

On the distance between convex-ordered random variables, with applications

MICHAEL V. BOUTSIKAS¹ AND EUTICHIA VAGGELATOU²

¹*University of Piraeus, Department of Statistics and Insurance Science,*

²*National Technical University of Athens, Department of Applied Mathematics and Physical Sciences*

Abstract

Simple approximation techniques are developed exploiting relationships between generalized convex orders and appropriate probability metrics. In particular, the distance between s -convex ordered random variables is investigated. Results connecting positive/negative dependence concepts and convex ordering are also presented. These results lead to approximations and bounds for the distributions of sums of positive/negative dependent random variables. Applications and extensions of the main results pertaining to compound Poisson, normal and exponential approximation are provided as well.

Key words and phrases: stochastic orders of convex type, probability metrics, positive/negative dependence, compound Poisson approximation, rate of convergence in CLT, exponential approximation, geometric convolutions, aging distributions.

AMS 2000 subject classification: Primary 60E15, 62E17; Secondary 60F05, 62E20.

1 Introduction

The study of many stochastic models may sometimes be so complicated that an explicit calculation of their characteristics turns out to be practically unfeasible. In order to overcome such unfortunate situations and be able to extract as much information as possible, many techniques that lead to useful bounds or approximations have been developed. Very often, the problem under investigation requires the approximation of the distributions of certain variables. In the last decades, the use of appropriate probability metrics and their properties has been recognized as a very efficient tool for the development of such approximations (see e.g. Rachev (1991)).

Another indirect way to extract useful information may be offered by a comparison of the random variables (r.v.'s) involved in the original "complex" model with variables related to an appropriately modified model (in order to be computationally more tractable). The most convenient method to study such comparisons is provided by the elegant theory of stochastic orders. For a comprehensive treatment of this subject, including a variety of applications, we refer to the books of Shaked and Shanthikumar (1994) and Szekli (1995).

Although the above two approaches may seem quite unrelated, there exists a close relation between probability metrics and stochastic orders. A first attempt to systematically study relationships between probability metrics and stochastic orders was carried out by Lefèvre and Utev (1998). In a very recent work, Denuit and Bellegem (2001) exploited such relationships, in order to derive upper bounds for the distances between random sums.

The primary purpose of this paper is to exploit possible relationships between certain classes of stochastic orders and appropriate probability metrics, so as to develop simple and rather general approximation techniques. In particular, we present several results concerning the distance (in terms of appropriate probability metrics) between stochastically ordered r.v.'s. Our study will be mainly focused on a very important class of stochastic ordering, the convex ordering. Numerous applications of this ordering can be found in the fields of reliability, economics, insurance, comparison of experiments, queueing, epidemics and so on. To be more specific, we shall exploit an extension of the usual convex ordering, namely the s -convex ordering, introduced by Denuit, Lefèvre and Shaked (1998). We shall also resort to some useful properties of Zolotarev's ideal metric ζ_s (cf. Zolotarev (1983)) and stop-loss metric of order s (cf. Rachev and Rüschendorf (1990)), which are proved to be very closely related to the s -convex ordering.

A secondary aim of this work is to provide results connecting positive/negative dependence concepts with convex ordering. Apart from their independent interest, these results, combined with the aforementioned study, readily lead to approximations and bounds for the distributions of sums of positive/negative dependent r.v.'s. Some of these approximations and bounds can be considered as an extension of the results presented by Boutsikas and Koutras (2000), which concerned sums of integer-valued associated r.v.'s.

Finally, an important part of this work comprises applications and extensions of the main results related to three major research areas of distribution approximation. More specifically, we present several interesting results pertaining to compound Poisson, normal and exponential approximation.

In the course of completing this article we discovered that a relevant work was carried out independently by Denuit, Lefèvre and Utev (2002). In their article they study (in the actuarial context) connections between probabilistic distances and stochastic orders for arithmetic r.v.'s, while our results mainly concern real-valued r.v.'s.

The organization of the paper is as follows: Section 2 reviews preliminaries on stochastic orders and probability metrics that are essential for our exposition. In Sections 3.1 and 3.2 we present our main results relating to the distance between s -convex ordered r.v.'s and to situations that yield s -convex ordering respectively. In Section 3.3, we focus on positive/negative dependence concepts and derive approximation results for sums of positively/negatively dependent r.v.'s. In Section 4 we specialize and extend the results of Section 3 in various areas of interest. Particularly, in the first paragraph of Section 4 we investigate approximations for distributions of sums of positively/negatively dependent r.v.'s by a suitable compound Poisson distribution. In the second paragraph we establish normal convergence results for sums of positively / negatively dependent r.v.'s. Finally, in the third paragraph we study exponential approximations for distributions that belong to certain classes of aging distributions or can be represented as geometric convolutions.

2 Preliminaries - Notation

Let $\mathcal{C}^s(\mathcal{I})$ denote the class of all the functions $\phi : \mathcal{I} \rightarrow \mathbb{R}$ such that their s -th derivative, $\phi^{(s)}$, exists and is continuous on the subinterval $\mathcal{I} \subseteq \mathbb{R}$. The univariate stochastic order relations that will be used in this paper belong to the class of integral stochastic orders (see e.g. Müller (1997)). Such stochastic orders are defined by a reference to a class $\mathcal{U}^{\mathcal{I}}$ of measurable functions $\phi : \mathcal{I} \rightarrow \mathbb{R}$ satisfying some desirable properties. More specifically, the r.v. X is said to be $\mathcal{U}^{\mathcal{I}}$ -smaller than the

r.v. Y (denoted by $X \preceq_{\mathcal{U}^{\mathcal{I}}} Y$) if

$$\mathbb{E}\phi(X) \leq \mathbb{E}\phi(Y) \quad (1)$$

for all $\phi \in \mathcal{U}^{\mathcal{I}}$, such that the expectations exist. Many of the usual stochastic orders (cf. Shaked and Shanthikumar (1994)) can be defined through (1). In particular, the usual *stochastic dominance* \preceq_{st} is obtained when $\mathcal{U}^{\mathcal{I}} = \mathcal{U}_{st}^{\mathbb{R}}$, the class of all non-decreasing functions over \mathbb{R} . Moreover, we get the usual *convex* (resp. *increasing convex*, *concave*, *increasing concave*) ordering, denoted by \preceq_{cx} (resp. \preceq_{icx} , \preceq_{cv} , \preceq_{icv}) when $\mathcal{U}^{\mathcal{I}} = \mathcal{U}_{cx}^{\mathbb{R}}$ (resp. $\mathcal{U}_{icx}^{\mathbb{R}}$, $\mathcal{U}_{cv}^{\mathbb{R}}$, $\mathcal{U}_{icv}^{\mathbb{R}}$), the class of all convex (resp. increasing convex, concave, increasing concave) functions over \mathbb{R} . Note that if $X \preceq_{icx} Y$ and $\mathbb{E}X = \mathbb{E}Y$ then $X \preceq_{cx} Y$.

Recently, Denuit, Lefèvre and Shaked (1998), introduced new classes of stochastic order relations, called *s*-convex (resp. *s*-concave) orders, that can be seen as extensions of the usual convex (resp. concave) order. Furthermore, they studied the *s*-increasing convex order (which already appears in Fishburn (1976, 1980a,b) and Stoyan (1983)), also called stop loss dominance of order $s - 1$ in the actuarial literature. Specifically, given any positive integer s , the *s*-convex (resp. *s*-concave) ordering $\preceq_{s-cx}^{\mathcal{I}}$ (resp. $\preceq_{s-cv}^{\mathcal{I}}$) is generated by the class $\mathcal{U}_{s-cx}^{\mathcal{I}}$ (resp. $\mathcal{U}_{s-cv}^{\mathcal{I}}$) of all the regular *s*-convex (rep. *s*-concave) functions on \mathcal{I} which is defined by

$$\mathcal{U}_{s-cx}^{\mathcal{I}} = \left\{ \phi \in \mathcal{C}^s(\mathcal{I}) : \phi^{(s)} \geq 0 \text{ on } \mathcal{I} \right\} \quad \left(\text{resp. } \mathcal{U}_{s-cv}^{\mathcal{I}} = \left\{ \phi \in \mathcal{C}^s(\mathcal{I}) : (-1)^{s+1} \phi^{(s)} \geq 0 \text{ on } \mathcal{I} \right\} \right).$$

Finally, the *s*-increasing convex $\preceq_{s-icx}^{\mathcal{I}}$ and the *s*-increasing concave ordering $\preceq_{s-icv}^{\mathcal{I}}$ are generated by the classes $\mathcal{U}_{s-icx}^{\mathcal{I}}$ and $\mathcal{U}_{s-icv}^{\mathcal{I}}$ of the regular *s*-increasing convex and *s*-increasing concave functions respectively, defined by

$$\mathcal{U}_{s-icx}^{\mathcal{I}} = \bigcap_{k=1}^s \mathcal{U}_{k-cx}^{\mathcal{I}} \quad \text{and} \quad \mathcal{U}_{s-icv}^{\mathcal{I}} = \bigcap_{k=1}^s \mathcal{U}_{k-cv}^{\mathcal{I}}.$$

It is of interest to note that there may be different classes of functions $\mathcal{U}^{\mathcal{I}}$ that generate the same integral stochastic order, e.g. $\preceq_{\mathcal{U}^{\mathcal{I}}}$ can also be generated by any set dense in $\mathcal{U}^{\mathcal{I}}$ with respect to some suitable topology (cf. Müller (1997)). For example, the \preceq_{icx} order can also be generated by $\mathcal{U}_{2-icx}^{\mathbb{R}}$ ($\mathcal{U}_{icx}^{\mathbb{R}}$ is the closure of $\mathcal{U}_{2-icx}^{\mathbb{R}}$ with respect to the uniform convergence). Thus, $\preceq_{2-icx} \Leftrightarrow \preceq_{icx}$ and by similar reasoning, $\preceq_{2-cx} \Leftrightarrow \preceq_{cx}$, $\preceq_{2-cv} \Leftrightarrow \preceq_{cv}$, $\preceq_{2-icv} \Leftrightarrow \preceq_{icv}$, $\preceq_{1-icx} \Leftrightarrow \preceq_{1-cx} \Leftrightarrow \preceq_{st}$.

From now on, we shall focus on the *s*-convex and the *s*-increasing convex orders, since, for any X, Y real valued r.v.'s, we have that $X \preceq_{s-icx} Y \Leftrightarrow -Y \preceq_{s-icv} -X$ and

$$X \preceq_{s-cx} Y \Leftrightarrow \begin{cases} X \preceq_{s-cv} Y & \text{when } s \text{ is odd} \\ Y \preceq_{s-cv} X & \text{when } s \text{ is even.} \end{cases} \quad (2)$$

Moreover, we confine ourselves to the case $\mathcal{I} = \mathbb{R}$ for the *s*-convex ordering and $\mathcal{I} = \mathbb{R}_+$ for the *s*-increasing convex ordering, using the simplified notation \preceq_{s-cx} instead of $\preceq_{s-cx}^{\mathbb{R}}$ and \preceq_{s-icx} instead of $\preceq_{s-icx}^{\mathbb{R}_+}$. The next characterization of the *s*-convex orders, (cf. Denuit, Lefèvre and Shaked (1998); see also a related result of Fishburn (1980a)) which extends relations (1.A.1) and (2.A.5) of Shaked and Shanthikumar (1994), will be proved very useful in the sequel. Henceforth we shall be using the notation $y_+^s := (\max\{y, 0\})^s$.

Theorem 1 (a) If X, Y are two real valued r.v.'s such that $\mathbb{E}|X|^{s-1}, \mathbb{E}|Y|^{s-1} < \infty$, then

$$X \preceq_{s-cx} Y \Leftrightarrow \begin{cases} \mathbb{E}X^k = \mathbb{E}Y^k, k = 1, 2, \dots, s-1, \text{ and} \\ \mathbb{E}(X-t)_+^{s-1} \leq \mathbb{E}(Y-t)_+^{s-1} \text{ for all } t \in \mathbb{R}. \end{cases}$$

(b) If X, Y are two non-negative real valued r.v.'s, such that $\mathbb{E}X^{s-1}, \mathbb{E}Y^{s-1} < \infty$ then

$$X \preceq_{s-icx} Y \Leftrightarrow \begin{cases} \mathbb{E}X^k \leq \mathbb{E}Y^k, k = 1, 2, \dots, s-1, \text{ and} \\ \mathbb{E}(X-t)_+^{s-1} \leq \mathbb{E}(Y-t)_+^{s-1} \text{ for all } t \in \mathbb{R}. \end{cases}$$

A very useful simple criterion for the verification of the increasing convex order is the well known Karlin-Novikoff cut-criterion (cf. Karlin and Novikoff (1963)). For the statement of this criterion, we need to introduce first the following notation. Define the number of sign-changes of ϕ on \mathbb{R} by

$$S^-(\phi) = \sup\{S^-[\phi(x_1), \phi(x_2), \dots, \phi(x_n)] : x_1 < x_2 < \dots < x_n \in \mathbb{R}, n \in \mathbb{N}\},$$

where $S^-[y_1, y_2, \dots, y_n]$ denotes the number of sign changes in the sequence y_1, y_2, \dots, y_n (zero terms are being discarded). Two real functions ϕ_1, ϕ_2 are said to have k crossing points (or cross each other k times) if $S^-(\phi_1 - \phi_2) = k$. According to Karlin-Novikoff cut-criterion, if $S^-(F_X - F_Y) = 1, \mathbb{E}X \leq \mathbb{E}Y$ and the last sign of $F_X - F_Y$ is a $+$, then $X \preceq_{icx} Y$. As usual, F_X denotes the cumulative distribution function (c.d.f.) of a r.v. X .

An analogous criterion for the s -convex order is offered by the next proposition.

Proposition 2 (Denuit, Lefèvre and Shaked (1998)). Let X, Y be two r.v.'s such that $\mathbb{E}(X^j - Y^j) = 0, j = 1, 2, \dots, s-1$. If $S^-(F_X - F_Y) = s-1$ and the last sign of $F_X - F_Y$ is a $+$, then $X \preceq_{s-cx} Y$.

Denuit, Lefèvre and Shaked (1998) generalized some well known properties of the usual convex orders (see e.g. Shaked and Shanthikumar (1994) or Szekli (1995)) in the s -convex orders. In particular, they proved that the s -convex orders are closed under mixtures, convolutions, compounding and they are preserved under limits. Moreover, they proved that, if $X \preceq_{s-cx} Y$, then $X =_{st} Y \Leftrightarrow \mathbb{E}X^s = \mathbb{E}Y^s$. Here, and in what follows, $=_{st}$ denotes equality in distribution. This last result leads to the following fertile question which can be considered as the initial motivation of the present paper: if $X \preceq_{s-cx} Y$ and $\mathbb{E}X^s$ is close to $\mathbb{E}Y^s$, then how close are the respective distributions of X, Y ?

Apparently, in order to answer the above question, we shall be needing the use of some appropriate metrics on the space of probability measures over $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. For typographical convenience, we shall allow an abuse of the notation and write $d(X, Y)$ instead of $d(F_X, F_Y)$ for any probability metric d .

Initially, we remind some well known probability metrics:

-The *uniform* or the *Kolmogorov* metric: $d_K(X, Y) = \sup_{x \in \mathbb{R}} |F_X(x) - F_Y(x)|$.

-The *Wasserstein* or *Kantorovich* or L_1 metric: $d_W(X, Y) = \int_{-\infty}^{\infty} |F_X(x) - F_Y(x)| dx$.

A first crucial step towards the solution of an approximation problem is the choice of the most appropriate probability metric. For the problem we attempt to study, the use of the following probability metrics will be proved quite efficient:

- *Stop-loss metrics of order $s \in \mathbb{N}$* , (cf. Rachev and Rüschendorf (1990))

$$\mathbf{d}_s(X, Y) = \sup_{t \in \mathbb{R}} \frac{1}{s!} |\mathbb{E}(X-t)_+^s - \mathbb{E}(Y-t)_+^s|, \quad \mathbb{E}|X|^s, \mathbb{E}|Y|^s < \infty.$$

- Zolotarev's ideal metric ζ_s , $s \in \mathbb{N}$, (cf. Zolotarev (1983))

$$\zeta_s(X, Y) = \frac{1}{(s-1)!} \int_{-\infty}^{\infty} |\mathbb{E}(X-t)_+^{s-1} - \mathbb{E}(Y-t)_+^{s-1}| dt, \quad \mathbb{E}|X|^{s-1}, \mathbb{E}|Y|^{s-1} < \infty.$$

For a thorough investigation of probability metrics and their properties we refer to Rachev (1991). It is worth stressing that the Wasserstein distance d_W coincides with ζ_1 . The stop-loss metrics originally appeared in risk theory (cf. Gerber (1981), p.97 for $s = 1$) for the estimation of the difference between two stop-loss premiums. The metric \mathbf{d}_1 is often denoted by d_{sl} (stop-loss distance) in the actuarial literature. In order to be in tune with the notations used by most authors, in the sequel we shall prefer to write d_W and d_{sl} instead of ζ_1 and \mathbf{d}_1 respectively.

Apparently, \mathbf{d}_s and ζ_s are closely related to the s -convex orders. We summarize below some useful properties of the above metrics (cf. Rachev and Rüschendorf (1990), Rachev (1991)) that will be utilized in the sequel.

Proposition 3 (i) If $\zeta_s(X, Y) < \infty$ then $\mathbb{E}(X^j - Y^j) = 0$, $j = 1, 2, \dots, s-1$.

(ii) If $x^s(F_X(x) - F_Y(x)) \rightarrow_{x \rightarrow +\infty} 0$, then $\mathbf{d}_s(X, Y) \leq \zeta_s(X, Y)$.

(iii) If Y has a bounded Lebesgue density f_Y , then

$$d_K(X, Y) \leq (1 + \sup_{x \in \mathbb{R}} f_Y(x)) \left(\frac{(2s)! \sqrt{2s+1}}{s!} \zeta_s(X, Y) \right)^{\frac{1}{s+1}}.$$

(iv) If $\mathbb{E}X^j = \mathbb{E}Y^j$, $j = 1, 2, \dots, s-1$, then $\zeta_s(X, Y) \leq \frac{1}{s!} (\mathbb{E}|X|^s + \mathbb{E}|Y|^s)$.

3 Main results

3.1 On the distance between s -convex ordered r.v.'s

As already mentioned, if $X \preceq_{s-cx} Y$ then $X =_{st} Y \Leftrightarrow \mathbb{E}X^s = \mathbb{E}Y^s$. From this fact, it seems plausible to expect that e.g. the distance (in terms of some appropriate probability metric) between the distributions of convex ordered r.v.'s, with "almost" identical second moments, is "almost" zero. This simple approximation principle has been implicitly used for a number of years (e.g. see Chacon and Walsh (1976), Meilijson (1983), Machina and Pratt (1997) for the convex/concave case) but only very recently it was explicitly expressed by Denuit, Lefèvre and Utev (2002) for the case of arithmetic r.v.'s. The results that follow intend to mathematically translate this principle in the class of s -convex ordered real-valued r.v.'s. In the sequel, we shall denote by $\mathfrak{X}^s(U)$, $U \subseteq \mathbb{R}$ the space of all r.v.'s defined on a probability space $(\Omega, \mathcal{A}, \Pr)$ and taking values in U with $\mathbb{E}|X|^s < \infty$.

Theorem 4 If $X, Y \in \mathfrak{X}^s(\mathbb{R})$ and $X \preceq_{s-cx} Y$, then

$$\zeta_s(X, Y) = \frac{1}{s!} (\mathbb{E}Y^s - \mathbb{E}X^s). \quad (3)$$

Moreover, if $t^s(F_X(t) - F_Y(t)) \rightarrow 0$ as $t \rightarrow +\infty$, then $\mathbf{d}_s(X, Y) \leq \frac{1}{s!} (\mathbb{E}Y^s - \mathbb{E}X^s)$.

Proof. Since $X \preceq_{s-cx} Y$, Theorem 1 implies that

$$\mathbb{E}(X^k - Y^k) = 0, \quad k = 1, 2, \dots, s-1, \quad \text{and} \quad \mathbb{E}(X-t)_+^{s-1} \leq \mathbb{E}(Y-t)_+^{s-1} \quad \text{for all } t \in \mathbb{R}.$$

Invoking Lemma 2.2. of Rachev and Rüschendorf (1990), we obtain that

$$\mathbb{E}(X^s - Y^s) = s \int_{-\infty}^{+\infty} (\mathbb{E}(X-t)_+^{s-1} - \mathbb{E}(Y-t)_+^{s-1}) dt = s! \zeta_s(X, Y),$$

and the proof of the first result is completed. The second result follows immediately from Proposition 3(ii). ■

Note that a similar theorem was proved by Denuit, Lefèvre and Utev (2002) for arithmetic r.v.'s. The next corollary provides a result concerning the uniform rate of convergence of F_X to F_Y when $X \preceq_{s-cx} Y$ and $\mathbb{E}Y^s - \mathbb{E}X^s \rightarrow 0$.

Corollary 5 *If $X, Y \in \mathfrak{X}^s(\mathbb{R})$ such that $X \preceq_{s-cx} Y$ or $Y \preceq_{s-cx} X$ and Y has a bounded Lebesgue density f_Y , then*

$$d_K(X, Y) \leq \left(1 + \frac{1}{s}\right) \left(\frac{K(s) \cdot M_Y^s}{(s-1)!} |\mathbb{E}X^s - \mathbb{E}Y^s|\right)^{\frac{1}{s+1}},$$

where $K(s) = \frac{(2s)!\sqrt{2s+1}}{s!}$, $M_Y = \sup_{x \in \mathbb{R}} f_Y(x)$.

Proof. Employing a standard minimizing technique (cf. Rachev (1991)) for the upper bound in Proposition 3(iii) we have that, for every $c > 0$,

$$d_K(X, Y) = d_K(cX, cY) \leq (1 + \sup_{x \in \mathbb{R}} f_{cY}(x)) (K(s) \zeta_s(cX, cY))^{\frac{1}{s+1}} \leq (1 + \frac{M_Y}{c}) (K(s) c^s \zeta_s(X, Y))^{\frac{1}{s+1}}.$$

The desired result follows by minimizing the right-hand-side with respect to c (i.e. taking $c = M_Y/s$), and applying Theorem 4. ■

Next, let us turn our attention to the usual (increasing) convex order ($s = 2$) and the usual stochastic dominance ($s = 1$). These orders are widely known, possess many useful properties and have a great number of interesting applications in diverse areas (see e.g. Shaked and Shanthikumar (1994) or Szekli (1995)). A straightforward application of Theorem 4 reveals that if $X, Y \in \mathfrak{X}^2(\mathbb{R})$ and $X \preceq_{cx} Y$ then

$$\zeta_2(X, Y) = \frac{1}{2} (\mathbb{E}Y^2 - \mathbb{E}X^2) = \frac{1}{2} (\text{Var}(Y) - \text{Var}(X)), \quad (4)$$

(cf. also Kaas (1993)) while if $X, Y \in \mathfrak{X}^1(\mathbb{R})$ and $X \preceq_{st} Y$ then $d_W(X, Y) = \zeta_1(X, Y) = \mathbb{E}Y - \mathbb{E}X$, a result which is rather straightforward and has been effectively used in two applications by Lefèvre and Utev (1998). Moreover, Corollary 5 states that if $X \preceq_{cx} Y$ or $Y \preceq_{cx} X$ and Y has a bounded Lebesgue density f_Y , then $d_K(X, Y) \leq 4.5 M_Y^{2/3} |\mathbb{E}X^2 - \mathbb{E}Y^2|^{1/3}$. It is worth stressing that a slightly better result can be extracted if we exploit the inequality $d_K(X, Y) \leq 3 M_Y^{2/3} (\zeta_2(X, Y))^{1/3}$ (see Rachev (1991), relation (14.1.16)), namely

$$d_K(X, Y) \leq 3 \cdot 2^{-1/3} M_Y^{2/3} |\mathbb{E}Y^2 - \mathbb{E}X^2|^{1/3}. \quad (5)$$

It is useful to mention that, if $X, Y \in \mathfrak{X}^2(\mathbb{R})$ and $\mathbb{E}X = \mathbb{E}Y$, then $\zeta_2(X, Y) < \infty$ (see Proposition 3(iv)) and moreover,

$$\zeta_2(X, Y) = \int_{-\infty}^{\infty} \left| \int_x^{\infty} (F_X(t) - F_Y(t)) dt \right| dx = \int_{-\infty}^{\infty} \left| \int_{-\infty}^x (F_X(t) - F_Y(t)) dt \right| dx \quad (6)$$

since $\mathbb{E}(X - t)_+ = \int_t^{\infty} (1 - F_X(u)) du$. Furthermore, if $X_n, X \in \mathfrak{X}^2(\mathbb{R})$, then $\zeta_2(X_n, X) \rightarrow 0 \Rightarrow X_n \rightarrow_{st} X$.

Another interesting remark concerns the distance between the stationary renewal distributions of two convex ordered r.v.'s. More specifically, let X be a nonnegative r.v with survival function $\bar{F}_X = 1 - F_X$ and finite mean, and denote by X^* a r.v. whose survival function \bar{F}_{X^*} is the *stationary renewal* distribution of X , i.e. $\bar{F}_{X^*}(x) = \frac{1}{\mathbb{E}X} \int_x^{\infty} \bar{F}_X(t) dt$ (X^* is often called the *stationary forward recurrence time* associated with X). Suppose now that $X, Y \in \mathfrak{X}^2(\mathbb{R}_+)$ such that $\mathbb{E}X = \mathbb{E}Y$. From representation (6) we get that,

$$\zeta_2(X, Y) = \mathbb{E}X \int_0^{\infty} |\bar{F}_{X^*}(x) - \bar{F}_{Y^*}(x)| dx = \mathbb{E}X \cdot d_W(X^*, Y^*). \quad (7)$$

Hence, if $Y \succeq_{cx} X$ then, (4) implies that

$$d_W(X^*, Y^*) = \frac{1}{\mathbb{E}X} \zeta_2(X, Y) = \frac{\text{Var}(Y) - \text{Var}(X)}{2\mathbb{E}X}.$$

Note also that, in correspondence with (7), we can easily check that $d_K(X^*, Y^*) = \frac{1}{\mathbb{E}X} \mathbf{d}_1(X, Y)$.

Next, we proceed to the establishment of some results connecting the metrics ζ_2 , d_{sl} and d_W . Apart from their independent interest, these results may be used in conjunction with equation (4) to produce bounds for the Wasserstein or the stop-loss distance between convex ordered r.v.'s.

Proposition 6 *If $X, Y \in \mathfrak{X}^2(\mathbb{R})$ and $\mathbb{E}X = \mathbb{E}Y$ then*

$$d_{sl}(X, Y) \leq \sqrt{\zeta_2(X, Y)}.$$

Moreover, if $X \succeq_{cx} Y$ then, for every $d \in \mathbb{R}$, $0 \leq \mathbb{E}(X - d)_+ - \mathbb{E}(Y - d)_+ \leq \sqrt{\frac{1}{2}(\mathbb{E}X^2 - \mathbb{E}Y^2)}$.

Proof. Set $H(x) := F_X(x) - F_Y(x)$. For every $a, t \in \mathbb{R}$ we have that

$$\begin{aligned} \left| \left| \int_a^{\infty} H(u) du \right| - \left| \int_t^{\infty} H(u) du \right| \right| &\leq \left| \int_a^{\infty} H(u) du - \int_t^{\infty} H(u) du \right| = \left| \int_{\min\{a, t\}}^{\max\{a, t\}} H(u) du \right| \\ &\leq \int_{\min\{a, t\}}^{\max\{a, t\}} |H(u)| du \leq \int_{\min\{a, t\}}^{\max\{a, t\}} du = |t - a| \end{aligned}$$

Hence, $|\int_t^{\infty} H(u) du| \geq |\int_a^{\infty} H(u) du| - |t - a|$, $a, t \in \mathbb{R}$, and therefore, for every $\epsilon > 0$,

$$\begin{aligned} \zeta_2(X, Y) &= \int_{-\infty}^{\infty} \left| \int_t^{\infty} H(u) du \right| dt \geq \int_{a-\epsilon}^{a+\epsilon} \left| \int_t^{\infty} H(u) du \right| dt \\ &\geq \int_{a-\epsilon}^{a+\epsilon} \left| \int_a^{\infty} H(u) du \right| dt - \int_{a-\epsilon}^{a+\epsilon} |t - a| dt = 2\epsilon \left| \int_a^{\infty} H(u) du \right| - \epsilon^2. \end{aligned}$$

In order to minimize the RHS, we take $\epsilon = \left| \int_a^\infty H(u) du \right|$ obtaining that

$$\zeta_2(X, Y) \geq \left(\int_a^\infty H(u) du \right)^2, \text{ for every } a \in \mathbb{R}.$$

Finally,

$$d_{sl}(X, Y) = \sup_{a \in \mathbb{R}} |\mathbb{E}(X - a)_+ - \mathbb{E}(Y - a)_+| = \sup_{a \in \mathbb{R}} \left| \int_a^\infty H(u) du \right| \leq \sqrt{\zeta_2(X, Y)}.$$

The second part is immediate from the above inequality and (4). ■

In insurance, the second outcome of the above proposition can be translated as follows: when $X \succeq_{cx} Y$, the difference between two stop-loss premiums $\mathbb{E}(X - d)_+$ and $\mathbb{E}(Y - d)_+$ is bounded above by $\sqrt{\frac{1}{2}(\mathbb{E}X^2 - \mathbb{E}Y^2)}$ for all deductibles d . A useful relation between the Wasserstein and the stop-loss distance is offered by the following proposition.

Proposition 7 *If $X, Y \in \mathfrak{X}^1(\mathbb{R})$ and F_X, F_Y have $k \geq 1$ crossing points, then*

$$d_W(X, Y) \leq 2kd_{sl}(X, Y) + |\mathbb{E}(X - Y)| \leq (2k + 1)d_{sl}(X, Y).$$

Proof. Since $S^-(F_X - F_Y) = k$, there exist $x_1 < x_2 < \dots < x_k \in \mathbb{R}$ ($x_0 = -\infty, x_{k+1} = \infty$) such that, for every $i = 0, 1, \dots, k$, either $H(x) := F_X(x) - F_Y(x) \geq 0, \forall x \in (x_i, x_{i+1})$ or $H(x) \leq 0, \forall x \in (x_i, x_{i+1})$. For the Wasserstein distance we get that

$$\begin{aligned} d_W(X, Y) &= \int_{-\infty}^{\infty} |H(x)| dx = \sum_{i=1}^{k+1} \int_{x_{i-1}}^{x_i} |H(x)| dx = \sum_{i=1}^{k+1} \left| \int_{x_{i-1}}^{x_i} H(x) dx \right| \\ &\leq \sum_{i=1}^k \left(\left| \int_{x_{i-1}}^{\infty} H(x) dx \right| + \left| \int_{x_i}^{\infty} H(x) dx \right| \right) + \left| \int_{x_k}^{\infty} H(x) dx \right| \end{aligned}$$

and since $\int_{x_0}^{\infty} H(x) dx = \int_{-\infty}^{\infty} H(x) dx = \mathbb{E}(Y - X)$, we finally conclude that

$$d_W(X, Y) \leq 2k \sup_{t \in \mathbb{R}} \left| \int_t^{\infty} H(x) dx \right| + |\mathbb{E}(X - Y)| = 2kd_{sl}(X, Y) + |\mathbb{E}(X - Y)|$$

The proof is completed by observing that $|\mathbb{E}(X - Y)| \leq d_{sl}(X, Y)$. ■

The above result remains valid even if F_X, F_Y have $k = 0$ crossing points. But in this special case, we may say even more; if $S^-(F_X - F_Y) = 0$, then either $F_X \geq F_Y$ (i.e. $X \preceq_{st} Y$) or $F_Y \geq F_X$ (i.e. $Y \preceq_{st} X$). Thus, it can be easily verified (see also Lefèvre and Utev (1998)) that either $d_W(X, Y) = d_{sl}(X, Y) = \mathbb{E}(Y - X)$ or $d_W(X, Y) = d_{sl}(X, Y) = \mathbb{E}(X - Y)$ respectively.

Suppose now that F_X, F_Y have exactly one crossing point and $\mathbb{E}X = \mathbb{E}Y$. Karlin-Novikoff cut-criterion guarantees that either $X \succeq_{cx} Y$ (if $Var(X) \geq Var(Y)$) or $Y \succeq_{cx} X$ (if $Var(Y) \geq Var(X)$). A combined use of Propositions 6, 7 and (4) implies that

$$d_W(X, Y) \leq 2d_{sl}(X, Y) \leq 2\sqrt{\zeta_2(X, Y)} = \sqrt{2|Var(Y) - Var(X)|}.$$

The next result reveals a relationship between the Wasserstein distance and ζ_2 for the case of integer-valued r.v.'s.

Proposition 8 Let $X, Y \in \mathfrak{X}^2(\mathbb{Z})$. If $\mathbb{E}X = \mathbb{E}Y$, then $d_W(X, Y) \leq \frac{2}{\sqrt{2}-1} \zeta_2(X, Y)$. If moreover $Y \succeq_{cx} X$, then

$$d_W(X, Y) \leq 2\zeta_2(X, Y) = \mathbb{E}Y^2 - \mathbb{E}X^2.$$

Proof. Denoting again $H(x) := F_X(x) - F_Y(x)$ and recalling representation (6) of ζ_2 we get

$$\begin{aligned} \zeta_2(X, Y) &= \int_{-\infty}^{\infty} \left| \int_t^{\infty} H(u) du \right| dt = \sum_{i=-\infty}^{\infty} \int_i^{i+1} \left| \int_t^{i+1} H(u) du + \int_{i+1}^{\infty} H(u) du \right| dt \\ &= \sum_{i=-\infty}^{\infty} \int_i^{i+1} \left| (i+1-t)H(i) + \sum_{u=i+1}^{\infty} H(u) \right| dt = \sum_{i=-\infty}^{\infty} \int_0^1 \left| -tH(i) + \sum_{u=i}^{\infty} H(u) \right| dt \end{aligned} \quad (8)$$

It is not hard to prove that, if $a, b \in \mathbb{R}$ then $\int_0^1 |at + b| dt \geq (\sqrt{2} - 1)|b|$. Using this inequality for $a = -H(i), b = \sum_{u=i}^{\infty} H(u)$ and relation (8) we are led to

$$\zeta_2(X, Y) \geq (\sqrt{2} - 1) \sum_{i=-\infty}^{\infty} \left| \sum_{u=i}^{\infty} H(u) \right|. \quad (9)$$

For the Wasserstein distance between the distributions of X and Y we get that

$$\begin{aligned} d_W(X, Y) &= \int_{-\infty}^{\infty} |H(x)| dx = \sum_{i=-\infty}^{\infty} |H(i)| \leq \sum_{i=-\infty}^{\infty} \left| \sum_{u=i}^{\infty} H(u) - \sum_{u=i+1}^{\infty} H(u) \right| \\ &\leq \sum_{i=-\infty}^{\infty} \left| \sum_{u=i}^{\infty} H(u) \right| + \sum_{i=-\infty}^{\infty} \left| \sum_{u=i+1}^{\infty} H(u) \right| = 2 \sum_{i=-\infty}^{\infty} \left| \sum_{u=i}^{\infty} H(u) \right| \end{aligned} \quad (10)$$

which, combined with (9), ascertains the first inequality. If moreover $Y \succeq_{cx} X$, then $\int_t^{\infty} H(u) du \geq 0$ for all $t \in \mathbb{R}$ and following the same procedure as above, we get that

$$\zeta_2(X, Y) = \int_{-\infty}^{\infty} \int_t^{\infty} H(u) du dt = \sum_{i=-\infty}^{\infty} \int_0^1 \left(\sum_{u=i}^{\infty} H(u) - tH(i) \right) dt = \sum_{i=-\infty}^{\infty} \sum_{u=i}^{\infty} H(u) - \frac{1}{2} \sum_{i=-\infty}^{\infty} H(i).$$

But $\sum_{i=-\infty}^{\infty} H(i) = \mathbb{E}Y - \mathbb{E}X = 0$ and thus, using (10) and relation (4), the proof of the second inequality is completed. ■

3.2 A connection between upper orthant and s -convex orders.

An interesting case where s -convex ordering appears, is between sums of upper orthant ordered r.v.'s. We first recall the well known notion of the upper orthant order. The random vector $\mathbf{X} = (X_1, \dots, X_n)$ is smaller than $\mathbf{Y} = (Y_1, \dots, Y_n)$ in the upper orthant order (i.e. $\mathbf{X} \preceq_{uo} \mathbf{Y}$) if

$$\Pr(X_1 > x_1, X_2 > x_2, \dots, X_n > x_n) \leq \Pr(Y_1 > x_1, Y_2 > x_2, \dots, Y_n > x_n) \text{ for all } x_1, x_2, \dots, x_n \in \mathbb{R}.$$

Note that $\mathbf{X} \preceq_{uo} \mathbf{Y}$ if and only if (see e.g. Shaked and Shanthikumar (1994))

$$\mathbb{E} \left(\prod_{i=1}^n g_i(X_i) \right) \leq \mathbb{E} \left(\prod_{i=1}^n g_i(Y_i) \right), \quad (11)$$

for every collection $\{g_1, g_2, \dots, g_n\}$ of univariate nonnegative increasing functions. Next, we prove an interesting result connecting the upper orthant order and the s -(increasing) convex order. Two auxiliary lemmas are initially presented.

Lemma 9 *If $X_1, X_2, \dots, X_n \in \mathfrak{X}^{n-1}(\mathbb{R})$ then, for any $t \in \mathbb{R}$,*

$$\begin{aligned} & \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} \Pr(X_1 > u_1, X_2 > u_2, \dots, X_{n-1} > u_{n-1}, X_n > t - u_1 - \dots - u_{n-1}) du_{n-1} \dots du_1 \\ &= \frac{1}{(n-1)!} \mathbb{E}(X_1 + X_2 + \dots + X_n - t)_+^{n-1}. \end{aligned} \quad (12)$$

Proof. For $r \in \mathbb{N}$, $x_1, x_2 \in \mathbb{R}$ set

$$H_{r,t}(x_1, x_2) := \int_{-\infty}^{+\infty} I(x_1 > u) \cdot (x_2 - (t - u))_+^r du.$$

where, as usual, $I(x > y) = I_{(y, \infty)}(x) = 1$ if $x > y$ and 0 otherwise. If $x_1 + x_2 > t$ then

$$H_{r,t}(x_1, x_2) = \int_{t-x_2}^{x_1} (x_2 - (t - u))^r du = \frac{(x_1 + x_2 - t)^{r+1}}{r+1},$$

whereas if $x_1 + x_2 \leq t$ then $H_{r,t}(x_1, x_2) = 0$. Thus, for every $t \in \mathbb{R}$,

$$\int_{-\infty}^{+\infty} I(x_1 > u) \cdot (x_2 - (t - u))_+^r du = \frac{(x_1 + x_2 - t)_+^{r+1}}{r+1}. \quad (13)$$

For $r = 0$ we get that

$$\int_{-\infty}^{+\infty} I(x_1 > u) \cdot I(x_2 > t - u) du = (x_1 + x_2 - t)_+.$$

We shall use induction to generalize the above equality. Assume that the following relation

$$\begin{aligned} & \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} I(x_1 > u_1) I(x_2 > u_2) \dots I(x_k > t - u_1 - \dots - u_{k-1}) du_{k-1} \dots du_1 \\ &= \frac{1}{(k-1)!} (x_1 + x_2 + \dots + x_k - t)_+^{k-1}, \quad \forall t \in \mathbb{R}, \end{aligned} \quad (14)$$

is valid for $k = n \geq 2$. We shall prove that (14) is also valid for $k = n + 1$. Indeed,

$$\begin{aligned} & \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} I(x_1 > u_1) I(x_2 > u_2) \dots I(x_{n+1} > t - u_1 - \dots - u_n) du_n \dots du_1 \\ &= \int_{-\infty}^{+\infty} I(x_1 > u_1) \left\{ \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} I(x_2 > u_2) \dots I(x_{n+1} > t - u_1 - u_2 - \dots - u_n) du_n \dots du_2 \right\} du_1 \\ &= \frac{1}{(n-1)!} \int_{-\infty}^{+\infty} I(x_1 > u) (x_2 + \dots + x_{n+1} - t + u)_+^{n-1} du = \frac{(x_1 + x_2 + \dots + x_{n+1} - t)_+^n}{n!}, \end{aligned}$$

where the last equality is achieved by virtue of (13). Replacing x_i with X_i and applying the mean value operator on both sides of (14), the Fubini theorem completes the proof. ■

Lemma 10 If $X_i, Y_i \in \mathfrak{X}^{s-1}(\mathbb{R}), i = 1, 2, \dots, s$ and $(Y_1, \dots, Y_s) \succeq_{uo} (X_1, \dots, X_s)$ then,

$$\mathbb{E}(Y_1 + Y_2 + \dots + Y_s - t)_+^{s-1} - \mathbb{E}(X_1 + X_2 + \dots + X_s - t)_+^{s-1} \geq 0, \text{ for every } t \in \mathbb{R}.$$

Proof. It is an immediate consequence of Lemma 9 and the definition of the upper orthant order. ■

Theorem 11 (a) If $X_i, Y_i \in \mathfrak{X}^{s-1}(\mathbb{R}_+), i = 1, 2, \dots, s$ and $(Y_1, \dots, Y_s) \succeq_{uo} (X_1, \dots, X_s)$ then

$$Y_1 + \dots + Y_s \succeq_{s-icx} X_1 + \dots + X_s.$$

(b) If $X_i, Y_i \in \mathfrak{X}^{s-1}(\mathbb{R}), i = 1, 2, \dots, s$ such that $(Y_1, \dots, Y_s) \succeq_{uo} (X_1, \dots, X_s)$ and

$$\mathbb{E}(Y_1 + Y_2 + \dots + Y_s)^k = \mathbb{E}(X_1 + X_2 + \dots + X_s)^k \quad (15)$$

for every $k = 1, 2, \dots, s - 1$, then $Y_1 + \dots + Y_s \succeq_{s-cx} X_1 + \dots + X_s$.

Proof. (a) Using relation (11) we deduce that

$$\begin{aligned} \mathbb{E}(Y_1 + Y_2 + \dots + Y_s)^k &= \mathbb{E} \left(\sum_{i=1}^s Y_i \right)^k = \sum_{i_1=1}^s \sum_{i_2=1}^s \dots \sum_{i_k=1}^s \mathbb{E}(Y_{i_1} Y_{i_2} \dots Y_{i_k}) \\ &\geq \sum_{i_1=1}^s \sum_{i_2=1}^s \dots \sum_{i_k=1}^s \mathbb{E}(X_{i_1} X_{i_2} \dots X_{i_k}) = \mathbb{E}(X_1 + X_2 + \dots + X_s)^k \end{aligned}$$

for every $k = 1, 2, \dots, s - 1$. The proof of (a) is completed by recalling Lemma 10 and Theorem 1. The proof of (b) follows from (15), Lemma 10 and Theorem 1. ■

It is of interest to note that, if $X_i, Y_i \in \mathfrak{X}^{s-1}(\mathbb{R}_+)$ and $(Y_1, \dots, Y_s) \succeq_{uo} (X_1, \dots, X_s)$, then obviously $(a_1 Y_1, \dots, a_s Y_s) \succeq_{uo} (a_1 X_1, \dots, a_s X_s)$ for every $a_i \geq 0, i = 1, 2, \dots, s$ and therefore

$$\sum_{i=1}^s a_i Y_i \succeq_{s-icx} \sum_{i=1}^s a_i X_i \text{ for every } a_i \geq 0, i = 1, 2, \dots, s. \quad (16)$$

Relation (16) was proved in the special case $s = 2$ by Dhaene and Goovaerts (1996,1997) for $(X_1, X_2), (Y_1, Y_2)$ such that $X_i =_{st} Y_i, i = 1, 2$, using the terminology *correlation order* and *stop-loss order* instead of *upper orthant order* and *convex order* respectively.

3.3 Positive/negative dependence and convex orders

In this paragraph we obtain results which connect convex orders and various types of positive/negative dependence of r.v.'s. Apart from their independent interest, these results enable us to derive approximations for distributions of sums of r.v.'s that exhibit some form of positive/negative dependence (see Theorem 15). Applications of such approximations will be presented in Sections 4.1 and 4.2. Initially, it is necessary to recall some well known concepts of positive/negative dependence.

A collection of r.v.'s X_1, X_2, \dots, X_n is said to be (positively) *associated* (cf. Esary, Proschan and Walkup (1967)), if $Cov(f(\mathbf{X}), g(\mathbf{X})) \geq 0, \mathbf{X} = (X_1, X_2, \dots, X_n)$ for every pair of coordinatewise

nondecreasing functions f and g such that the covariance exists. A weaker concept of (positive) association is offered by assuming that, for every pair of disjoint subsets A_1, A_2 of $\{1, 2, \dots, n\}$,

$$\text{Cov}(f(X_i, i \in A_1), g(X_i, i \in A_2)) \geq 0, \quad (17)$$

for every pair of coordinatewise nondecreasing functions f, g of $\{x_i, i \in A_1\}, \{x_i, i \in A_2\}$ respectively. In this case X_1, X_2, \dots, X_n are called *weakly* (positively) *associated*. If (17) holds true for all f, g with the inequality sign reversed, the r.v.'s X_1, X_2, \dots, X_n are called *negatively associated* (cf. Joag-Dev and Proschan (1983)). In the special case $n = 2$, (17) is equivalent to

$$\Pr(X_1 \geq x_1, X_2 \geq x_2) \geq \Pr(X_1 \geq x_1) \Pr(X_2 \geq x_2) \text{ for all } x_1, x_2, \quad (18)$$

which is Lehmann's (1966) definition for *positively quadrant dependence* (PQD). If (18) (or equivalently (17) with $n = 2$) holds true with the inequality sign reversed, the r.v.'s X_1, X_2 are called *negatively quadrant dependent* (NQD). A collection X_1, X_2, \dots, X_n is said to be *linearly positively* (resp. *negatively*) *quadrant dependent*, LPQD (resp. LNQD), if for every pair A_1, A_2 of disjoint subsets of $\{1, 2, \dots, n\}$ and every $a_i \geq 0, i \in A_1, b_j \geq 0, j \in A_2$, the pair $\sum_{i \in A_1} a_i X_i, \sum_{i \in A_2} b_i X_i$ is PQD (resp. NQD) (cf. Joag-Dev (1983)). An even weaker notion of positive/negative multivariate dependence based on PQD/NQD pairs of r.v.'s was introduced by Boutsikas and Koutras (2000); a collection X_1, X_2, \dots, X_n is said to be *positively/negatively cumulative dependent* (PCD/NCD), if for every $i = 2, 3, \dots, n$ the r.v.'s X_i and $\sum_{j=1}^{i-1} X_j$ are PQD or NQD respectively. Manifestly, association (resp. negative association) implies LPQD (resp. LNQD) which in turn implies PCD (resp. NCD). In what follows, for every collection X_1, X_2, \dots, X_n of r.v.'s we shall denote by $X_1^\perp, X_2^\perp, \dots, X_n^\perp$ a collection of independent r.v.'s (also independent of X_i 's) such that $X_i =_{st} X_i^\perp, i = 1, 2, \dots, n$.

A simple way to connect positive dependence and convex ordering is described by the following observation. If $X_1, X_2 \in \mathfrak{X}^1(\mathbb{R})$ are PQD (resp. NQD) then equivalently $(X_1, X_2) \succeq_{uo} (X_1^\perp, X_2^\perp)$ (resp. $(X_1^\perp, X_2^\perp) \succeq_{uo} (X_1, X_2)$) and hence (see Theorem 11),

$$X_1 + X_2 \succeq_{cx} X_1^\perp + X_2^\perp \text{ (resp. } X_1^\perp + X_2^\perp \succeq_{cx} X_1 + X_2). \quad (19)$$

This result can also be derived from Tchen (1980) (see also Dhaene and Goovaerts (1996) for nonnegative r.v.'s). An immediate question that arises here is whether an analogous to (19) result is valid for n positively/negatively dependent r.v.'s. In other words, is a sum of PCD or NCD (e.g. associated or negatively associated) r.v.'s and a sum of their independent duplicates, convex ordered? In view of Theorem 11, we suspect that we could prove an even more general result yielding s -convex ordering. In this more general case, the summands must exhibit an appropriately extended notion of positive/negative dependence. Accordingly, we introduce the following definition.

Definition 12 *A collection X_1, X_2, \dots, X_n of real valued r.v.'s will be called s -PCD (resp. s -NCD), if for every $i \in \{2, \dots, n\}$ there exists a partition of nonempty sets $A_1, A_2, \dots, A_{k_i}, k_i \leq s - 1$ of $\{1, 2, \dots, i - 1\}$, such that*

$$\left(\sum_{j \in A_1} X_j, \dots, \sum_{j \in A_{k_i}} X_j, X_i \right) \succeq_{uo} \text{ (resp. } \preceq_{uo} \text{)} \left(\sum_{j \in A_1} X_j, \dots, \sum_{j \in A_{k_i}} X_j, X_i^\perp \right). \quad (20)$$

Obviously, 2-PCD (resp. 2-NCD) coincides with PCD (resp. NCD). Now we can prove the following result which states that a sum of s -PCD (rep. s -NCD) r.v.'s and a sum of their independent duplicates are s -(increasing) convex ordered.

Theorem 13 (a) Let X_1, X_2, \dots, X_n be s -PCD (resp. s -NCD) r.v.'s in $\mathfrak{X}^{s-1}(\mathbb{R}_+)$. Then

$$X_1 + X_2 + \dots + X_n \succeq_{s-icx} \text{ (resp. } \preceq_{s-icx} \text{) } X_1^\perp + X_2^\perp + \dots + X_n^\perp.$$

(b) If X_1, X_2, \dots, X_n are s -PCD (resp. s -NCD) r.v.'s in $\mathfrak{X}^{s-1}(\mathbb{R})$ and

$$\mathbb{E}(X_1 + \dots + X_{i-1} + X_i)^m = \mathbb{E}(X_1 + \dots + X_{i-1} + X_i^\perp)^m, \quad m = 1, 2, \dots, s-1, \quad (21)$$

for every $i = 2, 3, \dots, n$, then $X_1 + X_2 + \dots + X_n \succeq_{s-cx}$ (resp. \preceq_{s-cx}) $X_1^\perp + X_2^\perp + \dots + X_n^\perp$.

Proof. (a) Since X_1, X_2, \dots, X_n are s -PCD r.v.'s, it follows that (20) is valid for every $i \in \{2, \dots, n\}$. Hence, Theorem 11(a) yields

$$X_1 + \dots + X_{i-1} + X_i \succeq_{k_i+1-icx} X_1 + \dots + X_{i-1} + X_i^\perp, \quad i = 2, \dots, n$$

and since $k_i + 1 \leq s$, we also conclude that $(\mathcal{U}_{s-icx}^{\mathbb{R}_+} \subseteq \mathcal{U}_{k_i+1-icx}^{\mathbb{R}_+})$

$$\sum_{j=1}^i X_j \succeq_{s-icx} \sum_{j=1}^{i-1} X_j + X_i^\perp, \quad i = 2, \dots, n. \quad (22)$$

Recall that if Z is a r.v. independent of X, Y , then $X \preceq_{s-cx} Y$ implies that $X + Z \preceq_{s-cx} Y + Z$ (cf. Denuit, Lefèvre and Shaked (1998)). By a proper modification of the proof of this result we get that if $Z \geq 0$ is a r.v. independent of $X, Y \geq 0$, then $X \preceq_{s-icx} Y$ implies that $X + Z \preceq_{s-icx} Y + Z$. Hence, from (22) it follows that

$$\sum_{j=1}^i X_j + \sum_{j=i+1}^n X_j^\perp \succeq_{s-icx} \sum_{j=1}^{i-1} X_j + \sum_{j=i}^n X_j^\perp, \quad \text{for } i = 2, \dots, n.$$

Thus,

$$\sum_{j=1}^n X_j \succeq_{s-icx} \sum_{j=1}^{n-1} X_j + \sum_{j=n}^n X_j^\perp \succeq_{s-icx} \sum_{j=1}^{n-2} X_j + \sum_{j=n-1}^n X_j^\perp \succeq_{s-icx} \dots \succeq_{s-icx} X_1 + \sum_{j=2}^n X_j^\perp =_{st} \sum_{j=1}^n X_j^\perp$$

and the proof for s -PCD r.v.'s is completed. The proof for s -NCD r.v.'s is analogous.

(b) Relation (20) is again valid for every $i \in \{2, \dots, n\}$. In view of condition (21), Theorem 11(b) yields

$$X_1 + \dots + X_{i-1} + X_i \succeq_{k_i+1-cx} X_1 + \dots + X_{i-1} + X_i^\perp, \quad i = 2, 3, \dots, n.$$

If $k_i = s - 1$ then

$$X_1 + \dots + X_{i-1} + X_i \succeq_{s-cx} X_1 + \dots + X_{i-1} + X_i^\perp, \quad (23)$$

whereas if $k_i < s - 1$ then, invoking (21) for $m = k_i + 1$ and Theorem 4, we conclude that $X_1 + \dots + X_{i-1} + X_i =_{st} X_1 + \dots + X_{i-1} + X_i^\perp$. Thus, (23) is trivially also valid. The proof is now completed by a reasoning similar to (a). The proof for s -NCD r.v.'s is analogous. ■

A stronger but more convenient than (21) condition is to require X_1, \dots, X_n to be $s-1$ independent, i.e. every subset $\{X_{i_1}, \dots, X_{i_{s-1}}\}$ of $\{X_1, \dots, X_n\}$ consists of independent r.v.'s. Indeed, denoting by S_{i-1} the partial sum $\sum_{j=1}^{i-1} X_j$, in this case we get

$$\begin{aligned} & \mathbb{E}(S_{i-1} + X_i)^m - \mathbb{E}(S_{i-1} + X_i^\perp)^m = \sum_{j=0}^m \binom{m}{j} \mathbb{E}(X_i^{m-j} S_{i-1}^j) - \sum_{j=0}^m \binom{m}{j} \mathbb{E}X_i^{m-j} \mathbb{E}S_{i-1}^j \\ & = \sum_{j=1}^{m-1} \binom{m}{j} \text{Cov}(X_i^{m-j}, S_{i-1}^j) = \sum_{j=1}^{m-1} \sum_{a_1=1}^{i-1} \dots \sum_{a_j=1}^{i-1} \binom{m}{j} \text{Cov}(X_i^{m-j}, X_{a_1} X_{a_2} \dots X_{a_j}) = 0. \end{aligned}$$

Thus, if n "positively"/"negatively" dependent r.v.'s (in particular s -PCD/NCD) are $s-1$ -independent, $n \geq s$ (but of course not mutually independent), then their sum is greater/smaller than the sum of their independent duplicates in the s -convex order. A very interesting special case of Theorem 13(b) is obtained for $s = 2$. More specifically, we get the following corollary.

Corollary 14 *If X_1, X_2, \dots, X_n are PCD (resp. NCD) r.v.'s in $\mathfrak{X}^1(\mathbb{R})$, then*

$$X_1 + X_2 + \dots + X_n \succeq_{cx} \text{ (resp. } \preceq_{cx} \text{) } X_1^\perp + X_2^\perp + \dots + X_n^\perp.$$

Consequently, sums of associated or LPQD (resp. negatively associated or LNQD) r.v.'s are greater (resp. smaller) than the sums of their independent duplicates in the convex order. In the context of risk theory, Denuit, Dhaene and Ribas (2001) proved an analogous to Corollary 14 result for nonnegative PCD r.v.'s resorting to the result of Dhaene and Goovaerts (1996) mentioned after (19) above. Note, though, that their definition for PCD r.v.'s is slightly different from ours resulting to a slightly smaller class. Moreover, the above result was proved by Shao (2000) for the case of negatively associated r.v.'s.

Corollary 14 now readily leads to the following result which concerns the distance between the distribution of a sum of positively/negatively dependent r.v.'s and the distribution of a sum of independent r.v.'s with the same marginals.

Theorem 15 *Let X_1, X_2, \dots, X_n be PCD or NCD r.v.'s in $\mathfrak{X}^2(\mathbb{R})$. Then*

$$(a) \quad \zeta_2\left(\sum_{i=1}^n X_i, \sum_{i=1}^n X_i^\perp\right) = \left| \sum_{i<j} \text{Cov}(X_i, X_j) \right|, \quad (b) \quad d_{sl}\left(\sum_{i=1}^n X_i, \sum_{i=1}^n X_i^\perp\right) \leq \left| \sum_{i<j} \text{Cov}(X_i, X_j) \right|^{1/2}.$$

(c) *Moreover, if $\sum_{i=1}^n X_i^\perp$ has a bounded Lebesgue density f_Y , then*

$$d_K\left(\sum_{i=1}^n X_i, \sum_{i=1}^n X_i^\perp\right) \leq 3M_Y^{2/3} \left| \sum_{i<j} \text{Cov}(X_i, X_j) \right|^{1/3},$$

where $M_Y = \sup_{x \in \mathbb{R}} f_Y(x)$.

(d) *If the distributions of $\sum_{i=1}^n X_i, \sum_{i=1}^n X_i^\perp$ have $k < \infty$ crossing points, then*

$$d_W\left(\sum_{i=1}^n X_i, \sum_{i=1}^n X_i^\perp\right) \leq 2k \left| \sum_{i<j} \text{Cov}(X_i, X_j) \right|^{1/2}.$$

Proof. If X_1, X_2, \dots, X_n is a collection of PCD r.v.'s, Corollary 14 implies $\sum_{i=1}^n X_i \succeq_{cx} \sum_{i=1}^n X_i^\perp$ and invoking (4), we deduce that

$$\zeta_2\left(\sum_{i=1}^n X_i, \sum_{i=1}^n X_i^\perp\right) = \frac{1}{2} \left(\text{Var}\left(\sum_{i=1}^n X_i\right) - \text{Var}\left(\sum_{i=1}^n X_i^\perp\right) \right) = \sum_{i<j} \text{Cov}(X_i, X_j).$$

If X_1, X_2, \dots, X_n are NCD r.v.'s, then similarly we get that the above distance equals $-\sum_{i<j} \text{Cov}(X_i, X_j)$ and the proof of (a) is completed. Inequalities (b), (c) and (d) follow readily from (a) combined with Proposition 6, relation (5) and Proposition 7 respectively. ■

It is worth mentioning that Boutsikas and Koutras (2000) proved an analogous to Theorem 15 result for integer valued r.v.'s. More specifically, they proved that if $X_1, X_2, \dots, X_n \in \mathbb{Z}$ are PCD or NCD r.v.'s then,

$$d_W\left(\sum_{i=1}^n X_i, \sum_{i=1}^n X_i^\perp\right) \leq 2 \left| \sum_{i<j} \text{Cov}(X_i, X_j) \right|. \quad (24)$$

Obviously, (24) can be seen as a simple consequence of Theorem 15(a) and Proposition 8.

4 Applications

In this section we present three applications of the above outcomes pertaining to compound Poisson, normal and exponential approximation. Our aim is to develop interesting results that extend some already known ones and also to illustrate how the above analysis can be considered as a general framework offering a unified approach for many approximation problems in diverse areas.

If $X \sim \mathcal{D}$ for some known distribution \mathcal{D} , then, in some cases, we shall allow an abuse of the notation and write $\mathcal{D} \preceq Y$ instead of $X \preceq Y$ for some stochastic ordering \preceq , and $d(\mathcal{D}, Y)$ instead of $d(X, Y)$ for some probability metric d .

4.1 Compound Poisson approximation

In this paragraph we are going to investigate compound Poisson approximations for sums $\sum X_i$ of positively/negatively dependent r.v.'s. Such approximations are most suitable when the summands X_i are "weakly" or "locally" dependent and their masses are concentrated on 0. Many results of this type have appeared in the literature in the last decade mainly for binary-valued r.v.'s by the use of the Stein-Chen method (cf. Barbour, Holst and Janson (1992)). For a review on the recent developments of this method for compound Poisson approximation refer to Barbour and Chryssaphinou (2001).

In a recent work, Boutsikas and Koutras (2000) proceeded to the investigation of error bounds for compound Poisson approximations by taking a completely different approach than that of the Stein-Chen method. They proved that if X_1, X_2, \dots, X_n are non-negative, integer-valued, PCD or NCD r.v.'s with $\mathbb{E}(X_i), \mathbb{E}(X_i X_j) < \infty$ for $i, j = 1, 2, \dots, n, i \neq j$, then

$$d_W\left(\sum_{i=1}^n X_i, CP(\lambda, F)\right) \leq 2 \left| \sum_{i<j} \text{Cov}(X_i, X_j) \right| + \sum_{i=1}^n (\mathbb{E}X_i)^2, \quad (25)$$

where $\lambda = \sum_{i=1}^n \Pr(X_i > 0)$ and $F(x) = \frac{1}{\lambda} \sum_{i=1}^n \Pr(0 < X_i \leq x)$. In (25) and in what follows $CP(\lambda, F)$ denotes the (compound Poisson) distribution of the random sum $\sum_{i=1}^N Y_i$ where N is a Poisson r.v. with mean λ and Y_i are independent r.v.'s with distribution function F . We shall also use the notation $Po(\lambda)$ and $Be(p)$ for the ordinary Poisson distribution with mean λ and the Bernoulli distribution with mean p respectively.

The main purpose of this paragraph is the extension of (25) to the case of real valued r.v.'s. This extension, accomplished by exploiting Theorem 15, is presented below in Theorem 18. Initially, some intermediate results are needed.

Lemma 16 *If $X \in \mathfrak{X}^1(\mathbb{R}_+)$ and $p = \Pr(X > 0) > 0$ then $X \preceq_{cx} CP(p, F)$ where $F(x) = \Pr(X \leq x | X > 0)$.*

Proof. Let $I \sim Be(p)$, $N \sim Po(p)$ and Y_1, Y_2, \dots be a sequence of nonnegative independent r.v.'s (also independent of I, N) such that $Y_i \sim F, i = 1, 2, \dots$. Since $F_I(x) = 1 - p < e^{-p} = F_N(x)$, for $0 \leq x < 1$ and $F_I(x) = 1 > F_N(x)$, for $x \geq 1$, Karlin-Novikovoff cut-criterion along with the fact that $\mathbb{E}I = p = \mathbb{E}N$ yields $I \preceq_{cx} N$. Hence (see e.g. Shaked and Shanthikumar (1994)), $\sum_{i=1}^I Y_i \preceq_{cx} \sum_{i=1}^N Y_i$. The proof is completed by observing that $X =_{st} \sum_{i=1}^I Y_i$ and $\sum_{i=1}^N Y_i \sim CP(p, F)$. ■

Next we present a result, concerning the distance between the distribution of a sum of independent r.v.'s and a compound Poisson distribution.

Theorem 17 *If X_1, X_2, \dots, X_n are independent r.v.'s in $\mathfrak{X}^1(\mathbb{R}_+)$, then*

$$\sum_{i=1}^n X_i \preceq_{cx} CP(p, \frac{1}{p} \sum_{i=1}^n p_i F_i) \quad \text{and} \quad \zeta_2(\sum_{i=1}^n X_i, CP(p, \frac{1}{p} \sum_{i=1}^n p_i F_i)) = \frac{1}{2} \sum_{i=1}^n (\mathbb{E}X_i)^2,$$

where $p_i = \Pr(X_i > 0)$, $p = \sum_{i=1}^n p_i$, $F_i(x) = \Pr(X_i \leq x | X_i > 0)$.

Proof. Let Y_1, Y_2, \dots be a sequence of independent r.v.'s such that $Y_i \sim CP(p_i, F_i)$, $i = 1, 2, \dots, n$. It can be easily checked that $Var(Y_i) = \mathbb{E}X_i^2$. Invoking Lemma 16 we get that $Y_i \succeq_{cx} X_i$ and since the convex order is closed under convolutions (see e.g. Shaked and Shanthikumar (1994), Theorem 2.A.6) we deduce that $\sum_{i=1}^n X_i \preceq_{cx} \sum_{i=1}^n Y_i$. Therefore, by virtue of (4), we may write

$$2\zeta_2(\sum_{i=1}^n X_i, \sum_{i=1}^n Y_i) = Var(\sum_{i=1}^n Y_i) - Var(\sum_{i=1}^n X_i) = \sum_{i=1}^n (\mathbb{E}X_i^2 - Var(X_i)) = \sum_{i=1}^n (\mathbb{E}X_i)^2.$$

The validity of the theorem is now evident by observing that $\sum_{i=1}^n Y_i \sim CP(p, \frac{1}{p} \sum_{i=1}^n p_i F_i)$. ■

We are now in possession of the machinery needed in order to state the main result of this paragraph.

Theorem 18 *If $X_1, X_2, \dots, X_n \in \mathfrak{X}^2(\mathbb{R}_+)$ are PCD or NCD r.v.'s, then*

$$\zeta_2(\sum_{i=1}^n X_i, CP(p, \frac{1}{p} \sum_{i=1}^n p_i F_i)) \leq \left| \sum_{i < j} Cov(X_i, X_j) \right| + \frac{1}{2} \sum_{i=1}^n (\mathbb{E}X_i)^2$$

where $p_i = \Pr(X_i > 0)$, $p = \sum_{i=1}^n p_i$, $F_i(x) = \Pr(X_i \leq x | X_i > 0)$.

Proof. The proof is immediate from Theorems 15, 17 and the triangle inequality for ζ_2 . ■

It is obvious that, if the non-negative PCD or NCD (e.g. associated or negatively associated) r.v.'s X_1, \dots, X_n are almost uncorrelated and their distributions are concentrated on zero (i.e. the events $\{X_i > 0\}, i = 1, 2, \dots, n$ can be considered as "rare"), then their sum can be satisfactorily approximated by an appropriate compound Poisson distribution. Note that if X_1, X_2, \dots, X_n is a collection of integer-valued nonnegative PCD or NCD r.v.'s, then Theorem 18 and Proposition 8 readily reestablish (25) of Boutsikas and Koutras (2000).

Interesting applications of the above results arise from risk theory. For example, a problem of practical interest for actuaries is the approximation of the distribution of the aggregate claim in the individual risk model by an appropriate compound Poisson distribution (see e.g. Gerber (1984), Rachev and Rüschendorf (1990), de Pril and Dhaene (1992)). Theorem 18 along with Propositions 6, 7, 8 can be very useful in the case of dependent claim amounts (cf. Denuit, Levevre and Utev (2002) for an investigation of this matter concerning arithmetic r.v.'s).

4.2 Normal approximation for LPQD/LNQD sequences

Let X_1, X_2, \dots be a sequence of r.v.'s such that $\mathbb{E}X_i = 0, 0 < \mathbb{E}X_i^2 < \infty$ and set $S_n = X_1 + X_2 + \dots + X_n$. If the sequence of X_i 's is strictly stationary and consists of associated r.v.'s, the remarkable central limit theorem (CLT) proved by Newman (1980) states that

$$\text{if } \sigma^2 = \mathbb{E}X_1^2 + 2 \sum_{j=2}^{\infty} \mathbb{E}X_1 X_j < \infty, \text{ then } \frac{S_n}{\sqrt{n}} \rightarrow_{st} N(0, \sigma^2) \text{ as } n \rightarrow \infty.$$

This result inspired a series of limit theorems for associated r.v.'s. We mention the functional CLT (Newman and Wright (1981)), the Berry-Esseen inequality (Wood (1983), Dabrowski and Dehling (1988), Birkel (1988)), extensions to nonstationary cases (Cox and Grimmett (1984), Yu (1985)) and extensions to weakly associated sequences (Burton, Dabrowski and Dehling (1986)). Note that, instead of association, Newman's original CLT requires only that X_1, X_2, \dots are LPQD. Under the same dependence assumption, Birkel (1993) obtained a functional CLT.

The purpose of this paragraph is to show how the results of Section 3 can be employed in order to obtain a CLT for LPQD r.v.'s along with corresponding rates of convergence. It is remarkable that the approach we shall follow avoids the use of characteristic functions while its core is very similar to that of Wood (1983).

Theorem 19 *Let X_1, X_2, \dots be a strictly stationary sequence of LPQD r.v.'s such that $\mathbb{E}X_1 = 0, 0 < \mathbb{E}X_1^2 < \infty$. If $\sigma^2 := \mathbb{E}X_1^2 + 2 \sum_{j=2}^{\infty} \mathbb{E}X_1 X_j < \infty$ then, for $n = m \cdot k$,*

$$\zeta_2 \left(\frac{\sum_{i=1}^n X_i}{\sqrt{n}}, N(0, \sigma^2) \right) \leq \frac{a_k}{k} (\sigma^2 - \mathbb{E}X_1^2) + 2 \left(1 - \frac{a_k}{k} \right) u(a_k) + c \frac{\rho_k}{\sqrt{m}} \quad (26)$$

for some constant $c > 0$, where $\rho_k := \mathbb{E}|k^{-1/2} \sum_{i=1}^k X_i|^3$, $u(i) := \sum_{j=i+1}^{\infty} \mathbb{E}X_1 X_j \rightarrow_{i \rightarrow \infty} 0$ and $\{a_k\}$ is any sequence of positive integers such that $a_k \leq k$, $a_k \rightarrow \infty$, $a_k/k \rightarrow 0$ as $k \rightarrow \infty$.

Proof. For $m, k \in \mathbb{N}$ we write

$$\frac{\sum_{i=1}^{mk} X_i}{\sqrt{mk}} = \frac{Y_1 + Y_2 + \dots + Y_m}{\sqrt{m}} \quad \text{where } Y_j = \frac{1}{\sqrt{k}} \sum_{i=(j-1)k+1}^{jk} X_i.$$

It can be easily verified that the r.v.'s $Y_1/\sqrt{m}, \dots, Y_m/\sqrt{m}$ are also LPQD and thus Theorem 15 yields (taking into account stationarity and the fact that $Cov(Y_i, Y_j) \geq 0$),

$$\zeta_2 \left(\frac{\sum_{i=1}^m Y_i}{\sqrt{m}}, \frac{\sum_{i=1}^m Y_i^\perp}{\sqrt{m}} \right) = \frac{1}{m} \sum_{i=1}^{m-1} \sum_{j=i+1}^m Cov(Y_i, Y_j) \leq \sum_{j=2}^{\infty} Cov(Y_1, Y_j), \quad (27)$$

where $Y_1^\perp, Y_2^\perp, \dots, Y_m^\perp$ are independent r.v.'s such that $Y_i^\perp =_{st} Y_i, i = 1, 2, \dots, n$. The upper bound in (27) can be expressed as

$$\begin{aligned} \sum_{j=2}^{\infty} Cov(Y_1, Y_j) &= \sum_{j=2}^{\infty} Cov\left(\frac{1}{\sqrt{k}} \sum_{i=1}^k X_i, \frac{1}{\sqrt{k}} \sum_{i=(j-1)k+1}^{jk} X_i\right) = \frac{1}{k} \sum_{j=2}^{\infty} \sum_{i=1}^k \sum_{l=(j-1)k+1}^{jk} \mathbb{E}X_i X_l \\ &= \frac{1}{k} \sum_{i=1}^k \sum_{l=k+1}^{\infty} \mathbb{E}X_i X_l \leq \sum_{i=2}^{\infty} \min\left\{\frac{i-1}{k}, 1\right\} \mathbb{E}X_1 X_i \\ &= \sum_{i=2}^{a_k} \min\left\{\frac{i-1}{k}, 1\right\} \mathbb{E}X_1 X_i + \sum_{i=a_k+1}^{\infty} \min\left\{\frac{i-1}{k}, 1\right\} \mathbb{E}X_1 X_i \\ &\leq \frac{a_k}{k} \sum_{i=2}^{a_k} \mathbb{E}X_1 X_i + \sum_{i=a_k+1}^{\infty} \mathbb{E}X_1 X_i. \end{aligned} \quad (28)$$

Therefore,

$$\begin{aligned} \zeta_2 \left(\frac{\sum_{i=1}^{km} X_i}{\sqrt{km}}, \frac{\sum_{i=1}^m Y_i^\perp}{\sqrt{m}} \right) &\leq \frac{a_k}{k} \left(\frac{\sigma^2 - \mathbb{E}X_1^2}{2} - \sum_{j=a_k+1}^{\infty} \mathbb{E}X_1 X_j \right) + \sum_{j=a_k+1}^{\infty} \mathbb{E}X_1 X_j \\ &= \frac{a_k}{2k} (\sigma^2 - \mathbb{E}X_1^2) + \left(1 - \frac{a_k}{k}\right) u(a_k). \end{aligned} \quad (29)$$

Now, since $Y_1^\perp, Y_2^\perp, \dots, Y_m^\perp$ are i.i.d. r.v.'s, $\mathbb{E}Y_1^\perp = 0$ and $\mathbb{E}Y_1^{\perp 2} := \sigma_k^2 \leq \sigma^2$, Theorem 4 of Senatov (1980) implies that there exist constants $c_1, c_2 > 0$ such that

$$\zeta_2 \left(\sum_{i=1}^m \frac{Y_i^\perp}{\sigma_k \sqrt{m}}, N \right) \leq \frac{c_1}{\sqrt{m}} \left(\zeta_2 \left(\frac{Y_1}{\sigma_k}, N \right) + c_2 \max \left\{ \zeta_1 \left(\frac{Y_1}{\sigma_k}, N \right), \zeta_3 \left(\frac{Y_1}{\sigma_k}, N \right) \right\} \right), \quad (30)$$

where $N \sim N(0, 1)$. Note also that (cf. Proposition 3(iv))

$$\zeta_s \left(\frac{Y_1}{\sigma_k}, N \right) \leq \frac{1}{s!} \left(\frac{\mathbb{E}|Y_1|^s}{\sigma_k^s} + \mathbb{E}|N|^s \right), \quad s = 1, 2, 3$$

and thus,

$$\zeta_1 \left(\frac{Y_1}{\sigma_k}, N \right) \leq 1 + \sqrt{\frac{2}{\pi}}, \quad \zeta_2 \left(\frac{Y_1}{\sigma_k}, N \right) \leq 1, \quad \zeta_3 \left(\frac{Y_1}{\sigma_k}, N \right) \leq \frac{\rho_k}{6\sigma_k^3} + \frac{1}{3} \sqrt{\frac{2}{\pi}} \quad (31)$$

where $\rho_k = \mathbb{E}|Y_1|^3$. Hence,

$$\zeta_2 \left(\sum_{i=1}^m \frac{Y_i^\perp}{\sqrt{m}}, N(0, \sigma_k^2) \right) \leq \frac{\sigma_k^2 c_1}{\sqrt{m}} \left(1 + c_2 \max \left\{ 1 + \sqrt{\frac{2}{\pi}}, \frac{\rho_k}{6\sigma_k^3} + \frac{1}{3} \sqrt{\frac{2}{\pi}} \right\} \right) \leq c \frac{\rho_k + 1}{\sqrt{m}} \quad (32)$$

for some constant $c > 0$. Moreover Karlin-Novikoff cut-criterion and (4) imply that $\zeta_2(N(\mu, a_1^2), N(\mu, a_2^2)) = \frac{1}{2}|a_1^2 - a_2^2|$. Since, $\sigma_k^2 \leq \sigma^2$ we get (cf. (28)) that,

$$\zeta_2(N(0, \sigma_k^2), N(0, \sigma^2)) = \frac{1}{2}(\sigma^2 - \sigma_k^2) = \frac{1}{k} \sum_{i=1}^k \sum_{l=k+1}^{\infty} \mathbb{E}X_i X_l \leq \frac{a_k}{2k}(\sigma^2 - \mathbb{E}X_1^2) + \left(1 - \frac{a_k}{k}\right) u(a_k). \quad (33)$$

The proof is completed by combining (29), (32), (33) and the triangle inequality. ■

The above result provides rates of convergence in the CLT for LPQD summands in terms of the ζ_2 metric when $\rho_k < \infty, k = 1, 2, \dots$. It suffices to take $m, k \rightarrow \infty$ such that $m^{-1/2}\rho_k \rightarrow 0$ and in this case it can be easily checked that the upper bound in (26) tends to 0 (e.g. for $a_k = \lfloor \sqrt{k} \rfloor$ or $a_k = \lfloor \ln k \rfloor$). It goes without saying that, exploiting inequality (5), we could also get rates of convergence in terms of the Kolmogorov metric d_K . Note, though, that the rates of convergence provided by Theorem 19 may not be optimal. This can be understood by observing that, in view of Theorem 19 we also have that $\frac{Y_1}{\sigma_k} \rightarrow N(0, 1)$ as $k \rightarrow \infty$. Thus, the distances $\zeta_1, \zeta_2, \zeta_3$ appearing in (31) are not only bounded, but even better, they tend to 0. Therefore the convergence rate provided by (32) may be slower than the actual one. For associated sequences, Birkel (1988) obtained the convergence rate $O(n^{-1/2} \ln^2 n)$ for d_K provided that $u(n)$ exponentially decreases to 0.

Let us point out that by a similar approach (this time using Proposition 2.12 of Rachev and Rüschendorf (1990) and Proposition 6), we could also get a similar upper bound for the stop loss distance d_{sl} . Finally, it is worth mentioning that an analogous approach could be used in order to establish a similar bound for LNQD (e.g. negatively associated) r.v.'s, namely,

$$\zeta_2\left(\frac{\sum_{i=1}^n X_i}{\sqrt{n}}, N(0, \sigma^2)\right) \leq \frac{a_k}{k}(\mathbb{E}X_1^2 - \sigma^2) - 2\left(1 - \frac{a_k}{k}\right)u(a_k) + c\frac{\rho_k + 1}{\sqrt{m}}.$$

Note that in this case it is essential to assume $\sigma^2 > 0$ and not $\sigma^2 < \infty$ since $\sigma^2 \leq \mathbb{E}X_1^2 < \infty$.

4.3 Exponential approximation for aging distributions and geometric convolutions

a. Exponential approximation for aging distributions. In this paragraph we are going to study approximations and bounds for distributions that belong to certain classes of aging distributions. Such distributions arise quite naturally in many applied probability models such as queuing theory (see e.g. Szekli (1995)) or reliability theory (see e.g. Barlow and Proschan (1981)).

A non-negative r.v. X with c.d.f. F_X and $\mathbb{E}X < \infty$ is said to be HNBUE (*harmonic new better than used in expectation*) if $\int_x^\infty (1 - F_X(t))dt \leq \mathbb{E}X e^{-x/\mathbb{E}X}$ for all $x \geq 0$ or equivalently if $\mathbb{E}(X - x)_+ \leq \mathbb{E}(Y - x)_+$ for all $x \geq 0$, where Y follows an exponential distribution $\mathcal{E}(\mathbb{E}X)$ with mean value $\mathbb{E}X$. Hence (see Theorem 1), X is HNBUE iff $X \preceq_{cx} \mathcal{E}(\mathbb{E}X)$. Analogously, X is said to be HNWUE (*harmonic new worse than used in expectation*) if $X \succeq_{cx} \mathcal{E}(\mathbb{E}X)$. The class of HNBUE (HNWUE) distributions include all the standard aging (anti-aging) classes.

The exponential distribution is often used as an approximation for the unknown distribution of a r.v. X that is known to belong to a certain aging class. A straightforward application of (4) and (5) readily lead to the following result. If $X \in \mathfrak{X}^2(\mathbb{R}_+)$ is HNBUE or HNWUE r.v., then

$$\zeta_2(X, \mathcal{E}(\mathbb{E}X)) = \left| (\mathbb{E}X)^2 - \frac{\mathbb{E}X^2}{2} \right| \quad \text{and} \quad d_K(X, \mathcal{E}(\mathbb{E}X)) \leq 3|\rho_X|^{1/3} \quad (34)$$

where $\rho_X = 1 - \mathbb{E}X^2/2(\mathbb{E}X)^2$. The proof is immediate; since X is HNBUE or HNWUE, it follows that $X \preceq_{cx} \mathcal{E}(\mathbb{E}X)$ or $Y \preceq_{cx} \mathcal{E}(\mathbb{E}X)$ respectively. So, invoking relation (4) we are led to the first result. The inequality for d_K follows from (5).

A slightly better result for d_K was proved by Daley (1988) employing an entirely different approach. Note that, relations (34) were also proved by Rachev (1991) (cf. Chapter 14). His approach was essentially the same with the above, except that he did not explicitly identify that it is a consequence of the fact that X and $\mathcal{E}(\mathbb{E}X)$ are convex ordered.

If now X is a NBUE r.v., $\mathbb{E}X^2 < \infty$, then better error estimates can be extracted by the use of the following proposition. We remind that a nonnegative r.v. X with c.d.f. F_X and $\mathbb{E}X < \infty$ is NBUE (*new better than used in expectation*) if, for all $x \geq 0$,

$$\int_x^\infty (1 - F_X(t))dt \leq \mathbb{E}X(1 - F_X(x)). \quad (35)$$

Similarly, X is NWUE (*new worse than used in expectation*) if (35) is valid for all $x \geq 0$ with the inequality sign reversed.

Proposition 20 *If X is NBUE, Y is NWUE and $\mathbb{E}X = \mathbb{E}Y = \mu$, $\mathbb{E}X^2, \mathbb{E}Y^2 < \infty$, then*

$$d_W(X, Y) \leq \frac{2}{\mu} \zeta_2(X, Y) = \frac{1}{\mu} (\mathbb{E}Y^2 - \mathbb{E}X^2) \quad (36)$$

and

$$F_X(x) - F_Y(x) \leq \frac{1}{\mu} d_{sl}(X, Y) \leq \frac{1}{\mu^2} (\mathbb{E}Y^2 - \mathbb{E}X^2) \text{ for every } x \geq 0. \quad (37)$$

Proof. Since X is NBUE, Y is NWUE, it follows that X is HNBUE, Y is HNWUE and hence $X \preceq_{cx} Z$ and $Z \preceq_{cx} Y$, where $Z \sim \mathcal{E}(\mu)$ and therefore $X \preceq_{cx} Y$. If $A = \{x \in \mathbb{R}_+ : F_X(x) \geq F_Y(x)\}$, representation (6) of ζ_2 along with (35) guarantees that,

$$\begin{aligned} d_W(X, Y) &= \int_0^\infty |F_X(x) - F_Y(x)| dx = 2 \int_A (F_X(x) - F_Y(x)) dx \\ &= \frac{2}{\mu} \int_A (\mathbb{E}Y \bar{F}_Y(x) - \mathbb{E}X \bar{F}_X(x)) dx \leq \frac{2}{\mu} \int_A \left(\int_x^\infty \bar{F}_Y(t) dt - \int_x^\infty \bar{F}_X(t) dt \right) dx \\ &\leq \frac{2}{\mu} \int_0^\infty \left(\int_x^\infty \bar{F}_Y(t) dt - \int_x^\infty \bar{F}_X(t) dt \right) dx = \frac{2}{\mu} \zeta_2(X, Y), \end{aligned}$$

where $\bar{F} = 1 - F$. The proof of (36) is completed by invoking relation (4). For the proof of (37) observe that

$$\mu(F_X(x) - F_Y(x)) = \mu(\bar{F}_Y(x) - \bar{F}_X(x)) \leq \int_x^\infty \bar{F}_Y(t) dt - \int_x^\infty \bar{F}_X(t) dt \leq d_{sl}(X, Y)$$

for every $x \geq 0$. Note also that

$$d_{sl}(X, Y) = \sup_x \left| \int_x^\infty \bar{F}_X(t) dt - \int_x^\infty \bar{F}_Y(t) dt \right| dx \leq \int_0^\infty |F_X(x) - F_Y(x)| dx = d_W(X, Y)$$

and hence, for every $x \geq 0$,

$$F_X(x) - F_Y(x) \leq \frac{1}{\mu} d_{sl}(X, Y) \leq \frac{1}{\mu} d_W(X, Y) \leq \frac{2}{\mu^2} \zeta_2(X, Y) = \frac{1}{\mu^2} (\mathbb{E}Y^2 - \mathbb{E}X^2).$$

■

By the use of the above proposition we get the following corollary.

Corollary 21 *If X is NBUE or NWUE, then*

$$d_W(X, \mathcal{E}(\mathbb{E}X)) \leq 2\mathbb{E}X |\rho_X| \quad \text{and} \quad d_K(X, \mathcal{E}(\mathbb{E}X)) \leq 2^{3/2} |\rho_X|^{1/2}$$

where $\rho_X = 1 - \mathbb{E}X^2/2(\mathbb{E}X)^2$. Moreover, if X is NBUE then $F_X(x) - (1 - e^{-x/\mathbb{E}X}) \leq 2\rho_X$, while if X is NWUE then $(1 - e^{-x/\mathbb{E}X}) - F_X(x) \leq -2\rho_X$, for every $x \geq 0$.

Proof. The exponential distribution is trivially both NBUE and NWUE and hence Proposition 20 leads to

$$d_W(X, \mathcal{E}(\mathbb{E}X)) \leq \frac{2}{\mathbb{E}X} \zeta_2(X, \mathcal{E}(\mathbb{E}X)) = 2\mathbb{E}X |\rho_X|. \quad (38)$$

If a r.v. Y has a density f_Y then $d_K(X, Y) \leq (1 + \sup_{x \in \mathbb{R}} f_Y(x)) d_W(X, Y)^{1/2}$ (cf. Rachev (1991), relations (16.2.21), (16.2.13)). This inequality and a bound minimizing technique similar to the one used in the proof of Corollary 5 yields

$$d_K(X, Y) \leq 2 \left(\sup_{x \in \mathbb{R}} f_Y(x) \right)^{1/2} d_W(X, Y)^{1/2}, \quad (39)$$

which, in view of (38), implies that

$$d_K(X, \mathcal{E}(\mathbb{E}X)) \leq 2 \sqrt{\frac{1}{\mathbb{E}X}} \sqrt{2\mathbb{E}X |\rho_X|} = 2^{3/2} |\rho_X|^{1/2}.$$

The last part of the proposition follows by a straightforward application of (37). ■

Again, a slightly better result for d_K was proved by Daley (1988) engaging an entirely different reasoning.

b. Exponential approximation for geometric convolutions. Next, let us turn our attention to the exponential approximation of geometric convolutions. More specifically, suppose that X_1, X_2, \dots is a sequence of nonnegative i.i.d. r.v.'s with $0 < \mathbb{E}X_i^2 < \infty$ and let $S_n = \sum_{i=1}^n X_i$. The random sum S_{N_0} , where N_0 is independent of the summands X_i and follows the geometric distribution $\Pr(N_0 = k) = q^k p$, $q = 1 - p$, $k \in \{0, 1, 2, \dots\}$, is called a *geometric convolution*. Geometric convolutions arise in many applied fields such as risk theory (e.g. in ruin theory for the Cramer-Lundberg model), queueing (e.g. the waiting time distribution in a G/G/1 queue in equilibrium), reliability, regenerative models etc. It is well known that, under appropriate conditions, geometric convolutions converge in distribution to an exponentially distributed r.v. (see e.g. Szekli (1995), Brown (1990) and the references therein).

In this paragraph we investigate how the results of Section 3 readily lead to bounds for the distance between geometric convolutions and appropriate exponentially distributed r.v.'s.

Proposition 22 *Let $\gamma = \gamma_X = \mathbb{E}X^2/2(\mathbb{E}X)^2$. The following inequalities hold true,*

$$d_W \left(\frac{S_{N_0}}{\mathbb{E}S_{N_0}}, \mathcal{E}(1) \right) \leq \frac{2p}{q} \gamma, \quad d_K(S_{N_0}, \mathcal{E}(\mathbb{E}S_{N_0})) \leq 2^{3/2} \left(\frac{p\gamma}{q} \right)^{1/2}$$

and $(1 - e^{-x/\mathbb{E}S_{N_0}}) - \Pr(S_{N_0} \leq x) \leq 2\frac{p\gamma}{q}$ for every $x \geq 0$.

Proof. The r.v. S_{N_0} is NWU (*new worse than used*, cf. Brown (1990)) and hence, S_{N_0} is NWUE. Proposition 20 yields

$$d_W \left(\frac{S_{N_0}}{\mathbb{E}S_{N_0}}, \mathcal{E}(1) \right) \leq \frac{1}{\mathbb{E}S_{N_0}} d_W(S_{N_0}, \mathcal{E}(\mathbb{E}S_{N_0})) \leq \frac{1}{(\mathbb{E}S_{N_0})^2} \left(V(S_{N_0}) - (\mathbb{E}S_{N_0})^2 \right) = \frac{2p}{q} \gamma.$$

Moreover, using (39) and Proposition 20 we also deduce that

$$d_K(S_{N_0}, \mathcal{E}(\mathbb{E}S_{N_0})) = d_K \left(\frac{S_{N_0}}{\mathbb{E}S_{N_0}}, \mathcal{E}(1) \right) \leq 2d_W \left(\frac{S_{N_0}}{\mathbb{E}S_{N_0}}, \mathcal{E}(1) \right)^{1/2} \leq 2^{3/2} (pq^{-1}\gamma)^{1/2}$$

and

$$(1 - e^{-x/\mathbb{E}S_{N_0}}) - F_{S_{N_0}}(x) \leq \frac{1}{(\mathbb{E}S_{N_0})^2} \left(V(S_{N_0}) - (\mathbb{E}S_{N_0})^2 \right) = 2pq^{-1}\gamma \text{ for every } x \geq 0,$$

respectively. ■

Note that the above approximations are satisfactory only when $p \rightarrow 0$. In fact, Proposition 22 ascertains that $S_{N_0}/\mathbb{E}S_{N_0} \rightarrow_{st} \mathcal{E}(1)$ as $p \rightarrow 0$. For p relatively large though, the distribution of $S_{N_0}/\mathbb{E}S_{N_0}$ cannot be adequately approximated by $\mathcal{E}(1)$. This happens because we are trying to approximate a distribution having a point mass on 0 ($\Pr(S_{N_0} = 0) = \Pr(N_0 = 0) = p$) by the absolutely continuous $\mathcal{E}(\mathbb{E}S_{N_0})$. It is noteworthy that we could easily overcome this situation if X_i are HNBUE or HNWUE r.v.'s. More specifically, we have the following result.

Proposition 23 *If X_i are HNBUE or HNWUE r.v.'s then*

$$\zeta_2 \left(\frac{S_{N_0}}{\mathbb{E}S_{N_0}}, Y \right) = \frac{p}{q} |\rho_X|$$

where Y is a r.v. with distribution $\Pr(Y \leq x) = p + q(1 - e^{-qx})$, $x \geq 0$.

Proof. Let $Z = \sum_{i=1}^{N_0} Z_i$, where Z_i are i.i.d r.v.'s, independent of N_0 , such that $Z_i \sim \mathcal{E}(\mathbb{E}X)$. Since X_i are HNBUE r.v.'s then $X_i \preceq_{cx} Z_i$ and hence, $S_{N_0} = \sum_{i=1}^{N_0} X_i \preceq_{cx} \sum_{i=1}^{N_0} Z_i = Z$, which, by virtue of (4), leads to

$$\zeta_2(S_{N_0}, Z) = \frac{1}{2} \left(\text{Var} \left(\sum_{i=1}^{N_0} Z_i \right) - \text{Var} \left(\sum_{i=1}^{N_0} X_i \right) \right) = qp^{-1}(\mathbb{E}X)^2 \rho_X.$$

It can be easily verified that the distribution of Z is a mixture of the exponential $\mathcal{E}(\mathbb{E}X/p)$ and the point mass zero, i.e. $\Pr(Z \leq x) = p + q(1 - e^{-px/\mathbb{E}X})$ for $x \geq 0$ and $\Pr(Z \leq x) = 0$ for $x < 0$. Finally, we get that

$$\zeta_2 \left(\frac{S_{N_0}}{\mathbb{E}S_{N_0}}, \frac{Z}{\mathbb{E}S_{N_0}} \right) = \frac{1}{(\mathbb{E}S_{N_0})^2} \zeta_2(S_{N_0}, Z) = pq^{-1} \rho_X,$$

where $\Pr(Z/\mathbb{E}S_{N_0} \leq x) = \Pr(Z \leq xqp^{-1}\mathbb{E}X) = p + q(1 - e^{-qx})$. The proof for the case of HNWUE r.v.'s is analogous. ■

Hence, the distributions of S_{N_0} and Y are close to each other if $|\rho_X| = |1 - \mathbb{E}X^2/2(\mathbb{E}X)^2|$ is close to 0. Finally, we mention that similar results can be extracted for the geometric sum S_{N_0+1} .

References

- [1] BARBOUR, A.D., AND CHRYSAPHINO O. (2001) Compound Poisson approximation: A users guide. *The Annals of Applied Probability* **11**, 964-1002.
- [2] BARBOUR, A. D., HOLST, L. AND JANSON, S. (1992) *Poisson approximation*. Clarendon Press, Oxford.
- [3] BARLOW, R.E. AND PROSCHAN, F. (1981). *Statistical Theory of Reliability and Life Testing*. To Begin with, Silver Spring, MD.
- [4] BIRKEL T. (1988) On the convergence rate in the central limit theorem for associated processes. *The Annals of Probability* **16**, 1685-1698.
- [5] BIRKEL T. (1993) A functional central limit theorem for positively dependent random variables. *Journal of Multivariate Analysis* **44**, 314-320.
- [6] BOUTSIKAS, M.V. AND KOUTRAS, M.V. (2000) A bound for the distribution of the sum of discrete associated or negatively associated random variables. *The Annals of Applied Probability* **10**, 1137-1150.
- [7] BROWN, M. (1990) Error bounds for exponential approximations of geometric convolutions. *The Annals of Probability* **18**, 1388-1402.
- [8] BURTON, R., DABROWSKI, A. R. AND DEHLING, H. (1986) An invariance principle for weakly associated random vectors. *Stochastic Processes and their Applications* **23**, 301-306.
- [9] CHACON, R.V. AND WALSH, J.B. (1976) One-dimensional potential embedding. *Semin. Probab. X*, Univ. Strasbourg 1974/75, Lect. Notes Math. 511, 19-23.
- [10] COX., J. T. AND GRIMMETT G. (1984) Central limit theorems for associated random sequences and the percolation model. *The Annals of Probability* **12**, 514-528.
- [11] DABROWSKI, A.R. AND DEHLING, H. (1988) A Berry-Esseen theorem and a functional law of the iterated logarithm for weakly associated random variables. *Stochastic Processes and their Applications* **30**, 277-289.
- [12] DALEY D.J. (1988) Tight bounds on the exponential approximation of some aging distributions. *The Annals of Probability* **16**, 414-423.
- [13] DENUIT M. AND VAN BELLEGEM S. (2001) On the stop-loss and total variation distances between random sums. *Insurance: Mathematics and Economics* **53**, 153-165.
- [14] DENUIT M., DHAENE J. AND RIBAS, C. (2001) Does positive dependence between individual risks increase stop-loss premiums? *Insurance: Mathematics and Economics* **28**, 305-308.
- [15] DENUIT M., LEFEVRE C. AND SHAKED M. (1998) The s -convex orders among real random variables, with applications. *Mathematical Inequalities & Applications* **1**, 585-613.
- [16] DENUIT M., LEFEVRE C. AND UTEV S. (2002) Measuring the impact of dependence between claims occurrences. *Insurance: Mathematics and Economics* **30**, 1-19.

- [17] DHAENE J. AND GOOVAERTS M.J. (1996) Dependency of risks and stop-loss order. *ASTIN Bulletin* **26**, 201-212.
- [18] DHAENE J. AND GOOVAERTS M.J. (1997) On the dependency of risks in the individual life model. *Insurance: Mathematics and Economics* **19**, 243-253.
- [19] ESARY, J.D., PROSCHAN, F. AND WALKUP, D. (1967) Association of random variables with applications. *The Annals of Mathematical Statistics* **38**, 1466-1474.
- [20] FISHBURN P.C. (1976) Continua of stochastic dominance relations for bounded probability distributions. *Journal of Mathematical Economics* **3**, 295-311.
- [21] FISHBURN P.C. (1980a) Continua of stochastic dominance relations for unbounded probability distributions. *Journal of Mathematical Economics* **7**, 271-285.
- [22] FISHBURN P.C. (1980b) Stochastic dominance and moments of distributions. *Mathematics of Operations research* **5**, 94-100.
- [23] GERBER, H. (1981) *An Introduction to mathematical risk theory*. Huebner Foundation Monograph 8, Homewood, Illinois.
- [24] GERBER, H. (1984) Error bounds for the compound Poisson approximation. *Insurance: Mathematics and Economics* **3**, 191-194.
- [25] JOAG-DEV, K. (1983) Independence via uncorrelatedness under certain dependence structures. *The Annals of Probability* **11**, 1037-1041.
- [26] JOAG-DEV, K. AND PROSCHAN, F. (1983) Negative association of random variables, with applications. *The Annals of Statistics* **11**, 286-295.
- [27] KAAS, R. (1993) How to and how not to compute stop-loss premiums in practice. *Insurance: Mathematics And Economics* **13**, 241-254.
- [28] KARLIN, S. AND NOVIKOFF, A. (1963) Generalized convex inequalities. *Pacific Journal of Mathematics* **13**, 1251-1279.
- [29] LEFEVRE C. AND UTEV S. (1998) On order preserving properties of probability metrics. *Journal of Theoretical Probability* **11**, 907-920.
- [30] LEHMANN, E.L. (1966) Some concepts of dependence. *Ann. Math. Stat.* **37**, 1137-1153.
- [31] MACHINA M.J. AND PRATT J.W. (1997) Increasing Risk: Some Direct Constructions. *Journal of Risk and Uncertainty* **14**, 103-127.
- [32] MEILIJSON I. (1983) On the Azema-Yor stopping time. *Semin. de probabilites XVII*, Proc. 1981/82, Lect. Notes Math. 986, 225-226.
- [33] MULLER A. (1997) Stochastic orders generated by integrals: A unified study. *Advances in Applied probability* **29**, 414-428.
- [34] NEWMAN, C. M. (1980) Normal fluctuations and the FKG inequalities. *Communications in Mathematical Physics* **74**, 119-128.

- [35] NEWMAN, C. M. AND WRIGHT, A. L. (1981) An invariance principle for certain dependent sequences. *The Annals of Probability* **9**, 671-675.
- [36] DE PRIL, N. AND DHAENE, J. (1992) Error bounds for compound Poisson approximations of the individual risk model. *Astin Bulletin* **22**, 135-148.
- [37] RACHEV, S.T. (1991) *Probability metrics and the stability of stochastic models*, John Wiley, New York.
- [38] RACHEV S.T. AND RUSCHENDORF L. (1990) Approximation of sums by compound Poisson distributions with respect to stop-loss distances. *Advances in Applied Probability* **22**, 350-374.
- [39] SENATOV V.V. (1980) Uniform estimates of the rate of convergence in the multi-dimensional central limit theorem. *Theory of Probability and its Applications* **25**, 745-759.
- [40] SHAKED, M. AND SHANTHIKUMAR, J.G. (1994) *Stochastic orders and their applications*. Academic Press, New York.
- [41] SHAO, QI-MAN (2000) A comparison theorem on moment inequalities between negatively associated and independent random variables. *Journal of Theoretical Probability* **13**, 343-356.
- [42] STOYAN, D. (1983) *Comparison methods for queues and other stochastic models*. Ed. with rev. by Daryl J. Daley. (Transl. from the German). John Wiley & Sons Ltd., Chichester.
- [43] SZEKLI, R. (1995) *Stochastic Ordering and Dependence in Applied Probability*. Lecture notes in Statistics, Vol 97. Springer, Berlin.
- [44] TCHEN A.H. (1980) Inequalities for distributions with given marginals. *Ann. Prob.* **8**, 814-827.
- [45] WOOD T.E. (1983) A Berry-Esseen theorem for associated random variables. *Ann. Prob.* **11**, 1042-1047.
- [46] YU, H. (1985) An invariance principle for associated sequences of random variables. *J. Engrg. Math.* **2**, 55-60.
- [47] ZOLOTAREV, V.M. (1983) Probability metrics. *Th. Prob. Appl.* **28**, 278-302.