Stochastic ordering of variability measure
estimators

Stochastic ordering of variability measure estimators

Juan Baz
Department of Statistics and I.O. and Didactics of Mathematics
University of Oviedo, Spain
bazjuan@uniovi.es

Franco Pellerey
Dipartimento di Scienze Matematiche
Politecnico di Torino, Italy
franco.pellerey@polito.it

Irene Díaz
Department of Computer Science
University of Oviedo, Spain
sirene@uniovi.es

Susana Montes
Department of Statistics and I.O. and Didactics of Mathematics
University of Oviedo, Spain
montes@uniovi.es

January 20, 2024

Abstract

Variability measures, such as the variance or the Gini mean difference, are widely used to summarize the dispersion of random variables. In the statistical setting, it is quite natural to assume that if a random quantity has more variability, then the estimators of its variability measures should be greater in some stochastic sense. Stochastic orders can be used to give inequalities in this regard, confirming, or not, this suitable property. This paper is devoted to the stochastic comparison of some variability measure estimators; conditions such that some of these estimators are comparable in the usual stochastic order and in the increasing convex order whenever the involved random variables have different variability are provided. Special attention is devoted to the cases of sample variance and Gini mean differences, and to the case of simple random samples.

Keywords: Stochastic orders, Variability measures, Sample variance, Gini mean difference.

AMS 2020 Subject Classification: Primary 60E15, 62D99 Secondary 94A20

1 Introduction

Given a random quantity X whose distribution represents, for example, the distribution of a character of a given population, location and variability measures are often used to summarize its behavior. The most commonly considered variability measures include the variance, the range, the interquantile range and the Gini mean difference (see, e.g., Rohatgi and Saleh, 2015, Sordo et al., 2016 and La Haye and Zizler, 2019). Well established estimators of these variability measures based on samples of observations of X, such as the sample variance, are commonly considered in the statistical setting. For these estimators, it is quite natural to assume that if the random variable X increases variability, then they should correspondingly increase in some stochastic sense.

Consider for example two random quantities X and Y, and let $\mathbf{X} = \{X_1, X_2, \ldots, X_n\}$ and $\mathbf{Y} = \{Y_1, Y_2, \ldots, Y_n\}$ be two corresponding random samples (random vectors with independent and identically distributed components) of the same size n. Assume that $Var[X] \le Var[Y]$. Then it is well known that the expectations of the sample variances

$$S_X^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2 = \frac{1}{2n(n-1)}\sum_{i=1}^{n}\sum_{j=1}^{n}(X_i - X_j)^2 \tag{1.1}$$

and

$$S_Y^2 = \frac{1}{n-1}\sum_{i=1}^{n}(Y_i - \overline{Y})^2 = \frac{1}{2n(n-1)}\sum_{i=1}^{n}\sum_{j=1}^{n}(Y_i - Y_j)^2$$

are also ordered, i.e., $E[S_X^2] \le E[S_Y^2]$, since sample variances are unbiased.

However, one may wonder if they are also ordered in the usual stochastic order ($S_X^2 \le_{st} S_Y^2$), i.e. if the cumulative distribution function of S_X^2 is pointwise greater than the cumulative distribution function S_Y^2 (see next section for the formal definition of the usual stochastic order). Apart from obvious theoretical motivations, there are several reasons why a stochastic inequality like $S_X^2 \le_{st} S_Y^2$ can be more useful than an inequality between the corresponding expected values. For example, suppose one wants to determine the confidence intervals for the means of X and Y having two large samples of the same size n for both populations. Fixed the confidence level $1-\alpha$ and denoting the ranges of the two intervals with $R_\alpha(X) = 2z_{1-\alpha/2}\sqrt{S_X^2/n}$ and $R_\alpha(Y) = 2z_{1-\alpha/2}\sqrt{S_Y^2/n}$ (being the $z_{1-\alpha/2}$ the quantile of order $1-\alpha/2$ from the standard normal distribution), thanks to the closure with respect to increasing functions of the usual stochastic order, from $S_X^2 \le_{st} S_Y^2$ follows $R_\alpha(X) \le_{st} R_\alpha(Y)$, i.e., the first interval has higher probabilities to be smaller than the second one. An analogous conclusion cannot instead be obtained from the simple comparison $E[S_X^2] \le E[S_Y^2]$, not even in terms of comparison between the expected values of the two ranges, given that from $E[S_X^2] \le E[S_Y^2]$ does not necessarily follow that $E[R_\alpha(X)] \le E[R_\alpha(Y)]$.

Unfortunately, the condition $Var[X] \le Var[Y]$ is not sufficient for the comparison $S_X^2 \le_{st} S_Y^2$, as illustrated in the following example.

Example 1.1. *Let X_1, X_2 be two independent random variables with standard uniform distribution and Y_1, Y_2 two independent random variables with exponential with parameter $\lambda = 3$ distribution, so that $Var[X_1] = \frac{1}{12}$ and $Var[Y_1] = \frac{1}{9}$. However, is not true that $S_X^2 \le_{st} S_Y^2$. In fact, note that $P[(X_1 - X_2)^2 \le x] = 1 - (1-\sqrt{x})^2,\ x \in [0,1]$ and $P[(Y_1 - Y_2)^2 \le x] = 1 - e^{-\lambda\sqrt{x}},\ x \in [0,\infty]$. If*

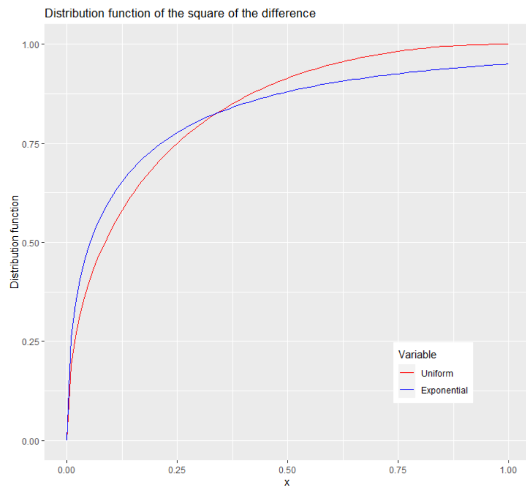

Figure 1: Distribution functions of the square of the difference of two independent random variables with standard uniform distribution and exponential distribution with parameter $\lambda = 3$.

we represent both distribution functions, see Figure 1, they cross. Noticing that $S_X^2 = \frac{1}{4}(X_1 - X_2)^2$ *and* $S_Y^2 = \frac{1}{4}(Y_1 - Y_2)^2$ *it is concluded that* $S_X^2 \not\leq_{st} S_Y^2$ *even with* $Var[X_1] \leq Var[Y_1]$ *being true.*

What has just been shown with a counterexample also applies to other kinds of estimators of variability or dispersion indices, such as for example the Gini mean difference, defined as $GMD(X) = E[|X_1 - X_2|]$ where X_1 and X_2 are independent copies of X, whose classical estimator is

$$G_n(X) = \frac{1}{n(n-1)} \sum_{i=1}^{n} \sum_{j=1}^{n} |X_i - X_j|. \tag{1.2}$$

With a direct calculation as the one described in Example 1.1, one can easily verify that the inequality $GMD(X) \leq GMD(Y)$ does not implies $G_n(X) \leq_{st} G_n(Y)$. However, one can be interested in the inequality $G_n(X) \leq_{st} G_n(Y)$ for different purposes. For example, recall that the Gini mean difference is a widely used measure of inequality in many fields such as economics and social sciences (the greater the GMD, the greater the inequality in the population). Given two populations X and Y with unknown distributions, suppose one is interested in evaluating whether $GMD(X)$ or $GMD(Y)$ are greater than a fixed threshold s. Then, from the inequality $G_n(X) \leq_{st} G_n(Y)$ one can affirms that $P[G_n(X) > s] \leq P[G_n(Y) > s]$, i.e., that the empirical version of the Gini mean difference has higher probability to be above the threshold for the population X rather than for Y, while the same can not be affirmed by the fact that $GMD(X) \leq GMD(Y)$.

This paper is devoted to the description of conditions which guarantee the comparison among variability estimators in terms of stochastic orderings stronger than the simple comparison of their expected values. In particular, it mainly deals with conditions for the usual and the increasing convex orders between sample variances and estimators of the Gini mean differences, providing conditions in a general setting by considering random vectors of observations with possibly dependent and differently distributed components. Variability measures that involve weight vectors or the order statistics will be considered as well. In addition, some results about the stochastic comparison of the differences of the components of the random samples are provided.

The remainder of the paper is organized as follows. In Section 2 the useful notions and preliminary results are briefly recalled . Section 3 is focused on giving conditions that lead to the usual stochastic order of the variability measure estimators. Similarly, conditions that lead to the increasing convex order are provided in Section 4. Section 5 is devoted to briefly present a practical application of our theoretical results in testing the dispersion and the convex orders between random vectors. Finally, the conclusions are discussed in Section 6.

2 Preliminaries

We devote this section to briefly introduce the concepts that are needed for the development of the paper and to recall some useful preliminary results. In particular, the definition of several stochastic orders, of copulas and of some basic notions of minimization problems in Statistics and Aggregation Theory are provided.

A huge amount of relations between distributions, known as stochastic orders, has been defined in the literature to compare the behavior of random quantities. We briefly recall here the orders considered in the rest of the paper. A detailed list of their characterizations, properties and applications may be found, e.g., in the monograph Shaked and Shanthikumar (2007).

The most prominent stochastic order is the usual stochastic order, which is applied to compare random variables and vectors in terms location (or magnitude), and defined by comparisons of expectations of increasing transformations of the random quantities.

Definition 2.1. *Let $\mathbf{X}$ and $\mathbf{Y}$ be two random vectors of dimension n. If $E[\varphi(\mathbf{X})] \leq E[\varphi(\mathbf{Y})]$ for any measurable increasing function $\varphi : \mathbb{R}^n \to \mathbb{R}$ such that $E[\varphi(\mathbf{X})]$ and $E[\varphi(\mathbf{Y})]$ exist, then $\mathbf{X}$ is said to be smaller than or equal to $\mathbf{Y}$ in the usual stochastic order and it is denoted as $\mathbf{X} \leq_{st} \mathbf{Y}$.*

For dimension $n = 1$, the usual stochastic order (also known as First Order Stochastic Dominance) is equivalent to a pointwise order of the distribution functions of the two random variables. A characterization of this order that will be used in the next sections is the following, based on construction in the same probability space.

Theorem 2.1. *Let $\mathbf{X}$ and $\mathbf{Y}$ be two random vectors of dimension n. Then $\mathbf{X} \leq_{st} \mathbf{Y}$ if and only if there exist two random vectors $\widehat{\mathbf{X}}$ and $\widehat{\mathbf{Y}}$ defined in the same probability space such that $\widehat{\mathbf{X}} =_{st} \mathbf{X}$, $\widehat{\mathbf{Y}} =_{st} \mathbf{Y}$ and $\widehat{\mathbf{X}} \leq_{a.s.} \widehat{\mathbf{Y}}$.*

Obviously, there exit cases where the random quantities can not be ordered in the usual stochastic order. In these cases, weaker stochastic orders can be used. The most common alternative to the usual stochastic order is the increasing convex stochastic order, defined as follows.

Definition 2.2. *Let $\mathbf{X}$ and $\mathbf{Y}$ be two random vectors of dimension n. If $E[\varphi(\mathbf{X})] \leq E[\varphi(\mathbf{Y})]$ for any measurable increasing convex function $\varphi : \mathbb{R}^n \to \mathbb{R}$ such that $E[\varphi(\mathbf{X})]$ and $E[\varphi(\mathbf{Y})]$ exist, then $\mathbf{X}$ is said to be smaller than or equal to $\mathbf{Y}$ in the increasing convex stochastic order and it is denoted as $\mathbf{X} \leq_{icx} \mathbf{Y}$.*

Trivially, the usual stochastic order implies the increasing convex stochastic order.

In some cases we are not interested in comparing the distributions in terms of location, but considering the variability. One of the classical variability stochastic comparisons is the dispersion order.

Definition 2.3. *Let X and Y be two random variables with distribution functions F and G, respectively, and let F^{-1} and G^{-1} denote their right continuous inverses. Then X is said to be smaller than or equal to Y in the dispersion order, denoted as $X \leq_{disp} Y$, if*

$$F^{-1}(\beta) - F^{-1}(\alpha) \leq G^{-1}(\beta) - G^{-1}(\alpha) \quad \text{for all } \; 0 < \alpha \leq \beta < 1. \tag{2.1}$$

Note that (2.1) holds (i.e., $X \leq_{disp} Y$ holds) if and only if the transformation $\phi = F^{-1} \circ G$, for which $X =_{st} \phi(Y)$, satisfies the properties to be increasing and such that

$$\phi(y') - \phi(y) \leq y' - y \quad \text{ whenever } \; y \leq y',$$

with both y and y' in the support of Y (see Shaked and Shanthikumar, 2007, for details). Also note that, in this case, the function ϕ is a contraction, i.e., it is such that $|\phi(x) - \phi(y)| \leq |x - y|$ for any $x, y \in \mathbb{R}$

Another common alternative to define stochastic comparisons based on variability is to use convexity. Recall that, given a function $\varphi : \mathbb{R}^n \to \mathbb{R}$, then φ is said to be convex if $\varphi(\lambda\mathbf{x} + (1-\lambda)\mathbf{y}) \leq \lambda\varphi(\mathbf{x}) + (1-\lambda)\varphi(\mathbf{y})$ for any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and $\lambda \in [0, 1]$, and φ is said to be componentwise convex if φ is convex in each argument when the other arguments are held fixed. It is well-known that componentwise convexity implies convexity.

Definition 2.4. *Let $\mathbf{X}$ and $\mathbf{Y}$ be two random vectors of dimension n. If $E[\varphi(\mathbf{X})] \leq E[\varphi(\mathbf{Y})]$ for any measurable convex [componentwise convex] function $\varphi : \mathbb{R}^n \to \mathbb{R}$ such that $E[\varphi(\mathbf{X})]$ and $E[\varphi(\mathbf{Y})]$ exist, then $\mathbf{X}$ is said to be smaller than or equal to $\mathbf{Y}$ in the convex [componentwise convex] stochastic order and it is denoted as*

$$\mathbf{X} \; \leq_{cx} \; [\leq_{ccx}] \; \mathbf{Y}.$$

Trivially, the componentwise convex stochastic order implies the convex stochastic order, while they are equivalent for $n = 1$, and both implies the increasing convex stochastic order.

In the following we will use $[\mathbf{X} \mid A]$ to denote the random vectors whose distribution is the distribution of $\mathbf{X}$ conditioned on an event A. Using conditional distributions, a characterization of the convex and increasing convex orders that will be specially useful in our results can be stated.

Theorem 2.2. *Let $\mathbf{X}$ and $\mathbf{Y}$ be two random vectors of dimension n. Then $\mathbf{X} \leq_{cx} [\leq_{icx}] \mathbf{Y}$ if and only if exist two random vectors $\widehat{\mathbf{X}}$ and $\widehat{\mathbf{Y}}$ defined in the same probability space such that $\widehat{\mathbf{X}} =_{st} \mathbf{X}$, $\widehat{\mathbf{Y}} =_{st} \mathbf{Y}$ and $E[\widehat{\mathbf{Y}} \mid \widehat{\mathbf{X}}] =_{a.s.} [\geq_{a.s.}] \widehat{\mathbf{X}}$.*

The distribution of a random vector can be separated in two parts. Firstly, one can consider the marginal distributions of the components of such vector, that describe the behavior of each component. Secondly, there is also the dependence structure between the components. The dependence structure can be described by using copula functions, briefly recalled here.

Definition 2.5. *Let $\mathbf{X}$ be a random vector with marginal distributions $F_1, \ldots, F_n$ and joint distribution function F. Then, a function $C : [0,1]^n \to [0,1]$ such that*

$$F(x_1, \ldots, x_n) = C\left(F_1(x_1), \ldots, F_n(x_n)\right), \; \forall (x_1, \ldots, x_n) \in \mathbb{R}^n$$

is said to be a copula for the vector $\mathbf{X}$.

It must be recalled that if the marginal distributions are continuous then the copula is unique. For this reason, we will assume here, and everywhere throughout the paper, continuity of the marginal distributions for the involved vectors. We address the readers to the monograph Nelsen (2006), or to the recent monograph Durante and Sempi (2015), for further details.

When the copula can be chosen to be the same for two random vectors, we will call them random vectors with the same copula, which means they share the same dependence structure. It is worth noting that if two random vectors have the same margins and copula, they also have the same distribution. Additionally, it's important to mention that an increasing transformation of the margins of a random vector does not change its copula (see, e.g., Theorem 2.4.3 in Nelsen, 2006). For our purposes it is also important to recall the following property.

Theorem 2.3. *Let $\mathbf{X}$ and $\mathbf{Y}$ be two random vectors of dimension n with the same copula. Then,*

- *If $X_i \leq_{st} Y_i$ for any $i \in \{1, \ldots, n\}$, then $\mathbf{X} \leq_{st} \mathbf{Y}$;*
- *If $X_i \leq_{cx} Y_i$ for any $i \in \{1, \ldots, n\}$ and the copula is the independence (product) copula, then $\mathbf{X} \leq_{cx} \mathbf{Y}$.*

Apart for the sample variance and the estimator of the Gini mean difference recalled in Section 1, other estimators of variability, of main interest in Aggregation Theory, will be considered through the paper. These are estimators defined as weighted averages of increasing transformations of the absolute difference between the argument and the elements of a sample, or between the ordered elements of the sample. Applications of these estimators may be found, for example, in Amemiya (1985) or Calvo and Beliakov (2010), while we address the reader to Grabisch et al. (2009) for an introduction on Aggregation Theory. For the proof of the results dealing with estimators defined as weighted averages we will make use of the following statement, whose proof be found, e.g., in Calvo and Beliakov (2010).

Theorem 2.4. *Consider a weight vector $\vec{w} \in [0,1]^n$ such that $\sum_{i=1}^n w_i = 1$ and $\mathbf{x} \in \mathbb{R}^n$. Then,*

(a) $\sum_{i=1}^n w_i x_i = \arg\min_{y \in \mathbb{R}} \sum_{i=1}^n w_i (x_i - y)^2$;

(b) $\sum_{i=1}^n w_i x_{(i)} = \arg\min_{y \in \mathbb{R}} \sum_{i=1}^n w_i (x_{(i)} - y)^2$ *with* $x_{(1)} \leq \cdots \leq x_{(n)}$.

3 Conditions for the usual stochastic order of variability measure estimators

The aim of this section is to give sufficient conditions such that the variability measures estimators mentioned above can be compared in terms of the usual stochastic order.

In the following we will consider two random vectors in which all the components have the same distribution, within each of them, and that share the same copula. They can be seen as random samples but with not necessarily independent variables. Let us start by proving that, under additional regularity conditions, one of the vectors has the same distribution than the other when applying a contraction to all its components. To this aim, given two random variables X and Y, the following two conditions will be considered in the next statements:

(P1) both X and Y have continuous distribution functions (say, F and G, respectively);

(P2) the transformation $\phi = F^{-1} \circ G$ such that $X =_{st} \phi(Y)$ is strictly monotone.

Note that condition (P2) is satisfied, for example, if both X and Y have non-zero density over an interval of the real line, that can be different for X and Y.

Lemma 3.1. *Let $\mathbf{X} = (X_1, \ldots, X_n)$ be a random vector with identically distributed components and $\mathbf{Y} = (Y_1, \ldots, Y_n)$ another random vector with also identically distributed components and with the same copula as $\mathbf{X}$. If $X_1 \leq_{disp} Y_1$ and X_1 and Y_1 satisfy the properties (P1) and (P2), then it holds $\mathbf{X} =_{st} (\phi(Y_1), \ldots, \phi(Y_n))$, where $\phi = F^{-1} \circ G$, being F and G the distributions of X_1 and Y_1, respectively.*

Proof. Since $X_1 =_{st} \cdots =_{st} X_n$ and $Y_1 =_{st} \cdots =_{st} Y_n$ and $X_1 \leq_{disp} Y_1$, then we have that $X_i =_{st} \phi(Y_i)$ for any $i \in \{1, \ldots, n\}$. In addition, since the marginal distributions are continuous by (P1) and ϕ is strictly increasing by (P2), we have that the copulas of the random vectors $(\phi(Y_1), \ldots, \phi(Y_n))$ and $\mathbf{Y}$ are the same. Therefore, $(\phi(Y_1), \ldots, \phi(Y_n))$ and $\mathbf{X}$ have the same copula. Moreover, since they have the same copula and the same marginal distributions, $\mathbf{X} =_{st} (\phi(Y_1), \ldots, \phi(Y_n))$. $\square$

We want to remark that the random vectors can have different margins, the margins inside the two random vectors are the ones which should be the same. A direct consequence of this result is that the vectors consisting of the absolute differences between the components are ordered in the usual stochastic order. On this aim, given a random vector $\mathbf{X}$, we will denote as $\{|X_i - X_j|\}_{i,j\in\{1,\ldots,n\}}$ the random vector consisting on the absolute difference between all the possible combinations of the components of $\mathbf{X}$.

Theorem 3.1. *Let $\mathbf{X} = (X_1, \ldots, X_n)$ be a random vector with identically distributed components and $\mathbf{Y} = (Y_1, \ldots, Y_n)$ another random vector with also identically distributed components and with the same copula as $\mathbf{X}$. If $X_1 \leq_{disp} Y_1$ and if X_1 and Y_1 satisfy the properties (P1) and (P2), then*

$$\{|X_i - X_j|\}_{i,j\in\{1,\ldots,n\}} \leq_{st} \{|Y_i - Y_j|\}_{i,j\in\{1,\ldots,n\}}.$$

Proof. Applying Lemma 3.1, we have that $\mathbf{X} =_{st} (\phi(Y_1), \ldots, \phi(Y_n))$. Define, in the same probability space where $\mathbf{Y}$ is defined, the vector $\widehat{\mathbf{X}} = (\phi(Y_1), \ldots, \phi(Y_n))$. Then, by using the fact that ϕ is a contraction, we have

$$\{|X_i - X_j|\}_{i,j\in\{1,\ldots,n\}} =_{st} \{|\phi(Y_i) - \phi(Y_j)|\}_{i,j\in\{1,\ldots,n\}} \leq_{a.s.} \{|Y_i - Y_j|\}_{i,j\in\{1,\ldots,n\}}.$$

Since $(|\widehat{X}_i - \widehat{X}_j|)_{i,j\in\{1,\ldots,n\}} =_{st} (|X_i - X_j|)_{i,j\in\{1,\ldots,n\}}$, the result holds by applying Theorem 2.1. $\square$

In particular, as an immediate consequence of Theorem 3.1 and equations (1.1) and (1.2), by applying Theorem 6.B.16 (a) in Shaked and Shanthikumar (2007), which states the closure of the $\leq_{st}$ order with respect to increasing functions, one gets the following conditions for the comparisons among sample variances and estimators of the Gini mean difference.

Corollary 3.1. *Let $(X_1, \ldots, X_n)$ and $(Y_1, \ldots, Y_n)$ be two simple random samples of, respectively, X and Y. If $X \leq_{disp} Y$ and if X and Y satisfy the properties (P1) and (P2), then $S_X^2 \leq_{st} S_Y^2$ and $G_n(X) \leq_{st} G_n(Y)$.*

Conditions for the stochastic ordering between weighted versions of the sample variances can be also stated. In particular, in the same condition than the latter results, the dispersion order of the random vectors implies the usual stochastic order between these quantities. In the following we consider a weight vector $\mathbf{w} \in [0,1]^n$ such that $\sum_{i=1}^n w_i = 1$.

Theorem 3.2. *Let $\mathbf{X} = (X_1, \ldots, X_n)$ be a random vector with identically distributed components and $\mathbf{Y} = (Y_1, \ldots, Y_n)$ another random vector with also identically distributed components and with the same copula as $\mathbf{X}$. Assume that X_1 and Y_1 satisfy the properties (P1) and (P2). Then, for any weight vector $\mathbf{w}$,*

$$X_1 \leq_{disp} Y_1 \implies \sum_{i=1}^n w_i \left(X_i - \sum_{j=1}^n w_j X_j \right)^2 \leq_{st} \sum_{i=1}^n w_i \left(Y_i - \sum_{j=1}^n w_j Y_j \right)^2 .$$

Proof. Applying Lemma 3.1, we have that $\mathbf{X} =_{st} (\phi(Y_1), \ldots, \phi(Y_n))$. Define, in the same probability space where $\mathbf{Y}$ is defined, the vector $\widehat{\mathbf{X}} = (\phi(Y_1), \ldots, \phi(Y_n))$, where $\phi = F^{-1} \circ G$. Then,

$$\sum_{i=1}^n w_i \left(\widehat{X}_i - \sum_{j=1}^n w_j \widehat{X}_j \right)^2 = \sum_{i=1}^n w_i \left(\phi(Y_i) - \sum_{j=1}^n w_j \phi(Y_j) \right)^2$$

$$\leq_{a.s} \sum_{i=1}^n w_i \left(\phi(Y_i) - \phi\left(\sum_{j=1}^n w_j Y_j \right) \right)^2 \leq \sum_{i=1}^n w_i \left(Y_i - \sum_{j=1}^n w_j Y_j \right)^2 ,$$

where the first inequality follows from Theorem 2.4 (a) and the second by the fact that ϕ is a contraction. Observing that

$$\sum_{i=1}^n w_i \left(\widehat{X}_i - \sum_{j=1}^n w_j \widehat{X}_j \right)^2 =_{st} \sum_{i=1}^n w_i \left(X_i - \sum_{j=1}^n w_j X_j \right)^2 ,$$

the result hold by applying Theorem 2.1. □

It must be observed that the conditions for the $\leq_{st}$ order between sample variances stated in Corollary 3.1 also follows from Theorem 3.2. Recall that, as mentioned in Section 2, estimators based on order statistics are relevant in many areas. A result similar to Theorem 3.2 can be stated also when working with ordered samples.

Theorem 3.3. *Let $\mathbf{X} = (X_1, \ldots, X_n)$ be a random vector with identically distributed components and $\mathbf{Y} = (Y_1, \ldots, Y_n)$ another random vector with also identically distributed components and with the same copula as $\mathbf{X}$. Assume that X_1 and Y_1 satisfy the properties (P1) and (P2). Then, for any weight vector $\mathbf{w}$,*

$$X_1 \leq_{disp} Y_1 \implies \sum_{i=1}^{n} w_i \left(X_{(i)} - \sum_{j=1}^{n} w_j X_{(j)} \right)^2 \leq_{st} \sum_{i=1}^{n} w_i \left(Y_{(i)} - \sum_{j=1}^{n} w_j Y_{(j)} \right)^2 .$$

Proof. Applying Lemma 3.1, we have that $\mathbf{X} =_{st} (\phi(Y_1), \ldots, \phi(Y_n))$ with ϕ an increasing contraction. Since ϕ is increasing, we have that the order statistics of $(\phi(Y_1), \ldots, \phi(Y_n))$ equal $\left(\phi(Y_{(1)}), \ldots, \phi(Y_{(n)})\right)$. Thus, we have that $\left(X_{(1)}, \ldots, X_{(n)}\right) =_{st} \left(\phi(Y_{(1)}), \ldots, \phi(Y_{(n)})\right)$. Then, we just need to consider Theorem 2.4 (b) and proceed exactly as in Theorem 3.2. $\square$

The next statements generalize Theorem 3.2 and Theorem 3.3 relaxing the condition of the same copula for $\mathbf{X}$ and $\mathbf{Y}$. In particular, the following results can be applied to random vectors having distributions that depend on a set of random parameters, such as, for example, vectors of lifetimes described by multivariate frailty models (see Marshall and Olkin, 1988); as pointed out in subsequent Example 3.1, the involved vectors can in fact have different copulas.

In the next statements we assume that the random vector $\mathbf{X}(\Theta_X)$ has distribution depending on the random parameter Θ_X, while the random vector $\mathbf{Y}(\Theta_Y)$ has distribution depending on the random parameter Θ_Y. The notations $[\mathbf{X}(\Theta_X) \mid \Theta_X = \theta]$ and $[X_i(\Theta_X) \mid \Theta_X = \theta]$ used below refers to the vector (or variable) whose distribution is the distribution of $\mathbf{X}(\Theta_X)$ (or $X_i(\Theta_X)$) given that $\Theta_X = \theta$.

Theorem 3.4. *Let $\mathbf{X}(\Theta_X)$ and $\mathbf{Y}(\Theta_Y)$ be two random vectors and Θ_X and Θ_Y be two random parameters with the same support $\mathcal{T}$ such that:*

(1) each one of $[\mathbf{X}(\Theta_X) \mid \Theta_X = \theta]$ and $[\mathbf{Y}(\Theta_Y) \mid \Theta_Y = \theta]$ is a random vector with identically distributed margins for any $\theta \in \mathcal{T}$;

(2) $[\mathbf{X}(\Theta_X) \mid \Theta_X = \theta_1]$, $[\mathbf{X}(\Theta_X) \mid \Theta_X = \theta_2]$, $[\mathbf{Y}(\Theta_Y) \mid \Theta_Y = \theta_1]$ and $[\mathbf{Y}(\Theta_Y) \mid \Theta_Y = \theta_2]$ have the same copula for any $\theta_1, \theta_2 \in \mathcal{T}$;

(3) $[X_i(\Theta_X) \mid \Theta_X = \theta_1] \leq_{disp} [X_i(\Theta_X) \mid \Theta_X = \theta_2]$, or $[Y_i(\Theta_Y) \mid \Theta_Y = \theta_1] \leq_{disp} [Y_i(\Theta_Y) \mid \Theta_Y = \theta_2]$, for all $i \in \{1, \ldots, n\}$ and for any $\theta_1 \leq \theta_2$, $\theta_1, \theta_2 \in \mathcal{T}$;

(4) $[X_i(\Theta_X) \mid \Theta_X = \theta] \leq_{disp} [Y_i(\Theta_Y) \mid \Theta_Y = \theta]$ for all $i \in \{1, \ldots, n\}$ and for any $\theta \in \mathcal{T}$;

(5) $\Theta_X \leq_{st} \Theta_Y$.

Also assume that $[X_i(\Theta_X) \mid \Theta_X = \theta]$ and $[Y_i(\Theta_Y) \mid \Theta_Y = \theta]$ satisfy the properties (P1) and (P2) for all $\theta \in \mathcal{T}$. Then, for any weight vector $\mathbf{w}$ we have that

$$\sum_{i=1}^{n} w_i \left(X_i(\Theta_X) - \sum_{j=1}^{n} w_j X_j(\Theta_X) \right)^2 \leq_{st} \sum_{i=1}^{n} w_i \left(Y_i(\Theta_Y) - \sum_{j=1}^{n} w_j Y_j(\Theta_Y) \right)^2$$

and

$$\sum_{i=1}^n w_i \left(X_{(i)}(\Theta_X) - \sum_{j=1}^n w_j X_{(j)}(\Theta_X) \right)^2 \leq_{st} \sum_{i=1}^n w_i \left(Y_{(i)}(\Theta_Y) - \sum_{j=1}^n w_j Y_{(j)}(\Theta_Y) \right)^2 .$$

Proof. Consider the case $[X_i(\Theta_X) \mid \Theta_X = \theta_1] \leq_{disp} [X_i(\Theta_X) \mid \Theta_X = \theta_2]$ for any $\theta_1 \leq \theta_2$ in assumption (3) (for the case $[Y_i(\Theta_Y) \mid \Theta_Y = \theta_1] \leq_{disp} [Y_i(\Theta_Y) \mid \Theta_Y = \theta_2]$ in (3) the proof is similar).

By assumptions (1), (2) and (3), applying Theorem 3.2 one has that

$$\begin{aligned} &\left[\textstyle\sum_{i=1}^n w_i \left(X_i(\Theta_X) - \sum_{j=1}^n w_j X_j(\Theta_X) \right)^2 \mid \Theta_X = \theta_1 \right] \\ &\quad \leq_{st} \left[\textstyle\sum_{i=1}^n w_i \left(X_i(\Theta_X) - \sum_{j=1}^n w_j X_j(\Theta_X) \right)^2 \mid \Theta_X = \theta_2 \right] \quad \forall \theta_1 \leq \theta_2, \end{aligned}$$

which in turns, by assumption (5) and Theorem 1.A.3(d) in Shaked and Shanthikumar (2007), implies

$$\sum_{i=1}^n w_i \left(X_i(\Theta_X) - \sum_{j=1}^n w_j X_j(\Theta_X) \right)^2 \leq_{st} \sum_{i=1}^n w_i \left(X_i(\Theta_X) - \sum_{j=1}^n w_j X_j(\Theta_X) \right)^2 . \tag{3.1}$$

By assumptions (1), (2) and (4), applying Theorem 3.2 one has that

$$\left[\sum_{i=1}^n w_i \left(X_i(\Theta_X) - \sum_{j=1}^n w_j X_j(\Theta_X) \right)^2 \middle| \ \Theta_X = \theta \right] \leq_{st} \left[\sum_{i=1}^n w_i \left(Y_i(\Theta_Y) - \sum_{j=1}^n w_j Y_j(\Theta_Y) \right)^2 \middle| \ \Theta_X = \theta \right] \ \ \forall \theta,$$

which in turns, by assumption (5) and Theorem 1.A.6 in Shaked and Shanthikumar (2007), implies

$$\sum_{i=1}^n w_i \left(X_i(\Theta_X) - \sum_{j=1}^n w_j X_j(\Theta_X) \right)^2 \leq_{st} \sum_{i=1}^n w_i \left(Y_i(\Theta_Y) - \sum_{j=1}^n w_j X_j(\Theta_Y) \right)^2 . \tag{3.2}$$

Thus the first statement follows from (3.1), (3.2) and transitivity of the usual stochastic order. For the second inequality the proof is the same but using Theorem 3.3 instead of Theorem 3.2. $\square$

The proof of the following statement is similar as the previous one, and therefore omitted.

Theorem 3.5. *Let* $\mathbf{X}(\Theta_X)$ *and* $\mathbf{Y}(\Theta_Y)$ *be two random vectors and* Θ_X *and* Θ_Y *be two random parameters with the same support* $\mathcal{T}$ *such that:*

(1) each one of $[\mathbf{X}(\Theta_X) \mid \Theta_X = \theta]$ *and* $[\mathbf{Y}(\Theta_Y) \mid \Theta_Y = \theta]$ *is a random vector with identically distributed margins for any* $\theta \in \mathcal{T}$*;*

(2) $[\mathbf{X}(\Theta_X) \mid \Theta_X = \theta_1]$, $[\mathbf{X}(\Theta_X) \mid \Theta_X = \theta_2]$, $[\mathbf{Y}(\Theta_Y) \mid \Theta_Y = \theta_1]$ *and* $[\mathbf{Y}(\Theta_Y) \mid \Theta_Y = \theta_2]$ *have the same copula for any* $\theta_1, \theta_2 \in \mathcal{T}$*;*

(3) $[X_i(\Theta_X) \mid \Theta_X = \theta_1] \geq_{disp} [X_i(\Theta_X) \mid \Theta_X = \theta_2]$*, or* $[Y_i(\Theta_Y) \mid \Theta_Y = \theta_1] \geq_{disp} [Y_i(\Theta_Y) \mid \Theta_Y = \theta_2]$*, for all* $i \in \{1, \ldots, n\}$ *and for any* $\theta_1 \leq \theta_2$*,* $\theta_1, \theta_2 \in \mathcal{T}$*;*

(4) $[X_i(\Theta_X) \mid \Theta_X = \theta] \geq_{disp} [Y_i(\Theta_Y) \mid \Theta_Y = \theta]$ *for all* $i \in \{1, \ldots, n\}$ *and for any* $\theta \in \mathcal{T}$*;*

(5) $\Theta_X \leq_{st} \Theta_Y$*.*

Also assume that $[X_i(\Theta_X) \mid \Theta_X = \theta]$ *and* $[Y_i(\Theta_Y) \mid \Theta_Y = \theta]$ *satisfy the properties (P1) and (P2) for all* $\theta \in \mathcal{T}$*. Then, for any weight vector* $\mathbf{w}$ *we have that*

$$\sum_{i=1}^n w_i \left(X_i(\Theta_X) - \sum_{j=1}^n w_j X_j(\Theta_X) \right)^2 \geq_{st} \sum_{i=1}^n w_i \left(Y_i(\Theta_Y) - \sum_{j=1}^n w_j Y_j(\Theta_Y) \right)^2 \tag{3.3}$$

and

$$\sum_{i=1}^n w_i \left(X(i)(\Theta_X) - \sum_{j=1}^n w_j X_{(j)}(\Theta_X) \right)^2 \geq_{st} \sum_{i=1}^n w_i \left(Y_{(i)}(\Theta_Y) - \sum_{j=1}^n w_j Y_{(j)}(\Theta_Y) \right)^2. \tag{3.4}$$

The following is an example where the stochastic order between weighted variability measures is guaranteed by the application of Theorem 3.5.

Example 3.1. *Let* $\mathbf{X}(\Theta_X)$ *and* $\mathbf{Y}(\Theta_Y)$ *be two vectors described by two multivariate frailty models having conditionally independent exponentially distributed underlying lifetimes with rate* $\lambda = 1$ *and frailties* Θ_X *and* Θ_Y*, respectively, with* $\Theta_X \sim \Gamma(1, \alpha_X)$ *and* $\Theta_Y \sim \Gamma(1, \alpha_Y)$*.*

That is, let $\mathbf{X}(\Theta_X)$ *and* $\mathbf{Y}(\Theta_Y)$ *have joint survival functions*

$$\begin{aligned}
\overline{F}_{\mathbf{X}(\Theta_X)}(t_1, \ldots, t_n) &= E_{\Theta_X}\left[\prod_{i=1}^n \left(e^{-t_i}\right)^{\Theta_X} \right] = \overline{W}_{\mathbf{X}}\left(\sum_{i=1}^n t_i \right), \\
\overline{F}_{\mathbf{Y}(\Theta_Y)}(t_1, \ldots, t_n) &= E_{\Theta_Y}\left[\prod_{i=1}^n \left(e^{-t_i}\right)^{\Theta_Y} \right] = \overline{W}_{\mathbf{Y}}\left(\sum_{i=1}^n t_i \right),
\end{aligned}$$

where $\overline{W}_{\mathbf{X}}(t) = E_{\Theta_X}[e^{-t\Theta_X}] = (1+t)^{-\alpha_X}$ *and* $\overline{W}_{\mathbf{Y}}(t) = E_{\Theta_Y}[e^{-t\Theta_Y}] = (1+t)^{-\alpha_Y}$*.*

Assume $\alpha_X < \alpha_Y$*. Conditions (1-5) in Theorems 3.5 are fulfilled because of the following facts:*

(1) both $[\mathbf{X}(\Theta_X) \mid \Theta_X = \theta]$ *and* $[\mathbf{Y}(\Theta_Y) \mid \Theta_Y = \theta]$ *are random vectors of independent and exponentially distributed variables with the same rate* θ*;*

(2) $[\mathbf{X}(\Theta_X) \mid \Theta_X = \theta]$ *and* $[\mathbf{Y}(\Theta_Y) \mid \Theta_Y = \theta]$*, have the independence copula for any* θ*;*

(3) $[X_i(\Theta_X) \mid \Theta_X = \theta_1] \geq_{disp} [X_i(\Theta_X) \mid \Theta_X = \theta_2]$ *for any* $\theta_1 \leq \theta_2$*, since both are vectors of independent and exponentially distributed variables with rates* θ_1 *and* θ_2*, respectively, and since* $Z_1 \geq_{disp} Z_2$ *whenever* $Z_i \sim Exp(\theta_i)$ *and* $\theta_1 \leq \theta_2$*;*

(4) the variables $[X_i(\Theta_X) \mid \Theta_X = \theta]$ *and* $[Y_i(\Theta_Y) \mid \Theta_Y = \theta]$ *have the same distribution for any* θ*;*

(5) $\Theta_X \leq_{st} \Theta_Y$ *by properties of the gamma distribution.*

The properties (P1) and (P2) for $[X_i(\Theta_X) \mid \Theta_X = \theta]$ and $[Y_i(\Theta_Y) \mid \Theta_Y = \theta]$ are clearly satisfied for all $\theta \in \mathcal{T}$.

Thus, Theorem 3.5 can be applied and, for any weight vector $\mathbf{w}$, the stochastic inequalities (3.3) and (3.4) hold.

Note that the vectors $\mathbf{X}(\Theta_X)$ and $\mathbf{Y}(\Theta_Y)$ have both different marginal survival functions, being $\overline{F}_{X_i(\Theta_X)}(t) = (1+t)^{-\alpha_X}$ and $\overline{F}_{Y_i(\Theta_Y)}(t) = (1+t)^{-\alpha_Y}$ with $t \geq 0$, and different copulas. In fact, $\mathbf{X}(\Theta_X)$ and $\mathbf{Y}(\Theta_Y)$ are special cases of the time transformed exponential model (TTE) whose survival copulas are $\widehat{C}_{\mathbf{X}}(u_1, \ldots, u_n) = \overline{W}_{\mathbf{X}}\left(\sum_{i=1}^n \overline{W}_{\mathbf{X}}^{-1}(u_i)\right)$ and $\widehat{C}_{\mathbf{Y}}(u_1, \ldots, u_n) = \overline{W}_{\mathbf{Y}}\left(\sum_{i=1}^n \overline{W}_{\mathbf{Y}}^{-1}(u_i)\right)$, i.e. Archimedean copulas having inverse generators $\overline{W}_{\mathbf{X}}$ and $\overline{W}_{\mathbf{Y}}$. In particular, they are Clayton copulas with parameters $\frac{1}{\alpha_X}$ and $\frac{1}{\alpha_Y}$, respectively. For more details in this regard, we refer to Mulero et al. (2010).

4 Conditions for the increasing convex order of variability measure estimators

As recalled already, other important variability orders are the convex and componentwise convex orders. Similar Results similar as the ones for the dispersion order can be proved using these orders, but obtaining the weaker increasing convex order between variability estimators.

The first statement is in the same spirit as Theorem 3.1. Here, $\{(X_i - X_j)\}_{i,j\in\{1,\ldots,n\}}$ denotes the random vector consisting on the difference between all the possible combinations of the components of $\mathbf{X}$.

Theorem 4.1. *Let $\mathbf{X}$ and $\mathbf{Y}$ be two random vectors such that $\mathbf{X} \leq_{ccx} \mathbf{Y}$. Then,*

$$\{(X_i - X_j)\}_{i,j\in\{1,\ldots,n\}} \leq_{cx} \{(Y_i - Y_j)\}_{i,j\in\{1,\ldots,n\}}.$$

Proof. By Theorem 3.2 in Müller (1997), it suffices to prove that $E[\varphi(\{(X_i - X_j)\}_{i,j\in\{1,\ldots,n\}})] \leq E[\varphi(\{(Y_i - Y_j)\}_{i,j\in\{1,\ldots,n\}})]$ for any twice differentiable convex function $\varphi : \mathbb{R}^{n^2} \to \mathbb{R}$. Thus, let φ be any twice differentiable convex function. Denote as φ_i and φ_{ij}, with $i, j \in \{1, \ldots, n\}$, its first and second derivatives. In addition, denote as H its Hessian, which is positive semi-definite. Now, consider the function $\phi : \mathbb{R}^n \to \mathbb{R}^{n^2}$ defined as $\phi(\mathbf{X}) = \{(x_i - x_j)\}_{i,j\in\{1,\ldots,n\}}$, and let us compute $\frac{\partial^2 \varphi\circ\phi}{\partial x_k^2}$ with $k \in \{1, \ldots, n\}$.

First, denote as I^+ the set of indices such that $i \in I^+$ if and only if $(\phi(\mathbf{X}))_i = x_k - x_j$ with $j \in \{1, \ldots, n\}$. Similarly, denote as I^- the set of indices such that $i \in I^-$ if and only if $(\phi(\mathbf{X}))_i = x_j - x_k$ with $j \in \{1, \ldots, n\}$. Then:

$$\frac{\partial \varphi\circ\phi}{\partial x_k} = \sum_{i\in I^+} \varphi_i - \sum_{i\in I^-} \varphi_i,$$
$$\frac{\partial^2 \varphi\circ\phi}{\partial x_k^2} = \sum_{i\in I^+\cup I^-} \varphi_{ii} + \sum_{i,j\in I^+} \varphi_{ij} + \sum_{i,j\in I^-} \varphi_{ij} - \sum_{i\in I^+, j\in I^-} \varphi_{ij} - \sum_{i\in I^-, j\in I^+} \varphi_{ij}.$$

Now, consider the vector $\mathbf{s}$ of dimension n^2 defined such that $s_i = 1$ if $i \in I^+$, $s_i = -1$ if $i \in I^-$ and $s_i = 0$ if $i \notin I^+ \cup I^-$. The last expression of the second derivative is equivalent to $\frac{\partial^2 \varphi\circ\phi}{\partial x_k^2} = \mathbf{s}^T H \mathbf{s}$.

Since H is positive semidefinite, we have $\frac{\partial^2 \varphi\circ\phi}{\partial x_k^2} \geq 0$. This holds for any $k \in \{1, \ldots, n\}$, thus $\varphi \circ \phi$ is a componentwise convex function.

Then, since $\mathbf{X} \leq_{ccx} \mathbf{Y}$, one has $E[\varphi\circ\phi(\mathbf{X})] \leq E[\varphi\circ\phi(\mathbf{Y})]$ and therefore $E[\varphi(\{(X_i - X_j)\}_{i,j\in\{1,\ldots,n\}})] \leq E[\varphi(\{(Y_i - Y_j)\}_{i,j\in\{1,\ldots,n\}})]$ for any twice differentiable convex function φ. $\square$

Unfortunately, from Theorem 4.1 does not follow a statement similar to Corollary 3.1 but having the convex order as assumption. The reason is that the composition between two convex functions (or between an increasing convex and a convex function) is not necessarily convex (or increasing convex). However, conditions to order sample variances in the increasing convex order, under the weaker assumption of convex order rather than the componentwise convex order, are provided in the following statement. Here, on the contrary as in Theorem 3.2, no additional conditions on the same marginal distribution or the same copula for the random vectors are required. Moreover, the result involves two possible different weighting vectors for the pondering of the mean and the squared differences.

Theorem 4.2. *Let $\mathbf{X}, \mathbf{Y}$ be two random vectors such that $\mathbf{X} \leq_{cx} \mathbf{Y}$. Then, for any pair of weight vectors $\mathbf{w}$ and $\mathbf{v}$ one has*

$$\sum_{i=1}^n w_i \left(X_i - \sum_{j=1}^n v_j X_j \right)^2 \leq_{icx} \sum_{i=1}^n w_i \left(Y_i - \sum_{j=1}^n v_j Y_j \right)^2 .$$

Proof. Since $\mathbf{X} \leq_{cx} \mathbf{Y}$, then there exist two random vectors $\widehat{\mathbf{X}}$ and $\widehat{\mathbf{Y}}$ defined in the same probability space such that $\mathbf{X} =_{st} \widehat{\mathbf{X}}$, $\mathbf{Y} =_{st} \widehat{\mathbf{Y}}$ and $E[\widehat{\mathbf{Y}}|\widehat{\mathbf{X}}] =_{a.s.} \widehat{\mathbf{X}}$ (see Theorem 2.2).

Define the function $h : \mathbb{R} \to \mathbb{R}$ as

$$h(\lambda) = E\left[\sum_{i=1}^n w_i \left(\widehat{Y}_i - \sum_{j=1}^n v_j \widehat{Y}_j \right)^2 \,\middle|\, \sum_{i=1}^n w_i \left(\widehat{X}_i - \sum_{j=1}^n v_j \widehat{X}_j \right)^2 = \lambda \right].$$

Consider a vector $\mathbf{x} \in \mathbb{R}^n$ such that $\sum_{i=1}^n w_i \left(x_i - \sum_{j=1}^n v_j x_j \right)^2 = \lambda$. If $\widehat{\mathbf{X}} = \mathbf{x}$, then we can express $\widehat{\mathbf{Y}}$ as $\widehat{\mathbf{Y}} = \mathbf{x} + \boldsymbol{\epsilon}$, with $\boldsymbol{\epsilon} = (\epsilon_1, \ldots, \epsilon_n)$ being a random vector with a null mean vector. Then, we have

$$\begin{aligned}
E\left[\sum_{i=1}^n w_i \left(\widehat{Y}_i - \sum_{j=1}^n v_j \widehat{Y}_j \right)^2 \,\middle|\, \widehat{\mathbf{X}} = \mathbf{x} \right] &= E\left[\sum_{i=1}^n w_i \left(x_i - \sum_{j=1}^n v_j x_j + \epsilon_i - \sum_{j=1}^n v_j \epsilon_j \right)^2 \right] \\
= E\left[\sum_{i=1}^n w_i \left(\left(x_i - \sum_{j=1}^n v_j x_j \right)^2 + \left(\epsilon_i - \sum_{j=1}^n v_j \epsilon_j \right)^2 + 2 \left(x_i - \sum_{j=1}^n v_j x_j \right) \left(\epsilon_i - \sum_{j=1}^n v_j \epsilon_j \right) \right) \right] \\
= E\left[\sum_{i=1}^n w_i \left(\left(x_i - \sum_{j=1}^n v_j x_j \right)^2 + \left(\epsilon_i - \sum_{j=1}^n v_j \epsilon_j \right)^2 \right) \right] \geq \sum_{i=1}^n w_i \left(x_i - \sum_{j=1}^n v_j x_j \right)^2 = \lambda.
\end{aligned}$$

Define the set $C_\lambda = \left\{ \mathbf{x} \in \mathbb{R}^n : \sum_{i=1}^n w_i \left(x_i - \sum_{j=1}^n v_j x_j \right)^2 = \lambda \right\}$ and denote as f the density function of the conditional distribution of $\widehat{\mathbf{X}}$ given $\widehat{\mathbf{X}} \in C_\lambda$. Using it, we can establish the inequality

$$h(\lambda) = \int_{\mathbf{x} \in C_\lambda} E\left[\sum_{i=1}^n w_i \left(\widehat{Y}_i - \sum_{j=1}^n v_j \widehat{Y}_j \right)^2 \middle| \widehat{\mathbf{X}} = \mathbf{x} \right] f(\mathbf{x}) d\mathbf{x} \geq \int_{\mathbf{x} \in C_\lambda} \lambda f(\mathbf{x}) d\mathbf{x} = \lambda.$$

Using the latter inequality, we have that

$$h\left(\sum_{i=1}^n w_i \left(\widehat{X}_i - \sum_{j=1}^n v_j \widehat{X}_j \right)^2 \right) \geq_{a.s} \sum_{i=1}^n w_i \left(\widehat{X}_i - \sum_{j=1}^n v_j \widehat{X}_j \right)^2.$$

Then, it is concluded that

$$E\left[\sum_{i=1}^n w_i \left(\widehat{Y}_i - \sum_{j=1}^n v_j \widehat{Y}_j \right)^2 \middle| \sum_{i=1}^n w_i \left(\widehat{X}_i - \sum_{j=1}^n v_j \widehat{X}_j \right)^2 \right] \geq \sum_{i=1}^n w_i \left(\widehat{X}_i - \sum_{j=1}^n v_j \widehat{X}_j \right)^2.$$

In addition, since

$$\sum_{i=1}^n w_i \left(\widehat{X}_i - \sum_{j=1}^n v_j \widehat{X}_j \right)^2 =_{st} \sum_{i=1}^n w_i \left(X_i - \sum_{j=1}^n v_j X_j \right)^2$$

and

$$\sum_{i=1}^n w_i \left(\widehat{Y}_i - \sum_{j=1}^n v_j \widehat{Y}_j \right)^2 =_{st} \sum_{i=1}^n w_i \left(Y_i - \sum_{j=1}^n v_j Y_j \right)^2,$$

by applying Theorem 2.2 it holds that

$$\sum_{i=1}^n w_i \left(X_i - \sum_{j=1}^n v_j X_j \right)^2 \leq_{icx} \sum_{i=1}^n w_i \left(Y_i - \sum_{j=1}^n v_j Y_j \right)^2.$$

□

Unfortunately, a similar result cannot be stated when working with the order statistics, since the convex order is not preserved when the ordering is applied. However, it is possible to state and prove a similar result dealing with the Gini's mean difference.

Theorem 4.3. *Let $\mathbf{X}, \mathbf{Y}$ be two random vectors such that $\mathbf{X} \leq_{cx} \mathbf{Y}$. Then,*

$$G_n(X) = \frac{1}{n(n-1)} \sum_{i,j=1}^n |X_i - X_j| \leq_{icx} \frac{1}{n(n-1)} \sum_{i,j=1}^n |Y_i - Y_j| = G_n(Y)$$

Proof. Since $\mathbf{X} \leq_{cx} \mathbf{Y}$, then there exist two random vectors $\widehat{\mathbf{X}}$ and $\widehat{\mathbf{Y}}$ defined in the same probability space such that $\mathbf{X} =_{st} \widehat{\mathbf{X}}$, $\mathbf{Y} =_{st} \widehat{\mathbf{Y}}$ and $E[\widehat{\mathbf{Y}}|\widehat{\mathbf{X}}] =_{a.s.} \widehat{\mathbf{X}}$ (see Theorem 2.2).

Define the function $h : \mathbb{R} \to \mathbb{R}$ as

$$h(\lambda) = E\left[\frac{1}{n(n-1)}\sum_{i,j=1}^{n}|\widehat{Y}_i - \sum_{j=1}^{n} v_j\widehat{Y}_j| \,\middle|\, \frac{1}{n(n-1)}\sum_{i,j=1}^{n}|\widehat{X}_i - \sum_{j=1}^{n} v_j\widehat{X}_j| = \lambda\right].$$

Consider a vector $\mathbf{x} \in \mathbb{R}^n$ such that $\frac{1}{n(n-1)}\sum_{i,j=1}^{n}|x_i - x_j| = \lambda$. If $\widehat{\mathbf{X}} = \mathbf{x}$, then we can express $\widehat{\mathbf{Y}}$ as $\widehat{\mathbf{Y}} = \mathbf{x} + \boldsymbol{\epsilon}$, with $\boldsymbol{\epsilon} = (\epsilon_1, \ldots, \epsilon_n)$ being a random vector with a null mean vector. Then, we have

$$\begin{aligned} E\left[\frac{1}{n(n-1)}\sum_{i,j=1}^{n}|Y_i - Y_j| \,\middle|\, \widehat{\mathbf{X}} = \mathbf{x}\right] &= E\left[\frac{1}{n(n-1)}\sum_{i,j=1}^{n}|x_i + \epsilon_i - x_j - \epsilon_j|\right] \\ &= \frac{1}{n(n-1)}\sum_{i,j=1}^{n} E\left[|x_i + \epsilon_i - x_j - \epsilon_j|\right] \end{aligned}$$

If we focus on each of the terms, we can use that the expectation of the absolute value is always greater than or equal to the expectation of the original random variable to compute the following inequality,

$$\left(E\left[|X_i + \epsilon_i - X_j - \epsilon_j|\right]\right)^2 \geq \left(E\left[x_i + \epsilon_i - x_j - \epsilon_j\right]\right)^2 = (x_i - x_j)^2 = |x_i - x_j|^2$$

Then, one has that $E\left[|X_i + \epsilon_i - X_j - \epsilon_j|\right] \geq |X_i - X_j|$ and, therefore,

$$\frac{1}{n(n-1)}\sum_{i,j=1}^{n} E\left[|X_i + \epsilon_i - X_j - \epsilon_j|\right] \geq \frac{1}{n(n-1)}\sum_{i,j=1}^{n}|x_i - x_j| = \lambda$$

For the rest of the proof, proceed analogously as in Theorem 4.2. □

We end this section by remarking an important consequence of the results for the case of random simple samples, that follows applying Theorem 4.2, Theorem 4.3 and Theorem 2.3.

Corollary 4.1. *Let $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_n$ be two simple random samples from X and Y, respectively. If $X \leq_{cx} Y$, then $S_X^2 \leq_{icx} S_Y^2$ and $G_n(X) \leq_{icx} G_n(Y)$.*

We want to end this section by stating that the same results can be stated for the componentwise convex order, since it implies the convex order.

5 Testing dispersion and convex orders for random vectors

We aim this section to briefly describe a possible application in hypothesis testing of the results presented in the latter sections. In particular, in testing for stochastic orders between two distinct populations with unknown distributions, which is a common problem in many applicative fields.

The most usual case in applications deals with the usual stochastic order between the distributions of the two populations, for which different hypothesis testing procedures have been developed; see, e.g., MacFadden (1989), Barret and Donald (2003) and Tse and Zhang (2004). In the case of the increasing convex order in the univariate case, there exist also several alternatives; see, e.g., Liu and Wang (2003), Baringhaus and Gruebel (2009) or Zardasht (2015).

However, it is not possible to find in the literature hypothesis testing procedures for the dispersion order or the convex order between random vectors. By using our results, these tests can be constructed using as base the well-known tests for the usual stochastic order and the increasing convex order, respectively.

Suppose that we have two simple random samples of size N from two random vectors $\mathbf{X}$ and $\mathbf{Y}$ of the same dimension n for which we want to test for the convex order. We can determine the null hypothesis as $H_0 : \mathbf{X} \leq_{cx} \mathbf{Y}$. Working under this hypothesis, we can use Theorem 4.2 and Theorem 4.3 to conclude that the quantities S_X^2 and S_Y^2, or $G_n(X)$ and $G_n(Y)$, should be ordered with respect the increasing convex order. Then, we can estimate the distributions of such quantities and test for the increasing convex order among them. If we reject the increasing convex order between the variability estimators, then we can reject the initial null hypothesis.

A similar procedure can be done with the dispersion order. In this case, we can consider the statements of Theorem 3.2 and Theorem 3.3 (or of Corollary 3.1 when using the Gini mean difference) and then test the usual stochastic order between the variability measures. A similar procedure can be found in Sordo et al. (2016), in which the order between expectations of the Gini mean differences (which is a consequence of the here-presented Corollary 3.1) is used to test some variability orders. However, this procedure was presented there only for the case of independent margins. In addition, the usual stochastic order is a stronger condition than the order based on expectations, which allows to construct more powerful tests.

6 Conclusions

Conditions that lead to the stochastic comparison of some variability measure estimators are provided. In particular, the variability measures are related to the solution of some minimization problems involving the difference between a location measure and the components of the random vector, weight vectors and possibly the order statistics. In addition, the Gini mean difference is also considered.

This type of variability measures are proven to be ordered in the usual stochastic order if each of the random vectors have identically distributed components and share the same copula (see Theorems 3.2 and 3.3). These results are also generalized to the case of different copulas in Theorems 3.4 and 3.5, the latter results being useful in the case of TTE model (see Example 3.1).

In addition, the weighted versions of the sample variance and the Gini mean difference are ordered in the increasing convex order if the initial random vectors are ordered in the convex order (Theorem 4.2 and Theorem 4.3).

A scheme of this results can be found in Figure 2. In addition, conditions that lead to the stochastic comparison of the difference of the components, which are closely related to the variability measures, are given in Theorems 3.1 and 4.1. The particular case of simple random samples and sample variance, which is specially relevant is Statistics, is stated as a consequence of these results.

These results can be the starting point to define hypothesis testing procedures for the dispersion and the convex order for random vectors, as explained in Section 5. One of the main challenges is how to determine the choice of the variability measure and the test for the usual stochastic order or the increasing convex order, since there are many alternatives. A deeper study in this regard, along with numerical results, is left as a future research study.

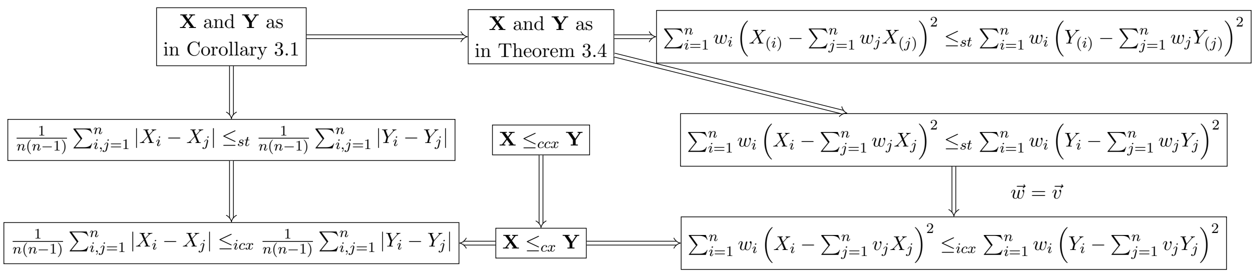

Figure 2: Scheme of the relation between the stochastic comparison of vectors and the stochastic ordering of the variability measures.

Acknowledgements

We sincerely thanks the two reviewers and the Associate Editor who made comments and suggestions that significantly contributed to the improvement of the paper.

Funding details

J. Baz is partially supported by Programa Severo Ochoa of Principality of Asturias (BP21042). F. Pellerey carried out this study within the FAIR - Future Artificial Intelligence Research and received funding from the European Union Next-GenerationEU (Piano Nazionale di Ripresa e Resilienza (PNRR) Missione 4 Componente 2, Investimento 1.3 D.D. 1555 11/10/2022, PE00000013). J. Baz, S. Montes and I. Díaz are been supported by the Ministry of Science and Innovation (PDI2022-139886NB-l00).

Disclosure statement

The authors report there are no competing interests to declare.

Biographical note

Juan Baz received the B.Sc. degree in mathematics and the B.Sc. degree in physics from the University of Oviedo, Spain, in 2020, the M.Sc. degree from the University of Oviedo, Spain, in 2021, and currently is a PhD student in University of Oviedo, Spain. Aggregation of random structures and graphical models are his main topics of research.

Franco Pellerey is Full Professor in Probability and Mathematical Statistics at the Department of Mathematical Sciences "Giuseppe Luigi Lagrange", Politecnico di Torino. His research activity is mainly focused on stochastic orderings and dependence modeling, with applications in reliability theory, actuarial sciences and engineering problems. He also boasts a long experience of research collaborations in areas related to social, medical and urban-territorial sciences. Author of more than 80 scientific publications mainly appearing in peer-reviewed international journals.

Irene Díaz received the M.Sc degree in mathematics, option applied mathematics and computation, from the University of Oviedo, Oviedo, Spain, in 1995, and the Ph.D. (cum laude) degree from the University Carlos III of Madrid, Spain, in 2001. She is currently a Full Professor of Computer Science and Artificial Intelligence at the Department of Computer Science of the University of Oviedo, where she is belongs to the research group UNIMODE. She has several publications in international journals and communications in international conferences, and she is participating in several national and international projects at the moment, some of them led by her.

Susana Montes received the M.Sc degree in mathematics, option statistics and operational research, from the University of Valladolid, Valladolid, Spain, in 1993, and the Ph.D. (cum laude) degree from the University of Oviedo, Gijón, Spain, in 1998. She is currently a Full Professor with the Department of Statistics and Operational Research, University of Oviedo, where she is the Leader of the research group UNIMODE. She has several publications in international journals and communications in international conferences, and she is participating in several national and international projects at the moment, some of them led by her. Dr. Montes received the Best Mathematics Ph.D. Thesis Award from the University of Oviedo. She is currently the President of EUSFLAT and Vice-President of IFSA.

References

[1] Amemiya, T. (1985). Asymptotic properties of extremum estimators. *Advanced econometrics* 105–158.

[2] Baringhaus, L. and Gruebel, R. (2009). Nonparametric two-sample tests for increasing convex order. *Bernoulli*, **15**(1), 99–123

[3] Barrett, G.F. and Donald, S.G. (2003). Consistent tests for stochastic dominance. *Econometrica*, **71**(1), 71–104.

[4] Calvo, T. and Beliakov, G. (2010) Aggregation functions based on penalties. *Fuzzy sets and Systems*, **161**(10), 1420–1436.

[5] Durante, F. and Sempi, C. (2015). *Principles of Copula Theory.* CRC Press, Boca Raton, FL.

[6] Grabisch, M., Marichal, J.L, Mesiar, R. and Pap, E. (2009), *Aggregation functions.* Cambridge University Press

[7] La Haye, R. and Zizler, P. (2019). The Gini mean difference and variance. *Metron*, **77**, 43–52.

[8] Liu, X. and Wang, J. (2003). Testing for increasing convex order in several populations. *Annals of the Institute of Statistical Mathematics*, **55**, 121–136.

[9] MacFadden, D. (1989). Testing for stochastic dominance. In: Fomby, T.B., Seo, T.K. (eds) *Studies in the economics of uncertainty: In honor of Josef Hadar*, 113-134. Springer, New York, NY.

[10] Marshall, A.W. and Olkin, I. (1988) Families of multivariate distributions. *Journal of the American statistical association* **83**(403), 834–841.

[11] Mulero, J., Pellerey, F. and Rodríguez-Grinolo, R. (2010). Stochastic comparisons for time transformed exponential models. *Insurance: Mathematics and Economics*, **46**(2), 328–333.

[12] Müller, A. (1997). Stochastic orders generated by integrals: a unified study. *Advances in Applied probability*, **29**(2), 414-428.

[13] Nelsen, R. B. (2006). *An Introduction to Copulas*, Springer-Verlag, New York.

[14] Rohatgi, V.K. and Saleh A.M.E. (2015). *An introduction to probability and statistics.* John Wiley & Sons.

[15] Shaked, M. and Shanthikumar, J.G. (2007), *Stochastic orders*, Springer Verlag, New York.

[16] Sordo, M.A., de Souza, M.C. and Suárez-Llorens, A. (2016). Testing variability orderings by using Gini's mean differences. *Statistical Methodology*, **32**, 63–76.

[17] Tse, Y. K. and Zhang, X. (2004). A Monte Carlo investigation of some tests for stochastic dominance. *Journal of Statistical Computation and Simulation*, **74**(5), 361–378.

[18] Zardasht, V. (2015). A test for the increasing convex order based on the cumulative residual entropy. *Journal of the Korean Statistical Society*, **44**, 491–497.