A coupling proof of convex ordering for compound distributions

Jean Bérard[*]   Nicolas Juillet[*]

Abstract

In this paper, we give an alternative proof of the fact that, when compounding a nonnegative probability distribution, convex ordering between the distributions of the number of summands implies convex ordering between the resulting compound distributions. Although this is a classical textbook result in risk theory, our proof exhibits a concrete coupling between the compound distributions being compared, using the representation of one-period discrete martingale laws as a mixture of the corresponding extremal measures.

1 Introduction

Let us first recall the definition of convex ordering: given two real-valued integrable random variables Y and Z, we say that $Y \prec_{\mathrm{cx}} Z$ if, for any convex function f, the inequality $\mathbb{E}f(Y) \leq \mathbb{E}f(Z)$ holds. Moreover, Strassen's condition ([17]) for convex ordering states that $Y \prec_{\mathrm{cx}} Z$ if and only if there exists a pair of random variables (A, B) such that

$$A \stackrel{d}{=} Y, \ B \stackrel{d}{=} Z, \text{ and } A = \mathbb{E}(B|\sigma(A)) \text{ a.s.} \tag{1.1}$$

We refer to e.g. [16, 13] for a detailed treatment of this and related notions.

Now consider an i.i.d. sequence of nonnegative random variables $\mathcal{X} = (X_i)_{i \geq 1}$, and two integer-valued random variables M, N, independent from the sequence $\mathcal{X}$. Assume that $\mathbb{E}(X_1) < +\infty$, $\mathbb{E}(M) < +\infty$, $\mathbb{E}(N) < +\infty$, and that a comparison between M and N holds with respect to the convex ordering: $M \prec_{\mathrm{cx}} N$. We then have the following comparison between the compound variables $X_1 + \cdots + X_M$ and $X_1 + \cdots + X_N$ with respect to the convex ordering:

$$X_1 + \cdots + X_M \prec_{\mathrm{cx}} X_1 + \cdots + X_N. \tag{1.2}$$

This is a classical result (see e.g. Theorem 4.A.9 in [16] or Theorem 4.3.6 in [13])[1], useful in the context of risk theory (see e.g. [10], chap. 7) where its interpretation is

---

[*]Université de Strasbourg, France. E-mail: jberard@math.unistra.fr,juillet@math.unistra.fr

[1]In [16, 13], the theorems are stated with respect to the increasing convex order, and involve two sequences of random variables instead of just one, so that (1.2) appears as a special case of these results. We refer to Section 4.3 for a discussion of how the more general case can be deduced from (1.2).

that compounding with a riskier frequency distribution leads to a riskier aggregated loss distribution.

The proof given in [16] is analytical in nature, and consists in showing that, given a non-decreasing convex function $f : [0, +\infty[\to \mathbb{R}$, the sequence $(u_n)_{n \geq 0}$ defined by $u_n = \mathbb{E}f(X_1 + \cdots + X_n)$ satisfies, for all $n \geq 0$, the condition $u_{n+2} - u_{n+1} \geq u_{n+1} - u_n$ (provided that u_{n+2}, u_{n+1}, u_n have a finite value)[2]. Here, we give an alternative proof of (1.2), based on a coupling between the two random variables being compared, which provides an explicit realization of Strassen's condition: we construct a pair of random variables (A, B) such that

$$A \stackrel{d}{=} X_1 + \cdots + X_M, \ B \stackrel{d}{=} X_1 + \cdots + X_N, \text{ and } A = \mathbb{E}(B|\sigma(A)) \text{ a.s.} \tag{1.3}$$

Our proof relies a representation result which we call a *diatomic representation* of convex ordering, stated as Theorem 2.1 in Section 2, where we review several approaches for proving the existence of this representation, including an explicit algorithm in the case of discrete distributions. Section 3 contains the coupling construction leading to (1.3) and the proof of (1.2). Finally, in Section 4, we discuss various extensions of these results.

2 Diatomic representation

Given two real numbers $x < y$, we define, for all z the normalized barycentric coordinates:
$$\alpha^{x,y}(z) = \frac{y - z}{y - x} \text{ and } \beta^{x,y}(z) = 1 - \alpha^{x,y}(z) = \frac{z - x}{y - x}.$$

In the case $x = y$, we extend the above definition by setting $\alpha^{x,y}(z) = 1$ and $\beta^{x,y}(z) = 0$. Whenever $x \leq z \leq y$, both $\alpha^{x,y}(z)$ and $\beta^{x,y}(z)$ lie in the interval $[0, 1]$, and z can be written as the convex combination of x and y:

$$z = \alpha^{x,y}(z) \cdot x + \beta^{x,y}(z) \cdot y.$$

Theorem 2.1. *Given two probability distributions μ, ν on $\mathbb{R}$ possessing a finite expectation, the comparison $\mu \prec_{cx} \nu$ holds if and only if there exists a triple of random variables (V_-, U, V_+) defined on the same probability space and such that:*

$$U \in [V_-, V_+] \tag{2.1}$$
$$Law(U) = \mu \tag{2.2}$$
$$\nu = \mathbb{E}\big[\alpha^{V_-, V_+}(U) \cdot \delta_{V_-} + \beta^{V_-, V_+}(U) \cdot \delta_{V_+}\big] \tag{2.3}$$

We call such a triple (V_-, U, V_+) a diatomic representation of the stochastic ordering $\mu \prec_{cx} \nu$.

A more concrete statement of (2.3) is that $Law(V) = \nu$, where V is a random variable whose conditional distribution with respect to V_-, U, V_+ is given by:

$$V = \begin{cases} V_- & \text{with probability } \alpha^{V_-, V_+}(U) \\ V_+ & \text{with probability } \beta^{V_-, V_+}(U) \end{cases}$$

Theorem 2.1 may not have been stated under this specific form in the mathematical literature, but its content is certainly not new. In the following subsections we review several ways of proving this result.

---

[2]The proof given in [13] is similar, using functions of the special form $f(x) = (x - t)^+$.

2.1 Proof of Theorem 2.1 via Choquet's and Douglas' theorems (for compactly supported measures ν)

We prove here the result only in the case of measures $\mu \prec_{\mathrm{cx}} \nu$ concentrated on some closed interval K.

Consider the following space of measures on $\mathbb{R}^2$:

$$S_K = \left\{ \pi \in \mathcal{M}_+(K^2) : \forall a \in \mathbb{R}, \forall b \in \mathcal{C}_b(\mathbb{R}), \iint a + b(u).(v-u)\mathrm{d}\pi(u,v) = a \right\}.$$

Here $\mathcal{M}_+(K^2)$ is the space of finite positive Borel measures on $\mathbb{R}^2$ with support contained in K^2, and $\mathcal{C}_b(\mathbb{R})$ the space of real-valued continuous and bounded functions on $\mathbb{R}$. In other words, denoting by F the space of functions $f_{a,b} : (u,v) \in \mathbb{R}^2 \to a + b(u)(v-u)$, the above definition reads:

$$S_K = \{\pi \in \mathcal{M}_+(K^2) : \forall f_{a,b} \in F, \iint f_{a,b}\mathrm{d}\pi = a\}.$$

It turns out that S_K is a non-empty space of probability measures, known as martingale measures in the literature since for $(X_1, X_2) \sim \pi \in S_K$, the process $(X_i)_{i=1,2}$ is a martingale on its natural filtration.

Our task amounts to proving that a measure $\pi \in S_K$ is represented as a mixture of 'triplet' measures of the form $\delta_z \otimes [\alpha\delta_x + \beta\delta_y]$, where $z = \alpha x + \beta y$ and α, β are nonnegative and satisfy $\alpha + \beta = 1$.

We admit without proof that S_K is convex and compact for the weak topology. Choquet's Theorem (see [3] or e.g. [14]) then implies that every measure in S_K can be represented as a mixture of extremal measures in S_K. So we shall be done as soon as we can prove that the extremal measures in S_K are triplet measures concentrated on S_K. Let η be such an extremal measure and μ, ν its marginals. We aim at proving that μ is supported on a single point and that ν is supported on at most two points. Striving for a contradiction suppose that μ is not a Dirac measure. Hence, there exists a Borel set E such that $\mu(E) \notin \{0,1\}$. We consider the $L^1(\eta)$ distance between $f : (u,v) \to \mathbb{1}_E(u)$ and the functions $f_{a,b} \in F$. Letting (U,V) be distributed according to η we find that[3]

$$\|f - f_{a,b}\|_{L^1(\eta)} \geq \mathbb{E}\left(\left|\mathbb{E}(f(U,V) - f_{a,b}(U,V)|U)\right|\right)$$
$$= \mathbb{E}\left(\left|\mathbb{1}_E(U) - a\right|\right)$$
$$= \mathbb{P}(U \in E)|1 - a| + \mathbb{P}(U \notin E)|a|$$
$$\geq \min(\mathbb{P}(U \in E), \mathbb{P}(U \notin E)).$$

Note that this lower bounds only depends on f. We have thus proved that F is not dense in $L^1(\eta)$. According to Douglas's theorem [4] (see also [15, Chapter V]) this in contradiction with the fact that η is extremal. Thus η is of type $\delta_{u_0} \otimes \nu$ where u_0 is the barycenter of ν, i.e $\int v\,\mathrm{d}\nu(v) = u_0$. Next, again striving for a contradiction suppose that there exists a partition $(A_i)_{i \in \{1,2,3\}}$ of $\mathbb{R}$ such that $\nu(A_i) > 0$ holds for every i. From Douglas's Theorem again the set F of functions is dense in $L^1(\eta)$. In particular any function $g_{c_1,c_2,c_3} : (u,v) \mapsto \sum_{i=1}^3 c_i \mathbb{1}_{A_i}(v)$ can be approximated in $L^1(\eta)$ by functions $f_{a,b}$. The linear map $(c_1,c_2,c_3) \mapsto (\int_{A_1} g_{c_1,c_2,c_3}\,\mathrm{d}\eta, \int_{A_2} g_{c_1,c_2,c_3}\,\mathrm{d}\eta, \int_{A_3} g_{c_1,c_2,c_3}\,\mathrm{d}\eta)$ is clearly linear and onto. It follows that the linear map

$$(a,b) \in \mathbb{R}^2 \mapsto \left(\int_{A_1} f_{a,b}\,\mathrm{d}\eta, \int_{A_2} f_{a,b}\,\mathrm{d}\eta, \int_{A_3} f_{a,b}\,\mathrm{d}\eta\right) \in \mathbb{R}^3$$

is onto as well, a contradiction. Therefore, extremal measures of S_K are of type $\delta_z \otimes [\alpha\delta_x + \beta\delta_y]$ where $z = \alpha x + \beta y$, as it was required.

---

[3]To go from the first to the second line, we use the fact that $\mathbb{E}(f_{a,b}(U,V)|U) = a$ a.s. since $\mathbb{E}(V\,|\,U) = U$ a.s.

2.2 Proof of Theorem 2.1 via Strassen's theorem (general case)

Another approach to proving Theorem 2.1 is to use Strassen's theorem instead of Choquet's and Douglas's theorems: there exists a kernel $k : \mathbb{R} \to \mathcal{P}(\mathbb{R})$ such that μ almost surely $k^u = k(u)$ has mean u and it holds $\mu \cdot k = \nu$. (Such kernels are known as dilations or martingale kernels in the literature.) Hence the mixture with weight μ of the measures $\delta_u \otimes k^u$ defines a probability measure π on $\mathbb{R}^2$ whose marginals are μ and ν. To complete the proof, it remains to check that each measure k^u can be represented as a mixture of diatomic measures with mean u. This last fact is a classical step in the proof of Skorokhod's representation theorem (see e.g. [5, Theorem 8.1.1]): every probability measure on $\mathbb{R}$ with mean u can be represented as a mixture of diatomic measures with mean u. See [9, §5.1] for another approach.

Remark 2.2. The search for martingale kernels $k : u \mapsto k^u$ is a key question in the field of martingale optimal transport. The first completely canonical method seems to be the left-curtain coupling by Beiglböck and Juillet [1] that is also of particular interest to us. Not only is k^u canonical but when μ is diffuse its kernels k^u are automatically diatomic (this also holds for the former coupling by Hobson and Neuberger [6] under more general assumptions). This is not the case if μ possesses atoms. However a quantile version of the left-curtain coupling is described in a second paper by the same authors [2] where the martingale measure π directly appears as a mixture over the set $[0, 1]$ of quantile levels $\omega \in [0, 1]$ of diatomic kernels $\delta_{z_\omega} \otimes (\alpha^{x_\omega, y_\omega}(z_\omega)\delta_{x_\omega} + \beta^{x_\omega, y_\omega}(z_\omega)\delta_{y_\omega})$ where z_ω is the ω-quantile of μ. Note that the same can be said of the recent coupling by Jourdain and Margheriti [8].

2.3 Algorithmic proof of Theorem 2.1 (for finitely supported μ, ν)

We now describe an explicit algorithmic construction, inspired by [1, 2], leading to a diatomic decomposition in the case where both μ and ν are finitely supported probability measures. This algorithm is used to produce the simulation shown in Fig. 1.

Let us write $\mu = \sum_{i=1}^p \mu(u_i)\delta_{u_i}$ and $\nu = \sum_{j=1}^q \nu(v_j)\delta_{v_j}$.

Initialization: $\mu_* \leftarrow \mu$, $\nu_* \leftarrow \nu$, $\theta_* \leftarrow 1$, $\mathcal{T} \leftarrow \emptyset$

Loop: Repeat the following steps while $\theta_* > 0$
Pick a triple (v_{j-}, u_i, v_{j+}) such that

    a. $\nu_*(v_{j-}) > 0$, $\mu_*(u_i) > 0$, $\nu_*(v_{j+}) > 0$
    b. $v_{j-} \leq u_i \leq v_{j+}$
    c. $\nu_*((v_j-, v_j+)) = 0$

$s \leftarrow \min\left(\mu_*(u_i), \nu_*(v_{j-})/\alpha^{v_j-, v_j+}(u_i), \nu_*(v_{j+})/\beta^{v_j-, v_j+}(u_i)\right)$
$\mu_* \leftarrow \mu_* - s\delta_{u_i}$
$\nu_* \leftarrow \nu_* - s\alpha^{v_j-, v_j+}(u_i)\delta_{v_{j-}} - s\beta^{v_j-, v_j+}(u_i)\delta_{v_{j+}}$
$\theta_* \leftarrow \theta_* - s$
$\mathcal{T} \leftarrow \mathcal{T} \cup \{((v_{j-}, u_i, v_{j+}), s)\}$

Result: return the set $\mathcal{T}$

The probability distribution of $(V-, U, V+)$ is then deduced from $\mathcal{T}$ as

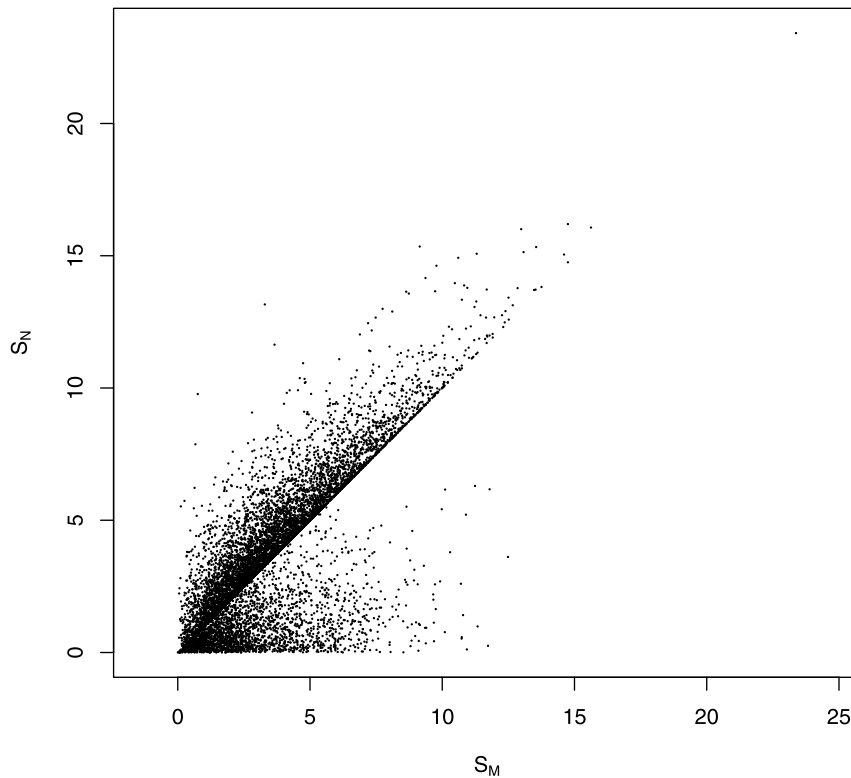$$\sum_{(v-, u, v+, s) \in \mathcal{T}} s\delta_{(v-, u, v+)}.$$

Figure 1: The figure shows 10^4 simulated pairs (A, B), with $\mu = \sum_{n=2}^{5} \frac{1}{4}\delta_n$ and $\nu = \frac{1}{4}\delta_1 + \frac{1}{4}\delta_3 + \sum_{n=4}^{6} \frac{1}{6}\delta_n$, $\mathrm{Law}(X_i) = \mathscr{E}(1)$, and the algorithm of Section 2.3 to produce the diatomic decomposition.

The reason why the above algorithms stops lies in the fact that the transformation performed on μ_* and ν_* keeps the comparison $\mu_* \prec_{\mathrm{cx}} \nu_*$ valid throughout the execution of the algorithm (see the proof of Lemma 2.8 in [1]), with μ_* and ν_* having equal total mass. As a consequence, as long as μ_* and ν_* do not have zero total mass, the comparaison $\mu_* \prec_{\mathrm{cx}} \nu_*$ ensures that a triple (v_{j-}, u_i, v_{j+}) satisfying conditions a.-b.-c. exists. Finally, since at each step at least one of the three numbers $\nu_*(v_{j-}), \mu_*(u_i), \nu_*(v_{j+})$ is set to zero, the total mass of both μ_* and ν_* must reach zero after a finite number of steps.

Remark 2.3. If, in the loop part of the algorithm, one systematically choses the unique triple (v_-, u, v_+) such that u is the leftmost point in the support of μ_* (such a choice is always possible, see the proof of Lemma 2.8 in [1]), the end-result of the algorithm is the so-called left-curtain coupling.

3 Coupling construction

We now describe the coupling construction leading to our proof of (1.2). Consider a triple (N_-, M', N_+) as in Theorem 2.1, with $\mu = \mathrm{Law}(M)$ and $\nu = \mathrm{Law}(N)$, and an i.i.d. sequence $\mathscr{X} = (X_i)_{i \geq 1}$ independent from (N_-, M', N_+). For all integer $k \geq 1$, we let $S_k = \sum_{i=1}^{k} X_i$, with the convention that $S_0 = 0$. Finally, we define two σ-algebras $\mathcal{F} = \sigma(M', N_-, N_+, S_{N_-}, S_{N_+})$ and $\mathcal{G} = \sigma(M', N_-, N_+, S_{M'}, S_{N_-}, S_{N_+})$. Note that $\mathcal{G} = \mathcal{F} \vee \sigma(S_{M'})$.

We start our construction by setting:

$$A = S_{M'}. \tag{3.1}$$

Next, we specify N' by the requirement that, conditional upon $\mathcal{G}$, the distribution of N' is:

$$\alpha^{S_{N_-}, S_{N_+}}(A) \cdot \delta_{N_-} + \beta^{S_{N_-}, S_{N_+}}(A) \cdot \delta_{N_+}, \tag{3.2}$$

and we let

$$B = S_{N'}. \tag{3.3}$$

Note that $\alpha^{S_{N_-}, S_{N_+}}(A)$ and $\beta^{S_{N_-}, S_{N_+}}(A)$ do indeed lie in the interval $[0, 1]$ thanks to the assumption that the random variables X_i are nonnegative, so that $S_{N_-} \leq A = S_{M'} \leq S_{N_+}$, since $N_- \leq M' \leq N_+$.

We now proceed to checking that all three properties listed in (1.3) are satisfied by the random variables A and B. From (3.1), it is immediate that A has the required distribution. Moreover, from the definition of α and β, one has the identity

$$\alpha^{S_{N_-}, S_{N_+}}(A) \cdot S_{N_-} + \beta^{S_{N_-}, S_{N_+}}(A) \cdot S_{N_+} = A,$$

which rewrites as:

$$\mathbb{E}[B|\mathcal{G}] = A \text{ a.s.,}$$

whence, taking the conditional expectation $\mathbb{E}(\cdot|\sigma(A))$ on both sides, and using the fact that $\sigma(A) \subset \mathcal{G}$ since $A = S_{M'}$ is $\mathcal{G}$-measurable,

$$\mathbb{E}[B|\sigma(A)] = A \text{ a.s.,}$$

as required by (1.3).

To conclude the proof, it remains to check that $B \stackrel{d}{=} S_N$. By construction, the conditional distribution of N' given $\mathcal{F}$ can be written as:

$$\mathbb{E}\left[\alpha^{S_{N_-}, S_{N_+}}(A) \,\Big|\, \mathcal{F}\right] \cdot \delta_{N_-} + \mathbb{E}\left[\beta^{S_{N_-}, S_{N_+}}(A) \,\Big|\, \mathcal{F}\right] \cdot \delta_{N_+}. \tag{3.4}$$

Now observe that, by symmetry, given integers $n_- \leq m \leq n_+$ such that $n_- < n_+$, we have:

$$\mathbb{E}\left[S_m - S_{n_-} \,\Big|\, \sigma(S_{n_-}, S_{n_+})\right] = \frac{m - n_-}{n_+ - n_-}(S_{n_+} - S_{n_-}) \text{ a.s.,}$$

from which we deduce that[4]

$$\mathbb{E}\left[\alpha^{S_{n_-}, S_{n_+}}(S_m) \,\Big|\, \sigma(S_{n_-}, S_{n_+})\right] = \alpha^{n_-, n_+}(m) \text{ a.s.}$$

and

$$\mathbb{E}\left[\beta^{S_{n_-}, S_{n_+}}(S_m) \,\Big|\, \sigma(S_{n_-}, S_{n_+})\right] = \beta^{n_-, n_+}(m) \text{ a.s.}$$

Since the sequence $\mathcal{X} = (X_i)_{i \geq 1}$ is independent from (N_-, M', N_+), the above identities imply that

$$\mathbb{E}\left[\alpha^{S_{N_-}, S_{N_+}}(S_{M'}) \,\Big|\, \mathcal{F}\right] = \alpha^{N_-, N_+}(M') \text{ a.s.}$$

and

$$\mathbb{E}\left[\beta^{S_{N_-}, S_{N_+}}(S_{M'}) \,\Big|\, \mathcal{F}\right] = \beta^{N_-, N_+}(M') \text{ a.s.}$$

Using (3.4), we see that the conditional distribution of N' given $\mathcal{F}$ is none but:

$$\alpha^{N_-, N_+}(M') \cdot \delta_{N_-} + \beta^{N_-, N_+}(M') \cdot \delta_{N_+}, \tag{3.5}$$

---

[4]In the case where $n_- = m = n_+$, these identities are still (obviously) valid.

so that the conditional distribution of $B = S_{N'}$ given $\mathcal{F}$ is:

$$\alpha^{N_-,N_+}(M') \cdot \delta_{S_{N_-}} + \beta^{N_-,N_+}(M') \cdot \delta_{S_{N_+}}. \tag{3.6}$$

Now let N'' denote a random variable independent from $\mathscr{X} = (X_i)_{i\geq 1}$ and whose conditional distribution given $\sigma(M', N_-, N_+)$ is precisely (3.5). From (2.3), we see that $N'' \stackrel{d}{=} N$. Since both N'' and N are assumed to be independent from $\mathscr{X} = (X_i)_{i\geq 1}$, we also have that $S_{N''} \stackrel{d}{=} S_N$. Moreover, (3.6) coincides with the conditional distribution of $S_{N''}$ given $\sigma(M', N_-, N_+, \mathscr{X})$ (hence given $\mathcal{F}$), so that, getting back to unconditional distributions, we have that:

$$B = S_{N'} \stackrel{d}{=} S_{N''} \stackrel{d}{=} S_N,$$

which concludes the proof.

4 Final remarks and extensions

4.1 Continuous time

We note that the coupling construction described in Section 3 can be extended in continuous time. For instance, let $\mathscr{N} = (N_t)_{t\geq 0}$ be a standard Poisson process and $S \prec_{\mathrm{cx}} T$ nonnegative integrable random variables independent from $\mathscr{N}$. Then one has $N_S \prec_{\mathrm{cx}} N_T$, and it is straightforward to extend our approach to define a pair of random variables (A, B) such that

$$A \stackrel{d}{=} N_S, \ B \stackrel{d}{=} N_T, \text{ and } A = \mathbb{E}(B|\sigma(A)) \text{ a.s.} \tag{4.1}$$

The same approach still works in exactly the same way if we consider an integrable subordinator instead of a Poisson process.

4.2 Exchangeable random variables

If the sequence of random variables $(X_i)_{i\geq 1}$ is assumed to be exchangeable instead of i.i.d., the coupling construction described in Section 3 works in exactly the same way. (The classical proof found in [16, 13] also works in this case.) Note that, in the case of an infinite exchangeable sequence of random variables, one can directly deduce (1.2) from the i.i.d. case, using the De Finetti representation of such a sequence as a mixture of (distributions of) i.i.d. sequences, and the characterization of (1.2) through the inequality

$$\mathbb{E}f(X_1 + \cdots + X_M) \leq \mathbb{E}f(X_1 + \cdots + X_N) \tag{4.2}$$

for all convex functions f. On the other hand, if M and N are assumed to have finite support, say $\{0, 1, \ldots, q\}$, and one considers a finite exchangeable sequence of random variables $X_1, \ldots, X_q$, the extension of De Finetti's theorem to this case (see [12, 7]) leads in general to a signed mixture of i.i.d. sequences, so one cannot integrate the inequality (4.2) with respect to the mixing measure in order to directly deduce (1.2).

4.3 Increasing convex ordering

Assume that a comparison between M and N holds with respect to the increasing convex ordering[5]: $M \prec_{\mathrm{icx}} N$. We then have the following modified version of (1.2):

$$X_1 + \cdots + X_M \prec_{\mathrm{icx}} X_1 + \cdots + X_N. \tag{4.3}$$

---

[5]The definition is similar to convex ordering, with the class of convex functions replaced by the class of non-decreasing convex functions.

To deduce (4.3) from (1.2), we note that the comparison $M \prec_{\mathrm{icx}} N$ implies that there exists an integer-valued[6] random variable N_0 such that $M \prec_{\mathrm{st}} N_0$ and $N_0 \prec_{\mathrm{cx}} N$, where $\prec_{\mathrm{st}}$ denotes the usual stochastic ordering. Given an i.i.d. sequence $\mathscr{X} = (X_i)_{i \geq 1}$ independent from M, N_0, N, the fact that the X_is are nonnegative random variables, combined with $M \prec_{\mathrm{st}} N_0$, yields the comparison $X_1 + \cdots + X_M \prec_{\mathrm{st}} X_1 + \cdots + X_{N_0}$. Then, using (1.2), we deduce that $X_1 + \cdots + X_{N_0} \prec_{\mathrm{cx}} X_1 + \cdots + X_N$, so that (4.3) is proved.

Now consider two sequences $\mathscr{X} = (X_i)_{i \geq 1}$ and $\mathscr{Y} = (Y_i)_{i \geq 1}$ of i.i.d. nonnegative random variables, and, in addition to $M \prec_{\mathrm{icx}} N$, assume that $X_i \prec_{\mathrm{icx}} Y_i$. We then have the following extension of (4.3):

$$X_1 + \cdots + X_M \prec_{\mathrm{icx}} Y_1 + \cdots + Y_N, \tag{4.4}$$

which is a straightforward consequence of (4.3) and of the fact that the increasing convex ordering is preserved by convolution.

4.4 Counterexample to (1.2) when the X_i are not positive random variables

We let $M \equiv 1$ and $N \sim \frac{1}{2}(\delta_0 + \delta_2)$, so that the condition $M \prec_{\mathrm{cx}} N$ is satisfied. Then let X_1, X_2 be i.i.d. random variables, independent from N (and M), with $X_i \sim \frac{1}{2}(\delta_{-1} + \delta_1)$. We find that $S_M \sim \frac{1}{2}(\delta_{-1} + \delta_1)$, while $S_N \sim \frac{1}{8}(\delta_{-2} + 6\delta_0 + \delta_2)$. Hence $\frac{1}{2} = \mathbb{E}(|S_N|) < \mathbb{E}(|S_M|) = 1$ so that $S_M \prec_{\mathrm{cx}} S_N$ cannot hold.

References

[1] Mathias Beiglböck and Nicolas Juillet. On a problem of optimal transport under marginal martingale constraints. *Ann. Probab.*, 44(1):42–106, 2016. MR-3456332

[2] Mathias Beiglböck and Nicolas Juillet. Shadow couplings. *ArXiv e-prints*, page arXiv:1609.03340, September 2016.

[3] Gustave Choquet. Existence des représentations intégrales au moyen des points extrémaux dans les cônes convexes. *C. R. Acad. Sci. Paris*, 243:699–702, 1956. MR-0080264

[4] Ronald G. Douglas. On extremal measures and subspace density. *Michigan Math. J.*, 11:243–246, 1964. MR-0185427

[5] Rick Durrett. *Probability—theory and examples*, volume 49 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge, 2019. Fifth edition. MR-3930614

[6] David Hobson and Anthony Neuberger. Robust bounds for forward start options. *Math. Finance*, 22(1):31–56, 2012. MR-2881879

[7] Svante Janson, Takis Konstantopoulos, and Linglong Yuan. On a representation theorem for finitely exchangeable random vectors. *J. Math. Anal. Appl.*, 442(2):703–714, 2016. MR-3504021

[8] Benjamin Jourdain and William Margheriti. A new family of one dimensional martingale couplings. *arXiv e-prints*, page arXiv:1808.01390, Aug 2018.

[9] Nicolas Juillet. Peacocks parametrised by a partially ordered set. In *Séminaire de Probabilités XLVIII*, volume 2168 of *Lecture Notes in Math.*, pages 13–32. Springer, Cham, 2016. MR-3618125

[10] Rob Kaas, Marc Goovaerts, Jan Dhaene, and Michel Denuit. *Modern actuarial risk theory: using R*, volume 128. Springer Science & Business Media, 2008.

---

[6] The existence of a random variable N_0 such that $M \prec_{\mathrm{st}} N_0$ and $N_0 \prec_{\mathrm{cx}} N$ is a classical and easily proved decomposition result for the increasing convex order. That N_0 can, in addition, be chosen to be integer-valued is less standard. One possible proof of this fact is that, when $\mu \prec_{\mathrm{icx}} \nu$, there exists a kernel $k : \mathbb{R} \to \mathcal{P}(\mathbb{R})$ such that μ almost surely $k^u = k(u)$ has mean $\geq u$ and it holds $\mu \cdot k = \nu$. In turn, k^u can be written as a mixture $k^u = \alpha_u k_1^u + (1 - \alpha_u) k_2^u$, where μ almost surely $\alpha_u \in [0, 1]$, k_1^u has mean u, and the support of k_2^u is contained in $[u, +\infty[$. For $k'(u) = \alpha_u \delta_u + (1 - \alpha_u) k_2^u$ we have $\mu \prec_{\mathrm{st}} (\mu \cdot k') \prec_{\mathrm{cx}} \nu$ and a coupling of the corresponding random variables can also be easily deduced. Alternatively for another proof one can consider Kellerer's kernels defined in [11, §2.1] for connecting $\mu \prec_{\mathrm{icx}} \nu$.

[11] Hans G. Kellerer. Markov-Komposition und eine Anwendung auf Martingale. *Math. Ann.*, 198:99–122, 1972. MR-0356250

[12] G. Jay Kerns and Gábor J. Székely. De Finetti's theorem for abstract finite exchangeable sequences. *J. Theoret. Probab.*, 19(3):589–608, 2006. MR-2280511

[13] Alfred Müller and Dietrich Stoyan. *Comparison methods for stochastic models and risks.* Wiley Series in Probability and Statistics. John Wiley & Sons, Ltd., Chichester, 2002. MR-1889865

[14] Robert R. Phelps. *Lectures on Choquet's theorem*, volume 1757 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, second edition, 2001. MR-1835574

[15] Daniel Revuz and Marc Yor. *Continuous martingales and Brownian motion*, volume 293 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, third edition, 1999. MR-1725357

[16] Moshe Shaked and J. George Shanthikumar. *Stochastic orders.* Springer Series in Statistics. Springer, New York, 2007. MR-2265633

[17] Volker Strassen. The existence of probability measures with given marginals. *Ann. Math. Statist.*, 36:423–439, 1965. MR-0177430

Electronic Journal of Probability
Electronic Communications in Probability

Advantages of publishing in EJP-ECP

- Very high standards

- Free for authors, free for readers

- Quick publication (no backlog)

- Secure publication (LOCKSS[1])

- Easy interface (EJMS[2])

Economical model of EJP-ECP

- Non profit, sponsored by IMS[3], BS[4] , ProjectEuclid[5]

- Purely electronic

Help keep the journal free and vigorous

- Donate to the IMS open access fund[6] (click here to donate!)

- Submit your best articles to EJP-ECP

- Choose EJP-ECP over for-profit journals

---

[1] LOCKSS: Lots of Copies Keep Stuff Safe http://www.lockss.org/
[2] EJMS: Electronic Journal Management System http://www.vtex.lt/en/ejms.html
[3] IMS: Institute of Mathematical Statistics http://www.imstat.org/
[4] BS: Bernoulli Society http://www.bernoulli-society.org/
[5] Project Euclid: https://projecteuclid.org/
[6] IMS Open Access Fund: http://www.imstat.org/publications/open.htm