

# Does the Convex Order Between the Distributions of Linear Functionals Imply the Convex Order Between the Probability Distributions Over $\mathbb{R}^d$ ?

Iosif Pinelis  
Michigan Technological University  
Houghton, Michigan USA  
Email: ipinelis@mtu.edu

**Abstract.** It is shown that the convex order between the distributions of linear functionals does not imply the convex order between the probability distributions over  $\mathbb{R}^d$  if  $d \geq 2$ . This stands in contrast with the well-known fact that any probability distribution in  $\mathbb{R}^d$ , for any  $d \geq 1$ , is determined by the corresponding distributions of linear functionals. By duality, it follows that, for any  $d \geq 2$ , not all convex functions from  $\mathbb{R}^d$  to  $\mathbb{R}$  can be represented as the limits of sums  $\sum_{i=1}^k g_i \circ \ell_i$  of convex functions  $g_i$  of linear functionals  $\ell_i$  on  $\mathbb{R}^d$ .

Let  $\mu$  and  $\nu$  be probability measures over  $\mathbb{R}^d$  with

$$\int_{\mathbb{R}^d} \|x\| \mu(dx) < \infty \quad \text{and} \quad \int_{\mathbb{R}^d} \|x\| \nu(dx) < \infty, \quad (1)$$

where  $\|\cdot\|$  denotes the Euclidean norm. It is said that  $\mu$  is dominated by  $\nu$  in the convex order sense and written  $\mu \preceq \nu$  if

$$\int_{\mathbb{R}^d} f d\mu \leq \int_{\mathbb{R}^d} f d\nu \quad (2)$$

for all convex functions  $f: \mathbb{R}^d \rightarrow \mathbb{R}$ . Note that, given (1), the integrals in (2) always exist for any convex function  $f: \mathbb{R}^d \rightarrow \mathbb{R}$  but may take the value  $\infty$ —because for any such  $f$  there is an affine function  $a: \mathbb{R}^d \rightarrow \mathbb{R}$  such that  $f \geq a$ .

For any affine function  $a: \mathbb{R}^d \rightarrow \mathbb{R}$ , both  $a$  and  $-a$  are convex functions. So, if  $\mu \preceq \nu$ , then the barycenters  $x_\mu := \int_{\mathbb{R}^d} x \mu(dx)$  and  $x_\nu := \int_{\mathbb{R}^d} x \nu(dx)$  of  $\mu$  and  $\nu$  are the same.

Informally, the convex order relation  $\mu \preceq \nu$  means that, while the barycenters  $x_\mu$  and  $x_\nu$  of the distributions  $\mu$  and  $\nu$  are the same, the distribution  $\nu$  is more spread out than  $\mu$ . For instance, one has  $\delta_y \prec \nu$  for any  $y \in \mathbb{R}^d$  and any probability measure  $\nu$  with barycenter  $y$ , where  $\delta_y$  is the Dirac probability measure supported on the singleton set  $\{y\}$ .

The convex order has been widely studied and/or used; see, e.g., Jovan Karamata [1], David Blackwell [2], Paul-André Meyer [3], and Robert Ralph Phelps [4]. For stochastic orders in general, see Moshe Shaked and J. George Shanthikumar [5]. The notion of the decreasing concave order—which is a reverse of the increasing convex order and is also known as the second-order stochastic dominance—was introduced by Michael Rothschild and Joseph Eugene Stiglitz [6] and has been very widely used in economics literature afterwards.

For any probability measure  $\lambda$  over  $\mathbb{R}^d$  and any  $v \in \mathbb{R}^d$ , let  $\lambda_v$  denote the “ $v$ -projection” of  $\lambda$ —that is, the pushforward of  $\lambda$  under the linear map

$$\mathbb{R}^d \ni x \mapsto p_v(x) := v \cdot x \in \mathbb{R}, \quad (3)$$

where  $\cdot$  stands for the dot product.

It is well known that any probability measure  $\lambda$  over  $\mathbb{R}^d$  is determined by the family  $(\lambda_v)_{v \in \mathbb{R}^d}$  of its one-dimensional projections; see, e.g., [7, Theorem 26.2]. Using this result, one can establish its extension, due to Cramér and Wold [8], stating that the weak convergence of multivariate distributions is determined by the weak convergence of their one-dimensional projections; see also, e.g., [7, Theorem 29.4].

One may then ask whether the convex order is similarly determined by the one-dimensional projections—that is, whether

$$\mu \preceq \nu \stackrel{(?)}{\iff} (\mu_v \preceq \nu_v \text{ for all } v \in \mathbb{R}^d). \quad (4)$$

The implication  $\implies$  in (4) is obvious, because the composition  $g \circ p_v$  is convex for all convex functions  $g: \mathbb{R} \rightarrow \mathbb{R}$  and all  $v \in \mathbb{R}^d$ ; of course, for the same reason, an appropriate version of this implication holds for any real topological vector space in place of  $\mathbb{R}^d$ .

So, the question is only about the implication  $\impliedby$  in (4). This implication is trivially true if  $d = 1$ . However, we will show that the answer to the question is negative if  $d \geq 2$ . In fact, a counterexample to the implication  $\impliedby$  in (4) will be given explicitly by two probability measures  $\mu$  and  $\nu$  that have finite support sets and take rational values. Of course, it follows that the answer to the corresponding question for any real topological vector space of dimension  $d \geq 2$  in place of  $\mathbb{R}^d$  will also be negative.

Indeed, already for  $d = 2$ , let

$$\nu = \frac{1}{6} (3\delta_{(x+y+z)/3} + \delta_x + \delta_y + \delta_z), \quad \mu = \frac{1}{3} (\delta_{(x+y)/2} + \delta_{(y+z)/2} + \delta_{(x+z)/2}) \quad (5)$$

for some  $x, y, z$  in  $\mathbb{R}^2$ . Then  $\mu \preceq \nu$  means that

$$3f\left(\frac{x+y+z}{3}\right) + f(x) + f(y) + f(z) \geq 2f\left(\frac{x+y}{2}\right) + 2f\left(\frac{y+z}{2}\right) + 2f\left(\frac{x+z}{2}\right) \quad (6)$$

for all convex  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ . Similarly, for each  $v \in \mathbb{R}^2$ , the relation  $\mu_v \preceq \nu_v$  means that (6) holds for all functions  $f$  of the form  $g \circ p_v$ , where  $g$  is any convex function from  $\mathbb{R}$  to  $\mathbb{R}$  and  $p_v$  is as defined in (3). So, for each  $v \in \mathbb{R}^2$ , the relation  $\mu_v \preceq \nu_v$  means that

$$3g\left(\frac{r+s+t}{3}\right) + g(r) + g(s) + g(t) \geq 2g\left(\frac{r+s}{2}\right) + 2g\left(\frac{s+t}{2}\right) + 2g\left(\frac{r+t}{2}\right) \quad (7)$$

for  $r = p_v(x)$ ,  $s = p_v(y)$ , and  $t = p_v(z)$  and all convex functions  $g: \mathbb{R} \rightarrow \mathbb{R}$ .

In turn, inequality (7) does hold for all such convex functions  $g$  and all real  $r, s, t$ —being the simplest instance of an inequality due to Tiberiu Popoviciu [9]; see also e.g. [10, Theorem 1.1.8] and [11, p. 74]. Inequality (7) is also a special case (with  $p = q = r = 1/3$ ) of [12, inequality (6.2)].

It is easy to prove (7) directly as well. To do this, let us first recall the definition and a basic characterization of majorization. For a vector  $x = (x_1, \dots, x_n)$  in  $\mathbb{R}^n$ , let  $(x_{[1]}, \dots, x_{[n]})$  be the nonincreasing rearrangement of the  $x_i$ 's, so that  $x_{[1]} \geq \dots \geq x_{[n]}$ . Let us then say that a vector  $x = (x_1, \dots, x_n) \in \mathbb{R}^n$  is majorized by a vector  $y =$

$(y_1, \dots, y_n) \in \mathbb{R}^n$  (and write  $x \prec y$ ) if  $\sum_{i=1}^k x_{[i]} \leq \sum_{i=1}^k y_{[i]}$  for all  $k = 1, \dots, n-1$  and  $\sum_{i=1}^n x_{[i]} = \sum_{i=1}^n y_{[i]}$ —see e.g. [13, Definition 1.A.1].

A most important characterization of majorization is as follows: For vectors  $x$  and  $y$  as above, one has  $x \prec y$  if and only if  $\sum_{i=1}^n g(x_i) \leq \sum_{i=1}^n g(y_i)$  for all convex functions  $g: \mathbb{R} \rightarrow \mathbb{R}$  (cf., e.g., [13, Proposition 4.B.1], where it is additionally required that  $g$  be continuous; however, any convex function from  $\mathbb{R}$  to  $\mathbb{R}$  is automatically continuous—see, e.g., [14, Corollary 10.1.1]).

So, to prove (7), it suffices to check that

$$\left(\frac{r+s}{2}, \frac{r+s}{2}, \frac{s+t}{2}, \frac{s+t}{2}, \frac{r+t}{2}, \frac{r+t}{2}\right) \prec \left(\frac{r+s+t}{3}, \frac{r+s+t}{3}, \frac{r+s+t}{3}, r, s, t\right) \quad (8)$$

for any real  $r, s, t$ . To do this, first note that here, in view of permutation symmetry, without loss of generality  $r \leq s \leq t$ . Also, in view of the reflection symmetry  $u \leftrightarrow -u$ , without loss of generality  $s \leq \frac{r+t}{2}$  and hence

$$r \leq \frac{r+s}{2} \leq s \leq \frac{r+s+t}{3} \leq \frac{r+t}{2} \leq \frac{s+t}{2} \leq t.$$

Now the majorization (8) is easy to check just by the definition.

So, for any  $x, y, z$  in  $\mathbb{R}^2$ , we do have  $\mu_v \preceq \nu_v$  for all  $v \in \mathbb{R}^2$ .

However, (6) fails to hold if, e.g.,

$$f(w) = \max(0, \xi_1, \xi_2) \text{ for all } w = (\xi_1, \xi_2) \in \mathbb{R}^2 \quad (9)$$

and

$$x = (0, -1), \quad y = (-1, 0), \quad z = (2, 2). \quad (10)$$

So,  $\mu \not\preceq \nu$ . ■

**Remark 1.** One may note that in this counterexample  $\nu$  is the uniform distribution on the multiset consisting of the vertices  $x, y, z$  of a triangle (each vertex taken with multiplicity 1) and the barycenter  $\frac{x+y+z}{3}$  (taken with multiplicity 3), whereas  $\mu$  is the uniform distribution on the set of the midpoints of the sides of that triangle.

**Remark 2.** Since  $\mathbb{R}^2$  can be linearly embedded into  $\mathbb{R}^d$  for any natural  $d \geq 2$ , clearly (4) fails to hold for any such  $d$ .

Simple duality arguments lead to the following corollary.

**Corollary 3.** *For any integer  $d \geq 2$ , there is a convex function  $f: \mathbb{R}^d \rightarrow \mathbb{R}$  that is not in the closed (say, with respect to the topology of pointwise convergence) convex hull  $\overline{\text{conv } F}$  of the set  $F$  of all functions of the form  $g \circ p_v$ , where  $g: \mathbb{R} \rightarrow \mathbb{R}$  is a convex function and  $v \in \mathbb{R}^d$ .*

Indeed, as in Remark 2, here without loss of generality  $d = 2$ . Let then  $\mu$  and  $\nu$  be as in (5) with  $x, y, z$  as in (10). Then, as was shown above, (2) holds for all  $f \in F$  and hence for all  $f \in \overline{\text{conv } F}$ . It was also shown that (2) does not hold for the convex function  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$  defined by (9). So, the latter function  $f$  is not in  $\overline{\text{conv } F}$ . ■

---

This note is related to the previous one [15], sharing with it references [10, 11].

---

**Disclosure of interest** No conflict of interest to declare.

**Disclosure of funding** No funding was received for this work.

## REFERENCES

1. Karamata J. Sur une inégalité relative aux fonctions convexes. *Publ Math Univ Belgrade*. 1932;1:145-8.
2. Blackwell D. Equivalent comparisons of experiments. *Ann Math Stat*. 1953;24:265-72.
3. Meyer PA. Probability and potentials; 1966.
4. Phelps RR. Lectures on Choquet's theorem. D. Van Nostrand Co., Inc., Princeton, N.J.-Toronto, Ont.-London; 1966.
5. Shaked M, Shanthikumar JG. Stochastic orders. Springer Series in Statistics. New York: Springer; 2007. Available from: <http://dx.doi.org/10.1007/978-0-387-34675-5>.
6. Rothschild M, Stiglitz JE. Increasing risk: I. A definition. *Journal of Economic Theory*. 1970;2(3):225-43. Available from: <https://www.sciencedirect.com/science/article/pii/0022053170900384>.
7. Billingsley P. Probability and measure. 3rd ed. Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York; 1995. A Wiley-Interscience Publication.
8. Cramér H, Wold H. Some Theorems on Distribution Functions. *Journal of the London Mathematical Society*. 1936;s1-11(4):290-4. Available from: <https://londmathsoc.onlinelibrary.wiley.com/doi/abs/10.1112/jlms/s1-11.4.290>.
9. Popoviciu T. Sur certaines inégalités qui caractérisent les fonctions convexes. *An Ști Univ "Al I Cuza" Iași Secț I a Mat (NS)*. 1965;11B:155-64.
10. Niculescu CP, Persson LE. Convex functions and their applications. CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC, 23. Springer, New York; 2006. A contemporary approach. Available from: <http://dx.doi.org/10.1007/0-387-31077-0>.
11. Fechner W. Hlawka's functional inequality. *Aequationes Math*. 2014;87(1-2):71-87. Available from: <http://dx.doi.org/10.1007/s00010-012-0178-2>.
12. Pečarić JE, Proschan F, Tong YL. Convex functions, partial orderings, and statistical applications. vol. 187 of *Mathematics in Science and Engineering*. Academic Press, Inc., Boston, MA; 1992.
13. Marshall AW, Olkin I, Arnold BC. Inequalities: theory of majorization and its applications. 2nd ed. Springer Series in Statistics. New York: Springer; 2011.
14. Rockafellar RT. Convex analysis. Princeton Landmarks in Mathematics. Princeton, NJ: Princeton University Press; 1997. Reprint of the 1970 original, Princeton Paperbacks.
15. Pinelis I. Explicit Additive Decomposition of Norms on  $\mathbb{R}^2$ . *Amer Math Monthly*. 2016;123(5):491-6. Available from: <http://dx.doi.org/10.4169/amer.math.monthly.123.5.491>.