqrt{(X-a)(b-X)})$ we have

$$\left\| \boldsymbol{A} \right\| = \prod_{i=0}^{k-1} \sqrt{(X_i - a)(X_i - b)} \prod_{0 \le j < l \le k-1} (X_j - X_l),$$

and (A.4) follows. □

B Moment Matrices and Linear Regression

Let $\boldsymbol{Y} = (Y_{n+1}, Y_n, \ldots, Y_1, Y_0)^t$ be an $(n+2)$-dimensional random vector and consider the best linear predictor $\widehat{Y}_{n+1}$ of Y_{n+1} by $Y_0, Y_1, \ldots, Y_n$; that is, $\widehat{Y}_{n+1} = \sum_{j=0}^{n} \beta_j^* Y_{n-j}$, where $\boldsymbol{\beta}^* = (\beta_0^*, \beta_1^*, \ldots, \beta_n^*)^t$ is the solution of

$$\boldsymbol{\beta}^* = \arg \inf_{\boldsymbol{\beta} \in \mathbb{R}^{n+1}} E\Big[Y_{n+1} - \sum_{j=0}^{n} \beta_j Y_{n-j}\Big]^2.$$

Let $\boldsymbol{\Sigma} = E[\boldsymbol{Y}\boldsymbol{Y}^t]$ be the second order moment matrix of $\boldsymbol{Y}$, and let $\boldsymbol{\Sigma}_{i:j}$ be $\boldsymbol{\Sigma}$ without its ith row and jth column. Then $\boldsymbol{\beta}^*$ is the solution of the Cramér system

$$\boldsymbol{\beta}^t \boldsymbol{\Sigma}_{1:1} = (EY_{n+1}Y_n, EY_{n+1}Y_{n-1}, \ldots, EY_{n+1}Y_0)^t.$$

Assuming that $\|\boldsymbol{\Sigma}_{1:1}\| \neq 0$ we get

$$\beta_k^* = (-1)^k \frac{\|\boldsymbol{\Sigma}_{1:k+2}\|}{\|\boldsymbol{\Sigma}_{1:1}\|}, \qquad k = 0, 1, \ldots, n.$$

The residual variance σ_r^2 in the linear prediction of Y_{n+1} by $Y_0, Y_1, \ldots, Y_n$ is

$$\begin{aligned}\sigma_r^2 &= E(Y_{n+1} - \widehat{Y}_{n+1})^2 \\ &= EY_{n+1}^2 - \sum_{i=0}^{n} \beta_i^* E(Y_{n+1}Y_{n-i}) - \sum_{j=0}^{n} \beta_j^* \Big[E(Y_{n+1}Y_{n-j}) - \sum_{i=0}^{n} \beta_i^* E(Y_{n-i}Y_{n-j})\Big].\end{aligned}$$

Now, since

$$E(Y_{n+1}Y_{n-j}) - \sum_{i=0}^{n} \beta_i^* E(Y_{n-i}Y_{n-j}) = 0,$$

for $j = 0, 1, \ldots, n$, we get

$$\sigma_r^2 = \frac{EY_{n+1}^2 \|\boldsymbol{\Sigma}_{1:1}\| + \sum_{i=0}^{n} (-1)^{i+1} E(Y_{n+1}Y_{n-i}) \|\boldsymbol{\Sigma}_{1:i+2}\|}{\|\boldsymbol{\Sigma}_{1:1}\|} = \frac{\|\boldsymbol{\Sigma}\|}{\|\boldsymbol{\Sigma}_{1:1}\|}. \tag{B.1}$$

Proposition B.1. *Let X be a random variable valued in $[a, b]$, with moments μ_k, $k = 0, 1, \ldots$. Then,*
(i)

$$\frac{\|\boldsymbol{P}_{n+1}\|}{\|\boldsymbol{P}_n\|} = \inf_{\boldsymbol{\beta} \in \mathbb{R}^{n+1}} E\Big[X^{n+1} - \sum_{j=0}^{n} \beta_j X^j\Big]^2,$$

with the convention that $X^0 \equiv 1$;
(ii)

$$\frac{\| - ab\boldsymbol{P}_{n+1} + (a+b)\boldsymbol{Q}_{n+1} - \boldsymbol{R}_{n+1}\|}{\| - ab\boldsymbol{P}_n + (a+b)\boldsymbol{Q}_n - \boldsymbol{R}_n\|} = \inf_{\boldsymbol{\beta} \in \mathbb{R}^{n+1}} E\Big[(X-a)(b-X)\Big(X^{n+1} - \sum_{j=0}^{n} \beta_j X^j\Big)^2\Big];$$

(iii)

$$\frac{\|\boldsymbol{Q}_{n+1} - a\boldsymbol{P}_{n+1}\|}{\|\boldsymbol{Q}_n - a\boldsymbol{P}_n\|} = \inf_{\boldsymbol{\beta}\in\mathbb{R}^{n+1}} E\Big[(X-a)\Big(X^{n+1} - \sum_{j=0}^{n} \beta_j X^j\Big)^2\Big];$$

(iv)

$$\frac{\|b\boldsymbol{P}_{n+1} - \boldsymbol{Q}_{n+1}\|}{\|b\boldsymbol{P}_n - \boldsymbol{Q}_n\|} = \inf_{\boldsymbol{\beta}\in\mathbb{R}^{n+1}} E\Big[(b-X)\Big(X^{n+1} - \sum_{j=0}^{n} \beta_j X^j\Big)^2\Big].$$

Proof. It suffices to apply (B.1) with $Y_j = X^j$, $j = 0, 1, \ldots, n+1$, for (i); with $Y_j = X^j\sqrt{(X-a)(b-X)}$, $j = 0, 1, \ldots, n+1$, for (ii); with $Y_j = X^j\sqrt{X-a}$, $j = 0, 1, \ldots, n+1$, for (iii); and with $Y_j = X^j\sqrt{b-X}$, $j = 0, 1, \ldots, n+1$, for (iv). $\square$

References

[1] Denuit, M., De Vylder, F. E. & Lefèvre, Cl. (1999). Extremal generators and extremal distributions for the continuous s-convex stochastic orderings. *Insurance: Mathematics & Economics* **24**, 201–217.

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

[3] Denuit, M., Lefèvre, Cl. & Mesfioui, M. (1999). On s-convex stochastic extrema for arithmetic risks. *Insurance: Mathematics & Economics*, to appear.

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

[5] Denuit, M., Lefèvre, Cl. & Shaked, M. (1999). Some contributions to the theory of high convexity stochastic orders. *Statistics & Probability Letters*, to appear.

[6] Denuit, M., Lefèvre, Cl. & Utev, S. (1999). Stochastic orderings of convex/concave-type on an arbitrary grid. *Mathematics of Operations Research*, to appear.

[7] Dette, H. & Studden, W. (1997). *The Theory of Canonical Moments with Applications in Statistics, Probability, and Analysis*. Wiley. New York.

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

[9] Eckberg, A. E. (1977). Sharp bounds on Laplace-Stieltjes transforms, with applications to various queueing problems. *Mathematics of Operations Research* **2**, 135–142.

[10] Harris, B. (1959). Determining bounds on integrals with applications to cataloging problems. *Annals of Mathematical Statistics* **30**, 521–548.

[11] Harris, B. (1962). Determining bounds on expected values of certain functions. *Annals of Mathematical Statistics* **33**, 1454–1457.

[12] Karlin, S. & Shapley, L. S. (1953). *Geometry of moment spaces*. Memoirs **12** of the American Mathematical Society. Providence.

[13] Karlin, S. & Studden, W. J. (1966). *Tchebycheff Systems: with Applications in Analysis and Statistics*. Wiley. New York.

[14] Lau, T.-S. (1996). On the heterogeneity of Fecundability. *Lifetime Data Analysis* **2**, 403–415.

[15] Lefèvre, Cl., Hallin, M. & Narayan, P. (1986). On fractional linear bounds for probability generating functions. *Journal of Applied Probability* **23**, 904–913.

[16] Lindsay, B. G. (1989a). On the determinants of moment matrices. *Annals of Statistics* **17**, 711–721.

[17] Lindsay, B. G. (1989b). Moment matrices: applications in mixtures. *Annals of Statistics* **17**, 722–740.

[18] Lindsay, B. G. & Roeder, K. (1997). Moment-based oscillation properties of mixture models. *Annals of Statistics* **25**, 378–386.

[19] Mammana, C. (1954). Sul problema algebraico dei momenti. *Ann. Scuola Norm. Sup. Pisa Cl. Sci.* **8**, 133–140.

[20] Menezes, C., Geiss, C. & Tressler, J. (1980). Increasing downside risk. *The American Economic Review* **70**, 921–932.

[21] Pečarić, J. E., Proschan, F. & Tong, Y. L. (1992). *Convex Functions, Partial Orderings, and Statistical Applications*. Academic Press. New York.

[22] Popoviciu, T. (1933). Sur quelques propriétés des fonctions d'une ou de deux variables réelles. *Mathematica* **7**, 1–85.

[23] Rolski, T. (1976). Order relations in the set of probability distribution functions and their applications in queueing theory. *Dissertationes Mathematicae* **132**, 5–47.

[24] Scarsini, M. (1994). Comparing risk and risk aversion. In *Stochastic Orders and their Applications* (M. Shaked and J. G. Shanthikumar, eds.). Academic Press. New York, 351–378.

[25] Shaked, M. & Shanthikumar, J. G. (1994). *Stochastic Orders and Their Applications*. Academic Press. New York.

[26] Stoyan, D. (1983). *Comparison Methods for Queues and Other Stochastic Models*. Wiley. New York.

[27] Studden, W. J. (1982). Optimal designs for weighted polynomial regression using canonical moments. In *Statistical Decision Theory and Related Topics III* **2** (S. S. Gupta and J. O. Berger, eds.). Academic Press. New York, 335–350.

[28] Von Neumann, J. & Morgenstern, O. (1947). *Theory of Games and Economic Behavior* (Second edition). Princeton University Press, Princeton.

[29] Whitt, W. (1983). Untold horrors of the waiting room: what the equilibrium distribution will never tell about the queue-length process. *Management Science* **29**, 395–408.

[30] Wilks, S. S. (1960). Multidimensional statistical scatter. In *Contributions to Probability and Statistics* (I. Olkin et al., eds.). Stanford University Press. Stanford, California, 486–503.