Coherent Distortion Risk Measures - A Pitfall

Grzegorz Darkiewicz *† Jan Dhaene † Marc Goovaerts †

July 15, 2003

Abstract

Concave distortion risk measures were introduced in the actuarial literature by Wang (1996). Loosely speaking, such a risk measure assigns a "distorted expectation" to any distribution function. The expectation is distorted by a so-called "distortion function". Concavity of the distortion function ensures that the risk measure preserves stop-loss order.

Consider two random couples with identical marginal distributions but of which the dependency structure differs. Assume that the covariance of the second couple exceeds the covariance of the first one. Let us now consider a risk measure for the sum of the components of each couple. One would expect that any reasonable risk measure will lead to a smaller real number for the sum of the components of the first couple. However, we will demonstrate that this property does not hold in general for concave distortion risk measures. Moreover, for any such risk measure, it is possible to construct an example where the correlation order is not preserved.

Despite this theoretical result, some simulation-based testing indicates that most well-known concave distortion risk measures for sums of random couples with given marginals frequently do preserve the order of the correlations.

Keywords: Distortion risk measures, orderings of risks, premium principles, comonotonicity.

*Corresponding author: grzegorz.darkiewicz@econ.kuleuven.ac.be

†Faculty of Economics and Applied Economics, K.U.Leuven

1 Introduction

Distortion risk measures were introduced in the actuarial literature by Wang (1996). For a given nondecreasing function $g : [0,1] \to [0,1]$ such that $g(0) = 0$ and $g(1) = 1$, the distorted expectation $H_g(X)$ of any nonnegative random variable X is defined as follows:

$$H_g(X) = \int_0^\infty g(1 - F_X(t))dt = \int_0^1 F_X^{-1}(1-q)dg(q), \tag{1}$$

where $F_X(t)$ denotes the cumulative distribution function of X. The function g is called a distortion function. Distortion risk measures have several properties, such as positive homogeneity (PH), translation invariance (TI), monotonicity (M) and additivity for comonotonic risks (CA). In insurance applications one often assumes the distortion functions g to be concave, which implies that the risk measure $H_g(\cdot)$ is subadditive (SA). The above mentioned properties of concave distortion risk measures mean that they are "coherent" in the sense of Artzner *et al.* (1999).

Definition 1 *Let (X_1, Y_1) and (X_2, Y_2) be elements of $R(F_X, F_Y)$ (i.e. both random couples have marginal distributions equal to F_X and F_Y respectively). Then we say that (X_1, Y_1) precedes (X_2, Y_2) in the correlation order sense, notation $(X_1, Y_1) \leq_{corr} (X_2, Y_2)$, if either of the following equivalent conditions holds:*
(a) For all non-decreasing functions f and g, we have that

$$\text{Cov}(f(X_1), g(Y_1)) \leq \text{Cov}(f(X_2), g(Y_2)),$$

provided the covariance functions exist.
(b) For any non-negative real numbers x and y, we have that

$$F_{(X_1,Y_1)}(x,y) \leq F_{(X_2,Y_2)}(x,y).$$

Consider two bivariate random variables (X_1, Y_1) and (X_2, Y_2) being elements of $R(F_X, F_Y)$ and such that (X_1, Y_1) precedes (X_2, Y_2) in the correlation order sense. Obviously in this case, we consider the sum $S_1 = X_1 + Y_1$ as "less risky" than $S_2 = X_2 + Y_2$ (see [22])).
The following theorem states that distortion risk measures order the riskiness of the two sums properly.

Theorem 1 *Let g be a nondecreasing concave function, such that $g(0) = 0$ and $g(1) = 1$. Let (X_1, Y_1) and (X_2, Y_2) be elements of $R(F_X, F_Y)$, such that*

$$(X_1, Y_1) \leq_{corr} (X_2, Y_2).$$

Then we have that for $S_1 = X_1 + Y_1$ and $S_2 = X_2 + Y_2$ the following inequality holds:

$$H_g(S_1) \leq H_g(S_2).$$

The proof for this result can be found in Wang & Dhaene (1998). In fact, they prove that S_1 precedes S_2 in the stop-loss order sense, which immediately implies the stated result. All these properties make that the class of distortion risk measures is often seen as a very powerful class of premium principles or, more generally, of risk measures. However, we must realize that correlation order is only a partial order. In order to compare the riskiness of the sums S_i of bivariate random variables $(X_i, Y_i) \in R(F_X, F_Y)$ also in the case these couples are not ordered in the correlation sense, we will look at the ordering of the variances $\mathrm{Var}(S_i)$, or, equivalently, of the correlation coefficients $\mathrm{corr}(X_i, Y_i)$ of the summands. We will investigate if (concave) distortion risk measures preserve this ordering.

Example 1 The concave distortion risk measure with distortion function given by $g_p(x) = \min\left(\frac{x}{p}, 1\right)$ for a value of p in $(0, 1)$ is called Tail Value-at-Risk at level p, which is denoted by $\mathrm{TVaR}_p(X)$. In this example we will illustrate that $\mathrm{corr}(X_1, Y_1) \leq \mathrm{corr}(X_2, Y_2)$ does not always imply that $\mathrm{TVaR}_p(S_1) \leq \mathrm{TVaR}_p(S_2)$.
Let X and Y be two random variables with probabilities $\Pr(X = i) = p_i$ and $\Pr(Y = i) = q_i$ given by:

$$p_0 = p_1 = \frac{1 - \sqrt{p}}{2}, \quad p_2 = \sqrt{p} \tag{2}$$

and

$$q_0 = 1 - \sqrt{p}, \quad q_1 = \sqrt{p}. \tag{3}$$

Now let (X_1, Y_1) and (X_2, Y_2) be two elements of $R(F_X, F_Y)$. Concerning the dependency structure of the couples, we assume that X_1 and Y_1 are mutually independent, while the distribution of (X_2, Y_2) is given in the Table 1

	X_2		
Y_2	0	1	2
0	$p_0q_0 + 2\varepsilon$	$p_1q_0 - 3\varepsilon$	$p_2q_0 + \varepsilon$
1	$p_0q_1 - 2\varepsilon$	$p_1q_1 + 3\varepsilon$	$p_2q_1 - \varepsilon$

Table 1: The distribution of (X_2, Y_2).

In this definition ε denotes an arbitrary positive number satisfying the condition

$$\varepsilon \leq \min\left(\frac{p_0q_1}{2}, \frac{p_1q_0}{3}, p_2q_1\right). \tag{4}$$

One immediately can verify that $(X_2, Y_2) \in R(F_X, F_Y)$. Furthermore, $\mathrm{Cov}(X_2, Y_2) = \varepsilon$, and hence

$$0 = \mathrm{corr}(X_1, Y_1) < \mathrm{corr}(X_2, Y_2) = \varepsilon.$$

For the decumulative distribution functions of the sums $S_i = X_i + Y_i$ we find:

$$\overline{F}_{S_1}(t) = \begin{cases} 1 & \text{for } t < 0, \\ p + \upsilon + \vartheta & \text{for } 0 \leq t < 1, \\ p + \upsilon & \text{for } 1 \leq t < 2, \\ p & \text{for } 2 \leq t < 3, \\ 0 & \text{for } t \geq 3. \end{cases}$$

and

$$\overline{F}_{S_2}(t) = \begin{cases} 1 & \text{for } t < 0, \\ p + \upsilon + \vartheta - 2\varepsilon & \text{for } 0 \leq t < 1, \\ p + \upsilon + 3\varepsilon & \text{for } 1 \leq t < 2, \\ p - \varepsilon & \text{for } 2 \leq t < 3, \\ 0 & \text{for } t < 0. \end{cases}$$

(for simplicity of notation we denote $\Pr(S_1 = 2)$ by υ and $\Pr(S_1 = 1)$ by ϑ).
From (1) we find that $\text{TVaR}_p(S_1)$ and $\text{TVaR}_p(S_2)$ are given by:

$$\text{TVaR}_p(S_1) = g_p(p + \upsilon + \vartheta) + g_p(p + \upsilon) + g_p(p) = 3, \tag{5}$$

and

$$\text{TVaR}_p(S_2) = g_p(p + \upsilon + \vartheta - 2\varepsilon) + g_p(p + \upsilon + 3\varepsilon) + g_p(p - \varepsilon) = 3 - \frac{\varepsilon}{p}. \tag{6}$$

Hence, we can conclude that TVaR does not always preserve the order induced by the correlation between the components of random couples.

■

In Section 2 we present the general result of this paper which states that there does not exist any concave distortion function g such that the corresponding concave distortion risk measure H_g preserves the correlation coefficient of summands in all cases. In Section 3 we investigate for some well-known distortion risk measures how likely it is that they preserve the desired order. Section 4 concludes the paper. The proof for our main result can be found in the Appendix.

2 Construction of a general counterexample

In the following theorem, we will state our main result. We will prove the result for distortion functions $g : [0,1] \to [0,1]$, which satisfy following conditions:
(i) $g(0) = 0$ and $g(1) = 1$,
(ii) g is nondecreasing and piecewise continuous,
(iii) g is piecewise continuously differentiable and $g'(p)$ is nonincreasing,
(iv) $\exists_{p_0}\ g(p_0) \neq p_0$.

Theorem 2 *Let g be an arbitrary function satisfying conditions (i)-(iv). Then there exist univariate distributions $F_{X^{(g)}}$, $F_{Y^{(g)}}$ and random couples $\left(X_1^{(g)}, Y_1^{(g)}\right)$, $\left(X_2^{(g)}, Y_2^{(g)}\right)$ belonging to $R(F_{X^{(g)}}, F_{Y^{(g)}})$ such that*
(i) $\operatorname{corr}\left(X_1^{(g)}, Y_1^{(g)}\right) < \operatorname{corr}\left(X_2^{(g)}, Y_2^{(g)}\right)$,
(ii) $H_g\left(X_1^{(g)} + Y_1^{(g)}\right) > H_g\left(X_2^{(g)} + Y_2^{(g)}\right)$.

PROOF. In the Appendix. ■

Obviously if g satisfies conditions (i)-(iii), then g is a concave distortion function. The theorem can be generalized to general concave distortion functions by using appropriate limit theorems. We skip this rather technical proof, because all distortion functions encountered in practical applications do normally satisfy conditions (i)-(iii).
The condition (iv) excludes the case where $g(p) = p$ for all $p \in (0, 1)$. The distortion risk measure related to this distortion function is the expectation, for which the result of the theorem obviously cannot hold.
Intuitively, it is clear that the assumption of concavity of g is somehow critical. If it was released, it should be easier to find counterexamples. However in the proof we use this assumption explicitly. In fact, when one releases the assumption of concavity, the situation is much easier in view of the following representation theorem, originally proved by Schmeidler (1986).

Theorem 3 *Let BV be a set of bounded random variables. If the functional $H : BV \to [0, \infty)$*
(i) is additive for comonotonic risks,
(ii) preserves first stochastic dominance (i.e. $\forall_t\, F_X(t) \le F_Y(t) \Rightarrow H(X) \le H(Y)$),
(iii) satisfies H(1)=1,
then there exists a distortion function h such that $H(X) = H_h(X)$ for all $X \in BV$. Moreover, $H(X + Y) \le H(X) + H(Y)$ holds for all X, $Y \in BV$ if and only if h is concave.

PROOF. See Dennenberg (1994). ■

Consider a distortion risk measure H_g generated by the distortion function g which is not concave. Clearly, H_g obeys (i), (ii) and (iii) in the theorem above, so that we find the following corrolary.

Corollary 1 *Let H_g denote a distortion risk measure generated by the distortion function g which is not concave. Then there exists a bivariate random variable (X, Y) with bounded marginals such that $H_g(X + Y) > H_g(X) + H_g(Y)$.*

Now it is straightforward to prove the general theorem.

Theorem 4 *Let g be an arbitrary distortion function, piecewise continuously differentiable. Then there exist univariate distributions $F_{X^{(g)}}$, $F_{Y^{(g)}}$ and bivariate distributions $\left(X_1^{(g)}, Y_1^{(g)}\right)$, $\left(X_2^{(g)}, Y_2^{(g)}\right)$ belonging to $R(F_{X^{(g)}}, F_{Y^{(g)}})$ such that*
(i) $\operatorname{corr}\left(X_1^{(g)}, Y_1^{(g)}\right) < \operatorname{corr}\left(X_2^{(g)}, Y_2^{(g)}\right)$,
(ii) $H_g\left(X_1^{(g)} + Y_1^{(g)}\right) > H_g\left(X_2^{(g)} + Y_2^{(g)}\right)$.

PROOF. If g is concave the conclusion follows immediately from Theorem 2.
If g is not concave, we find from Corollary 1 that there exists a random couple (X, Y) such that

$$H_g(X + Y) > H_g(X) + H_g(Y). \tag{7}$$

On the other hand, for the couple (X^c, Y^c) with the same marginals as the couple (X, Y), but with the comonotonic dependency structure, we have that

$$H_g(X^c + Y^c) = H_g(X) + H_g(Y). \tag{8}$$

Combining (7) with (8), one gets

$$H_g(X + Y) > H_g(X^c + Y^c). \tag{9}$$

On the other hand we have that $\operatorname{Var}(X + Y) < \operatorname{Var}(X^c + Y^c)$ and thus $\operatorname{corr}(X, Y) < \operatorname{corr}(X^c, Y^c)$ (see Dhaene *et al.* (2002)).

Hence, taking $\left(X_1^{(g)}, Y_1^{(g)}\right) = (X, Y)$ and $\left(X_2^{(g)}, Y_2^{(g)}\right) = (X^c, Y^c)$, we find the desired result. ∎

Remark 1 *The assumption of piecewise continuous differentiability of g is used only for consistency with assumptions of Theorem 2. In fact it is not used for functions not being concave. As mentioned before, for concave functions this assumption can also be released easily.*

3 Distortion risk measures and correlation coefficient - an experimental test of consistency

In this section we describe a simple methodology to test the consistency of distortion risk measures of sums of random variables with the order induced by the correlation coefficient between the marginals. We want to emphasize that the test presented here is just a first attempt to test this form of consistency.

3.1 Description of the experiment

First, we will select 50,000 couples (X_{1k}, Y_{1k}) in the class of bivariate random variables with support $\{(i,j) \mid i,j = 0,\ldots,9\}$. For each of the selected couples, we will also consider a random couple (X_{2k}, Y_{2k}) with the same marginals as (X_{1k}, Y_{1k}), but of which X_{2k} and Y_{2k} are mutually independent. Finally, we will check how many of these couples (X_{1k}, Y_{1k}) and (X_{2k}, Y_{2k}) satisfy the following relation:

$$\text{sign}\Big(\text{corr}(X_{1k}, Y_{1k}) - \text{corr}(X_{2k}, Y_{2k})\Big) = \text{sign}\Big(H_g(X_{1k}+Y_{1k}) - H_g(X_{2k}+Y_{2k})\Big). \tag{10}$$

In order to select (the distribution function of) the couple (X_{1k}, Y_{1k}), we start by generating 99 random numbers U_{ik} in the interval $(0,1)$. Let

$$V_0 = 0,$$

$$V_i = U_i' \text{ for } i = 1,\ldots,99,$$

$$V_{100} = 1, \tag{11}$$

where U_i' denotes i-th order statistic of the sequence $\{U_i\}$. Next, we consider the differences $a_{ik} = V_i - V_{i-1}$ for $i = 1,\ldots,100$. In this way, we get 100 identically distributed random numbers such that

$$a_1 + \ldots + a_{100} = 1. \tag{12}$$

Now we define the probability distribution of (X_{1k}, Y_{1k}) as follows:

$$\Pr(X_{1k} = i, Y_{1k} = j) = a_{i+1+10j}. \tag{13}$$

The marginal distributions of X_{1k} and Y_{1k} are given by $P(X_{1k} = i) = \sum_{j=0}^{9} a_{i+1+10j}$ and $P(Y_{1k} = j) = \sum_{i=0}^{9} a_{i+1+10j}$.
The related random couple (X_{2k}, Y_{2k}) is defined as the independent counterpart of (X_{1k}, Y_{1k}), hence

$$P(X_{2k} = i, Y_{2k} = j) = P(X_{1k} = i)\ P(Y_{1k} = j). \tag{14}$$

Finally, we compute $\Big(\text{corr}(X_{1k}, Y_{1k}), \text{corr}(X_{2k}, Y_{2k})\Big)$ and $\Big(H_g(X_{1k}+Y_{1k})\ ,\ H_g(X_{2k}+Y_{2k})\Big)$ and then verify whether the equation (10) is satisfied.
This procedure is repeated for every $k = 1,\ldots,50000$.
Then, for any particular choice of the distortion risk measure g we determine the frequency

$$r_g = \frac{N_g}{50,000}, \tag{15}$$

with N_g defined by

$$N_g = \#\Big\{\big((X_{1k}, Y_{1k}), (X_{2k}, Y_{2k})\big) \mid (10) \text{ holds}\Big\}, \tag{16}$$

We will call r_g the correlation consistency coefficient of the risk measure H_g, for the particular set of constructed bivariate distributions.

3.2 Results of the test and conclusions

3.2.1 Description of tested one-parameter families of distortion functions

We have performed the procedure described above for the following one-parameter families of distortion functions. Note that although Value-at-Risk is a nonconcave distortion risk measure, we have included it because of its importance in practice. Most of these distortion risk measures were introduced in Wang (1996). For each family the parameter p satisfies $p \in (0, 1)$.

- **Value at Risk:** $g_p(x) = \mathbf{1}_{(\mathbf{p},\mathbf{1}]}(\mathbf{x})$
- **Tail Value at Risk:** $g_p(x) = \min\left(\frac{x}{p}, 1\right)$
- **Proportional hazard transform:** $g_p(x) = x^p$
- **Dual-power transform:** $g_p(x) = 1 - (1 - x)^{\frac{1}{p}}$
- **Dennensberg's absolute deviation principle:**

$$g_p(x) = \begin{cases} (1+p)x & \text{for } 0 \le x \le \frac{1}{2} \\ p + (1-p)x & \text{for } \frac{1}{2} \le x \le 1 \end{cases}$$

- **Gini's principle:** $g_p(x) = (1 + p)x - px^2$
- **Square-root transform:** $g_p(x) = \frac{\sqrt{1-\ln(p)x}-1}{\sqrt{1-\ln(p)}-1}$
- **Exponential transform:** $g_p(x) = \frac{1-p^x}{1-p}$
- **Logarithmic transform:** $g_p(x) = \frac{\ln(1-\ln(p)x)}{\ln(1-\ln(p))}$

	Parameter p						
Risk measure	0.01	0.1	0.25	0.5	0.75	0.9	0.99
Value at Risk	84.18%	92.87%	94.28%	88.95%	75.07%	69.08%	74.51%
Tail Value at Risk	66.83%	71.25%	82.38%	89.68%	82.19%	70.93%	58.65%
PH transform	70.11%	71.72%	74.78%	80.55%	85.65%	88.09%	89.49%
Dual-power	60.01%	77.98%	89.37%	96.90%	93.55%	90.96%	89.66%
Dennenberg	89.68%	89.68%	89.68%	89.68%	89.68%	89.68%	89.68%
Gini	96.90%	96.90%	96.90%	96.90%	96.90%	96.90%	96.90%
Square-root	92.11%	94.14%	95.28%	96.22%	96.78%	96.88%	96.89%
Exponential	87.11%	92.57%	94.91%	96.32%	96.82%	96.88%	96.90%
Logarithmical	89.63%	92.33%	94.17%	95.76%	96.64%	96.89%	96.89%

Table 2: The results of the correlation consistency tests for different families of distortion risk measures.

3.2.2 Results

In the Table 2 we show the correlation consistency coefficient r_g for different distortion functions g.

From this table we can draw the overall conclusion, that the correlation coefficient is preserved in the majority of cases, for many tested distortion risk measures more frequently than nine times out of ten, for some of them even more than than nineteen times out of twenty. Favorite risk measures, such as Value-at-Risk, Tail Value-at-Risk and Proportional Hazard do not perform very well. We also observe that correlation consistency differs not only between different families of distortion risk measures, but also between different parameters within the same family. In this respect, the dispersion of the correlation consistency seems to be the worst for the Dual-power transform.

Risk measures such as the square root transform, the exponential transform and logarithmical transform perform very well. For these distortion risk measures, the correlation consistency coefficient does not seem to be very dispersed and tends to increase monotonically together with the parameter p.

From the table, it seems that the Dennenberg principle and the Gini principle have a very stable correlation consistency coefficient. In our test this coefficient is even identical for all parameters p. This is not accidental, because both risk measures can be expressed as the sum of the expectation and a summand proportional to some dispersion measures independent from the parameter p. Interested readers are referred to [2].

The table leads to the conclusion that Gini's principle performs the best in terms of correlation consistency.

4 Summary

The most well-know measure of riskiness of a random variable is without doubt its variance. Obviously the variance cannot be viewed as the universal risk measure, because it cannot detect sources of uncertainty such as skewness and heavy tails in an adequate way. However, when we consider random couples (X_1, Y_1) and (X_2, Y_2), both elements of $R(F_X, F_Y)$, relative riskiness of the sums $S_i = X_i + Y_i$ results from dependency structure between the summands, and then the variance seems to be reasonable for the purpose of risk ordering.
In this paper we investigated if the risk ordering generated by comparing the variances of the sums S_i (or equivalently the correlation coefficient between the summands) is preserved by (concave) distortion risk measures. We found that for each distortion risk measure one can find random couples for which the order is not preserved. In the last section we tested the consistency between both orderings. We found that for some distortion transforms the consistency seems to be very high (especially for Gini's risk measure), but that the consistency varies significantly between different risk measures.

5 Acknowledgements

Grzegorz Darkiewicz, Jan Dhaene and Marc Goovaerts acknowledge the financial support of the Onderzoeksfonds K.U.Leuven (GOA/02: Actuariële, financiële en statistische aspecten van afhankelijkheden in verzekerings- en financiële portefeuilles).

Appendix

In this Appendix we give a proof for Theorem 2. We start with proving a result which will turn out to be helpful for the proof of the theorem.

Lemma 1 *Let g be an arbitrary function satisfying conditions (i)-(iv). Then there exist real numbers $p_1 < p_2$ in $(0,1)$ such that $g'(p_1) > g'(p_2)$ and $g'(p_1) + 2g'_-(1) > 3g'(p_2)$.*

Proof. Note that g' cannot be constant in view of condition (iv), and thus, from the concavity and nondecreasingness of g, one can find that there exist a $p_1 \in (0,1)$ satisfying

$$g'(p_1) > g'_-(1) > 0. \tag{17}$$

This leads to

$$3g'_-(1) < g'(p_1) + 2g'_-(1). \tag{18}$$

Define $\varepsilon = \frac{g'(p_1) - g'_-(1)}{3} > 0$. The left continuity of g' in 1 implies that we can choose a point p_2 such that

$$g'(p_2) - g'_-(1) < \varepsilon. \tag{19}$$

Now we easily find that

$$3g'(p_2) < g'(p_1) + 2g'_-(1). \tag{20}$$

Moreover, also the following inequality holds:

$$g'(p_2) - g'_-(1) < \varepsilon = \frac{g'(p_1) - g'_-(1)}{3} < g'(p_1) - g'_-(1), \tag{21}$$

and hence,

$$g'(p_1) > g'(p_2), \tag{22}$$

which completes the proof of Lemma 1. ■

Now we are able to prove Theorem 2.

Proof of Theorem 2. Consider two points $0 < p_1 < p_2 < 1$ satisfying the conditions of Lemma 1. Consider the random variables $X^{(g)}$and $Y^{(g)}$ with respective distribution functions:

$$\Pr\left(X^{(g)} = 0\right) = \Pr\left(X^{(g)} = 1\right) = \frac{1 - \sqrt{p_2}}{2},\ P\left(X^{(g)} = 2\right) = \sqrt{p_2} \tag{23}$$

and

$$\Pr\left(Y^{(g)}=0\right)=1-\frac{p_1}{\sqrt{p_2}},\ \Pr\left(Y^{(g)}=1\right)=\frac{p_1}{\sqrt{p_2}}. \tag{24}$$

Furthermore, let $\left(X_1^{(g)},Y_1^{(g)}\right)$ be the independent pair with marginals as defined in (23) and (24), i.e.:

$$\Pr\left(X_1^{(g)}=i,Y_1^{(g)}=j\right)=p_{ij}=\Pr\left(X^{(g)}=i\right)\ \Pr\left(Y^{(g)}=j\right). \tag{25}$$

The joint distribution of $\left(X_2^{(g)},Y_2^{(g)}\right)$ is defined in the Table 3

	$X_2^{(g)}$		
$Y_2^{(g)}$	0	1	2
0	$p_{00}+2\varepsilon$	$p_{10}-3\varepsilon$	$p_{20}+\varepsilon$
1	$p_{01}-2\varepsilon$	$p_{11}+3\varepsilon$	$p_{21}-\varepsilon$

Table 3: The distribution of $\left(X_2^{(g)},Y_2^{(g)}\right)$.

In this definition ε denotes an arbitrary positive number satisfying the condition

$$\varepsilon\le\min\left(\frac{p_{01}}{2},\frac{p_{10}}{3},p_{21}\right). \tag{26}$$

Note that $(X_2^{(g)},Y_2^{(g)})\in R(F_{X^{(g)}},F_{Y^{(g)}})$.
The computation of the covariance is then straightforward. One gets that

$$\mathrm{Cov}\left(X_2^{(g)},Y_2^{(g)}\right)=\varepsilon>0=\mathrm{Cov}\left(X_1^{(g)},Y_1^{(g)}\right). \tag{27}$$

Let us define $S_1^{(g)}=X_1^{(g)}+X_1^{(g)}$ and $S_2^{(g)}=X_2^{(g)}+Y_2^{(g)}$. To complete the proof of Theorem 2, it suffices to prove that

$$H_g\left(S_1^{(g)}\right)>H_g\left(S_2^{(g)}\right). \tag{28}$$

We compute the distribution of $S_1^{(g)}$ as follows:

$$f_1(2)=\Pr\left(S_1^{(g)}>2\right)=p_{21}=\sqrt{p_2}\frac{p_1}{\sqrt{p_2}}=p_1, \tag{29}$$

$$\begin{aligned}f_1(1)=\Pr\left(S_1^{(g)}>1\right)&=p_{21}+p_{11}+p_{20}=p_1+\frac{1-\sqrt{p_2}}{2}\frac{p_1}{\sqrt{p_2}}+\sqrt{p_2}\left(1-\frac{p_1}{\sqrt{p_2}}\right)\\&>\sqrt{p_2}>p_2,\end{aligned} \tag{30}$$

$$f_1(0) = \Pr\left(S_1^{(g)} > 0\right) = 1 - p_{00} < 1. \tag{31}$$

One finds the following expression for the decumulative distribution function:

$$\overline{F}_{S_1^{(g)}}(t) = \begin{cases} 1 & \text{for } t < 0 \\ f_1(k) & \text{for } k \le t < k+1 \text{ and } k = 0,1,2 \\ 0 & \text{for } t \ge 3 \end{cases}$$

Now using formula (1), we find

$$H_g\left(S_1^{(g)}\right) = g\left(f_1(0)\right) + g\left(f_1(2)\right) + g\left(f_1(2)\right). \tag{32}$$

Analogously, we define values $f_2(k) = \Pr\left(S_2^{(g)} > k\right)$ for $k = 0, 1, 2$. We get the following identities:

$$f_2(2) = f_1(2) - \varepsilon, \tag{33}$$

$$f_2(1) = f_1(1) + 3\varepsilon, \tag{34}$$

$$f_2(0) = f_1(0) - 2\varepsilon. \tag{35}$$

Thus

$$H_g\left(S_2^{(g)}\right) = g\left(f_1(0) - 2\varepsilon\right) + g\left(f_1(1) + 3\varepsilon\right) + g\left(f_1(2) - \varepsilon\right). \tag{36}$$

Combining (32) with (36), we see that in order to complete the proof of inequality (28), it suffices to prove that

$$g\left(f_1(2)\right) - g\left(f_1(2) - \varepsilon\right) + g\left(f_1(0)\right) - g\left(f_1(0) - 2\varepsilon\right) > g\left(f_1(1) + 3\varepsilon\right) - g\left(f_1(1)\right). \tag{37}$$

Now let us take a closer insight in differences occurring in inequality (37). From the Lagrange Theorem it follows that there exist $0 < \varepsilon_0, \varepsilon_1, \varepsilon_2 < \varepsilon$ such that the following identities hold:

$$g\left(f_1(0)\right) - g\left(f_1(0) - 2\varepsilon\right) = 2\varepsilon \cdot g'\left(f_1(0) - 2\varepsilon_0\right) > 2\varepsilon \cdot g'_-(1), \tag{38}$$

$$g\left(f_1(1) + 3\varepsilon\right) - g\left(f_1(0)\right) = 3\varepsilon \cdot g'\left(f_1(1) + 3\varepsilon_1\right) < 3\varepsilon \cdot g'(p_2), \tag{39}$$

$$g\left(f_1(2)\right) - g\left(f_1(2) - \varepsilon\right) = \varepsilon \cdot g'\left(f_1(2) - \varepsilon_2\right) > \varepsilon \cdot g'(p_1). \tag{40}$$

However, from Lemma 1, we find that

$$g'(p_1) + 2g'_-(1) > 3g'(p_2). \tag{41}$$

Multiplying both sides of (41) by ε and combining with the inequalities (38), (39) and (40), we get the sequence of inequalities:

$$g\Big(f_1(2)\Big) - g\Big(f_1(2) - \varepsilon\Big) + g\Big(f_1(0)\Big) - g\Big(f_1(0) - 2\varepsilon\Big) > \varepsilon\ g'(p_1) + 2\varepsilon\ g'_-(1) >$$

$$> 3\varepsilon\ g'(p_2) > g\Big(f_1(1) + 3\varepsilon\Big) - g\Big(f_1(1)\Big). \tag{42}$$

This proves inequality (37), and thus also Theorem 2. ■

References

[1] Artzner, P.; Delbaen, F.; Eber, J.M.; Heath, D. (1999). "Coherent measures of risk", *Mathematical Finance* 9, 203-228.

[2] Denneberg, D. (1990) "Premium Calculation: Why Standard Deviation should be replaced by Absolute Deviation", *ASTIN Bulletin* 20(2), 181-190.

[3] Denneberg, D. (1994). "Non-additive measure and integral", Kluwer Academic Publishers, Boston.

[4] Dhaene, J.; Denuit, M.; Goovaerts, M.J.; Kaas, R.; Vyncke, D. (2002a). "The concept of comonotonicity in actuarial science and finance: theory", *Insurance: Mathematics & Economics* 31(1), 3-33.

[5] Dhaene, J.; Denuit, M.; Goovaerts, M.J.; Kaas, R.; Vyncke, D. (2002b). "The concept of comonotonicity in actuarial science and finance: applications", *Insurance: Mathematics & Economics,* 31(2), 133-161.

[6] Dhaene, J.; Wang, S.; Young, V; Goovaerts, M. (2000). "Comonotonicity and maximal stop-loss premiums", *Mitteilungen der Schweiz. Aktuarvereinigung* 2000(2), 99-113.

[7] Goovaerts, M.J.; De Vylder, F.; Haezendonck, J. (1984). "Insurance Premiums", North-Holland, Amsterdam.

[8] Goovaerts, M.J.; Kaas, R.; Van Heerwaarden, A.E. & Bauwelinkx, T. (1990). "Effective Actuarial Methods", North-Holland, Amsterdam.

[9] Kaas, R.; Dhaene, J.; Goovaerts; M.J. (2000). "Upper and lower bounds for sums of random variables", *Insurance: Mathematics & Economics* 23, 151-168.

[10] Kaas, R.; Dhaene, J.; Vyncke, D.; Goovaerts, M.J.; Denuit, M. (2002). "A simple geometric proof that comonotonic risks have the convex-largest sum", *ASTIN Bulletin* 32(1), 71-80.

[11] Kaas, R.; Goovaerts, M.J., Dhaene, J.; Denuit, M. (2001). "Modern Actuarial Risk Theory", Kluwer Academic Publishers, Dordrecht.

[12] Kaas, R.; Van Heerwaarden, A.E.; Goovaerts, M.J. (1994). "Ordering of Actuarial Risks", Caire Education Series, Amsterdam.

[13] Panjer, H.H. (2002). "Measurement of risk, solvency requirements and allocation of capital within financial conglomerates", *Institute of Insurance and Pension Research*, University of Waterloo, Research Report 01-15.

[14] Ramsay, C.M. (1994). "Loading gross premiums for risk without using utility theory, with Discussions", *Transactions of the Society of Actuaries* XLV, 305-349.

[15] Shaked, M.; Shanthikumar, J.G. (1994). "Stochastic orders and their applications", Academic Press, New York.

[16] Schmeidler, D. (1986). "Integral representation without additivity", *Proceedings of the American Mathematical Society* 97, 225-261.

[17] Von Neumann, J; Morgenstern, O. (1947). "Theory of games and economic behavior", Princeton University Press, Princeton.

[18] Wang, S. (1995). "Insurance pricing and increased limits ratemaking by proportional hazard transforms", *Insurance: Mathematics & Economics* 17, 43-54.

[19] Wang, S. (1996). "Premium calculation by transforming the layer premium density", *ASTIN Bulletin* 26, 71-92.

[20] Wang, S. (2000). "A class of distortion operators for pricing financial and insurance risks", *Journal of Risk and Insurance* 67(1), 15-36.

[21] Wang, S. (2001). "A risk measure that goes beyond coherence", *Institute of Insurance and Pension Research*, University of Waterloo, Research Report 01-18.

[22] Wang, S.; Dhaene, J. (1998). "Comonotonicity, correlation order and premium principles", *Insurance: Mathematics & Economics* 22, 235-242.

[23] Wang, S. and Young, V.R. (1998). "Ordering risks: expected utility theory versus Yaari's dual theory of risk", *Insurance: Mathematics & Economics* 22, 235-242.

[24] Wirch, J.L.; Hardy, M.R. (2000). "Ordering of risk measures for capital adequacy", *Institute of Insurance and Pension Research*, University of Waterloo, Research Report 00-03.

[25] Yaari, M.E. (1987). "The dual theory of choice under risk", *Econometrica* 55, 95-115.