KATHOLIEKE
UNIVERSITEIT
LEUVEN

DEPARTEMENT TOEGEPASTE ECONOMISCHE WETENSCHAPPEN

RESEARCH REPORT 0049

UPPER AND LOWER BOUNDS FOR SUMS OF RANDOM VARIABLES

by
R. KAAS
J. DHAENE
M. GOOVAERTS

D/2000/2376/49

Upper and Lower Bounds for Sums of Random Variables*

Rob Kaas Jan Dhaene Marc J. Goovaerts

November 16, 2000

Abstract

In this contribution, the upper bounds for sums of dependent random variables $X_1 + X_2 + ... + X_n$ derived by using comonotonicity are sharpened for the case when there exists a random variable Z such that the distribution functions of the X_i, given $Z = z$, are known. By a similar technique, lower bounds are derived. A numerical application for the case of lognormal random variables is given.

1 Introduction

In some recent articles, Goovaerts, Denuit, Dhaene, Müller and several others have applied theory originally studied by Fréchet in the previous century to derive upper bounds for sums $S = X_1 + X_2 + ... + X_n$ of random variables $X_1, X_2, ..., X_n$ of which the marginal distribution is known, but the joint distribution of the random vector $(X_1, X_2, ..., X_n)$ is either unspecified or too cumbersome to work with. These upper bounds are actually suprema in the sense of convex order. The concept of convex order is closely related to the notion of stop-loss order which is more familiar in actuarial circles. Both express which of two risks is the more risky one. Assuming that only the marginal distributions of the X_i are given (or used), the riskiest instance S_n of S occurs when the risks $X_1, X_2, ..., X_n$ are comonotonous. This means that they

*Paper presented at the 4th Conference on Insurance: Mathematics & Economics, Barcelona, July 24-26, 2000.

are all non-decreasing functions of one uniform(0,1) random variable U, and since the marginal distribution must be $\Pr[X_i \leq x] = F_i(x)$, the comonotonous distribution is that of the vector $(F_1^{-1}(U), F_2^{-1}(U), ..., F_n^{-1}(U))$.

In this contribution we assume that the marginal distribution of each random variable $X_1, X_2, ..., X_n$ is known. In addition, we assume that there exists some random variable Z, with a known distribution function, such that for any i and for any z in the support of Z, the conditional distribution function of X_i, given $Z = z$, is known. We will derive upper and lower bounds in convex order for $S = X_1 + X_2 + ... + X_n$, based on these conditional distribution functions. Two extreme situations are possible here. One is that $Z = S$, or some one-to-one function of it. Then the convex lower bound for S, which equals $E[S|Z]$, will just be S itself. The other is that Z is independent of all $X_1, X_2, ..., X_n$. In this case we actually do not have any extra information at all and the upper bound for S is just the same comonotonous bound as before, while the lower bound reduces to the trivial bound $E[S]$. But in some cases, and the lognormal discount process of section 5 is a good example, a random variable Z can be found with the property that by conditioning on it we can actually compute a non-trivial lower bound and a sharper upper bound than S_u for S.

In section 2, we will present a short exposition of the theory we need. Section 3 gives upper bounds, section 4 improved upper bounds, as well as lower bounds, both applied to the case of lognormal distributions in section 5. Section 6 gives numerical examples of the performance of these bounds, and section 7 concludes.

2 Some theory on comonotonous random variables

Let $F_1, F_2, ..., F_n$ be univariate cumulative distribution functions (cdf's for short). Fréchet studied the class of all n-dimensional cdf's $F_{\mathbf{X}}$ of random vectors $\mathbf{X} \equiv (X_1, X_2, ..., X_n)$ with given marginal cdf's $F_1, F_2, ..., F_n$, where for any real number x we have $\Pr[X_i \leq x] = F_i(x)$, $i = 1, 2, ..., n$. In this paper, we will consider the problem of determining stochastic lower and upper bounds for the cdf of the random variable $X_1 + X_2 + ... + X_n$, without restricting to independence between the terms X_i. We will always assume that the marginals cdf's of the X_i are given, and that all cdf's involved have a finite mean.

The stochastic bounds for random variables will be in terms of "convex order", which is defined as follows:

Definition 1 *Consider two random variables X and Y. Then X is said to precede Y in the convex order sense, notation $X \leq_{cx} Y$, if and only if for all convex real functions v such that the expectations exist, we have*

$$E[v(X)] \leq E[v(Y)].$$

It can be proven, see e.g. Shaked & Shanthikumar (1994), that the condition in this definition is equivalent with the following condition:

$$\begin{aligned} E[X] &= E[Y], \\ E[X-d]_+ &\leq E[Y-d]_+ \text{ for all } d, \end{aligned}$$

where $E[Z]_+$ is a notation for $E[\max\{Z, 0\}]$.

Using an integration by parts, the ordering condition between the stop-loss premiums $E[X-d]_+$ and $E[Y-d]_+$ can also be expressed as

$$\int_d^\infty (1 - F_X(x))\, dx \leq \int_d^\infty (1 - F_Y(x))\, dx \text{ for all } d.$$

In case $X \leq_{cx} Y$, extreme values are more likely for Y than for X. In terms of utility theory, $X \leq_{cx} Y$ entails that loss X is preferred to loss Y by all risk averse decision makers, i.e., $E[u(-X)] \geq E[u(-Y)]$ for all concave non-decreasing utility functions u. This means that replacing the (unknown) distribution function of a loss X by the distribution function of a loss Y can be considered as an actuarially prudent strategy, for example when determining reserves.

From the relation above, we see immediately that

$$\frac{d}{dx}\{E[X-x]_+ - E[Y-x]_+\} = F_X(x) - F_Y(x).$$

Thus, two random variables X and Y with equal mean are convex ordered if their cdf's cross once. This last condition can be observed to hold in most conceivable examples, but it is easy to construct instances with $X \leq_{cx} Y$ where the cdf's cross more than once.

It follows immediately that $X \leq_{cx} Y$ implies $Var[X] \leq Var[Y]$. The reverse implication does not hold in general; for a counterexample, see e.g. Brockett and Garven (1998). Also note that $X \leq_{cx} Y$ is equivalent to $-X \leq_{cx} -Y$. This means that it makes no difference if we interpret the random variables as losses or as gains.

For any random vector $\mathbf{X}$ with marginal cdf's $F_1, F_2, ..., F_n$ the following convex order relation holds:

$$X_1 + X_2 + ... + X_n \leq_{cx} S_u \stackrel{def}{=} F_1^{-1}(U) + F_2^{-1}(U) + ... + F_n^{-1}(U),$$

where U is a uniform(0,1) random variable, and where the p-th quantile of a random variable X with cdf F_X is, as usual, defined by

$$F_X^{-1}(p) \stackrel{def}{=} \inf \{x \epsilon R | F_X(x) \geq p\}, \quad p \in [0,1].$$

Goovaerts, Dhaene & De Schepper (2000) prove this order relation directly, while Müller (1997) derives it as a special case of the concept of supermodular ordering. This relation can be interpreted as follows: the most risky random vector with given marginals (in the sense that the sum of their components is largest in the convex order sense) has the *comonotonous* joint distribution, which means that it has the joint distribution of $\left(F_1^{-1}(U), F_2^{-1}(U), ..., F_n^{-1}(U)\right)$. The components of this random vector are maximally dependent, all components being non-decreasing functions of the same random variable.

The inverse cdf of a sum of comonotonous random variables can easily be computed. Indeed, if $S_u =_d F_1^{-1}(U) + F_2^{-1}(U) + ... + F_n^{-1}(U)$, where $=_d$ means equality in distribution, then

$$F_{S_u}^{-1}(p) = \sum_{i=1}^{n} F_i^{-1}(p), \quad p \in [0,1].$$

Recently, the concept of comonotonicity has been considered in many actuarial papers, see e.g. Müller (1997), Wang & Dhaene (1998), Dhaene, Wang, Young & Goovaerts (1998). Dependence in portfolios and related stochastic orders are also considered in Dhaene & Goovaerts (1996), Denuit & Lefèvre (1997), Dhaene & Goovaerts (1997), Bäuerle & Müller (1998), Wang & Young (1998), Goovaerts & Redant (1999), Denuit, De Vylder & Lefèvre (1999), Dhaene & Denuit (1999), and others.

3 Comonotonous Upper Bounds for Sums of Random Variables

The usual definition of the inverse of a cdf is the left-continuous function $F_X^{-1}(p) = \inf\{x \in R | F_X(x) \geq p\}$. But if $F_X(x) = p$ holds for an interval of values for x, any element of it could serve as $F_X^{-1}(p)$. In this paper, we introduce a more sophisticated definition which enables us to choose that particular inverse cdf with the property that for a certain d, the relation $F_X^{-1}(F_X(d)) = d$ holds.

For $p \in [0,1]$, a possible choice for the inverse of F_X in p is any point in the interval

$$[\inf\{x \in R | F_X(x) \geq p\}\,;\ \sup\{x \in R | F_X(x) \leq p\}]\,.$$

Here we take $\inf \emptyset = +\infty$ and $\sup \emptyset = -\infty$. Taking the left hand border of this interval to be the value of the inverse cdf at p, we get $F_X^{-1}(p)$. Similarly, we define $F_X^{-1\bullet}(p)$ as the right hand border of the interval:

$$F_X^{-1\bullet}(p) = \sup\{x \in R | F_X(x) \leq p\}\,, \quad p \in [0,1]\,.$$

Note that $F_X^{-1}(0) = -\infty$ and $F_X^{-1\bullet}(1) = +\infty$, while $F_X^{-1}(p)$ and $F_X^{-1\bullet}(p)$ are finite for all $p \in (0,1)$. For any α in $[0,1]$, we define the α-inverse of F_X as follows:

$$F_X^{-1(\alpha)}(p) = \alpha\, F_X^{-1}(p) + (1-\alpha)\, F_X^{-1\bullet}(p), \qquad p \in (0,1)\,.$$

For a comonotonous random vector $(X_1, X_2, ..., X_n)$, it follows that for all α in $[0,1]$:

$$F_{X_1+X_2+...+X_n}^{-1(\alpha)}(p) = \sum_{i=1}^{n} F_{X_i}^{-1(\alpha)}(p), \quad p \in (0,1)\,.$$

The following result was already mentioned in Section 1. We give a new proof for it, based on the α-inverse just introduced, because this method of proof leads to new results that we will need in the sequel of this paper.

Proposition 2 *Let U be a uniform(0,1) random variable. For any random vector $(X_1, X_2, ..., X_n)$ with marginal cdf's $F_1, F_2, ..., F_n$, we have*

$$X_1 + X_2 + ... + X_n \leq_{cx} F_1^{-1}(U) + F_2^{-1}(U) + ... + F_n^{-1}(U).$$

Proof. Let S and S_u be defined by $S = X_1 + X_2 + ... + X_n$ and $S_u = F_1^{-1}(U) + F_2^{-1}(U) + ... + F_n^{-1}(U)$, respectively, with U uniform(0,1). Then obviously $E[S] = E[S_u]$. To prove the stop-loss inequalities needed to establish convex order, consider an arbitrary fixed real number d, with $0 < F_{S_u}(d) < 1$. Let $\alpha \in [0,1]$ be determined such that

$$F_{S_u}^{-1(\alpha)}(F_{S_u}(d)) = d.$$

Then we have

$$\begin{aligned} E[S-d]_+ &= E\left[S - F_{S_u}^{-1(\alpha)}[F_{S_u}(d)]\right]_+ = E\left[\sum_{i=1}^{n}\left(X_i - F_i^{-1(\alpha)}(F_{S_u}(d))\right)\right]_+ \\ &\leq \sum_{i=1}^{n} E\left[X_i - F_i^{-1(\alpha)}(F_{S_u}(d))\right]_+ . \end{aligned}$$

On the other hand we find

$$\begin{aligned} E[S_u - d]_+ &= E\left[F_{S_u}^{-1}(U) - d\right]_+ = \int_0^1 \left(F_{S_u}^{-1}(p) - d\right)_+ dp \\ &= \int_{F_{S_u}(d)}^{1} \left(F_{S_u}^{-1}(p) - F_{S_u}^{-1(\alpha)}(F_{S_u}(d))\right) dp \\ &= \sum_{i=1}^{n} \int_{F_{S_u}(d)}^{1} \left(F_i^{-1}(p) - F_i^{-1(\alpha)}(F_{S_u}(d))\right) dp. \end{aligned}$$

One can verify that for any $p \in \left(F_{S_u}(d); F_i\left(F_i^{-1(\alpha)}(F_{S_u}(d))\right)\right)$ we have

$$F_i^{-1}(p) = F_i^{-1(\alpha)}(F_{S_u}(d)).$$

This implies

$$\begin{aligned} E[S_u - d]_+ &= \sum_{i=1}^{n} \int_{F_i\left(F_i^{-1(\alpha)}(F_{S_u}(d))\right)}^{1} \left(F_i^{-1}(p) - F_i^{-1(\alpha)}(F_{S_u}(d))\right) dp \\ &= \sum_{i=1}^{n} \int_0^1 \left(F_i^{-1}(p) - F_i^{-1(\alpha)}(F_{S_u}(d))\right)_+ dp \\ &= \sum_{i=1}^{n} E\left[F_i^{-1}(U) - F_i^{-1(\alpha)}(F_{Su}(d))\right]_+ \\ &= \sum_{i=1}^{n} E\left[X_i - F_i^{-1(\alpha)}(F_{S_u}(d))\right]_+ , \end{aligned}$$

so we have proven that $E[S-d]_+ \leq E[S_u-d]_+$ holds for all retentions d with $0 < F_{S_u}(d) < 1$.

As the stop-loss transform is a continuous, non-increasing function of the retention d, we find that the result above implies

$$E\left[S - F_{S_u}^{-1\bullet}(0)\right]_+ \leq E\left[S_u - F_{S_u}^{-1\bullet}(0)\right]_+,$$

as well as

$$E\left[S - F_{S_u}^{-1}(1)\right]_+ \leq E\left[S_u - F_{S_u}^{-1}(1)\right]_+.$$

So $E[S-d]_+ \leq E[S_u-d]_+$ also holds for retentions d with $F_{S_u}(d) = 0$ or $F_{S_u}(d) = 1$. ■

If $d \in \left(F_{S_u}^{-1\bullet}(0), F_{S_u}^{-1}(1)\right)$, then $0 < F_{S_u}(d) < 1$, so we find the following Corollary from the proof of Proposition 1.

Corollary 3 *Let U be uniform(0,1) and let $S_u = F_1^{-1}(U)+...+F_n^{-1}(U)$. If $d \in \left(F_{S_u}^{-1\bullet}(0), F_{S_u}^{-1}(1)\right)$, then the stop-loss premium at retention d of S_u is given by*

$$E[S_u-d]_+ = \sum_{i=1}^{n} E\left[X_i - F_i^{-1(\alpha)}\left(F_{S_u}(d)\right)\right]_+$$

with $\alpha \in [0,1]$ determined by $F_{S_u}^{-1(\alpha)}\left(F_{S_u}(d)\right) = d$.

The expression for the stop-loss premiums of a comonotonous sum S_u can also be written in terms of the usual inverse cdf's. Indeed, for any retention $d \in \left(F_{S_u}^{-1\bullet}(0), F_{S_u}^{-1}(1)\right)$, we have

$$\begin{aligned} E\left[X_i - F_i^{-1(\alpha)}\left(F_{S_u}(d)\right)\right]_+ &= E\left[X_i - F_i^{-1}\left(F_{S_u}(d)\right)\right]_+ \\ &\quad - \left(F_i^{-1(\alpha)}\left(F_{S_u}(d)\right) - F_i^{-1}\left(F_{S_u}(d)\right)\right)\left(1 - F_{S_u}(d)\right). \end{aligned}$$

Summing over i, and taking into account the definition of α, we find the expression derived in Dhaene, Wang, Young & Goovaerts (1999), where the random variables are assumed to be non-negative:

$$E[S_u-d]_+ = \sum_{i=1}^{n} E\left[X_i - F_i^{-1}\left(F_{S_u}(d)\right)\right]_+ - \left(d - F_{S_u}^{-1}\left(F_{S_u}(d)\right)\right)\left(1 - F_{S_u}(d)\right).$$

From Corollary 1, we can conclude that any stop-loss premium of a sum of comonotonous random variables can be written as the sum of stop-loss premiums for the individual random variables involved. Corollary 1 provides an algorithm for directly computing stop-loss premiums of sums of comonotonous random variables, without having to compute the entire cdf of the sum itself. Indeed, in order to compute the stop-loss premium with retention d, we only need to know $F_{S_u}(x)$ for x equal to d. The cdf at x follows from

$$F_{S_u}(x) = \sup \left\{ p \in [0,1] \mid \sum_{i=1}^{n} F_i^{-1}(p) \leq x \right\}.$$

Now assume that the marginal cdf's F_i are continuous on R and strictly increasing on $\left(F_i^{-1\bullet}(0), F_i^{-1}(1)\right)$. Then one can verify that F_{S_u} is also continuous on R and strictly increasing on $\left(F_{S_u}^{-1\bullet}(0), F_{S_u}^{-1}(1)\right)$, and that $F_{S_u}^{-1}$ is strictly increasing and continuous on $(0,1)$. Hence, for any $x \in \left(F_{S_u}^{-1\bullet}(0), F_{S_u}^{-1}(1)\right)$, the value $F_{S_u}(x)$ can be obtained unambiguously from

$$\sum_{i=1}^{n} F_i^{-1}\left(F_{S_u}(x)\right) = x.$$

In this case, we also find

$$E[S_u - d]_+ = \sum_{i=1}^{n} E\left[X_i - F_i^{-1}\left(F_{S_u}(d)\right)\right]_+,$$

which holds for any retention $d \in \left(F_{S_u}^{-1\bullet}(0), F_{S_u}^{-1}(1)\right)$.

Corollary 1 can be used for deriving upper bounds for the price of an Asian option, see Simon, Goovaerts & Dhaene (2000).

4 Improved Bounds for Sums of Random Variables

4.1 Upper Bounds

As $\left(F_1^{-1}(U), F_2^{-1}(U), ..., F_n^{-1}(U)\right)$ is a random vector with marginals $F_1, ..., F_n$, the upper bound $S_u = F_1^{-1}(U) + F_2^{-1}(U) + ... + F_n^{-1}(U)$ is the best that can be derived under the conditions stated in Proposition 1; it is a supremum in terms of convex order. Let us now assume that we have complete (or partial) information, more than just the marginal distributions, concerning the dependence structure of the random vector $(X_1, X_2, ..., X_n)$, but that exact computation of the cdf of the sum $S = X_1 + X_2 + ... + X_n$ is not feasible. In this case, we will show that it is possible to derive improved upper bounds for S, and also non-trivial lower bounds, based on the information we have on the dependence structure. This is accomplished by conditioning on a random variable Z which is assumed to be some function of the random vector X. We will assume that we know the distribution of Z, and also the conditional cdf's, given $Z = z$, of the random variables X_i. A suitable example is to use $Z = \sum \log X_i$ when the X_i are lognormal. In the following proposition, we introduce the notation $F_{X_i|Z}^{-1}(U)$ for the random variable $f_i(U, Z)$, where the function f_i is defined by $f_i(u, z) = F_{X_i|Z=z}^{-1}(u)$.

Proposition 4 *Let U be uniform(0,1), and consider a random variable Z which is independent of U. Then we have*

$$X_1 + X_2 + ... + X_n \leq_{cx} S_u' \stackrel{def}{=} F_{X_1|Z}^{-1}(U) + F_{X_2|Z}^{-1}(U) + ... + F_{X_n|Z}^{-1}(U).$$

Proof. From Proposition 1, we get for any convex function v,

$$\begin{aligned} E\left[v\left(X_1 + ... + X_n\right)\right] &= \int_{-\infty}^{\infty} E\left[v\left(X_1 + ... + X_n\right)|Z = z\right] \, dF_Z(z) \\ &\leq \int_{-\infty}^{\infty} E\left[v\left(f_1(U, z) + ... + f_n(U, z)\right)\right] dF_Z(z) \\ &= E\left[v\left(f_1(U, Z) + ... + f_n(U, Z)\right)\right] \end{aligned}$$

from which the stated result follows directly. ■

Note that the random vector $\left(F_{X_1|Z}^{-1}(U), F_{X_2|Z}^{-1}(U), ..., F_{X_n|Z}^{-1}(U)\right)$ has marginals $F_1, F_2, ..., F_n$, because

$$\begin{aligned} F_i(x) &= \Pr[X_i \leq x] \\ &= \int_{-\infty}^{\infty} \Pr\left[X_i \leq x \,|\, Z = z\right] dF_Z(z) \\ &= \int_{-\infty}^{\infty} \Pr\left[F_{X_i|Z=z}^{-1}(U) \leq x\right] dF_Z(z) \end{aligned}$$

$$
\begin{aligned}
&= \int_{-\infty}^{\infty} \Pr\left[f_i(U,z) \leq x\right] dF_Z(z) \\
&= \Pr\left[f_i(U,Z) \leq x\right].
\end{aligned}
$$

In view of Proposition 1 this implies

$$F^{-1}_{X_1|Z}(U) + F^{-1}_{X_2|Z}(U) + \ldots + F^{-1}_{X_n|Z}(U) \leq_{cx} F_1^{-1}(U) + F_2^{-1}(U) + \ldots + F_n^{-1}(U).$$

The left hand side of this relation is $S_u^{'}$; the right hand side is S_u. In order to obtain the distribution function of $S_u^{'}$, observe that given the event $Z = z$, this random variable is a sum of comonotonous random variables. Hence,

$$F^{-1}_{S_u^{'}|Z=z}(p) = \sum_{i=1}^{n} F^{-1}_{X_i|Z=z}(p), \quad p \in [0,1].$$

If the marginal cdf's $F_{X_i|Z=z}$ are strictly increasing and continuous, so is $F_{S_u^{'}|Z=z}$, and then $F_{S_u^{'}|Z=z}(x)$ follows by solving

$$\sum_{i=1}^{n} F^{-1}_{X_i|Z=z}\left(F_{S_u^{'}|Z=z}(x)\right) = x.$$

The cdf of $S_u^{'}$ then follows from

$$F_{S_u^{'}}(x) = \int_{-\infty}^{\infty} F_{S_u^{'}|Z=z}(x)\, dF_Z(z).$$

Application of Proposition 2 to lognormal marginals X_i is considered in Section 5, but see also the simple examples with $n = 2$ at the end of this section. Note that if Z is independent of all $X_1, X_2, ..., X_n$, upper bound $S_u^{'}$ reduces to S_u.

4.2 Lower Bounds

Let $\mathbf{X}$ be a random vector with marginals $F_1, F_2, ..., F_n$, and assume that we want to find a lower bound S_l, in the sense of convex order, for $S = X_1 + X_2 + ... + X_n$. We can obtain such a bound by conditioning on some random variable Z, again assumed to be a function of the random vector $\mathbf{X}$.

Proposition 5 *For any random vector* $\mathbf{X}$ *and random variable* Z*, we have*

$$S_l \overset{def}{=} E[X_1|Z] + E[X_2|Z] + ... + E[X_n|Z] \leq_{cx} X_1 + X_2 + ... + X_n.$$

Proof. By Jensen's inequality, we find that for any convex function v, the following inequality holds:

$$\begin{aligned} E[v(X_1 + X_2 + ... + X_n)] &= E_Z\, E[v(X_1 + X_2 + ... + X_n)|Z] \\ &\geq E_Z\, [v(E[X_1 + X_2 + ... + X_n|Z])] \\ &= E_Z\, [v(E[X_1|Z] + ... + E[X_n|Z])]. \end{aligned}$$

This proves the stated result. ■

Note that if Z and S are mutually independent, Proposition 3 leads to the trivial lower bound $E[S] \leq_{cx} S$. On the other hand, if Z and S have a one-to-one relation, the lower bound in Proposition 3 coincides with S. Note further that $E[E[X_i|Z]] = E[X_i]$ always holds, but $Var[E[X_i|Z]] < Var[X_i]$ unless $E[Var[X_i|Z]] = 0$ which means that X_i, given $Z = z$, is a constant for each z. This implies that the random vector $(E[X_1|Z], E[X_2|Z], ..., E[X_n|Z])$ will in general not have $F_1, F_2, ..., F_n$ as its marginal distribution functions. But if the conditioning random variable Z has the property that all random variables $E[X_i|Z]$ are non-increasing functions of Z (or all are non-decreasing functions of Z), the lower bound in Proposition 3 has the form of a sum of n comonotonous random variables. The cdf of this sum is then obtained by the results of Section 2. An application of Proposition 3 in the case of lognormal marginals X_i is considered in Section 5.

With $S = X_1 + X_2 + ... + X_n$, the lower bound S_l in Proposition 3 can be written as $E[S|Z]$. To judge the quality of this stochastic bound, we might look at its variance. To maximize it, the mean value of $Var[S|Z = z]$ should be minimized. Thus, for the best lower bound, Z and S should be as alike as possible.

Let's further assume that the random variable Z is such that all $E[X_i|Z]$ are non-increasing continuous functions of Z. The quantiles of the random variable $E[S|Z]$ then follow from

$$F^{-1}_{E[S|Z]}(p) = \sum_{i=1}^{n} F^{-1}_{E[X_i|Z]}(p) = \sum_{i=1}^{n} E[X_i|Z = F_Z^{-1}(1-p)], \quad p \in (0,1).$$

In order to derive the result above, we used the fact that for a non-increasing continuous function f, we have

$$F_{f(S)}^{-1}(p) = f\left(F_S^{-1}(1-p)\right), \quad p \in (0,1).$$

Similarly, for a non-decreasing continuous function f, we have

$$F_{f(S)}^{-1}(p) = f\left(F_S^{-1}(p)\right), \quad p \in (0,1).$$

If we now in addition assume that the cdf's of the random variables $E[X_i|Z]$ are strictly increasing and continuous, then the cdf of $E[S|Z]$ is also strictly increasing and continuous, and we get for all $x \in \left(F_{E[S|Z]}^{-1\bullet}(0)\,,\ F_{E[S|Z]}^{-1}(1)\right)$,

$$\sum_{i=1}^{n} F_{E[X_i|Z]}^{-1}\left(F_{E[S|Z]}(x)\right) = x,$$

or equivalently,

$$\sum_{i=1}^{n} E\left[X_i|Z = F_Z^{-1}\left(1 - F_{E[S|Z]}(x)\right)\right] = x,$$

which unambiguously determines the cdf of the convex order lower bound $E[S|Z]$ for S in case all $E[X_i|Z=z]$ are non-increasing in z.

The stop-loss premiums of $E[S|Z]$ can be computed as follows:

$$E\left[E[S|Z] - d\right]_+ = \sum_{i=1}^{n} \left\{E[X_i|Z] - E\left[X_i|Z = F_Z^{-1}\left(1 - F_{E[S|Z]}(d)\right)\right]\right\}_+,$$

which holds for all retentions $d \in \left(F_{E[S|Z]}^{-1\bullet}(0)\,,\ F_{E[S|Z]}^{-1}(1)\right)$.

The technique for deriving lower bounds as explained in this section is also considered (for some special cases) in Vyncke, Goovaerts & Dhaene (2000). The idea of this technique stems from mathematical physics, and was applied by Rogers & Shi (1995) to derive approximate values for the value of Asian options.

4.3 Some simple examples

Let X, Y be independent $N(0,1)$ random variables, and consider random variables of the type $Z = X + aY$ for some real a. We want to derive stochastic bounds for $S = X + Y$. The conditional distribution of X, given $Z = z$, is, as is well-known,

$$N\left(\mu_X + \frac{\rho_{X,Z}\,\sigma_X}{\sigma_Z}(z - \mu_Z), \sigma_X^2(1 - \rho_{X,Z}^2)\right) = N\left(\frac{z}{1+a^2}, \frac{a^2}{1+a^2}\right).$$

But this means that for the conditional expectation $E[X|Z]$ and for the random variable $F_{X|Z}^{-1}(U)$, with U uniform(0,1) and independent of Z, we get

$$E[X|Z] = \frac{Z}{1+a^2} \text{ and } F_{X|Z}^{-1}(U) = E[X|Z] + \frac{|a|\Phi^{-1}(U)}{\sqrt{1+a^2}}.$$

In line with $E[X + aY|Z] \equiv Z$, we also get

$$E[Y|Z] = \frac{aZ}{1+a^2} \text{ and } F_{Y|Z}^{-1}(U) = E[Y|Z] + \frac{\Phi^{-1}(U)}{\sqrt{1+a^2}}.$$

It can be shown that both $F_{X|Z}^{-1}(U)$ and $F_{Y|Z}^{-1}(U)$ have $N(0,1)$ distributions. Their U-dependent parts are comonotonous. For the lower and upper bounds derived above we get

$$\begin{aligned}
S = X + Y &\sim N(0,2), \\
S_l = E[X + Y|Z] = \frac{1+a}{1+a^2}Z &\sim N\left(0, \frac{(1+a)^2}{1+a^2}\right), \\
S_u' = \frac{1+a}{1+a^2}Z + \frac{1+|a|}{\sqrt{1+a^2}}\Phi^{-1}(U) &\sim N\left(0, \frac{(1+a)^2 + (1+|a|)^2}{1+a^2}\right), \\
S_u =_d 2X &\sim N(0,4).
\end{aligned}$$

For some special choices of a, we get the following distributions for the lower and upper bounds S_l and S_u':

$$\begin{aligned}
a = 0 &: N(0,1) \leq_{cx} S \leq_{cx} N(0,2), \\
a = 1 &: N(0,2) \leq_{cx} S \leq_{cx} N(0,4), \\
a = -1 &: N(0,0) \leq_{cx} S \leq_{cx} N(0,2), \\
|a| \to \infty &: N(0,1) \leq_{cx} S \leq_{cx} N(0,2).
\end{aligned}$$

Note that the actual distribution of S is $N(0,2)$, so the best convex lower bound ($a=1$) and the best upper bound ($a \leq 0$ or $a \to \infty$) coincide with S. Of course taking $|a| \to \infty$ gives the same results as taking $Z=Y$. The variance of S_l can be seen to have a maximum at $a=+1$, a minimum at $a=-1$. On the other hand, $Var[S_u']$ also has a maximum at $a=1$, and minima at $a \leq 0$ and $a \to \infty$. So the best lower bound in this case is attained for $Z=S$, the worst for Z and S independent. The best improved upper bound is found by taking $Z=X$, $Z=Y$, or any $a<0$, including the case $a=-1$ with Z and S independent; the worst, however, by taking $Z=S$.

To compare the variance of the stochastic upper bound S_u' with the variance of S boils down to comparing $cov\left(F_{X|Z}^{-1}(U), F_{Y|Z}^{-1}(U)\right)$ with $cov\,(X,Y)$. It is clear that, in general, the optimal choice for the conditioning random variable Z will depend on the correlation of X and Y. If this correlation equals 1, any Z results in $S =_d S_u' =_d S$. In our case where X and Y are mutually independent, the optimal choice proves to be taking $Z \equiv X$ or $Z \equiv Y$, thus ensuring that S and S_u' coincide.

As a second example, consider a simple special case of the theory dealt with in the next section. We present it here for the reader's convenience, just as an illustration. Take Y_1 and Y_2 independent $N(0,1)$ random variables. Look at the sum of $X_1 = e^{Y_1} \sim$ lognormal$(0,1)$, and $X_2 = e^{Y_1+Y_2} \sim$ lognormal$(0,2)$. Take $Z = Y_1 + Y_2$. For the lower bound S_l, note that $E[X_2|Z] = e^Z$, while $Y_1|Y_1+Y_2 = z \sim N(\frac{1}{2}z, \frac{1}{2})$, hence

$$E[e^{Y_1}|Y_1+Y_2=z] = m(1; \frac{1}{2}z, \frac{1}{2}),$$

where $m(t;\mu,\sigma^2) = e^{\mu t + \frac{1}{2}\sigma^2 t^2}$ is the $N(\mu,\sigma^2)$ moment generating function. This leads to

$$E[e^{Y_1}|Z] = e^{\frac{1}{2}Z+\frac{1}{4}}.$$

So the lower bound is

$$S_l = E[X_1+X_2|Z] = e^{\frac{1}{2}Z+\frac{1}{4}} + e^Z.$$

Upper bound S_u has $(X_1, X_2) =_d (e^W, e^{\sqrt{2}W})$ for $W \sim N(0,1)$. The improved upper bound S_u' has as a second term again e^Z, and as first term $e^{\frac{1}{2}Z+\frac{1}{2}\sqrt{2}W}$, with Z and W mutually independent. All terms occurring in the bounds given above are lognormal random variables, so the variances of the bounds are easy to compute. Note that to compare variances is meaningful when comparing stop-loss premiums of stop-loss ordered random variables, see, e.g., Kaas et al. (1994, p. 68). The following relation, which can be proven using partial integration, links variances and stop-loss premiums:

$$\frac{1}{2}Var[X] = \int_{-\infty}^{\infty} \{E[X-t]_+ - (E[X]-t)_+\}\,dt,$$

from which we deduce that if $X \leq_{cx} Y$, thus $E[Y-t]_+ \geq E[X-t]_+$ for all t, then

$$\frac{1}{2}\{Var[Y] - Var[X]\} = \int_{-\infty}^{\infty} \{E[Y-t]_+ - E[X-t]_+\}\,dt$$

Thus, half the variance difference between two convex ordered random variables equals the integrated difference of their stop-loss premiums. This implies that if $X \leq_{cx} Y$ and in addition $Var[X] = Var[Y]$, then X and Y must necessary be equal in distribution. Moreover, the ratio of the variances is roughly equal to the ratio of the stop-loss premiums, minus their minimal possible value for random variables with the same mean. We have, as the reader may verify,

$$\begin{aligned}
E[S]^2 &= e^1 + 2e^{\frac{5}{2}} + e^4, \\
E[S_l^2] &= e^{\frac{3}{2}} + 2e^{\frac{5}{2}} + e^4, \\
E[S^2] = E[S_u'^2] &= e^2 + 2e^{\frac{5}{2}} + e^4, \\
E[S_u^2] &= e^2 + 2e^{\frac{3}{2}+\sqrt{2}} + e^4.
\end{aligned}$$

Hence,

$$\begin{aligned}
Var[E[S]] &= 0, \\
Var[S_l] &= 1.763, \\
Var[S] = Var[S_u'] &= 4.671, \\
Var[S_u] &= 17.174.
\end{aligned}$$

So an improved stochastic lower bound S_l for S is obtained by conditioning on Y_1+Y_2, and the improved upper bound $S_u^{'}$ for this case proves to be very good indeed, having in fact the same distribution as S.

5 Present Values - Lognormal Discount Process

5.1 General Result

Consider a series of deterministic payments $\alpha_1, \alpha_2, ..., \alpha_n$, of arbitrary sign, that are due at times $1, 2, ..., n$ respectively. The present value of this series of payments equals:

$$S = \sum_{i=1}^{n} \alpha_i \; e^{-(Y_1+Y_2+...+Y_i)}.$$

Assume that $(Y_1, Y_2, ..., Y_n)$ has a multivariate normal distribution. We introduce the random variables X_i and $Y(i)$ defined by

$$Y(i) = Y_1 + Y_2 + ... + Y_i$$

and $S = X_1 + X_2 + ... + X_n$, where

$$X_i = \alpha_i \; e^{-Y(i)}.$$

For some given choice of the β_i, consider a conditioning random variable Z defined as follows:

$$Z = \sum_{i=1}^{n} \beta_i \; Y_i,$$

For a multivariate normal distribution, every linear function of its components has a univariate normal distribution, so Z is normally distributed. Also, $(Y(i),\ Z)$ has a bivariate normal distribution. Conditionally given $Z = z$, $Y(i)$ has a univariate normal distribution with mean and variance given by .

$$E\left[Y(i)|\ Z = z\right] = E\left[Y(i)\right] + \rho_i \frac{\sigma_{Y(i)}}{\sigma_Z} \left(z - E\left[Z\right]\right)$$

and

$$Var\left[Y(i)|\ Z = z\right] = \sigma^2_{Y(i)}\ \left(1 - \rho_i^2\right)$$

where ρ_i is the correlation between Z and $Y(i)$.

Proposition 6 *Let $S, S_l, S_u^{'}$ and S_u be defined as follows:*

$$S = \sum_{i=1}^{n} \alpha_i \ e^{-(Y_1+Y_2+\ldots+Y_i)},$$

$$S_l = \sum_{i=1}^{n} \alpha_i \ e^{-E[Y(i)]-\rho_i \ \sigma_{Y(i)} \ \Phi^{-1}(U)+\frac{1}{2}\left(1-\rho_i^2\right)\sigma_{Y(i)}^2},$$

$$S_u^{'} = \sum_{i=1}^{n} \alpha_i \ e^{-E[Y(i)]-\rho_i \ \sigma_{Y(i)} \ \Phi^{-1}(U)+sign(\alpha_i) \ \sqrt{1-\rho_i^2}\sigma_{Y(i)} \ \Phi^{-1}(V)},$$

$$S_u = \sum_{i=1}^{n} \alpha_i \ e^{-E[Y(i)]+sign(\alpha_i) \ \sigma_{Y(i)} \ \Phi^{-1}(U)},$$

where U and V are mutually independent uniform(0,1) random variables, and Φ is the cdf of the $N(0,1)$ distribution. Then we have

$$S_l \leq_{cx} S \leq_{cx} S_u^{'} \leq_{cx} S_u.$$

Proof. (a) If a random variable X is lognormal(μ, σ^2), then $E[X] = e^{\mu+\frac{1}{2}\sigma^2}$. Hence, for $Z = \sum_{i=1}^{n} \beta_i \ Y_i$, we find that, taking $U = \Phi\left(\frac{Z-E[Z]}{\sigma_Z}\right)$, so $U \sim$ uniform(0,1),

$$E[X_i|Z] = \alpha_i \ e^{-E[Y(i)]-\rho_i \ \sigma_{Y(i)} \ \Phi^{-1}(U)+\frac{1}{2}\left(1-\rho_i^2\right)\sigma_{Y(i)}^2},$$

From Proposition 3, we find $S_l \leq_{cx} S$.

(b) If a random variable X is lognormal(μ, σ^2), then we have $F_{\alpha X}^{-1}(p) = \alpha \ e^{\mu+sign(\alpha) \ \sigma \ \Phi^{-1}(p)}$. Hence, we find that

$$F_{X_i|Z}^{-1}(p) = \alpha_i \ e^{-E[Y(i)]-\rho_i \ \sigma_{Y(i)} \ \Phi^{-1}(U)+sign(\alpha_i) \ \sqrt{1-\rho_i^2}\sigma_{Y(i)} \ \Phi^{-1}(p)}.$$

From Proposition 2 we find that $S \leq_{cx} S_u^{'}$.

(c) The stochastic inequality $S_u^{'} \leq_{cx} S_u$ follows from Proposition 1. ■

In order to compare the cdf of $S = \sum_{i=1}^{n} \alpha_i \ e^{-(Y_1+Y_2+\ldots+Y_i)}$ with the cdf's of $S_l, S_u^{'}$ and S_u, especially their variances, we need the correlations of the different random variables involved. We find the following results for the lognormal discount process considered in this section:

$$
\begin{aligned}
corr\left[X_i,\ X_j\right] &= \frac{e^{cov[Y(i),Y(j)]}-1}{\sqrt{e^{\sigma^2_{Y(i)}}-1}\sqrt{e^{\sigma^2_{Y(j)}}-1}}; \\
corr\left[E\left(X_i|Z\right),\ E\left(X_j|Z\right)\right] &= \frac{e^{\rho_i.\rho_j\sigma_{Y(i)}\sigma_{Y(j)}}-1}{\sqrt{e^{\rho_i^2\ \sigma^2_{Y(i)}}-1}\sqrt{e^{\rho_j^2\ \sigma^2_{Y(j)}}-1}}; \\
corr\left[F^{-1}_{X_i|Z}(U),\ F^{-1}_{X_j|Z}(U)\right] &= \frac{e^{\left[\rho_i.\rho_j+sign(\alpha_i.\alpha_j)\sqrt{1-\rho_i^2}\sqrt{1-\rho_j^2}\right]\sigma_{Y(i)}\sigma_{Y(j)}}-1}{\sqrt{e^{\sigma^2_{Y(i)}}-1}\sqrt{e^{\sigma^2_{Y(j)}}-1}}; \\
corr\left[F^{-1}_{X_i}(U),\ F^{-1}_{X_j}(U)\right] &= \frac{e^{sign(\alpha_i.\alpha_j)\sigma_{Y(i)}\sigma_{Y(j)}}-1}{\sqrt{e^{\sigma^2_{Y(i)}}-1}\sqrt{e^{\sigma^2_{Y(j)}}-1}}.
\end{aligned}
$$

From these correlations, we can for instance deduce that if all payments α_i are positive and $corr\left[Y(i), Y(j)\right] = 1$ for all i and j, then $S =_d S_u$. In practice, the discount factors will not be perfectly correlated. But for any realistic discount process, $corr\left[Y(i), Y(j)\right] = corr\left[Y_1 + ... + Y_i,\ Y_1 + ... + Y_j\right]$ will be close to 1 provided that i and j are close to each other. This gives an indication that the cdf of S_u might perform well as approximation for the cdf of S for such processes. This is indeed the case in the numerical illustrations in Goovaerts, Dhaene & De Schepper (2000). A similar reasoning leads to the conclusion that the cdf of S_u will not perform well as a convex upper bound for the cdf of S if the payments α_i have mixed signs. This phenomenon will indeed be observed in the numerical illustrations in Section 6.

It remains to derive expressions for the cdf's of S_l, $S_u^{'}$ and S_u.

5.2 The cdf and the stop-loss premiums of S_u

The quantiles of S_u follow from Goovaerts, Dhaene & De Schepper (2000):

$$F^{-1}_{S_u}(p) = \sum_{i=1}^{n} \alpha_i\ e^{-E[Y(i)]+sign(\alpha_i)\ \sigma_{Y(i)}\ \Phi^{-1}(p)},\ p \in (0,1).$$

Also, $F_{S_u}(x)$ follows implicitly from solving

$$\sum_{i=1}^{n} \alpha_i\ e^{-E[Y(i)]+sign(\alpha_i)\ \sigma_{Y(i)}\ \Phi^{-1}(F_{S_u}(x))} = x.$$

It is straightforward to derive expressions for the stop-loss premiums in this case:

$$E\left[S_u - d\right]_+ = \sum_{i=1}^{n} |\alpha_i| \; e^{-E[Y(i)]} E\left[sign(\alpha_i) \; \left(Z_i - e^{sign(\alpha_i)\; \sigma_{Y(i)} \; \Phi^{-1}(F_{S_u}(d))}\right)\right]_+ ,$$

where the Z_i are lognormal$(0,\, \sigma^2_{Y(i)})$ random variables.

In order to derive an explicit expression for the stop-loss premiums $E\left[S_u - d\right]_+$, we first mention the following result, which can easily be proven, e.g. by using $\frac{d}{dt} E\left[X - t\right]_+ = F_X(t) - 1$.

Proposition 7 *If Y is lognormal$(\mu,\, \sigma^2)$, then for any $d > 0$ we have*

$$\begin{aligned} E\left[Y - d\right]_+ &= e^{\mu + \frac{\sigma^2}{2}} \; \Phi(d_1) - d \; \Phi(d_2), \\ E\left[Y - d\right]_- &= e^{\mu + \frac{\sigma^2}{2}} \; \Phi(-d_1) - d \; \Phi(-d_2), \end{aligned}$$

where d_1 and d_2 are determined by

$$d_1 = \frac{\mu + \sigma^2 - \ln(d)}{\sigma}, \quad d_2 = d_1 - \sigma.$$

At $d \leq 0$, the stop-loss premiums are trivially equal to $E[Y] - d$. The following expression results for the stop-loss premiums at $d > 0$:

$$\begin{aligned} E\left[S_u - d\right]_+ &= \sum_{i=1}^{n} \alpha_i \; e^{-E[Y(i)]} . \\ & \left\{ e^{\frac{\sigma^2_{Y(i)}}{2}} \; \Phi\left(sign(\alpha_i)\; d_{i,1}\right) - e^{sign(\alpha_i)\; \sigma_{Y(i)}\; \Phi^{-1}(F_{S_u}(d))} \; \Phi\left(sign(\alpha_i)\; d_{i,2}\right) \right\} \end{aligned}$$

with $d_{i,1}$ and $d_{i,2}$ given by

$$\begin{aligned} d_{i,1} &= \sigma_{Y(i)} - sign(\alpha_i) \quad \Phi^{-1}(F_{S_u}(d)), \\ d_{i,2} &= -sign(\alpha_i) \quad \Phi^{-1}(F_{S_u}(d)). \end{aligned}$$

Using the implicit definition for $F_{S_u}(d)$ leads to the following expression for the stop-loss premiums:

$$\begin{aligned} E\left[S_u - d\right]_+ &= \sum_{i=1}^{n} \alpha_i \; e^{-E[Y(i)] + \frac{\sigma^2_{Y(i)}}{2}} \; \Phi\left[sign(\alpha_i)\; \sigma_{Y(i)} - \Phi^{-1}(F_{S_u}(d))\right] \\ & -d \quad \left(1 - F_{S_u}(d)\right). \end{aligned}$$

5.3 The cdf and the stop-loss premiums of S_l

In general, S_l will not be a sum of n comonotonous random variables. But in the remainder of this subsection, we assume that all $\alpha_i \geq 0$ and all $\rho_i = \frac{cov[Y(i),Z]}{\sigma_{Y(i)}\ \sigma_Z} \geq 0$. These conditions ensure that S_l is the sum of n comonotonous random variables.
Taking into account that $Z = \sum_{i=1}^{n} \beta_i\ Y_i$ is normally distributed, we find that

$$F_Z^{-1}(1-p) = E\left[Z\right] - \sigma_Z\ \Phi^{-1}(p),$$

and hence

$$\begin{aligned} F_{S_l}^{-1}(p) &= \sum_{i=1}^{n} F_{E[X_i|Z]}^{-1}(p) = \sum_{i=1}^{n} E\left[X_i|Z = F_Z(1-p)\right] \\ &= \sum_{i=1}^{n} \alpha_i\ e^{-E[Y(i)]+\rho_i\ \sigma_{Y(i)}\ \Phi^{-1}(p)+\frac{1}{2}\sigma_{Y(i)}^2\ \left(1-\rho_i^2\right)}, \quad p\epsilon\,(0,1)\,. \end{aligned}$$

$F_{S_l}(x)$ can be obtained from

$$\sum_{i=1}^{n} \alpha_i\ e^{-E[Y(i)]+\rho_i\ \sigma_{Y(i)}\ \Phi^{-1}(F_{S_l}(x))+\frac{1}{2}\sigma_{Y(i)}^2\ \left(1-\rho_i^2\right)} = x.$$

We have

$$E\left[S_l - d\right]_+ = \sum_{i=1}^{n} E\left[E\left[X_i|Z\right] - F_{E[X_i|Z]}^{-1}\left(F_{S_l}(d)\right)\right]_+.$$

After some straightforward computations, one finds that an explicit expression for the stop-loss premiums is given by

$$\begin{aligned} E\left[S_l - d\right]_+ &= \sum_{i=1}^{n} \alpha_i\ e^{-E[Y(i)]+\frac{1}{2}\sigma_{Y(i)}^2}\Phi\left[\rho_i\ \sigma_{Y(i)} - \Phi^{-1}\left(F_{S_l}(d)\right)\right] \\ &\quad -d\ \left(1 - F_{S_l}(d)\right). \end{aligned}$$

5.4 The cdf of $S_u^{'}$

Since $F_{S_u^{'}|U=u}$ is a sum of n comonotonous random variables, we have

$$F_{S_u^{'}|U=u}^{-1}(p) = \sum_{i=1}^{n} F_{X_i|U=u}^{-1}(p)$$

$$= \sum_{i=1}^{n} \alpha_i \; e^{-E[Y(i)]-\rho_i \; \sigma_{Y(i)} \; \Phi^{-1}(u)+sign(\alpha_i) \; \sqrt{1-\rho_i^2}\sigma_{Y(i)}\sigma \; \Phi^{-1}(p)}.$$

$F_{S_u'|U=u}$ also follows implicitly from

$$\sum_{i=1}^{n} \alpha_i \; e^{-E[Y(i)]-\rho_i \; \sigma_{Y(i)} \; \Phi^{-1}(u)+sign(\alpha_i) \; \sqrt{1-\rho_i^2}\sigma_{Y(i)} \; \Phi^{-1}(F_{S_u'|U=u}(x))} = x.$$

The cdf of S_u' then follows from

$$F_{S_u'}(x) = \int_0^1 F_{S_u'|U=u}(x) \; du.$$

6 Numerical illustration

In this section, we will numerically illustrate the bounds we derived for $S = \sum_{i=1}^{n} \alpha_i \; e^{-(Y_1+Y_2+...+Y_i)}$. We will take $n = 20$. In order to be able to compare the distribution functions of the stochastic bounds S_l, S_u' and S_u with the distribution function of S, we will completely specify the multivariate distribution function of the random vector $(Y_1, Y_2, ..., Y_{20})$. In particular, we will assume that the random variables Y_i are i.i.d. and $N(\mu, \sigma^2)$. This will enable us to simulate the cdf's in case there is no way to compute them analytically. The conditioning random variable Z is defined as before:

$$Z = \sum_{i=1}^{20} \beta_i \; Y_i,$$

In this case, we find

$$\begin{aligned} E\left[Y(i)\right] &= i \; \mu, \\ Var\left[Y(i)\right] &= i \; \sigma^2, \\ Var\left[Z\right] &= \sigma^2 \sum_{k=1}^{20} \beta_k^2, \\ \rho_i = \frac{cov\left[Y(i), Z\right]}{\sigma_{Y(i)} \; \sigma_Z} &= \frac{\sum_{k=1}^{i} \beta_k}{\sqrt{i \sum_{k=1}^{20} \beta_k^2}}. \end{aligned}$$

In our numerical illustrations, we will choose the parameters of the normal distribution involved as follows:

$$\mu = 0.07; \ \sigma = 0.1.$$

We will compute the lower bound and the upper bounds for the following choice of the parameters β_i

$$\beta_i = \sum_{j=i}^{20} \alpha_j \, e^{-j\mu}, \quad i = 1, ..., 20.$$

By this choice, the lower bound will perform well in these cases. This is due to the fact that this choice makes Z a linear transformation of a first order approximation to S. This can be seen from the following computation, which depends on σ, and hence $Y_i - \mu$, being "small":

$$\begin{aligned} S &= \sum_{j=1}^{n} \alpha_j e^{-j\mu - \sum_{i=1}^{j}(Y_i-\mu)} \\ &\approx \sum_{j=1}^{n} \alpha_j e^{-j\mu}[1 - \sum_{i=1}^{j}(Y_i - \mu)] \\ &= C - \sum_{j=1}^{n} \alpha_j e^{-j\mu} \sum_{i=1}^{j} Y_i \\ &= C - \sum_{i=1}^{n} Y_i \sum_{j=i}^{n} \alpha_j e^{-j\mu}, \end{aligned}$$

where C is the appropriate constant. By the remarks in section 4, S_l will then be "close" to S.

Figure 1 shows the cdf's of $S, S_l, S_u^{'}$ and S_u for the following payments:

$$\alpha_k = 1, \quad k = 1, ..., 20.$$

Since $S_l \leq_{cx} S \leq_{cx} S_u^{'} \leq_{cx} S_u$, and the same ordering holds for the tails of their distribution functions which can be observed to cross only once, we can easily identify the cdf's. We see that the cdf of S_l is very close to the distribution of S, which was expected because of the choice of Z. Note that in this case S_l is a sum of comonotonous random variables, so its quantiles can be computed easily. The cdf of S_u also performs rather well, as was observed in Goovaerts, Dhaene & De Schepper (2000). We find that the improved upper bound S_u' is very close to the comonotonous upper bound S_u. This is due to the fact that $cov\left(F^{-1}_{X_i|Z}(U), F^{-1}_{X_j|Z}(U)\right)$ is close to $cov\,(X_i, X_j)$ for any pair (i, j) with $i \neq j$.

Figure 2 shows the cdf's of S, S_l, S_u' and S_u for the following payments:

$$\alpha_k = \begin{cases} -1, & k = 1, ..., 5, \\ 1 & k = 6, ..., 20. \end{cases}$$

Note that the cdf of the lower bound S_l cannot be computed exactly in this case; it is obtained by simulation. In this case, we see that the lower bound S_l still performs very well. The comonotonous upper bound S_u performs very badly in this case, as was to be expected from the observations in Section 5.1. The improved upper bound performs better.

In Figure 3, we consider the same series of payments as in Figure 2. We consider the cdf of the improved upper bound for a different choice of the conditioning random variable Z. We choose Z such that it is an approximation to the discounted total of the 5 negative payments:

$$\beta_i = \begin{cases} \sum_{j=i}^{5} \alpha_j \, e^{-j\mu}, & i = 1, ..., 5, \\ 0 & i = 6, ..., 20. \end{cases}$$

The (simulated) cdf of S is the dotted line. Note that the upper bound S_u' is much improved, the lower bound is worse.

7 Conclusions and related research

In this contribution we considered the problem of deriving stochastic lower and upper bounds, in the sense of convex order, for a sum $S = X_1 + X_2 + ... + X_n$ of possibly dependent random variables $X_1, X_2, ..., X_n$. We assumed that, as is often the case, the marginal distribution of each random variable $X_1, X_2, ..., X_n$ is known. The problem of deriving a convex upper bound without using additional information about the dependency structure was considered in Müller (1997) and Goovaerts, Dhaene & De Schepper (2000). In this paper, we additionally assumed that there exists some random variable Z, with a computable distribution, such that for any i and for any z in the support of Z, the conditional distribution function of X_i, given $Z = z$, is also computable. Based on this, we derived random variables S_l and $S_u^{'}$, the cdf's of which are known to be less and larger than the one of S in convex order, meaning that the tails of S_l are thinner, the ones of $S_u^{'}$ are thicker in general. Though it is not guaranteed that two convex ordered cdf's cross only once, in the majority of examples they do so. Thus, we obtain a band of possible values of $\Pr[S \leq x]$ which might provide more, and more reliable, information than a point estimate as obtained from a number of simulations. This is especially the case when the inverse cdf is sought, such as when one wants to determine fair values and supervisory values. But note that $\Pr[S \leq x]$ cannot be guaranteed to be between $\Pr[S_l \leq x]$ and $\Pr[S_u^{'} \leq x]$. It has been argued before, see e.g. Kaas (1994), that actuaries should not be focused on probabilities and quantiles, but rather on stop-loss premiums, since it is not the probability of exceeding a threshold d that matters, but the amount by which this happens, of which the expected value is just the stop-loss premium at d. And for stop-loss premiums, the property $E[S_l - d]_+ \leq E[S - d]_+ \leq E[S_u^{'} - d]_+$ does hold.

It should be noted that the upper bound $S_u^{'}$ is no longer a supremum (in the sense of convex order) over the set of all random vectors with fixed marginals, and that the lower bound S_l is not a sum of terms with the proper marginal distributions. This follows from the fact that the bounds that we derived take into account the dependency structure of the random vector under consideration.

It should also be noted that our results actually do not require the complete dependency structure, but only the distribution of Z and the conditional distributions of X_i given $Z = z$. In the numerical illustration section we chose an example where the distribution of the random vector was completely known, in order to be able to compare the bounds with the (simulated) exact cdf.

A topic for future research is the determination of the optimal conditioning random variable Z for the improved upper bound $S_u^{'}$, in the spirit of the remarks made at the end of section 4.3. Another item for future research is the extension of the results of this paper to the case where also the cash flows are stochastic, hence to find improved upper bounds and lower bounds for $S = X_1Y_1 + X_2Y_2 + \cdots + X_nY_n$. Another idea that we intend to pursue is conditioning on more than one random variable Z.

8 Acknowledgments

We would like to thank Etienne De Vylder for helpful and insightful comments on an earlier version of this paper, and David Vyncke for providing the numerical illustrations of Section 6. Marc Goovaerts and Jan Dhaene acknowledge the financial support of the Onderzoeksfonds K.U. Leuven (grant OT/97/6) and of FWO (grant "Actuarial ordering of dependent risks").

REFERENCES

Bäuerle, N.; Müller, A. (1998). "Modeling and comparing dependencies in multivariate risk portfolios", *ASTIN Bulletin* 28, 59-76.

Brockett, P.; Garven, J. (1998). "A reexamination of the relationship between preferences and moment orderings by rational risk averse investors", *Geneva Papers on Risk and Insurance Theory*, 23, 127-137.

Denneberg, D. (1994). "Non-additive measure and integral", Kluwer Academic Publishers, Boston, pp. 184.

Denuit, M.; Lefèvre, C. (1997). "Stochastic product orderings, with applications in actuarial sciences, *Bulletin Français d'Actuariat* 1, 61-82.

Denuit, M.; De Vylder, F.; Lefèvre, C. (1999). "Extrema with respect to s-convex orderings in moment spaces: a general solution", *Insurance: Mathematics & Economics* 24, 201-217.

Dhaene, J.; Denuit, M. (1999). "The safest dependency structure among risks", *Insurance: Mathematics & Economics* 25, 11-21.

Dhaene, J.; Goovaerts, M. (1996). "Dependency of risks and stop-loss order", *ASTIN Bulletin* 26, 201-212.

Dhaene, J.; Goovaerts, M.J. (1997). "On the dependency of risks in the individual life model", *Insurance: Mathematics & Economics* 19, 243-253.

Dhaene, J.; Wang, S.; Young, V.; Goovaerts, M.J. (1998). "Comonotonicity and maximal stop-loss premiums", D.T.E.W., K.U. Leuven, Research Report 9730, pp. 13, submitted.

Goovaerts, M.J.; Dhaene, J. De Schepper, A. (2000). "Stochastic Upper Bounds for Present Value Functions", *Journal of Risk and Insurance Theory*, 67.1, 1-14.

Goovaerts, M.J.; Redant, R. (1999). "On the distribution of IBNR reserves", *Insurance: Mathematics & Economics* 25, 1-9.

Kaas, R., Van Heerwaarden, A.E., Goovaerts, M.J. (1994). "Ordering of actuarial risks", Institute for Actuarial Science and Econometrics, University of Amsterdam, Amsterdam, pp. 144.

Kaas, R. (1994). "How to (and how not to) compute stop-loss premiums in practice", *Insurance: Mathematics & Economics*, 13, 241-254.

Müller, A. (1997). "Stop-loss order for portfolios of dependent risks", *Insurance: Mathematics & Economics*, 21, 219-223.

Rogers, L.C.G.; Shi, Z. (1995). "The value of an Asian option", *Journal of Applied Probability* 32, 1077-1088.

Shaked, M.; Shanthikumar, J.G. (1994). "Stochastic orders and their applications", Academic Press, pp. 545.

Simon, S.; Goovaerts, M.J.; Dhaene, J. (2000). "An easy computable upper bound for the price of an arithmetic Asian option", *Insurance: Mathematics & Economics*, 26.2-3, 175-184.

Vyncke, D.; Goovaerts, M.J.; Dhaene, J. (2000): "Convex upper and lower bounds for present value functions", submitted.

Wang, S.; Dhaene, J. (1998). "Comonotonicity, correlation order and stop-loss premiums", *Insurance: Mathematics & Economics*, 22(3), 235-243.

Wang, S.; Young, V. (1998). Ordering risks: expected utility versus Yaari's dual theory of choice under risk. *Insurance: Mathematics & Economics* 22, 145-162.

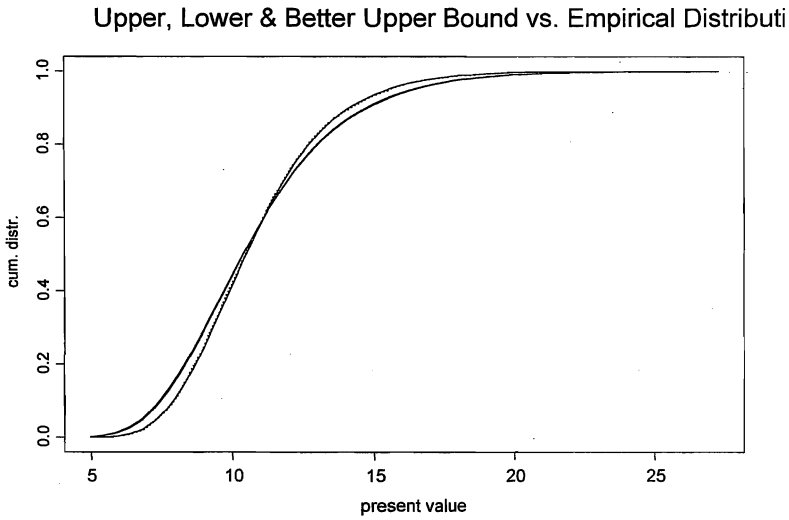

Figure 1: Payments: 20 x 1; Z such that the lower bound is optimized.

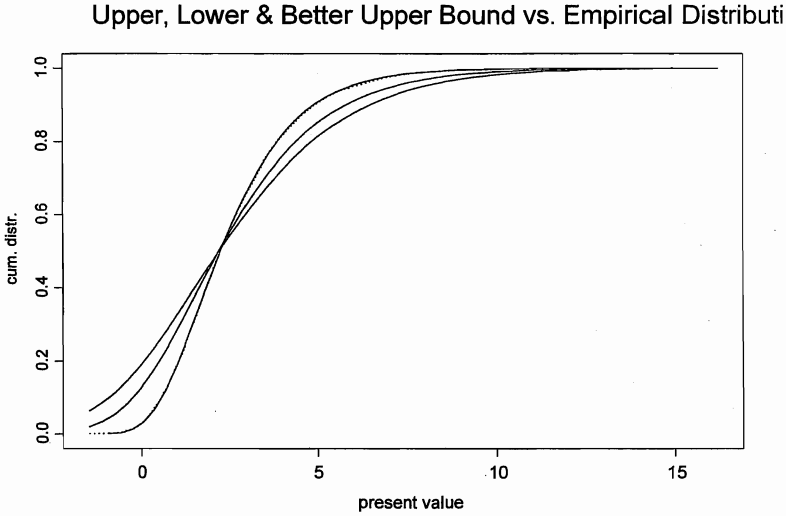

Figure 2: Payments: 5 x (-1), 15 x 1; Z such that the lower bound is optimized.

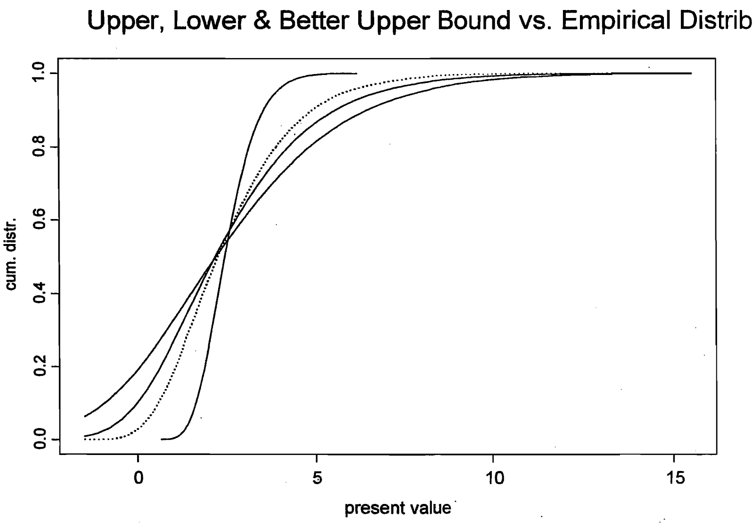

Figure 3: Payments: 5 x (-1), 15 x 1; Z is such that it is an approximation to the discounted total of the negative payments.