POLITECNICO DI TORINO
Repository ISTITUZIONALE

Input Importance in Aggregation Theory by Means of Dependence Stochastic Orders

Original
Input Importance in Aggregation Theory by Means of Dependence Stochastic Orders / Baz, J., Pellerey, F.. - ELETTRONICO. - 15883 - 1:(2025), pp. 258-269. (EUSFLAT 2025 - 14th Conference of the European Society for Fuzzy Logic and Technology Riga (Lituania) Luglio 21-25, 2025) [10.1007/978-3-031-97225-6_21].

Availability:
This version is available at: 11583/3001841 since: 2025-08-04T17:54:46Z

Publisher:
Springer

Published
DOI:10.1007/978-3-031-97225-6_21

Terms of use:

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

Publisher copyright
Springer postprint/Author's Accepted Manuscript (book chapters)

This is a post-peer-review, pre-copyedit version of a book chapter published in Advances in Fuzzy Logic and Technology. The final authenticated version is available online at: http://dx.doi.org/10.1007/978-3-031-97225-6_21

(Article begins on next page)

09 August 2026

Input importance in aggregation theory by means of dependence stochastic orders

Juan Baz[1][0000−0003−1142−6077] and Franco Pellerey[2][000−0002−8983−855X]

[1] Department of Statistics, O.R. and D.M., University of Oviedo, Fededrico García Lorca 18, 33007 Oviedo, Spain
bazjuan@uniovi.es

[2] Dipartimento di Scienze Matematiche, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Torino, Italy
franco.pellerey@polito.it

Abstract. Aggregation functions are usually used to summarize the information from different inputs into a unique value. Depending on the structure of the aggregation function and the behavior of the initial data, there are inputs that have a larger impact in the result of the aggregation. From a probabilistic approach, such an importance can be identified as the positive dependence between the input and the output. In this paper, sufficient conditions for the stochastic ordering with respect to positive dependence stochastic orders between bivariate random vectors consisting of an input and the output of aggregation functions are provided. In particular, quasi-arithmetic means and some OWA operators are considered, using as ordering the supermodular and concordance stochastic orders.

Keywords: Aggregation · Stochastic orders · Copulas · Dependence

1 Introduction

Aggregation functions [1, 4] are widely used for fusing the information from different sources in different applications such as decision making [5, 14] and prediction [7, 10]. In some contexts, it is interesting to identify the inputs that have a larger importance in the aggregation process, since they have more impact in the output value.

The most simple example is the case of weighted means, in which the value of the weight is usually associated with the importance of the input. This can be shown by considering the associated partial derivatives, which equal

$$\frac{\partial \sum_{i=1}^{n} w_i x_i}{\partial x_k} = w_k,\ \forall k \in \{1, \dots, n\},$$

so that the variations of an input change more drastically the value of the aggregation if the weight is greater.

From a probabilistic approach, the importance of an input can be measured by its dependence with the output of the aggregation. For instance, if one considers independent random variables $X_1, \dots, X_n$ with the same distribution as the inputs of a weighted mean, the correlation coefficients between each of them and the aggregation can be easily computed. Denoting the variance of X_1 as σ^2,

$$\rho_{X_k,\sum_{i=1}^n w_iX_i} = \frac{Cov(X_k, \sum_{i=1}^n w_iX_i)}{\sqrt{Var(X_k)Var(\sum_{i=1}^n w_iX_i)}} = \frac{Cov(X_k, w_kX_k)}{\sqrt{\sigma^2 \sum_{i=1}^n w_i^2\sigma^2}} =$$
$$= \frac{w_k\sigma^2}{\sigma^2\sqrt{\sum_{i=1}^n w_i^2}} = \frac{w_k}{||\boldsymbol{w}||_2}.$$

Therefore, the linear dependence increases with the value of the weight, since

$$w_j \le w_k \Rightarrow \rho_{X_j,\sum_{i=1}^n w_iX_i} \le \rho_{X_k,\sum_{i=1}^n w_iX_i}. \tag{1}$$

However, the case of independent and identically distributed random variables is quite limited and there are some kinds of dependence that are not linear [8]. In this direction, this paper is devoted to extend the comparison in Equation (1) by considering:

1. Dependence stochastic orders that not only take into account linear dependence,
2. Random variables that are dependent or have different distributions,
3. Other aggregation functions rather than the weighted mean.

More specifically, sufficient conditions for having the ordering

$$(X_i, A(\boldsymbol{X})) \le_{dp} (X_j, A(\boldsymbol{X})),$$

where $\le_{dp}$ is a positive dependence stochastic order and A an aggregation function are provided.

The rest of the paper is organized as follows. In Section 2, the basic notions about stochastic orders and copulas are provided. Then, Section 3 and 4 are devoted, respectively, to the cases of quasi-arithmetic means and OWA operators. Finally, the conclusions are discussed in Section 5.

2 Stochastic orders and copulas

Stochastic orders are relations that compare probability distributions and random vectors by means of their location, variability and dependence. The most known stochastic order is the usual stochastic order, which, for random variables, is defined as the pointwise comparison of the distribution functions.

Definition 1. *[9] Let X and Y be two random variables with distribution functions F_X and F_Y. If $F_X(t) \ge F_Y(t)$ for any $t \in \mathbb{R}$, it is said that X is smaller than or equal to Y in the usual stochastic order and it is denoted by $X \le_{st} Y$.*

The usual stochastic order can also be characterized by the comparison of the expectation of increasing functions. If $X \le_{st} Y$, the quantiles and the expectations (if they exist) are ordered.

Moving to variability stochastic orders, the dispersive order, which considers differences between quantiles, is usually considered.

Definition 2. *[9] Let X and Y be two random variables with quantile functions F^{-1} and G^{-1}. Then, X is said to be smaller than or equal to Y in the dispersive order, denoted as $X \leq_{disp} Y$, if*

$$F^{-1}(p) - F^{-1}(q) \leq G^{-1}(p) - G^{-1}(q) \text{ for any } 0 < q \leq p < 1.$$

In addition, variability can be also studied by using convex functions.

Definition 3. *[9] Let X and Y be two random variables. If $E[\varphi(\boldsymbol{X})] \leq E[\varphi(\boldsymbol{Y})]$ for any measurable convex function $\varphi : \mathbb{R}^n \to \mathbb{R}$ such that $E[\varphi(\boldsymbol{X})]$ and $E[\varphi(\boldsymbol{Y})]$ exist, then X is said to be smaller than or equal to Y in the convex stochastic order and it is denoted as $X \leq_{cx} Y$.*

Both the dispersive and the convex order imply the ordering of variances $Var(X) \leq Var(Y)$ and the convex order imposes the same expectation for X and Y. Moreover, if $E[X] = E[Y]$, then $X \leq_{disp} Y$ implies $X \leq_{cx} Y$. Regarding (positive) dependence stochastic orders, one of the most considered alternatives is the supermodular stochastic order. Recall that a function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be supermodular if $f(\boldsymbol{x})+f(\boldsymbol{y}) = f(\boldsymbol{x}\wedge\boldsymbol{y})+f(\boldsymbol{x}\vee\boldsymbol{y})$ for any $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$, where $\wedge$ and $\vee$ denote, respectively, the componentwise minimum and maximum.

Definition 4. *[9] Let $\boldsymbol{X}$ and $\boldsymbol{Y}$ be two random vectors. If $E[\varphi(\boldsymbol{X})] \leq E[\varphi(\boldsymbol{Y})]$ for any supermodular function $\varphi : \mathbb{R}^n \to \mathbb{R}$ such that $E[\varphi(\boldsymbol{X})]$ and $E[\varphi(\boldsymbol{Y})]$ exist, then $\boldsymbol{X}$ is said to be smaller than or equal to $\boldsymbol{Y}$ in the supermodular order, and it is denoted as $\boldsymbol{X} \leq_{sm} \boldsymbol{Y}$.*

If $\boldsymbol{X} \leq_{sm} \boldsymbol{Y}$, then both vectors have the same marginal distributions. This stochastic order implies the weaker PQD order.

Definition 5. *[9] Let $\boldsymbol{X}$ and $\boldsymbol{Y}$ be two random vectors that have the same marginal distributions. Then, $\boldsymbol{X}$ is said to be smaller than or equal to $\boldsymbol{Y}$ in the positive quadrant dependent order, and it is denoted as $\boldsymbol{X} \leq_{PQD} \boldsymbol{Y}$ if $P(\boldsymbol{X} \leq \boldsymbol{x}) \leq P(\boldsymbol{Y} \leq \boldsymbol{x})$ and $P(\boldsymbol{X} \geq \boldsymbol{x}) \leq P(\boldsymbol{Y} \geq \boldsymbol{x})$ for any $\boldsymbol{x} \in \mathbb{R}^n$.*

Moreover, the PQD order implies inequalities for the main correlation measures such as the Pearson's correlation coefficient, Kendall's τ, Spearman's ρ and the Blomquist's q [13].

When the marginal distributions are different, one can use copulas to compare the dependence. These functions capture the dependence structure of the random vector and coincide with the distribution functions of random vectors with standard uniform marginals. The multivariate distribution function, defined as $F(x_1, \ldots, x_n) = P(X_1 \leq x_1, \ldots, X_n \leq x_n)$ of any random vector can be expressed as the composition of a copula and the marginal distribution functions by the so-called Sklar Theorem.

Theorem 1. *[3, 11] Let $\boldsymbol{X}$ be a random vector with marginal distribution functions $F_1, \ldots, F_n$ and joint distribution function F. Then, there exists a copula $C : [0,1]^n \to [0,1]$ such that*

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

In addition, if $F_1, \ldots, F_n$ are continuous, the copula is unique.

Similarly, the survival function of $\boldsymbol{X}$, defined as $\hat{F}(x_1, \ldots, x_n) = P(X_1 > x_1, \ldots, X_n > x_n)$, can be decomposed as

$$\hat{F}(x_1, \ldots, x_n) = \hat{C}\left(\hat{F}_1(x_1), \ldots, \hat{F}_n(x_n)\right), \ \forall \boldsymbol{x} \in \mathbb{R}^n,$$

where $\hat{F}_1, \ldots, \hat{F}_n$ are the survival functions of the marginals and $\hat{C}$ is a copula, called the survival copula of $\boldsymbol{X}$.

Larger values of the copula and the survival copula are associated with a stronger positive dependence between the components of the random vector. In this direction, the concordance order is defined as follows.

Definition 6. *[6] Let $\boldsymbol{X}$ and $\boldsymbol{Y}$ be two random vectors with continuous marginal distributions with copulas C_X and C_Y and survival copulas $\hat{C}_X$ and $\hat{C}_Y$. If $C_X(\boldsymbol{x}) \leq C_Y(\boldsymbol{x})$ and $\hat{C}_X(\boldsymbol{x}) \leq \hat{C}_Y(\boldsymbol{x})$ for any $\boldsymbol{x} \in [0,1]^n$, it is said that $\boldsymbol{X}$ is smaller than or equal to $\boldsymbol{Y}$ in the concordance order and it is denoted as $\boldsymbol{X} \leq_c \boldsymbol{Y}$.*

If the marginal distributions are the same, then the concordance order is equivalent to the PQD order, thus implied by the supermodular order. Moreover, for bivariate random vectors, the concordance order and the supermodular are equivalent when the vectors have the same marginal distributions. The following useful result is Theorem 2.4.3. in [3].

Theorem 2. *[3] Let $\boldsymbol{X} = (X_1, \ldots, X_n)$ be a random vector with continuous components and copula C and consider n continuous and strictly monotone mappings $\varphi_j : RanX_j \to \mathbb{R}$ with $j \in \{1, \ldots, n\}$. Then, the copula of the random vector $\boldsymbol{Y} = (\varphi_1 X_1, \ldots, \varphi_n X_n)$ is given by $C \circ \sigma_{i_1} \circ \cdots \circ \sigma_{i_k}$, where $\sigma_{i_j} : [0,1]^n \to [0,1]^n$ is defined as $\sigma_j(x_1, \ldots, x_j, \ldots, x_n) = (x_1, \ldots, 1-x_j, \ldots, x_n)$ and the indices $i_1, \ldots, i_k$ correspond to those mappings $\varphi_{i_1}, \ldots, \varphi_{i_k}$ that are strictly decreasing.*

That is, applying strictly increasing functions to the variables remains the copula of the random vector invariant, while applying a strictly decreasing function to a variable transforms the copula by reflecting the associated component. In particular, one has that $C_{-X,-Y}(x,y) = \hat{C}(x,y)$.

3 Input importance for quasi-arithmetic means

Quasi-arithmetic means transform the inputs of the aggregation using a strictly increasing monotone transformation, perform a convex linear combination of the values and then inverse the transformation. They include, as particular cases, weighted arithmetic means, weighted geometric means and weighted harmonic means. The general definition is the following.

Definition 7. *[2] A function $f : I^n \to I$ is said to be a quasi-arithmetic weighted mean if can expressed as*

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

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

As already stated in the introduction, the value of the weight is usually interpreted as the importance of the associated input. However, the distribution of the input is also relevant. In the following, we will consider exchangeable random vectors, which are random vectors whose distribution is invariant when the order of their components is changed. In particular, all the marginal distributions of the input random vector have the same distribution. Therefore, it is reasonable to think that the following implication is true

$$w_i \leq w_j \Rightarrow \left(X_i, g^{-1}\left(\sum_{k=1}^n w_i g(x_i) \right) \right) \leq_{sm} \left(X_j, g^{-1}\left(\sum_{k=1}^n w_i g(x_i) \right) \right).$$

Let us start with a lemma that considers the bivariate case and a larger family of functions.

Lemma 1. *Let $f : \mathbb{R}^2 \to \mathbb{R}$ be a measurable function such that $f(x,y) \geq f(y,x)$ if $x \leq y$. Then, for any exchangeable bivariate random vector $\boldsymbol{X} = (X_1, X_2)$, it holds that*

$$(X_1, f(X_1, X_2)) \leq_{sm} (X_2, f(X_1, X_2)).$$

Proof. Let ϕ be a supermodular function. If $x_1 \leq x_2$, then $f(x_1,x_2) \geq f(x_2,x_1)$ and, therefore,

$$\phi(x_2, f(x_2,x_1)) + \phi(x_1, f(x_1,x_2)) \leq \phi(x_1, f(x_2,x_1)) + \phi(x_2, f(x_1,x_2)).$$

Similarly, if $x_1 \geq x_2$ then $f(x_1,x_2) \leq f(x_2,x_1)$ and latter equation also holds. Then, since $\boldsymbol{X}$ is exchangeable one has that $(X_1, f(X_2,X_1)) =_{st} (X_2, f(X_1,X_2))$ and $(X_1, f(X_1,X_2)) =_{st} (X_2, f(X_2,X_1))$. Therefore,

$$E\left[\phi(X_1, f(X_1,X_2))\right] = \frac{1}{2}E\left[\phi(X_1, f(X_1,X_2))\right] + \frac{1}{2}E\left[\phi(X_2, f(X_2,X_1))\right] \leq$$
$$\leq \frac{1}{2}E\left[\phi(X_1, f(X_2,X_1))\right] + \frac{1}{2}E\left[\phi(X_2, f(X_1,X_2))\right] = E\left[\phi(X_2, f(X_1,X_2))\right]$$

□

The inequality can be extended to random vectors of arbitrary dimensions by noting that exchangeability is preserved when considering some particular conditional distributions.

Theorem 3. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a measurable function such that $f(\boldsymbol{x}) \geq f(\pi_{i,j}\boldsymbol{x})$ if $x_i \leq x_j$ where $\pi_{i,j}$ denotes the permutation that changes the values of the i-th and j-th component. Then, for any exchangeable bivariate random vector $\boldsymbol{X}$, it holds that*

$$(X_i, f(\boldsymbol{X})) \leq_{sm} (X_j, f(\boldsymbol{X})).$$

Proof. Without lose of generality, consider $i = 1$ and $j = 2$. Define the random vector $\boldsymbol{Z} = (X_3, \ldots, X_n)$. For any possible value of $\boldsymbol{Z}$, $\boldsymbol{z}$, the exchangeability of $\boldsymbol{X}$ implies that the random vector

$$([X_1 \mid \boldsymbol{Z} = \boldsymbol{z}], [X_2 \mid \boldsymbol{Z} = \boldsymbol{z}]),$$

is exchangeable. Then, the result is a consequence of Lemma 1. □

The result for quasi-arithmetic means can be obtained by adapting the latter result.

Corollary 1. *Let $\boldsymbol{X}$ be an exchangeable random vector. Then, for any quasi-arithmetic mean*

$$f(\boldsymbol{x}) = g^{-1}\left(\sum_{k=1}^{n} w_k g(x_k)\right), \qquad \forall \boldsymbol{x} \in I^n,$$

it holds that

$$w_i \leq w_j \Rightarrow (X_i, f(\boldsymbol{X})) \leq_{sm} (X_j, f(\boldsymbol{X})).$$

Proof. If $x_i \leq x_j$, since g is increasing then $g(x_i) \leq g(x_j)$. Therefore, using that $w_i \leq w_j$,

$$\sum_{k=1}^{n} w_k g(x_k) \geq \left(\sum_{k=1, k\neq i,j}^{n} w_k g(x_k)\right) + w_i g(x_j) + w_j g(x_i).$$

Applying that g^{-1} is increasing, then $f(\boldsymbol{x}) \geq f(\pi_{i,j}\boldsymbol{x})$ where $\pi_{i,j}$ denotes the permutation that changes the values of the i-th and j-th component. Then, the conditions of Theorem 3 are fulfilled and the result holds. □

In words, Corollary 1 confirms the intuition that if $w_i < w_j$ then $f(\boldsymbol{X})$ is more dependent to the value assumed by X_j than the value assumed by X_i.

4 Input importance for OWA operators

Ordered Weighted Averagings (OWA) operators performs a convex linear combination of inputs of the aggregation after ordering them from the smallest to the greatest.

Definition 8. *[12] A function $f : I^n \to I$ is said to be an* Ordered Weighted Averaging *(OWA) operator if it can be expressed as*

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

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

OWA operators include the maximum and the minimum. In general, it is reasonable to think that larger inputs have more impact in the output of the maximum and smaller inputs have more impact in the output of the minimum. Therefore, it is intuitive to conjecture

$$X_i \leq_{st} X_j \Rightarrow (X_i, \max(\boldsymbol{X})) \leq_c (X_j, \max(\boldsymbol{X})), \, (X_i, \min(\boldsymbol{X})) \geq_c (X_j, \min(\boldsymbol{X})).$$

Notice that the supermodular order cannot be achieved, since the marginal distributions are not necessarily the same. In the following result, the concordance order is proved for the maximum by directly computing the associated copulas.

Theorem 4. *Let $\boldsymbol{X}$ be a random vector with continuous independent components. Then*

$$X_i \leq_{st} X_j \Rightarrow (X_1, \max(\boldsymbol{X})) \leq_c (X_2, \max(\boldsymbol{X})), \quad \forall\, i, j \in \{1, \dots, n\}.$$

Proof. Let us compute the copula of X_1 and $\max(\boldsymbol{X})$. The bivariate distribution function can be computed by considering conditional probabilities.

$$F_{X_1, \max(\boldsymbol{X})}(a, b) = P(X_1 \leq a) P(\max(\boldsymbol{X}) \leq b \mid X_1 \leq a).$$

If $a \leq b$, then $P(\max(\boldsymbol{X}) \leq b \mid X_1 \leq a) = P(\max(X_2, \dots, X_n) \leq b) = \prod_{i=1}^n F_i(b)$. If $a \geq b$, then $P(\max(\boldsymbol{X}) \leq b \mid X_1 \leq a) = P(\max(X_2, \dots, X_n) \leq b) P(X_1 \leq b \mid X_1 \leq a) = \left(\prod_{i=1}^n F_i(b)\right) \frac{F_1(b)}{F_1(a)}$. In addition, notice that $F_{\max(\boldsymbol{X})}(x) = \prod_{i=1}^n F_i(x)$, thus $\prod_{i=2}^n F_i(x) = \frac{F_{\max(\boldsymbol{X})}(x)}{F_1(x)}$. Therefore, the expression of the joint distribution function is

$$F_{X_1, \max(\boldsymbol{X})}(a, b) = \begin{cases} \frac{F_1(a)}{F_1(b)} F_{\max(\boldsymbol{X})}(b) & \text{if } a \leq b, \\ F_{\max(\boldsymbol{X})}(b) & \text{if } a \geq b, \end{cases}$$

which can be summarized as $F_{X_1, \max(\boldsymbol{X})}(a, b) = \min\left(1, \frac{F_1(a)}{F_1(b)}\right) F_{\max(\boldsymbol{X})}(b)$. Since the components of $\boldsymbol{X}$ are continuous and independent, $\max(\boldsymbol{X})$ is also continuous. Therefore, the copula has the following expression.

$$C_{X_1, \max(\boldsymbol{X})}(x, y) = \min\left(1, \frac{x}{F_1\left(F^{-1}_{\max(\boldsymbol{X})}(y)\right)}\right) y,$$

since

$$\begin{aligned} F_{X_1, \max(\boldsymbol{X})}(a, b) =& C_{X_1, \max(\boldsymbol{X})}(F_1(a), F_{\max(\boldsymbol{X})}(b)), \\ =& \min\left(1, \frac{F_1(a)}{F_1\left(F^{-1}_{\max(\boldsymbol{X})}\left(F_{\max(\boldsymbol{X})}(b)\right)\right)}\right) F_{\max(\boldsymbol{X})}(b), \\ =& \min\left(1, \frac{F_1(a)}{F_1(b)}\right) F_{\max(\boldsymbol{X})}(b). \end{aligned}$$

Similarly, the copulas of $(X_i, \max(\boldsymbol{X}))$ and $(X_j, \max(\boldsymbol{X}))$ with $i, j \in \{1, \ldots, n\}$ are of the form

$$C_{X_i,\max(\boldsymbol{X})}(x, y) = \min\left(1, \frac{x}{F_i\left(F^{-1}_{\max(\boldsymbol{X})}(y)\right)}\right) y,$$

$$C_{X_j,\max(\boldsymbol{X})}(x, y) = \min\left(1, \frac{x}{F_j\left(F^{-1}_{\max(\boldsymbol{X})}(y)\right)}\right) y.$$

If $X_i \leq_{st} X_j$, then $F_i(x) \geq F_j(x)$ for any $x \in \mathbb{R}$ and, therefore, it holds that $C_{X_i,\max(\boldsymbol{X})}(x, y) \leq C_{X_j,\max(\boldsymbol{X})}(x, y)$ for any $x, y \in [0, 1]$. □

The case of the minimum can be achieved by changing the sign of the variables.

Corollary 2. *Let* $\boldsymbol{X}$ *be a random vector with continuous independent components. Then*

$$X_i \leq_{st} X_j \Rightarrow (X_i, \min(\boldsymbol{X})) \geq_c (X_j, \min(\boldsymbol{X})), \quad \forall\, i, j \in \{1, \ldots, n\}.$$

Proof. Consider the random vector $-\boldsymbol{X}$ and apply Theorem 4. If $X_i \leq_{st} X_j$, then $-X_i \geq_{st} -X_j$ and

$$(-X_i, \max(-\boldsymbol{X})) \geq_c (-X_j, \max(-\boldsymbol{X})), (-X_i, -\min(\boldsymbol{X})) \geq_c (-X_j, -\min(\boldsymbol{X})).$$

Notice that, as a consequence of Theorem 2,

$$C_{-X_i,-\min(\boldsymbol{X})}(x, y) = \hat{C}_{X_i,\min(\boldsymbol{X})}(x, y),$$
$$C_{-X_i,-\min(\boldsymbol{X})}(x, y) = \hat{C}_{X_i,\min(\boldsymbol{X})}(x, y).$$

The result holds by applying that $(-X_i, -\min(\boldsymbol{X})) \geq_c (-X_j, -\min(\boldsymbol{X}))$ implies that $\hat{C}_{X_i,\min(\boldsymbol{X})}(x, y) \leq \hat{C}_{X_i,\min(\boldsymbol{X})}(x, y)$ for any $x, y \in [0, 1]^2$. □

A more general result without assuming independence cannot be proved, since it is easy to construct random vectors $\boldsymbol{X} = (X_1, X_2)$ such that $X_1 =_{st} X_2$ but with $(X_1, \max(X_1, X_2))$ and $(X_2, \max(X_1, X_2))$ having different copula. In addition, other inequalities for OWA operators are hard to prove. For instance, if one has a random vector $\boldsymbol{X} = (X_1, X_2, X_3)$ with independent components such that $X_1 \leq_{cx} X_2 \leq_{cx} X_3$, one could expect X_1 having more impact when computing the median, since it has the same location but smaller variability than X_2 and X_3, thus is in the middle position with larger probability. In symbols

$$X_1 \leq_{cx} X_2 \leq_{cx} X_3 \Rightarrow (X_1, X_{(2)}) \geq_c (X_2, X_{(2)}) \geq_c (X_1, X_{(3)}).$$

However, next example shows that the intuition is false in this case.

Example 1. Consider three independent random variables X_1, X_2 and X_3 such that X_1 has uniform distribution over $[-1, 1]$ and X_2 and X_3 have uniform distribution over $[-2, 2]$. It is clear that $E[X_1] = E[X_2] = E[X_3] = E[X_{(2)}] = 0$ and $Var[X_1] = \frac{1}{3}$, $Var[X_2] = Var[X_3] = \frac{4}{3}$. Moreover, $X_1 \leq_{cx} X_2, X_3$ and $X_1 \leq_{disp} X_2, X_3$.

Let us compute $\rho_{X_1,X_{(2)}}$ and $\rho_{X_2,X_{(2)}}$. Compute the covariance between X_1 and $X_{(2)}$ as follows,

$$Cov(X_1, X_{(2)}) = \int_{-1}^{1}\int_{-2}^{2}\int_{-2}^{2} \frac{1}{32} x_1 x_{(2)} dx_3 dx_2 dx_1,$$

where $x_{(2)}$ denotes the median value of x_1, x_2 and x_3.

Define $C = [-1, 1] \times [-2, 2]^2$ and consider the following subsets:

$$\begin{aligned}
C_1 &= \{\boldsymbol{x} \in C \mid x_2 \leq x_1 \leq x_3\}, \quad C_2 = \{\boldsymbol{x} \in C \mid x_3 \leq x_1 \leq x_2\},\\
C_3 &= \{\boldsymbol{x} \in C \mid x_1 \leq x_2 \leq x_3\}, \quad C_4 = \{\boldsymbol{x} \in C \mid x_1 \leq x_3 \leq x_2\},\\
C_5 &= \{\boldsymbol{x} \in C \mid x_2 \leq x_3 \leq x_1\}, \quad C_6 = \{\boldsymbol{x} \in C \mid x_3 \leq x_2 \leq x_1\}.
\end{aligned}$$

Notice that the intersection of any pair of the latter subsets has Lebesgue measure 0, thus the integral over it is negligible. In addition, since the permutation of X_2 and X_3 and all the random variables are symmetric with respect to 0, the integrals over C_1 and C_2 are the same, as well as the integrals over C_3, C_4, C_5 and C_6. Thus,

$$\begin{aligned}
Cov(X_1, X_{(2)}) &= \int_{-1}^{1}\int_{-2}^{2}\int_{-2}^{2} \frac{1}{32} x_1 x_{(2)} dx_3 dx_2 dx_1 = \\
= \frac{2}{32}\int_{-1}^{1}\int_{-2}^{x_1}\int_{x_1}^{2} x_1^2 dx_3 dx_2 dx_1 &+ \frac{4}{32}\int_{-1}^{1}\int_{-2}^{x_1}\int_{-2}^{x_2} x_1 x_2 dx_3 dx_2 dx_1.
\end{aligned}$$

Let us compute each of the integrals in the latter expression.

$$\begin{aligned}
\int_{-1}^{1}\int_{-2}^{x_1}\int_{x_1}^{2} x_1^2 dx_3 dx_2 dx_1 &= \int_{-1}^{1}\int_{-2}^{x_1} \left(2x_1^2 - x_1^3\right) dx_2 dx_1 = \\
= \int_{-1}^{1} \left(2x_1^3 - x_1^4 + 4x_1^2 - 2x_1^3\right) dx_1 &= -\frac{2}{5} + \frac{8}{3} = \frac{34}{15}.
\end{aligned}$$

$$\begin{aligned}
\int_{-1}^{1}\int_{-2}^{x_1}\int_{-2}^{x_2} x_1 x_2 dx_3 dx_2 dx_1 &= \int_{-1}^{1}\int_{-2}^{x_1} \left(x_1 x_2^2 + 2x_1 x_2\right) dx_1 dx_2 = \\
= \int_{-1}^{1} \left(\frac{x_1^4}{3} + x_1^3 + \frac{8}{3}x_1 - 4x_1\right) dx_1 &= \frac{2}{15}.
\end{aligned}$$

Then, the covariance is

$$Cov(X_1, X_{(2)}) = \frac{2}{32}\frac{34}{15} + \frac{4}{32}\frac{2}{15} = \frac{19}{120}.$$

For the covariance of X_2 and $X_{(2)}$, similarly as in the latter case, the integral can be decomposed in three different cases, the first associated to C_1 and C_2, the second to C_3 and C_6 and the last one to C_4 and C_5. Then, it is obtained

$$Cov(X_2, X_{(2)}) = \frac{2}{32}\int_{-1}^{1}\int_{-2}^{x_1}\int_{x_1}^{2} x_2 x_1 dx_3 dx_2 dx_1 +$$
$$+\frac{2}{32}\int_{-1}^{1}\int_{-2}^{x_1}\int_{-2}^{x_2} x_2^2 dx_3 dx_2 dx_1 + \frac{2}{32}\int_{-1}^{1}\int_{-2}^{x_1}\int_{-2}^{x_3} x_2 x_3 dx_2 dx_3 dx_1.$$

$$\int_{-1}^{1}\int_{-2}^{x_1}\int_{x_1}^{2} x_2 x_1 dx_3 dx_2 dx_1 = \int_{-1}^{1}\int_{-2}^{x_1} \left(2x_1x_2 - x_1^2 x_2\right) dx_2 dx_1 =$$
$$= \int_{-1}^{1} \left(x_1^3 - 4x_1 - \frac{x_1^4}{2} + 2x_1^2\right) dx_1 = -\frac{2}{10} + \frac{4}{3} = \frac{17}{15}.$$

$$\int_{-1}^{1}\int_{-2}^{x_1}\int_{-2}^{x_2} x_2^2 dx_3 dx_2 dx_1 = \int_{-1}^{1}\int_{-2}^{x_1} \left(x_2^3 + 2x_2^2\right) dx_2 dx_1 =$$
$$= \int_{-1}^{1} \left(\frac{x_1^4}{4} - 4 + \frac{2}{3}x_1^3 + \frac{16}{3}\right) dx_1 = \frac{2}{20} - 8 + \frac{32}{3} = \frac{83}{30}.$$

$$\int_{-1}^{1}\int_{-2}^{x_1}\int_{-2}^{x_3} x_2 x_3 dx_2 dx_3 dx_1 = \int_{-1}^{1}\int_{-2}^{x_1} \left(\frac{x_3^3}{2} - 2x_3\right) dx_3 dx_1 =$$
$$= \int_{-1}^{1} \left(\frac{x_1^4}{8} - 2 - x_1^2 + 4\right) dx_1 = \frac{2}{40} - 4 - \frac{2}{3} + 8 = \frac{203}{60}.$$

Thus,

$$Cov(X_2, X_{(2)}) = \frac{2}{32}\frac{17}{15} + \frac{2}{32}\frac{83}{30} + \frac{2}{32}\frac{203}{60} = \frac{437}{960}.$$

Finally, it is necessary to compute the variance of the median, which involves previously computed integrals.

$$Var(X_{(2)}) = \frac{2}{32}\int_{-1}^{1}\int_{-2}^{x_1}\int_{x_1}^{2} x_1^2 dx_3 dx_2 dx_1 + \frac{4}{32}\int_{-1}^{1}\int_{-2}^{x_1}\int_{-2}^{x_2} x_2^2 dx_3 dx_2 dx_1 =$$
$$= \frac{2}{32}\frac{34}{15} + \frac{4}{32}\frac{83}{30} = \frac{39}{80}.$$

Then, the correlation coefficients can be computed as follows.

$$\rho_{X_1,X_{(2)}} = \frac{\frac{19}{120}}{\sqrt{\frac{1}{3}\frac{39}{80}}} = \frac{304\sqrt{65}}{6240} \approx 0.3928,$$
$$\rho_{X_2,X_{(2)}} = \frac{\frac{437}{960}}{\sqrt{\frac{4}{3}\frac{39}{80}}} = \frac{437\sqrt{65}}{6240} \approx 0.5646.$$

Suppose that $(X_1, X_{(2)}) \geq_c (X_2, X_{(2)})$. Then, applying Theorem 2, $(X_1, X_{(2)})$ has the same copula as $(2X_1, X_{(2)})$ and, therefore, $(2X_1, X_{(2)}) \geq_c (X_2, X_{(2)})$. Since $2X_1 =_{st} X_2$, $(2X_1, X_{(2)})$ and $(X_2, X_{(2)})$ have the same marginal distributions and then $(2X_1, X_{(2)}) \geq_{sm} (X_2, X_{(2)})$. This cannot be possible, since $(2X_1, X_{(2)}) \geq_{sm} (X_2, X_{(2)})$ would imply $\rho_{X_1,X_{(2)}} \geq \rho_{X_1,X_{(2)}}$, which is not true. Therefore, it is concluded that $(X_1, X_{(2)}) \not\geq_c (X_2, X_{(2)})$.

5 Conclusions

In this paper, the importance of the inputs in aggregation functions has been studied by using positive dependence stochastic orders. In particular, it has been proved that a larger weight in quasi-arithmetic means implies a larger positive dependence between the associated variable and the aggregation, with respect to the supermodular order, when dealing with exchangeable random vectors. For random vectors with independent components, the usual stochastic order implies the concordance order between the vectors consisting on the variable and the maximum, and the opposite for the minimum. However, for general OWA operators no general results seem to be achievable, so further analyses on specific cases could be the subject of further studies.

There are many other aggregation functions that can be analyzed using this approach. For instance, larger variability could imply more dependence when dealing with some permutation symmetric aggregations, as shown in Example 1. In addition, the results for the maximum and minimum might be extended by imposing conditions to probabilities of the form $P(X_1 > X_2)$.

Funding details

J. Baz is supported by Programa Severo Ochoa of Principality of Asturias (BP21042) and by the Ministry of Science and Innovation (PID2022-139886NB-I00). 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).

References

1. Beliakov, G., Pradera, A., Calvo, T., et al.: Aggregation functions: A guide for practitioners, vol. 221. Springer (2007)
2. Bullen, P.S.: Handbook of means and their inequalities. Springer Science & Business Media, Berlin, Germany (2013)
3. Durante, F., Sempi, C., et al.: Principles of copula theory, vol. 474. CRC press Boca Raton, FL (2016)
4. Grabisch, M., Marichal, J.L., Mesiar, R., Pap, E.: Aggregation functions, vol. 127. Cambridge University Press (2009)
5. Mohd, W.R.W., Abdullah, L.: Aggregation methods in group decision making: A decade survey. Informatica **41**(1) (2017)
6. Nelsen, R.B.: An introduction to copulas. Springer, New York, USA (2006)
7. Nungesser, M.K., Joyce, L.A., McGuire, A.D.: Effects of spatial aggregation on predictions of forest climate change response. Climate research **11**(2), 109–124 (1999)
8. Rohatgi, V.K., Saleh, A.M.E.: An introduction to probability and statistics. John Wiley & Sons, New Jersey, USA (2015)
9. Shaked, M., Shanthikumar, J.G.: Stochastic orders. Springer, New York, USA (2007)
10. Shanmugam, D., Blalock, D., Balakrishnan, G., Guttag, J.: Better aggregation in test-time augmentation. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 1214–1223 (2021)
11. Sklar, A.: Fonctions de repartition an dimensions et leurs marges. Publ. inst. statist. univ. Paris **8**, 229–231 (1959)
12. Yager, R.R.: Families of owa operators. Fuzzy Sets and Systems **59**(2), 125–148 (1993)
13. Yanagimoto, T., Okamoto, M.: Partial orderings of permutations and monotonicity of a rank correlation statistic. Annals of the Institute of Statistical Mathematics **21**, 489–506 (1969)
14. Zahedi Khameneh, A., Kilicman, A.: Some construction methods of aggregation operators in decision-making problems: an overview. Symmetry **12**(5), 694 (2020)