POLITECNICO DI TORINO
Repository ISTITUZIONALE

On the property of reduction of variability for mean functions

Original
On the property of reduction of variability for mean functions / Baz Gonzalez, Juan; Pellerey, Franco; Diaz, Irene; Montes, Susana. - In: COMPUTATIONAL AND APPLIED MATHEMATICS. - ISSN 0101-8205. - ELETTRONICO. - 44:(2025), pp. 1-23. [10.1007/s40314-025-03250-7]

Availability:
This version is available at: 11583/3000413 since: 2025-06-27T08:39:41Z

Publisher:
Springer

Published
DOI:10.1007/s40314-025-03250-7

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)

13 May 2026

Computational and Applied Mathematics (2025) 44:297
https://doi.org/10.1007/s40314-025-03250-7

On the property of reduction of variability for mean functions

Juan Baz[1] · **Franco Pellerey**[2] · **Irene Díaz**[3] · **Susana Montes**[1]

Received: 7 December 2024 / Revised: 3 April 2025 / Accepted: 30 April 2025
© The Author(s) 2025

Abstract
Mean functions, formally introduced by Cauchy in 1821, are common tools both in the estimation of location measures of unknown distributions and in the aggregation of given data with the purpose of reducing their variability before application of prediction algorithms. However, apart for few specific cases, for this family of statistics there is a lack of existence of formal proofs confirming the reduction in variability with respect to original data. In this paper we provide several results dealing with this property, providing conditions such that the components of a sample and the corresponding mean function are comparable in the convex stochastic order. These results refer to different families of mean functions, such as Ordered Weighted Averagings (OWA), weighted quasi-arithmetic means and the nullnorms class of operators, which are aggregation functions of interest in the field of fuzzy neural networks and in machine learning algorithms. Additional results dealing with the closure under distortion of the convex stochastic order are provided as well.

Keywords Aggregation functions · Mean functions · Convex stochastic order · Copulas

Mathematics Subject Classification 60E15 · 62D99

✉ Juan Baz
bazjuan@uniovi.es

Franco Pellerey
franco.pellerey@polito.it

Irene Díaz
sirene@uniovi.es

Susana Montes
montes@uniovi.es

1 Department of Statistics and I.O. and Didactics of Mathematics, University of Oviedo, Calle Federico Garcia Lorca 18, Oviedo 33007, Asturias, Spain

2 Dipartimento di Scienze Matematiche, Politecnico di Torino, Corso Duca degli Abruzzi, 24, 10129 Torino, Piedmont, Italy

3 Department of Computer Science, University of Oviedo, Campus de Viesques, Gijón 33204, Asturias, Spain

Published online: 26 May 2025

1 Introduction

Given a sample $\vec{X} = (X_1, \ldots, X_n)$ extracted from a population X, a common procedure in inferential problems is to estimate location measures of X by means of statistics that are members of the family of mean functions, introduced by Cauchy (1821), which are functions that satisfy the property of assuming values between the minimum and the maximum of the initial values. Apart for the estimation of location measures (Crow and Siddiqui 1967; Lloyd 1952), mean functions are nowadays also widely used in a wide range of applied sciences to summarize the behavior of X through transformations of $\vec{X}$ whose final distributions have a reduced variability with respect to that of X (Ruta and Gabrys 2007; Leite and Škrjanc 2019; Melnikov and Hüllermeier 2016).

It is a well-know fact that for a number of specific mean functions, such as the arithmetic mean for which $\mathrm{Var}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \mathrm{Var}(X)/n$ when the observations are independent, the application of the function induces a reduction of the variance with respect to the original distribution of X. However, to the best knowledge of the authors, in the literature there are no results of this kind dealing with: (i) stronger notions of comparisons in terms of variability or dispersion, (ii) non-independent random variables, (iii) for more general families of mean functions. The purpose of this paper is to partially cover this gap, providing formal proofs of intuitions such as, for example, the fact that averages of central order statistics have less variability than averages of the extreme onesó. On this aim, the paper provides new results describing conditions for a random vector $\vec{X}$ and a mean function f in order to satisfy

$$f(\vec{X}) \leq_{cx} X_j, \forall j \in \{1, \ldots, n\},$$

where $\leq_{cx}$ stands for the convex stochastic order, which is a common tool to compare the dispersion of two random variables (see Definition 6).

In particular, the results provided here deals with the families of weighted quasi-arithmetic means, Ordered Weighted Averagings (OWA) operators, nullnorms and uninorms. Weighted quasi-arithmetic means are one of the most used family of mean functions, see (Bullen and Bullen 2003), and include the arithmetic mean, the geometric mean and the harmonic mean as particular cases. OWA operators, also known as convex L-statistics, are convex linear combinations of the ordered values of the sample (Hosking 1998) and are widely used in many applicative areas (Zhou and Chen 2011; Scherger et al. 2017). Trimmed means and sample quantiles, including the median, are examples of OWA operators. Finally, we are also interested in nullnorms and uninorms, which are generalizations of t-norms and t-conorms (Grabisch et al. 2009). They are crucial operators, for example, in the definition of fuzzy neural networks (Lemos et al. 2010; Campos Souza and Lughofer 2021).

Illustrative examples, as well as additional results dealing with the closure of the convex stochastic order with respect to distortions (Hürlimann 2004), which are relevant by their own in the theory of stochastic orders, are presented as well.

The paper is organized as follows. In Sect. 2 we recall some basic concepts about mean functions, stochastic orders, copulas and distortions, and we also provide a preliminary result applied in subsequent sections. Section 3 is focused on the property of reduction in variability for the weighted quasi-arithmetic means, while Sects. 4, 5 and 6 are devoted to OWA operators, and nullnorms and uninorms, respectively.

2 Preliminaries

We aim this section to introduce the preliminary notions needed for the development of the paper, i.e., mean functions, the convex stochastic order, copulas and distortions. A preliminary results dealing with the relation between distortions and the convex stochastic order, applied in subsequent sections, is also provided.

2.1 Mean functions

One of the first definitions of mean function is due to Cauchy Cauchy (1821), requiring the function to assume values between the minimum and the maximum of the arguments. The modern definition of mean function also requires increasing monotonicity (Grabisch et al. 2009). In the following, we consider I to be a closed non-degenerate interval of the real line.

Definition 1 (Grabisch et al. 2009) A function $f : I^n \to I$ is said to be a mean function if

- $\min(\vec{x}) \leq f(\vec{x}) \leq \max(\vec{x})$ for any $\vec{x} \in I^n$;
- $f(\vec{x}) \leq f(\vec{y})$ for any $\vec{x}, \vec{y} \in I^n$ such that $\vec{x} \leq \vec{y}$.

In the latter definition, $\vec{x} \leq \vec{y}$ stands for the component-wise order, that is, $x_i \leq y_i$ for any $i \in \{1, \ldots, n\}$. It can be proved that the definition is equivalent to the following characterization.

Proposition 1 (Grabisch et al. 2009) *A function $f : I^n \to I$ is a mean function if, and only if,*

- *$f(x, \ldots, x) = x$ for any $x \in I$;*
- *$f(\vec{x}) \leq f(\vec{y})$ for any $\vec{x}, \vec{y} \in I^n$ such that $\vec{x} \leq \vec{y}$.*

The property $f(x, \ldots, x) = x$ for any $x \in I$ is usually known as *idempotence*. Among the family of mean functions, the first one that will be considered in the sequel is the *quasi-arithmetic weighted means* family, whose definition is the following.

Definition 2 (Bullen and Bullen 2003) A function $f : I^n \to I$ is said to be a quasi-arithmetic weighted mean if can expressed as

$$f(\vec{x}) = g^{-1}\left(\sum_{i=1}^{n} w_i g(x_i)\right), \quad \forall \vec{x} \in I^n,$$

where $g : I \to \mathbb{R}$ is any real-valued strictly monotone function and $\vec{w} \in [0, 1]^n$ is a weighting vector such that $\sum_{i=1}^{n} w_i = 1$.

Quasi-arithmetic weighted averages are also known as Kolmogorov-Nagumo means (Bullen and Bullen 2003) and include, among others, weighed means, weighted geometric means, weighted harmonic means and weighted power means. The second family that will be considered through the paper is the family of OWA operators, which are functions of the vector of the order statistics corresponding to the given sample.

Definition 3 (Yager 1993) A function $f : I^n \to I$ is said to be an *Ordered Weighted Averaging* (OWA) operator if it can be expressed as

$$f(\vec{x}) = \sum_{i=1}^{n} w_i x_{\sigma(i)},$$

where $\sigma : \{1, \dots, n\} \to \{1, \dots, n\}$ denotes the (bijective) permutation such that $x_{\sigma(1)} \leq \dots \leq x_{\sigma(1)}$ and $\vec{w} \in [0, 1]^n$ is a weighting vector satisfying $\sum_{i=1}^{n} w_i = 1$.

In Statistics, these functions are commonly called *convex L-statistics* (Shao 1994). With the adequate election of the weights, they are robust to outliers. In particular, trimmed means, which are defined as the average of the central values of the ordered sample, are OWA operators. Moreover, the median is a trimmed mean and, therefore, an OWA operator.

Finally, we recall the definitions of *nullnorms* and *uninorms*. We recall them considering the interval I to be the unit interval $[0, 1] \subseteq \mathbb{R}$, as it is commonly considered in the literature (even if the definition can be extended to more general intervals of real numbers). Also, to simplify the notation, here and in the proofs appearing in the next sections, the definitions are given considering a flexible number of arguments: the functions are defined over $\cup_{m=1}^{\infty}[0, 1]^m$.

Definition 4 (Grabisch et al. 2009) A function $N : \cup_{m=1}^{\infty}[0, 1]^m \to [0, 1]$ is said to be a nullnorm if

- it is increasing in each of its arguments;
- $N(0, \dots, 0) = 0$ and $N(1, \dots, 1) = 1$;
- $N(x) = x$ for any $x \in [0, 1]$;
- $N(\vec{x}) = N(\vec{y})$ if $\vec{y}$ is any permutation of the components of $\vec{x}$;
- $N(\vec{x}) = N(N(x_1, \dots, x_k), N(x_{k+1}, \dots, x_n))$ for any $k, n \in \mathbb{N}$ such that $k < n$;
- there exists $a \in [0, 1]$ such that, if $a \in \{x_1, \dots, x_n\}$, then $N(\vec{x}) = a$.

The fourth and fifth properties are known as, respectively, *symmetry* and *associativity*. The element a in the sixth property is called an *annihilator element*. The definition for uninorms is the same but for the last property.

Definition 5 (Grabisch et al. 2009) A function $U : \cup_{m=1}^{\infty}[0, 1]^m \to [0, 1]$ is said to be a uninorm if

- it is increasing in each of its arguments;
- $U(0, \dots, 0) = 0$ and $U(1, \dots, 1) = 1$;
- $U(x) = x$ for any $x \in [0, 1]$;
- $U(\vec{x}) = U(\vec{y})$ if $\vec{y}$ is a permutation of the components of $\vec{x}$;
- $U(\vec{x}) = U(U(x_1, \dots, x_k), U(x_{k+1}, \dots, x_n))$ for any $k, n \in \mathbb{N}$ such that $k < n$;
- There exists $e \in [0, 1]$ such that,

$$U(x_1, \dots, x_{k-1}, e, x_{k+1}, \dots, x_n) = U(x_1, \dots, x_{k-1}, x_{k+1}, \dots, x_n)$$

  for any $k, n \in \mathbb{N}$ such that $k \leq n$.

In this case, the element e is known as a *neutral element* of the function. In general, nullnorms and uninorms do not have to be mean functions. When they are, one has the following characterizations.

Proposition 2 (Czogała and Drewniak 1984) *Let* $N : \cup_{m=1}^{\infty}[0, 1]^m \to [0, 1]$ *be an idempotent nullnorm with annihilator element* $a \in [0, 1]$. *Then,*

$$N(x, y) = \begin{cases} \max(x, y) & \textit{if } x, y < a \\ \min(x, y) & \textit{if } x, y > a \\ a & \textit{elsewhere} \end{cases}$$

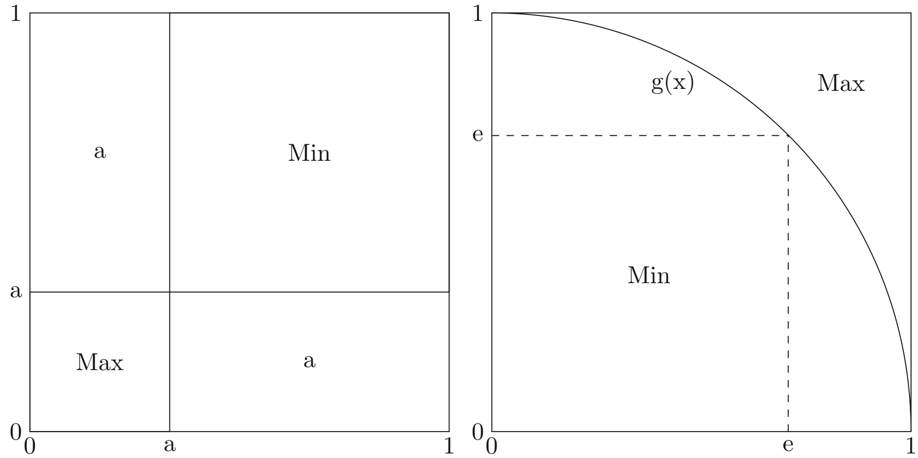

Fig. 1 Structure of an idempotent nullnorm with annihilator element a (left) and an idempotent uninorm with neutral element e (right)

Proposition 3 (Czogała and Drewniak 1984) *Let* $U : \cup_{m=1}^{\infty}[0,1]^m \to [0,1]$ *be an idempotent uninorm with neutral element* $e \in [0,1]$. *Then, there exists a decreasing function* $g : [0,1] \to [0,1]$ *with* $g(e) = e$ *such that:*

$$U(x, y) = \begin{cases} \min(x, y) & \textit{if } y < g(x) \\ \min(x, y) \textit{ or } \max(x, y) & \textit{if } y = g(x) \\ \max(x, y) & \textit{if } y > g(x) \end{cases}$$

Remark 1 The converse of Proposition 3 is not true. For instance, considering $e = 0.5$, the function $g(x) = 1$ if $x \in [0, 0.5)$, $g(0.5) = 0.5$ and $g(x) = 0$ if $x \in (0.5, 0]$ is decreasing and fulfills $g(e) = e$, but the resulting function $U(x, y)$ is not an uninorm. For instance, since $U(0.3, 0.7) = \min(0.3, 0.7) = 0.3$ and $U(0.7, 0.3) = \max(0.7, 0.3) = 0.7$, the symmetry is not fulfilled.

Notice that, even if the latter results only give the structure when the number of inputs is 2, the associativity allows to construct the value of the nullnorm or uninorm just by the recursive application of the bidimensional case. For instance, for dimension 3 one can write $N(x_1, x_2, x_3) = N(N(x_1, x_2), x_3)$ and $U(x_1, x_2, x_3) = U(U(x_1, x_2), x_3)$. A couple of examples of such structures can be found in Fig. 1.

2.2 The convex stochastic order

Stochastic orders are relations between probability distributions that compare them in terms of location, variability or dependence. Since any random variable induces a probability distribution over the real numbers, they can be used also in ordering random variables. In the following, given a random variable X, we will denote as $E[X]$ and $\mathrm{Var}(X)$ its expectation (or mean) and its variance (Rohatgi and Saleh 2015).

Along this paper, we are interested in the *convex stochastic order*, which is one of the most well-known stochastic orders comparing variability, whose definition is the following.

For it, recall that a function $\phi : \mathbb{R} \to \mathbb{R}$ is said to be convex if $\phi(\lambda x + (1-\lambda)y) \leq \lambda\phi(x) + (1-\lambda)\phi(y)$ for any $x, y \in \mathbb{R}$ and $\lambda \in [0, 1]$.

Definition 6 (Shaked and Shanthikumar 2007) Let X and Y be two random variables. If $E[\phi(X)] \leq E[\phi(Y)]$ for any convex function $\phi : \mathbb{R} \to \mathbb{R}$ such that $E[\phi(X)]$ and $E[\phi(Y)]$ exist, then X is said to be smaller than or equal to Y in the convex stochastic order, and it is denoted as $X \leq_{cx} Y$.

It is easy to verify that the convex stochastic order implies both $\mathrm{Var}(X) \leq \mathrm{Var}(Y)$ and $E[X] = E[Y]$, and this is one of the reasons it is interpreted as an order comparing dispersion of two random variables.

A related but weaker order that will be considered in the sequel is the *increasing convex stochastic order*, defined in a similar way but considering increasing convex functions.

Definition 7 (Shaked and Shanthikumar 2007) Let X and Y be two random variables. If $E[\phi(X)] \leq E[\phi(Y)]$ for any increasing convex function $\phi : \mathbb{R} \to \mathbb{R}$ such that $E[\phi(X)]$ and $E[\phi(Y)]$ exist, then X is said to be smaller than or equal to Y in the increasing convex stochastic order and it is denoted as $X \leq_{icx} Y$.

It has to be noted that both the convex and increasing convex orders are trivially reflexive and transitive, but they are not partial orders in the set of random vectors of the same dimensions (since there are different random vectors with the same distribution).

The increasing convex order is implied by the convex order and fulfills $X \leq_{icx} Y \implies E[X] \leq E[Y]$. The interpretation of the increasing convex order is not univocal, since it compare the variables both in terms of location and in terms of dispersion.

For some of the proofs in the following sections, we will use characterizations of the latter orders by means of integrals of the distribution and survival functions of the variables. We recall that the distribution function of a random variable X is defined as $F(t) = P(X \leq t)$ and the survival function is defined as $\bar{F}(t) = 1 - F(t) = P(X > t)$.

Theorem 1 (Shaked and Shanthikumar 2007) *Let X and Y be two random variables with distribution functions $F_X(t)$ and $F_Y(t)$. Then, the following statements are equivalent:*

- $X \leq_{cx} Y$,
- $E[X] = E[Y]$ *and*
$$\int_{-\infty}^{x} F_X(t)dt \leq \int_{-\infty}^{x} F_Y(t)dt, \ \forall x \in \mathbb{R},$$
- $E[X] = E[Y]$ *and*
$$\int_{x}^{\infty} \bar{F}_X(t)dt \leq \int_{x}^{\infty} \bar{F}_Y(t)dt, \ \forall x \in \mathbb{R}.$$

Theorem 2 (Shaked and Shanthikumar 2007) *Let X and Y two random variables with distribution functions $F_X(t)$ and $F_Y(t)$. Then, $X \leq_{icx} Y$ if and only if*

$$\int_{x}^{\infty} \bar{F}_X(t)dt \leq \int_{x}^{\infty} \bar{F}_Y(t)dt, \ \forall x \in \mathbb{R}.$$

We refer the reader to sections 3.A and 4.A in Shaked and Shanthikumar (2007) for properties and applications of the convex and the increasing convex stochastic orders.

2.3 Copulas and distortions

In some of the results, the notions of copulas and distortions are needed. Therefore, we briefly introduce these notions in this section.

Similarly to the univariate case, the distribution function of a random vector $\vec{X} = (X_1, \ldots, X_n)$ is defined as $F(t_1, \ldots, t_n) = P(X_1 \le t_1, \ldots, X_n \le t_n)$ and its survival function as $\bar{F}(t_1, \ldots, t_n) = P(X_1 > t_1, \ldots, X_n > t_n)$. *Copulas* appear as functions that entirely describes the dependence structure between the components of the random vector, and that relate the multivariate distribution function with the marginal distribution functions as follows.

Definition 8 (Nelsen 2007; Sklar 1959) Let $\vec{X}$ be a random vector with marginals distributions $F_1, \ldots, F_n$ and joint distribution function F. Then, a function $C : [0, 1]^n \to [0, 1]$ such that

$$F(t_1, \ldots, t_n) = C\left(F_1(t_1), \ldots, F_n(t_n)\right), \ \forall (t_1, \ldots, t_n) \in \mathbb{R}^n$$

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

Similarly, the survival function can be written as $\bar{F}(t_1, \ldots, t_n) = \bar{C}(\bar{F}_1(t_1), \ldots, \bar{F}_1(t_n))$, where $\bar{C}$ is the *survival copula* of the vector $\vec{X}$. The most famous copula is the product copula, $C(t_1, \ldots, t_n) = \prod_{i=1}^{n} t_i$, which is associated with the case of independent random variables. Moreover, if the marginal distributions are continuous, then the copula of the random vector is unique.

Another family of functions that will be used along the paper is the family of distortions functions. A *distortion* is a right-continuous increasing function $h : [0, 1] \to [0, 1]$ such that $h(0) = 0$ and $h(1) = 1$. One of the main characteristic of any distortion function is that the composition of a distribution function and a distortion is again a distribution function. In this regard, they have been used widely in reliability theory, see for instance (Arriaza and Sordo 2021; Navarro and Águila 2017), since the distribution functions of many coherent systems can be expressed as distortions of the distribution function of the lifetimes of its components, when they are identically distributed.

Given a random variable X with distribution function F, we will denote as X_h a random variable with distribution $h \circ F$. A special type of distortions will be of relevance for some the results presented in the next sections, since they preserve the symmetry of distributions.

Definition 9 A distortion h is said to be centrally symmetric if $h(t) = 1 - h(1 - t)$ for any $t \in [0, 1]$.

We use the term centrally symmetric to describe distortion satisfying this property since when it holds, then the graph of the function is symmetric with respect to the point $(0.5, 0.5)$, the center of the unit square.

Also related with symmetry, we recall that a random variable X is said to be *symmetric* if its distribution function fulfills $F(m + t) = 1 - F(m - t)$, being m the median of X, that is, $F(m) = 0.5$. For any symmetric random variable, the mean and the median equal the same value. In addition, it is clear that if a distortion h is centrally symmetric and X is symmetric, then X_h is also symmetric.

To prove the main statements for some of mean functions we are going to consider, we will make use of the following preliminary result, which describes conditions for the closure of the convex order with respect to centrally symmetric distortions.

Lemma 1 *Let h_1 and h_2 be two centrally symmetric distortions such that $h_1(t) \leq h_2(t)$ for any $t \in [0, 0.5]$, and let X be a symmetric random variable. Then,*

$$X_{h_1} \leq_{cx} X_{h_2}.$$

Proof Firstly, it is clear that, since h_1 and h_2 are centrally symmetric and X is symmetric, then both X_{h_1} and X_{h_2} are symmetric and have the same median m. Therefore, they have also the same mean m.

Let us consider the integrals

$$I_1(x) = \int_{-\infty}^{x} h_1(F(t))dt \quad \text{and} \quad I_2(x) = \int_{-\infty}^{x} h_2(F(t))dt.$$

Let $x \leq m$. In this case one has that $h_1(F(t)) \leq h_2(F(t))$ for any $t \in (-\infty, x]$, and therefore, $I_1(x) \leq I_2(x)$ for any $x \leq m$.

Let $x \geq m$. Note that, since X is symmetric, for all $t \geq m$ it holds $F(t) = \bar{F}(2m - t)$. Moreover, since h_1 is centrally symmetric, one also has $h_1(F(t)) = h_1(\bar{F}(2m - t)) = h_1(1 - F(2m - t)) = 1 - h_1(F(2m - t))$. It follows that

$$\begin{aligned} I_1(x) &= \int_{-\infty}^{x} h_1(F(t))dt = \int_{-\infty}^{m} h_1(F(t))dt + \int_{m}^{x} h_1(F(t))dt \\ &= \int_{-\infty}^{m} h_1(F(t))dt + \int_{m}^{x} (1 - h_1(F(2m - t)))\, dt = \int_{-\infty}^{2m-x} h_1(F(t))dt + (x - m) \end{aligned}$$

Similarly, one has that

$$I_2(x) = \int_{-\infty}^{2m-x} h_2(F(t))dt + (x - m).$$

Observing now that $2m - x \leq m$, being $x \geq m$, it follows that the inequality $I_1(x) \leq I_2(x)$ holds in this case as well.

It is concluded that $I_1(x) \leq I_2(x)$ for any $x \in \mathbb{R}$ and, as a consequence of Theorem 1 and the equality $E[X_{h_1}] = E[X_{h_2}]$, that $X_{h_1} \leq_{cx} X_{h_2}$. □

3 Reduction of variability for weighted quasi-arithmetic means

In this section the case of weighted quasi-arithmetic means is considered. As introduced in Definition 2, there is a linear convex combination involved in their computation, so one may expect a good behaviour with respect to the convex order. In fact, this is the case for weighted means.

Theorem 3 *Let $\vec{X}$ be a random vector of dimension n with identically distributed components and $\vec{w} \in [0, 1]^n$ a weighting vector such that $\sum_{i=1}^{n} w_i = 1$. Then,*

$$\sum_{i=1}^{n} w_i X_i \leq_{cx} X_j, \ \forall j \in \{1, \ldots, n\}$$

Proof Let ϕ be any convex function. Then,

$$E\left[\phi\left(\sum_{i=1}^{n} w_i X_i\right)\right] \leq E\left[\sum_{i=1}^{n} w_i \phi(X_i)\right] = \sum_{i=1}^{n} w_i E[\phi(X_i)]$$

Since $X_1, \ldots, X_n$ have the same distribution, it holds that $\phi(X_1), \ldots, \phi(X_n)$ also have the same distribution. Therefore,

$$E\left[\phi\left(\sum_{i=1}^{n} w_i X_i\right)\right] \le \sum_{i=1}^{n} w_i E[\phi(X_i)] = \sum_{i=1}^{n} w_i E[\phi(X_j)] = E[\phi(X_j)], \ \forall j \in \{1, \ldots, n\}$$

□

Notice that the result holds even in the case of non-independent components. However, it cannot be extended to other weighted quasi-arithmetic means, mainly because the same expectation for the compared variables is a necessary condition for the order to hold. Let us give an example in this regard.

Example 1 Let (X_1, X_2) be a random vector with standard uniform and independent variables. Then the inequality $\sqrt{X_1 X_2} \le_{cx} X_1$ cannot be true, since the two quantities have different expectations, being

$$E[\sqrt{X_1 X_2}] = E[\sqrt{X_1}]E[\sqrt{X_1}] = \frac{4}{9} \neq \frac{1}{2} = E[X_1]$$

Fortunately, if the adequate transformation is applied to both sides, one can prove a statement similar to that of Theorem 3.

Corollary 1 *Let f be a quasi-arithmetic weighted mean with associated real-valued strictly increasing function g. Then, for any random vector $\vec{X}$ with identically distributed components, it holds*

$$g(f(\vec{X})) \le_{cx} g(X_j), \forall j \in \{1, \ldots, n\}.$$

Proof Notice that $g(f(\vec{X})) = g(g^{-1}(\sum_{i=1}^{n} w_i g(X_i)) = \sum_{i=1}^{n} w_i g(X_i)$ and apply Theorem 3. □

We point out that the latter result can be applied for example to the geometric, harmonic and power means.

4 Reduction of variability for OWA operators

Ordered Weighted Averaging operators defined on samples of independent and identically distributed random variables are considered in this section. Given the sample $\vec{X}$, in the following the notation $X_{(i)}$ is used for the corresponding i-th order statistic of the components of $\vec{X}$, i.e. the i-th smaller value among $\vec{X}$.

Recall that the distribution function of an order statistic of a random vector with independent and identically distributed components is given by

$$F_{(i)}(t) = \sum_{k=i}^{n} \binom{n}{k} F(t)^k (1 - F(t))^{n-k},$$

where $F(t)$ is the distribution function of the components of $\vec{X}$ (David and Nagaraja 2004). Therefore, we can write $F_{(i)}(t) = h_{(i)}(F(t))$, being $h_{(i)}(t) = \sum_{k=i}^{n} \binom{n}{k} t^k (1-t)^{n-k}$.

First, we prove a statement showing that the average of the distribution functions is smaller in the convex order for central order statistics than for extreme order statistics. To better understand it, it must be pointed out the difference existing between the mixture of two order statistics and the convex linear combination of two order statistics, since they may be confusing. Consider two weights $w_i, w_j \in [0, 1]$ such that $w_i + w_j = 1$. The mixture of the order statistics $X_{(i)}$ and $X_{(j)}$ is a new random variable Y with distribution function $G(t) = w_i F_{(i)}(t) + w_j F_{(j)}(t)$ for any $t \in \mathbb{R}$. The linear convex combination is the random variable defined as $w_i X_{(i)} + w_j X_{(j)}$. That is, the first is a combination of the distributions and the second is a combination of the random variables.

Lemma 2 *Let $\vec{X}$ be a random vector of dimension n with independent, identically distributed and symmetric components. Let Y_i be a random variable with distribution function $G_i(t) = \frac{1}{2}F_{(i)}(t) + \frac{1}{2}F_{(n-i+1)}(t)$ for any $t \in \mathbb{R}$. Then,*

$$|n - 2i + 1| \le |n - 2j + 1| \implies Y_i \le_{cx} Y_j,$$

for any $i, j \in \{1, \ldots, n\}$

Proof Consider $i \le n - i + 1$ and $j \le n - j + 1$. For the other cases, notice that Y_j has the same distribution as Y_{n-j+1}. Since $|n - 2i + 1| \le |n - 2j + 1|$, then $j \le i$.

We can express the distortions associated to Y_i and Y_j as

$$\begin{aligned}
h_i(t) &= \frac{1}{2}h_{(i)}(t) + \frac{1}{2}h_{(n-i+1)}(t)\\
&= \frac{1}{2}\sum_{k=i}^{n}\binom{n}{k} t^k(1-t)^{n-k} + \frac{1}{2}\sum_{k=n-i+1}^{n}\binom{n}{k} t^k(1-t)^{n-k},\\
h_j(t) &= \frac{1}{2}h_{(j)}(t) + \frac{1}{2}h_{(n-j+1)}(t)\\
&= \frac{1}{2}\sum_{k=j}^{n}\binom{n}{k} t^k(1-t)^{n-k} + \frac{1}{2}\sum_{k=n-j+1}^{n}\binom{n}{k} t^k(1-t)^{n-k}.
\end{aligned}$$

It can be observed that h_i is centrally symmetric, being

$$\begin{aligned}
h_i(1-t) + h_i(t) &= \frac{1}{2}\sum_{k=i}^{n}\binom{n}{k}(1-t)^k t^{n-k} + \frac{1}{2}\sum_{k=n-i+1}^{n}\binom{n}{k}(1-t)^k t^{n-k}\\
&\quad+\frac{1}{2}\sum_{k=i}^{n}\binom{n}{k} t^k(1-t)^{n-k} + \frac{1}{2}\sum_{k=n-i+1}^{n}\binom{n}{k} t^k(1-t)^{n-k}\\
&= \frac{1}{2}\sum_{k=i}^{n}\binom{n}{k}(1-t)^k t^{n-k} + \frac{1}{2} - \frac{1}{2}\sum_{k=0}^{n-i}\binom{n}{k}(1-t)^k(t)^{n-k}\\
&\quad+\frac{1}{2}\sum_{k=i}^{n}\binom{n}{k} t^k(1-t)^{n-k} + \frac{1}{2} - \frac{1}{2}\sum_{k=0}^{n-i}\binom{n}{k} t^k(1-t)^{n-k}\\
&= \frac{1}{2}\sum_{k=i}^{n}\binom{n}{k}(1-t)^k t^{n-k} + \frac{1}{2} - \frac{1}{2}\sum_{k=i}^{n}\binom{n}{k}(1-t)^{n-k}t^k\\
&\quad+\frac{1}{2}\sum_{k=i}^{n}\binom{n}{k} t^k(1-t)^{n-k} + \frac{1}{2} - \frac{1}{2}\sum_{k=i}^{n}\binom{n}{k} t^{n-k}(1-t)^k = 1.
\end{aligned}$$

Similarly, the distortion h_j is also centrally symmetric. In addition, the difference between the distortions h_i and h_j is given by

$$h_i(t) - h_j(t) = \frac{1}{2} \sum_{k=n-i+1}^{n-j} \binom{n}{k} t^k (1-t)^{n-k} - \frac{1}{2} \sum_{k=j}^{i-1} \binom{n}{k} t^k (1-t)^{n-k}.$$

Notice that the index of the first summand can be reversed by the operation $n-k$, resulting in a unique sum over the same indices, thus obtaining

$$h_i(t) - h_j(t) = \sum_{k=i}^{j-1} \binom{n}{k} \left(t^{n-k}(1-t)^k - t^k(1-t)^{n-k} \right).$$

Since k is at most $j-1$, then $n-k$ is always greater than k. Therefore, if $t < 0.5$ all the summands are negative and the sum is negative. It is concluded that $h_i(t) \leq h_j(t)$ for any $t \in [0, 0.5]$. The statement now holds by Lemma 1. □

The latter result can be interpreted, when considering uniform mixtures, as the central order statistics having less variability than the extreme ones. Intuitively, one may think that, therefore, OWA operators having bigger central weights should reduce the dispersion. In the next result, we formalize this intuition.

Theorem 4 *Let $\vec{X}$ be a random vector of dimension n with symmetric independent and identically distributed components, and let $\vec{w} \in [0,1]^n$ be a weighting vector such that:*

1. $\sum_{i=1}^n w_i = 1$,
2. $|n-2i+1| \leq |n-2j+1| \implies w_i \geq w_j$,
3. $w_i = w_{n-i+1}$ *for any* $i, j \in \{1, \ldots, n\}$.

Then,

$$\sum_{i=1}^n w_i X_{(i)} \leq_{cx} X_j, \quad \forall j \in \{1, \ldots, n\}.$$

Proof Denote the distribution function of the components of $\vec{X}$ as F and consider $G_j(t) = \frac{1}{2}F_{(j)}(t) + \frac{1}{2}F_{(n-j+1)}(t)$ for any $t \in \mathbb{R}$, where $F_{(j)}$ and $F_{(n-j+1)}$ denote, respectively, the distribution functions of $X_{(j)}$ and $X_{(n-j+1)}$. Notice that any component of $\vec{X}$ can be written as a mixture of the order statistics of $\vec{X}$, i.e.,

$$F(t) = \frac{1}{n} \sum_{i=1}^n F_{(i)}(t) = \frac{1}{n} \sum_{i=1}^n G_i(t), \ \forall t \in \mathbb{R}. \tag{4.1}$$

Consider now a random variable Z with distribution function F_Z defined as

$$F_Z(t) = \sum_{i=1}^n w_i F_{(i)}(t) = \sum_{i=1}^n w_i G_i(t), \ \forall t \in \mathbb{R}, \tag{4.2}$$

where the second equality holds from the fact that $w_i = w_{n-i+1}$ for any $i \in \mathbb{N}$.

Thus, the variable Z and any component of $\vec{X}$ can be written as a mixture of the random variables Y_i defined as before, which are in turn mixtures of the order statistics. By applying Lemma 2 one has that $|n-2i+1| \leq |n-2j+1|$ implies $Y_i \leq_{cx} Y_j$. In addition, the weights fulfill $|n-2i+1| \le |n-2j+1| \implies w_i \ge w_j$. Then, for any convex function $\phi : \mathbb{R} \to \mathbb{R}$ one has

$$E[\phi(Z)] = \sum_{i=1}^{n} w_i E[\phi(Y_i)] \le \frac{1}{n}\sum_{i=1}^{n} E[\phi(Y_i)] = E[\phi(X_j)], \forall j \in \{1, \dots, n\}.$$

where the inequality holds by the fact that $E[\phi(Y_i)] \le E[\phi(Y_j)]$ if $w_i \ge w_j$, thus bigger weights are associated with smaller expectations, while the equalities follows from Eqs. (4.2) and (4.1), respectively. Thus $Z \le_{cx} X_j$ holds for any $j \in \{1, \dots, n\}$. Now observe that, again considering any convex function $\phi : \mathbb{R} \to \mathbb{R}$, it holds

$$E\left[\phi\left(\sum_{i=1}^{n} w_i X_{(i)}\right)\right] \le E\left[\sum_{i=1}^{n} w_i \phi\left(X_{(i)}\right)\right] = \sum_{i=1}^{n} w_i E\left[\phi\left(X_{(i)}\right)\right] = E[\phi(Z)]$$

where the equality follows from the convexity of ϕ and the last equality by Eq. (4.2).

It is concluded that $\sum_{i=1}^{n} w_i X_{(i)} \le_{cx} Z \le_{cx} X_j$, $j \in \{1, \dots, n\}$, thus the assertion. □

Let us illustrate the latter result with a simple example.

Example 2 Let $\vec{X} = (X_1, X_2, X_3)$ be a random vector with standard uniform components. Let us compute the integral of the distribution function of X_1 and $X_{(2)}$,

$$\int_0^x F_1(t)dt = \int_0^x t dt = \frac{1}{2}x^2, \quad \int_0^x F_{(2)}(t)dt = \int_0^x \left(3t^2(1-t) + t^3\right) dt = x^3 - \frac{1}{2}x^4,$$

for any $x \in [0, 1]$. Observing that $\frac{x^2}{2} \ge x^3 - \frac{x^4}{2}$ for any $x \in [0, 1]$ and that $E[X_1] = E[X_{(2)}] = 0.5$, it is concluded that $X_1 \ge_{cx} X_{(2)}$.

The required symmetry for the distribution and the weights is due to the necessity of having the same expectation for the convex order to hold. Having symmetric and bigger central weights is a condition that appears naturally in some contexts of mean estimation of particular location-scale families related with, for example, Laplace, Logistic of Hyperbolic secant distribution (see Figure 3 in Baz et al. 2024a for a representation of sums of such weights) and also constitute a wide family of OWA operators (Yager 2007). Unfortunately, if we remove the condition of having bigger central weights, the result is not longer true, as shown in the following example.

Example 3 Let X_1, X_2 and X_3 be three independent and identically distributed random variables such that $P(X_i = 0) = P(X_i = 2) = 0.1$ and $P(X_i = 1) = 0.8$ for any $i \in \{1, 2, 3\}$. Consider $Y = \frac{1}{2}X_{(1)} + \frac{1}{2}X_{(3)}$. The values of Y and the associated probability for any possible value of (X_1, X_2, X_3) can be found in Tables 1 and 2.

Table 1 Values of $\frac{1}{2}X_{(1)} + \frac{1}{2}X_{(3)}$ associated to the possible values of the random vector (X_1, X_2, X_3)

X_1	(X_2, X_3)								
	(0, 0)	(1, 0)	(0, 1)	(1, 1)	(2, 0)	(0, 2)	(2, 1)	(1, 2)	(2, 2)
0	0	0.5	0.5	0.5	1	1	1	1	1
1	0.5	0.5	0.5	1	1	1	1.5	1.5	1.5
2	1	1	1	1.5	1	1	1.5	1.5	2

Table 2 Probabilities associated to the possible values of the random vector (X_1, X_2, X_3)

X_1	(X_2, X_3)								
	$(0, 0)$	$(1, 0)$	$(0, 1)$	$(1, 1)$	$(2, 0)$	$(0, 2)$	$(2, 1)$	$(1, 2)$	$(2, 2)$
0	0.001	0.008	0.008	0.064	0.001	0.001	0.008	0.008	0.001
1	0.008	0.064	0.064	0.518	0.008	0.008	0.064	0.064	0.008
2	0.001	0.008	0.008	0.064	0.001	0.001	0.008	0.008	0.001

Therefore, summing all the probabilities for each value of Y, it holds that $P(Y = 0) = P(Y = 2) = 0.001$, $P(Y = 0.5) = P(Y = 1.5) = 0.216$ and $P(Y = 1) = 0.566$. By computing the expectations of the convex functions $|1 - t|$ and $(1 - t)^2$, one has that

$$E[|1 - Y|] = 0.218 > 0.2 = E\left[|1 - X_1|\right],$$

and

$$E\left[(1 - Y)^2\right] = 0.11 < 0.2 = E\left[(1 - X_1)^2\right].$$

Therefore, neither $\frac{1}{2}X_{(1)} + \frac{1}{2}X_{(3)} \leq_{cx} X_1$ nor $\frac{1}{2}X_{(1)} + \frac{1}{2}X_{(3)} \geq_{cx} X_1$ are satisfied.

We end this section by remarking that, as special cases, the latter theorem holds for the important family of trimmed means, in which the weight vector fulfills $w_i = 0$ if $i \in \{1, \ldots, k, n - k + 1, \ldots, n\}$ and $w_i = \frac{1}{n-2k}$ otherwise (i.e. it is the mean of the $n - 2k$ central order statistics) for a $k \in \mathbb{N}$ such that $2k < n$. Moreover, it holds for the median.

Corollary 2 *Let $\vec{X}$ a random vector of dimension n with symmetric independent and identically distributed components. Then, any sample trimmed mean is smaller in convex order than the components of $\vec{X}$.*

Corollary 3 *Let $\vec{X}$ a random vector of dimension n with symmetric independent and identically distributed components. Then, the sample median is smaller in convex order than the components of $\vec{X}$.*

5 Reduction in variability for idempotent nullnorms

We devote this section to discuss the reduction in variability of nullnorms. As already mentioned in the introduction, these functions have an interesting mathematical structure that generalizes the maximum and minimum. In particular, the associativity permits to define a sequence of random variables that are decreasing with respect to the convex order.

5.1 The symmetric case and the convex order

Nullnorms have an annihilator element $a \in [0, 1]$ which, if it is included between the input values, equals the value of the nullnorm. Some results proved in Baz et al. (2024b) imply that, when considering a sequence $\{X_n, n \in \mathbb{N}\}$ of independent and identically distributed continuous random variables, the sequence $\{N_a(X_1, \ldots, X_n\}, n \in \mathbb{N}\}$ converges to a, where N_a is a idempotent nullnorm with annihilator element a. Therefore, nullnorms, at least asymptotically, reduce dispersion.

In the following, we prove a statement in this regard considering sequences of random variables with the same distribution but non-necessarily independent. Similarly to the case of OWA operators, it is necessary an additional symmetry condition in order to preserve the same expectation. We restrict our study to case of the support [0, 1], since this is the usual consideration for nullnorms, but the following results can be easily extended to more general positive random variables.

Theorem 5 *Let N_a be an idempotent nullnorm with annihilator element $a \in [0,1]$. Let $\{X_i, i \in \mathbb{N}\}$ be a sequence of random variables taking values in $[0,1] \subseteq \mathbb{R}$ such that:*

1. *all random variables have the same distribution,*
2. *all random variables are symmetric with respect to a (i.e. a is the median),*
3. *for any $n \in \mathbb{N}$, the copula C_n and survival copula $\bar{C}_n$ of $(X_1,\dots,X_n)$ hold that $C_n(t,\dots,t) = \bar{C}_n(t,\dots,t)$ for all $t \in [0,1]$.*

Then,

$$N_a(X_1,\dots,X_n) \leq_{cx} N_a(X_1,\dots,X_{n-1}) \leq_{cx} X_j, \ \forall n, j \in \mathbb{N}, \ j \leq n.$$

Proof Denote as F_n and $\bar{F}_n$ the multivariate distribution function and survival function of $(X_1,\dots,X_n)$. Since all the components of $\vec{X}$ have the same distribution function F, one can write $F_n(t_1,\dots,t_n) = C_n(F(t_1),\dots,F(t_n))$ and $\bar{F}_1(t_1,\dots,t_n) = \bar{C}(1-F(t_1),\dots,1-F(t_n))$. In addition, denote as $F_{a,n}$ the distribution function of $N_a(X_1,\dots,X_n)$.

If $N_a(X_1,\dots,X_n) < a$, then $X_1,\dots,X_n < a$ and $N_a(X_1,\dots,X_n) = \max(X_1,\dots,X_n)$. Then, $F_{a,n}(t) = F_n(t,\dots,t) = C_n(F(t),\dots,F(t))$ for $t < a$. Similarly, if $N_a(X_1,\dots,X_n) \geq a$, then $X_1,\dots,X_n \geq a$ and $N_a(X_1,\dots,X_n) = \min(X_1,\dots,X_n)$. Then, $F_{a,n}(t) = 1 - \bar{F}_n(t,\dots,t) = 1 - \bar{C}_n(1-F(t),\dots,1-F(t))$ for $t \geq a$.

Since $F(a) = 0.5$, we can write $F_{a,n}(t) = h_n(F(t))$ with $h_n : [0,1] \to [0,1]$ being

$$h_n(t) = \begin{cases} C_n(t,\dots,t) & \text{if } t < 0.5 \\ 1 - \bar{C}_n(1-t,\dots,1-t) & \text{if } t \geq 0.5 \end{cases}$$

Consider a second distortion h'_n defined as

$$h'_n(t) = \begin{cases} C_n(t,\dots,t) & \text{if } t < 0.5 \\ 0.5 & \text{if } t = 0.5 \\ 1 - \bar{C}_n(1-t,\dots,1-t) & \text{if } t > 0.5 \end{cases}$$

Since $C_n(t,\dots,t) \leq C_n(t,\dots,t,1) = C_{n-1}(t,\dots,t)$ for any $t \in [0,1]$, where C_{n-1} is the copula of $(X_1,\dots,X_{n-1})$, then one has that $h'_n(t) \leq h'_{n-1}(t)$ for any $t \in [0,1]$. In addition, by using the assumption that $C_n(t,\dots,t) = \bar{C}_n(t,\dots,t)$ for any $t \in [0,1]$ and $n \in \mathbb{N}$, h'_n is centrally symmetric. Moreover, h'_n and h_n are equal almost everywhere for any $n \in \mathbb{N}$. Then, applying Theorem 1, one has the inequality

$$\begin{aligned} \int_{-\infty}^{x} F_{a,n}(t)dt &= \int_{-\infty}^{x} h_n(F(t))dt = \int_{-\infty}^{x} h'_n(F(t))dt \\ &\leq \int_{-\infty}^{x} h'_{n-1}(F(t))dt = \int_{-\infty}^{x} h_{n-1}(F(t))dt = \int_{-\infty}^{x} F_{a,n-1}(t)dt \end{aligned}$$

for any $x \in \mathbb{R}$.

Because of symmetry, it is clear that $E[N(X_1,\dots,X_n)] = E[N(X_1,\dots,X_{n-1})] = a$. Then, applying Theorem 1, it is concluded that $N_a(X_1,\dots,X_n) \leq_{cx} N_a(X_1,\dots,X_{n-1})$.

In addition, the inequality $N_a(X_1,\dots,X_n) \leq_{cx} X_1$ follows by the associativity of the nullnorm and the fact that all the random variables in the sequence $\{X_n, n \in \mathbb{N}\}$ have the same distribution. □

The last result implies that the sequence of random variables defined as $\{N_a(X_1, \ldots, X_n), n \in \mathbb{N}\}$ is decreasing with respect to the convex order. The condition over the copula may seem to be strongly restrictive. However, there are many examples of family copulas fulfilling this property.

Example 4 For a copula C_n and corresponding survival copula $\bar{C}_n$, a sufficient condition to fulfill the property $C_n(t, \ldots, t) = \bar{C}_n(t, \ldots, t)$ for any $t \in [0, 1]$ is to satisfy $C_n(x_1, \ldots, x_n) = \bar{C}_n(x_1, \ldots, x_n)$ for any $x_1, \ldots, x_n \in [0, 1]$, which is equivalent to have the density copula (when exists) fulfilling $c_n(x_1, \ldots, x_n) = c_n(1 - x_1, \ldots, 1 - x_n)$ for any $x_1, \ldots, x_n \in [0, 1]$. In particular,

- For the independent copula $c_n(x_1, \ldots, x_n) = 1$, so it is straightforward to show that $c_n(x_1, \ldots, x_n) = c(1 - x_1, \ldots, 1 - x_n)$.
- The density of a Gaussian copula has the expression

$$c_n(x_1, \ldots, x_n) = \frac{1}{\sqrt{|R|}} exp \left(-\frac{1}{2} \begin{pmatrix} \Phi^{-1}(x_1) \\ \ldots \\ \Phi^{-1}(x_n) \end{pmatrix}^T (R^{-1} - I) \begin{pmatrix} \Phi^{-1}(x_1) \\ \ldots \\ \Phi^{-1}(x_n) \end{pmatrix} \right),$$

  where R is a positive definite matrix, I is the identity matrix and Φ^{-1} is the quantile function of a standard Gaussian distribution (see Durante and Sempi 2016). Since the standard Gaussian distribution is symmetric with respect to 0, one has that $\Phi^{-1}(1-t) = -\Phi^{-1}(t)$ and therefore $c(x_1, \ldots, x_n) = c(1 - x_1, \ldots, 1 - x_n)$.
- The density of a T-copula has the expression

$$c(x_1, \ldots, x_n) = \frac{f_{\nu,R}(t_\nu^{-1}(x_1), \ldots, t_\nu^{-1}(x_n))}{\prod_{i=1}^n f_\nu(t_\nu^{-1}(x_i))},$$

  where R is a positive definite matrix with constant diagonal equal to 1 (being a correlation matrix), t^{-1} is the quantile function of a standard t-student distribution, f_ν is the density function of a standard t-student distribution and $f_{\nu,R}$ is the distribution function of a multivariate t-student distribution with mean vector $\vec{0}$ and dispersion matrix R (see Durante and Sempi 2016). Using again the symmetry of the standard t-student distribution with respect to 0 and the symmetry of any multivariate t-student with respect to its mean vector, it can be verified that $c_n(x_1, \ldots, x_n) = c_n(1 - x_1, \ldots, 1 - x_n)$.

As mentioned in the previous example, the condition $C_n(t, \ldots, t) = \bar{C}_n(t, \ldots, t)$ for any $t \in [0, 1]$ is fulfilled when $C_n(t_1, \ldots, t_n) = \bar{C}_n(t_1, \ldots, t_n)$ for any $t_1, \ldots, t_n \in [0, 1]$, which is the property known as *radial symmetry* in Copula Theory (see page 32 of Durante and Sempi 2016). However, there are cases in which the first condition holds without the necessity of the second.

Example 5 Consider the copula C with density c given by

$$c(x, y) = \begin{cases} \frac{4}{3} & \text{if } x \leq \frac{3}{4} \text{ and } y \geq \frac{1}{4}, \\ 4 & \text{if } x \geq \frac{3}{4} \text{ and } y \leq \frac{1}{4}. \end{cases}$$

Then, the expression of the copula and the survival copula are

$$C(x, y) = \begin{cases} 0 & \text{if } x \leq \frac{3}{4} \text{ and } y \leq \frac{1}{4}, \\ 4\left(x - \frac{3}{4}\right) y & \text{if } x \geq \frac{3}{4} \text{ and } y \leq \frac{1}{4}, \\ \frac{4}{3}x\left(y - \frac{1}{4}\right) & \text{if } x \leq \frac{3}{4} \text{ and } y \geq \frac{1}{4}, \\ 4\left(x - \frac{3}{4}\right) y + \frac{4}{3}x\left(y - \frac{1}{4}\right) + & \text{if } x \geq \frac{3}{4} \text{ and } y \geq \frac{1}{4}, \end{cases}$$

and

$$\bar{C}(x,y) = \begin{cases} 0 & \text{if } x \le \frac{1}{4} \text{ and } y \le \frac{3}{4}, \\ \frac{4}{3}\left(x - \frac{1}{4}\right) y & \text{if } x \ge \frac{1}{4} \text{ and } y \le \frac{3}{4}, \\ 4x\left(y - \frac{3}{4}\right) & \text{if } x \le \frac{1}{4} \text{ and } y \ge \frac{3}{4}, \\ \frac{4}{3}\left(x - \frac{1}{4}\right) y + 4x\left(y - \frac{3}{4}\right) + & \text{if } x \ge \frac{1}{4} \text{ and } y \ge \frac{3}{4}. \end{cases}$$

It is clear that $C(x, y) \neq \bar{C}(x, y)$ for any $x, y \in [0, 1]$ since, for instance, $C(\frac{1}{3}, \frac{1}{2}) = \frac{1}{9}$ and $\bar{C}(\frac{1}{3}, \frac{1}{2}) = \frac{1}{18}$. On the other hand, on the diagonal they assume the same values, i.e., it holds

$$C(t,t) = \bar{C}(t,t) = \begin{cases} 0 & \text{if } t \le \frac{1}{4}, \\ \frac{4}{3}t\left(t - \frac{1}{4}\right) & \text{if } \frac{1}{4} < t \le \frac{3}{4}, \\ \frac{4}{3}\left(t - \frac{1}{4}\right)t + 4t\left(t - \frac{3}{4}\right) & \text{if } \frac{3}{4} < t. \end{cases}$$

5.2 Non-symmetric case and the increasing convex order

Note that the symmetric distributions for the X_i is a restrictive assumption in Theorem 4 for many applicative purposes, as well as the assumption that the annihilator element a must be the median. This subsection is devoted to the case in which the distribution is non-symmetric, showing that in this case the weaker increasing convex stochastic order can hold under suitable assumptions. As for Theorem 5, we have restricted the result to variables taking values in the unit interval, however, it can be extended trivially to positive random variables. The following preliminary result will be used in the sequel.

Lemma 3 *Let N_a be an idempotent nullnorm with annihilator element $a \in [0, 1]$. Let $\{X_i, i \in \mathbb{N}\}$ be a sequence of random variables with the same distribution taking values in $[0, 1]$. Then,*

$$N_a(X_1, \ldots, X_n) \le_{icx} N_a(X_1, \ldots, X_{n-1}) \iff E[N_a(X_1, \ldots, X_n)] \le E[N_a(X_1, \ldots, X_{n-1})]$$

Proof Note that the implication $\Rightarrow$ holds by the definition of the increasing convex order.

For the second one, recall that the survival function of $N_a(X_1, \ldots, X_n)$, which will be denoted as $\bar{F}_{a,n}$, is

$$\bar{F}_{a,n}(t) = \begin{cases} 1 - C_n(F(t), \ldots, F(t)) & \text{if } t < a, \\ \bar{C}_n(1 - F(t), \ldots, 1 - F(t)) & \text{if } t \ge a, \end{cases}$$

where F denotes the distribution function of X_1.

Notice that, by the increasing monotonicity of C and $\bar{C}$, $C_n(t, \ldots, t) \le C_n(t, \ldots, t, 1) = C_{n-1}(t, \ldots, t)$ and $\bar{C}_n(t, \ldots, t) \le \bar{C}_n(t, \ldots, t, 1) = \bar{C}_{n-1}(t, \ldots, t)$ for any $t \in [0, 1]$.

Then, $\bar{F}_{a,n}(t) \ge \bar{F}_{a,n-1}(t)$ for $t < a$ and $\bar{F}_{a,n}(t) \le \bar{F}_{a,n-1}(t)$ for $t \ge a$.

Therefore, for $x \ge a$ then one has

$$\int_x^1 \bar{F}_{a,n}(t)dt \le \int_x^1 \bar{F}_{a,n-1}(t)dt.$$

In the case $x \le a$, one must first observe that $E[N_a(X_1, \ldots, X_n)] \le E[N_a(X_1, \ldots, X_{n-1})]$ by assumption, which implies that $\int_0^1 \bar{F}_{a,n}(t)dt \le \int_0^1 \bar{F}_{a,n-1}(t)dt$. Since $\int_0^x \bar{F}_{a,n}(t)dt \ge \int_0^x \bar{F}_{a,n-1}(t)dt$, then

$$\int_x^1 \bar{F}_{a,n}(t)dt = \int_0^1 \bar{F}_{a,n}(t)dt - \int_0^x \bar{F}_{a,n}(t)dt$$

$$\leq \int_0^1 \bar{F}_{a,n-1}(t)dt - \int_0^x \bar{F}_{a,n-1}(t)dt = \int_x^1 \bar{F}_{a,n-1}(t)dt$$

Finally, the result holds as a consequence of Theorem 2. □

The previous property affirms that the comparison in the increasing convex order between the nullnorms with n and $n-1$ inputs is equivalent to comparison among the corresponding expectations. In general, the latter comparison is not satisfied for any possible annihilator element a, however there always exists an upper bound for a such the inequality is satisfied, as shown in the next statement.

Theorem 6 *Let N_a be the idempotent nullnorm with annihilator element $a \in [0, 1]$. Let X_1 and X_2 be two random variables defined over $[0, 1]$. Then, there always exists $a' \in [0, 1]$ such that $N_a(X_1, X_2) \leq_{icx} X_1$ for any $a \in [0, 1]$ satisfying $a \leq a'$.*

Proof Denote the distribution functions of X_1 and X_2 as F_1 and F_2 and the copula and survival copula of (X_1, X_2) as C and $\bar{C}$. Then the survival function of $N_a(X_1, X_2)$, denoted as $\bar{F}_a$, is defined as

$$\bar{F}_a(t) = \begin{cases} 1 - C(F_1(t), F_2(t)) & \text{if } t < a, \\ \bar{C}(\bar{F}_1(t), \bar{F}_2(t)) & \text{if } t \geq a. \end{cases} \tag{5.1}$$

Therefore, the expectation of $N_a(X_1, X_2)$ can the written as a function of a as

$$E[N_a(X_1, X_2)] = \int_0^a (1 - C(F_1(t), F_2(t)))\, dt + \int_a^1 \bar{C}(\bar{F}_1(t), \bar{F}_2(t))dt.$$

Trivially, $E[N_0(X_1, X_2)] = E[\min(X_1, X_2)]$ and $E[N_1(X_1, X_2)] = E[\max(X_1, X_2)]$. Moreover, such a function of a is continuous with derivative

$$\begin{aligned} \frac{d}{da}E[N_a(X_1, X_2)] &= \frac{d}{da}\int_0^a (1 - C(F_1(t), F_2(t)))\, dt + \frac{d}{da}\int_a^1 \bar{C}(1 - F_1(t), 1 - F_2(t))dt \\ &= 1 - C(F_1(a), F_2(a)) - \bar{C}(\bar{F}_1(a), \bar{F}_2(a)). \end{aligned}$$

Since $C(F_1(a), F_2(a)) + \bar{C}(\bar{F}_1(a), \bar{F}_2(a)) = P(X_1 \leq a, X_2 \leq a) + P(X_1 > a, X_2 > a) \leq 1$ for any $a \in [0, 1]$, the latter derivative is positive, thus $E[N_a(X_1, X_2)]$ is increasing as a function of a.

Moreover, using that $E[N_0(X_1, X_2)] = E[\min(X_1, X_2)] \leq E[X_1] \leq E[\max(X_1, X_2)] = E[N_1(X_1, X_2)]$ and $E[N_a(X_1, X_2)]$ is continuous as a function of a, there exists $a' \in [0, 1]$ such that $E[N_{a'}(X_1, X_2)] = E[X_1]$, and $E[N_a(X_1, X_2)] \leq E[X_1]$ for any $a \leq a'$. The result holds by Lemma 3. □

Notice that, by using the associative property, the latter result can be extended to any number of possible inputs as follows.

Corollary 4 *Let N_a be the idempotent nullnorm with annihilator element $a \in [0, 1]$. Let $X_1, \ldots, X_n$ be random variables defined over $[0, 1]$. Then, there always exists $a' \in [0, 1]$ such that $N_a(X_1, \ldots, X_n) \leq_{icx} N_a(X_1, \ldots, X_{n-1})$ for any $a \in [0, 1]$ satisfying $a \leq a'$.*

Proof Notice that $N_a(X_1, \ldots, X_{n-1}, X_n) = N_a(N_a(X_1, \ldots, X_{n-1}), X_n)$. Then, apply Theorem 6. □

One of the main limitations of the latter statement is that it does not give a value for a'. In the next result we show that, when the inputs are independent and have the same distribution, under some additional conditions the value of a' is always bigger than the median of the X_i, thus that $N_a(X_1, \ldots, X_n) \leq_{icx} N_a(X_1, \ldots, X_{n-1})$ holds for any a smaller than the median of the X_i.

Theorem 7 *Let N_a be the idempotent nullnorm with annihilator element $a \in [0, 1]$. Let $(X_1, \ldots, X_n)$ be a random vector with independent and identically distributed continuous components having support $[0, 1]$. If the survival function $\bar{F}$ of the components is convex on $[0, 1]$ then*

$$N_a(X_1, \ldots, X_n) \leq_{icx} X_j, \forall j \in \{1, \ldots, n\}, \ \forall a \in [0, 1] \text{ such that } a \leq m,$$

where m denotes the median of the X_j.

Proof As a consequence of Lemma 3, to prove the statement it suffices to prove that $E[N_m(X_1, \ldots, X_n)] \leq E[X_j]$. Notice that the density function f is decreasing since it is the (almost everywhere) derivative of $F(t) = 1 - \bar{F}(t)$, which is concave. Since independence between the random variables is assumed, Equation (5.1) reads as

$$\bar{F}_m(t) = \begin{cases} 1 - F(t)^n & \text{if } t < m, \\ \bar{F}(t)^n & \text{if } t \geq m, \end{cases}$$

and, therefore, the expectations are ordered if and only if

$$\begin{aligned} E[N_m(X_1, \ldots, X_n)] &= \int_0^m \left(1 - F(t)^n\right) dt + \int_m^1 \bar{F}(t)^n dt \\ &\leq \int_0^m (1 - F(t))\, dt + \int_m^1 \bar{F}(t) dt = E[X_1]. \end{aligned} \tag{5.2}$$

The latter inequality can be rewritten as

$$\int_0^m \left(1 - F(t)^n - 1 - F(t)\right) dt \leq \int_m^1 \left(\bar{F}(t) - \bar{F}(t)^n\right) dt,$$

i.e.,

$$\int_0^m F(t) \left(1 - F(t)^{n-1}\right) dt \leq \int_m^1 \bar{F}(t) \left(1 - \bar{F}(t)^{n-1}\right) dt.$$

Since f is decreasing, the median m is smaller than 0.5. Computing the derivative of $F(m-t)\left(1 - F(m-t)^{n-1}\right)$ and of $\bar{F}(m+t)\left(1 - \bar{F}(m+t)^{n-1}\right)$ for $t \in [0, m]$, one gets

$$\begin{aligned} \frac{d}{dt}\left[F(m-t)\left(1 - F(m-t)^{n-1}\right)\right] &= -f(m-t)\left(1 - nF(m-t)^{n-1}\right), \\ \frac{d}{dt}\left[\bar{F}(m+t)\left(1 - \bar{F}(m+t)^{n-1}\right)\right] &= -f(m+t)\left(1 - n\bar{F}(m+t)^{n-1}\right). \end{aligned}$$

Notice that, since f is decreasing, then $f(m-t) \geq f(m+t)$ for any $t \in [0, m]$. In addition,

$$\begin{aligned} F(m-t) &= P(X \leq m) - P(X \in (m-t, m]) = 0.5 - P(X \in (m-t, m]), \\ \bar{F}(m+t) &= P(X > m) - P(X \in (m, m+t]) = 0.5 - P(X \in (m, m+t]). \end{aligned}$$

Observing again that f is decreasing, it follows that $P(X \in (m-t, m]) \geq P(X \in (m, m+t])$, so that also $F(m-t) \leq \bar{F}(m+t)$ and $\left(1 - nF(m-t)^{n-1}\right) \geq \left(1 - n\bar{F}(m+t)^{n-1}\right)$ hold for any $t \in [0, m]$. Therefore, it is concluded that

$$\frac{d}{dt}\left[F(m-t)\left(1 - F(m-t)^{n-1}\right)\right] \leq \frac{d}{dt}\left[\bar{F}(m+t)\left(1 - \bar{F}(m+t)^{n-1}\right)\right]$$

for any $t \in [0, m]$. Since $F(m) = \bar{F}(m) = 0.5$ it follows $F(m-t)\left(1 - F(m-t)^{n-1}\right) \leq \bar{F}(m+t)\left(1 - \bar{F}(m+t)^{n-1}\right)$ for any $t \in [0, m]$. Thus, since $2m \leq 1$,

$$\int_0^m F(t)\left(1 - F(t)^{n-1}\right) dt \leq \int_m^{2m} \bar{F}(t)\left(1 - \bar{F}(t)^{n-1}\right) dt \leq \int_m^1 \bar{F}(t)\left(1 - \bar{F}(t)^{n-1}\right) dt.$$

Finally, as a consequence of $E[N_a(X_1, \ldots, X_n)]$ being increasing in a, by Lemma 3 follows $N_a(X_1, \ldots, X_n) \leq_{icx} N_m(X_1, \ldots, X_n) \leq_{icx} X_j, \forall j \in \{1, \ldots, n\}$. □

Remark 2 Latter result is also valid when there is a probability mass in 0, i.e. $P(X = 0) > 0$. This change does not compromise the convexity of the survival function and the proof remains the same, since the formula for the expected value given in Eq. (5.2) still holds.

6 The case of uninorms

The structure of uninorms is more complex than the structure of nullnorms, mainly because of the choice of the function g in Proposition 3. One can think of uninorms as functions that return the maximum when the variables take big values and the minimum when the values are small (with some intermediate cases). In this sense, one can expect a convergence to 1 (or 0), if all variables take big (or small) enough values. In the following, we denote as $S(X)$ the support of a random variable X, i.e. the smallest closed set such that $P(X \in S(X)) = 1$.

Proposition 4 *Let U be a uninorm with neutral element e and $\{X_n, n \in \mathbb{N}\}$ a sequence of independent and identically distributed random variables over* $[0, 1]$ *such that* $P(X_1 > e) = 1$. *Then,* $U(X_1, \ldots, X_n) \rightarrow_{a.s.} \sup S(X_1)$.

Proof Notice that, if $X_1, \ldots, X_n > e$, then $U(X_1, \ldots, X_n) \geq \max(X_1, \ldots, X_n)$. Therefore, it is clear that $U(X_1, \ldots, X_n) \rightarrow_{a.s.} \sup S(X_1)$ □

We recall that $\rightarrow_{a.s}$ means almost sure convergence, see Rohatgi and Saleh (2015). A similar result can be stated for random variables assuming small values.

Proposition 5 *Let U be a uninorm with neutral element e and $\{X_n, n \in \mathbb{N}\}$ a sequence of independent and identically distributed random variables such that $P(X_1 < e) = 1$. Then,* $U(X_1, \ldots, X_n) \rightarrow_{a.s.} \inf S(X_1)$.

In both cases, although we do not have an inequality in terms of the convex order, at least asymptotically there is a reduction of the variability, since the limit is a degenerate random variable.

However, this is not the case in all situations. If the distribution is evenly distributed below and above the function g, some part of the distribution will increase, while the other part will decrease, leading to an increase of the variability. Although a general result for this behavior cannot be proven, it is possible to consider particular uninorms to illustrate this property.

Proposition 6 *Let X_1 and X_2 be two independent, continuous, symmetric with respect to* 0.5 *and identically distributed over* $[0, 1]$ *random variables. Let U be the uninorm defined by considering $g(x) = 1 - x$ in Proposition 3. Then,*

$$U(X_1, X_2) \geq_{cx} X_1.$$

Proof Notice that, since the random variables are continuous and $\{(x_1, x_2) \in [0, 1]^2 \mid x_1 = 1 - x_2\}$ has Lebesgue measure 0, the behaviour of the uninorm over this set is negligible.

Let us compute the density function f_U of $U(X_1, X_2)$ by means of the distribution function of X_1, denoted by f. Using Proposition 3, we can characterise the preimage of U as

$$U(x_1, x_2) = y \iff \begin{cases} y = x_1 < x_2 < 1 - y \text{ or } y = x_2 < x_1 < 1 - y & \text{if } y < 0.5, \\ y = x_1 > x_2 > 1 - y \text{ or } y = x_2 > x_1 > 1 - y & \text{if } y > 0.5. \end{cases}$$

Therefore, one has:

$$f_U(y) = \begin{cases} 2\int_y^{1-y} f(y)f(t)dt = f(t)\,(F(1-y) - F(y)) & \text{if } y < 0.5, \\ 2\int_{1-y}^{y} f(y)f(t)dt = f(t)\,(F(y) - F(1-y)) & \text{if } y > 0.5. \end{cases}$$

Using the symmetry of the random variables, $F(1-y)-F(y) = 1-2F(y)$ and, therefore, $f_U(y) = f(y)|1 - 2F(y)|$. The associated distribution function is

$$F_U(y) = \begin{cases} \frac{1}{2} - \frac{1}{2}(1 - 2F(y))^2 & \text{if } y < 0.5, \\ \frac{1}{2} + \frac{1}{2}(1 - 2F(y))^2 & \text{if } y \geq 0.5. \end{cases}$$

Then, $F_U(y) = h(F(y))$ with h being the distortion given by

$$h(t) = \begin{cases} \frac{1}{2} - \frac{1}{2}(1 - 2t)^2 & \text{if } t < 0.5, \\ \frac{1}{2} + \frac{1}{2}(1 - 2t)^2 & \text{if } t \geq 0.5. \end{cases}$$

Notice that h is centrally symmetric and fulfills $h(t) \geq t$ for any $t \in [0, 0.5]$. Then, applying Theorem 1, it is concluded that $U(X_1, X_2) \geq_{cx} X_1$. □

If the neutral element is different from 0.5, a modification of the latter result can be proven.

Corollary 5 *Let X_1 and X_2 be two independent, continuous, symmetric with respect to $e \in$* $[0, 1]$ *and identically distributed over* $[0, 1]$ *random variables. Let U be the uninorm defined by considering* $g(x) = \min(\max(0, 2e - x), 1)$ *in Proposition 3. Then,*

$$U(X_1, X_2) \geq_{cx} X_1.$$

Proof Notice that $g(x) = \min(\max(0, 2e - x), 1)$ is essentially the median of 0, $2e - x$ and 1. Let U' the uninorm of Proposition 6. Then, it holds that

$$U(X_1, X_2) = U'(X_1 - (e - 0.5), X_2 - (e - 0.5)) + (e - 0.5).$$

Since X_1 and X_2 are symmetric with respect to e, $X_1 - (e - 0.5)$ and $X_2 - (e - 0.5)$ are symmetric with respect to 0.5 (and they still take values over $[0, 1]$). Therefore, Proposition 6 can be applied and it holds that $U'(X_1 - (e - 0.5), X_2 - (e - 0.5)) \geq_{cx} X_1 - (e - 0.5)$. Since the convex order is closed under translations (Shaked and Shanthikumar 2007),

$$\begin{aligned} U(X_1, X_2) &= U'(X_1 - (e - 0.5), X_2 - (e - 0.5)) + (e - 0.5) \\ &\geq_{cx} X_1 - (e - 0.5) + (e + 0.5) = X_1. \end{aligned}$$

□

Unfortunately, for many cases, one cannot say anything about the relation between the inputs and the output with respect to the convex order, as shown in the following counterexample.

Example 6 Let (X_1, X_2) be two independent standard uniform random variables and U the uninorm given by

$$U(x_1, x_2) = \begin{cases} \min(x_1, x_2) & \text{if } x_1, x_2 < 0.5, \\ \max(x_1, x_2) & \text{elsewhere.} \end{cases}$$

Then, the survival function of $U(X_1, X_2)$ fulfills $\bar{F}_U(t) = 0.25(1-2t)^2 + 0.75$ if $t < 0.5$ and $\bar{F}_U(t) = 1 - t^2$ if $t \geq 0.5$. Thus,

$$E[U(X_1, X_2)] = \int_0^{0.5} \left(0.25(1-2t)^2 + 0.75\right) dt + \int_{0.5}^{1} \left(1 - t^2\right) dt = \frac{5}{8}$$

Keeping in mind that the expectation of a standard uniform is $\frac{1}{2}$, then neither $U(X_1, X_2) \leq_{cx} X_1$ nor $U(X_1, X_2) \geq_{cx} X_1$ hold.

7 Conclusions

In this paper, some intuitive ideas regarding the reduction of variability of some aggregation and mean functions have been proved. In summary,

1. For a random vector with identically distributed marginals, weighted arithmetic means reduce the variability with respect to the convex order. If an adequate transformation is applied, the same holds for other quasi-arithmetic weighted means.
2. If symmetry and independence between the components are assumed, then the same property is true for OWA operators with bigger central weights.
3. If symmetry with respect to the annihilator element and some assumptions over the diagonal of the copula are considered, the property holds for nullnorms. If the symmetry is removed, some additional results can be proved by considering the increasing convex order.
4. For uninorms, it is possible to have both reduction or an increasing of the variability.

We recall that the here-presented results formalized intuitive ideas that are applied when using means functions in Data Analysis as a way to reduce the variability of the inputs. In the particular case of nullnorms and uninorms, the reduction of variability is crucial for its use in fuzzy neural networks when working with noise information inputs. In addition, they contribute to the topic of behavioral analysis of aggregation functions, see Chapter 10 in Grabisch et al. (2009).

The main future line of research is try to replicate this results with other variability stochastic orders such as the dispersive order, the dilation or the Lorenz orders (Shaked and Shanthikumar 2007). They do not impose the same expectations for the involved random variables, so its use could allow to extend this results for wider family of aggregation functions.

Funding Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature. 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 (PID2022-139886NB-I00). S. Montes and I. Díaz are also been supported by the project UE Proyect SCIMIN-CRM-2024-00087-001.

Data availability No data was used for the research described in the article.

Declarations

Conflict of interest The authors have no Conflict of interest to declare that are relevant to the content of this article.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Arriaza A, Sordo MA (2021) Preservation of some stochastic orders by distortion functions with application to coherent systems with exchangeable components. Appl Stoch Models Bus Ind 37(2):303–317

Baz J, García-Zamora D, Díaz I, Montes S, Martínez L (2024) Flexible-dimensional l-statistic for mean estimation of symmetric distributions. Stat Papers 65:4001–4024

Baz J, Montes S, Díaz I (2024) Actualization probabilities of idempotent nullnorms and uninorms. In: Intelligent Management of Data and Information in Decision Making: Proceedings of the 16th FLINS Conference on Computational Intelligence in Decision and Control & the 19th ISKE Conference on Intelligence Systems and Knowledge Engineering (FLINS-ISKE 2024), pp. 113–120. World Scientific

Bullen PS, Bullen P (2003) Quasi-arithmetic means. Handbook of means and their inequalities. Springer, Berlin, pp 266–320

Campos Souza PV, Lughofer E (2021) An evolving neuro-fuzzy system based on uni-nullneurons with advanced interpretability capabilities. Neurocomputing 451:231–251

Cauchy ALB (1821) Cours D'analyse de l'École royale polytechnique. Imprimerie Royale, Paris

Crow EL, Siddiqui MM (1967) Robust estimation of location. J Am Stat Assoc 62(318):353–389

Czogała E, Drewniak J (1984) Associative monotonic operations in fuzzy set theory. Fuzzy Sets Syst 12(3):249–269

David HA, Nagaraja HN (2004) Order Stat. Wiley, Hoboken

Durante F, Sempi C et al (2016) Principles of copula theory, vol 474. CRC Press, Boca Raton

Grabisch M, Marichal J-L, Mesiar R, Pap E (2009) Aggregation functions, vol 127. Cambridge University Press, Cambridge

Hosking J (1998) L-estimation. Handbook of statistics, vol 17. Elsevier, Amsterdam, pp 215–235

Hürlimann W (2004) Distortion risk measures and economic capital. N Am Actuar J 8(1):86–95

Leite D, Škrjanc I (2019) Ensemble of evolving optimal granular experts, OWA aggregation, and time series prediction. Inf Sci 504:95–112

Lemos A, Caminhas W, Gomide F (2010) New uninorm-based neuron model and fuzzy neural networks. In: 2010 Annual Meeting of the North American Fuzzy Information Processing Society, pp. 1–6. IEEE

Lloyd E (1952) Least-squares estimation of location and scale parameters using order statistics. Biometrika 39(1/2):88–95

Melnikov V, Hüllermeier E (2016) Learning to aggregate using uninorms. In: Machine Learning and Knowledge Discovery in Databases: European Conference, ECML PKDD 2016, Riva del Garda, Italy, September 19–23, 2016, Proceedings, Part II 16, pp. 756–771. Springer

Navarro J, Águila Y (2017) Stochastic comparisons of distorted distributions, coherent systems and mixtures with ordered components. Metrika 80:627–648

Nelsen RB (2007) An introduction to copulas. Springer, Berlin

Rohatgi VK, Saleh AME (2015) An introduction to probability and statistics. Wiley, Hoboken

Ruta D, Gabrys B (2007) Neural network ensembles for time series prediction. In: 2007 International Joint Conference on Neural Networks, pp. 1204–1209. IEEE
Scherger V, Terceño A, Vigier H (2017) The OWA distance operator and its application in business failure. Kybernetes 46(1):114–130
Shaked M, Shanthikumar JG (2007) Stochastic orders. Springer, Berlin
Shao J (1994) L-statistics in complex survey problems. Ann Stat 22:946–967
Sklar M (1959) Fonctions de repartition an dimensions et leurs marges. Publ l'Inst Stat l'Univ Paris 8:229–231
Yager RR (1993) Families of OWA operators. Fuzzy Sets Syst 59(2):125–148
Yager RR (2007) Centered OWA operators. Soft Comput 11(7):631–639
Zhou L-G, Chen H-Y (2011) Continuous generalized OWA operator and its application to decision making. Fuzzy Sets Syst 168(1):18–34

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.