An Analysis of Winsorized Weighted Means

Bonifacio Llamazares[1] ![ORCID]

Abstract

The Winsorized mean is a well-known robust estimator of the population mean. It can also be seen as a symmetric aggregation function (in fact, it is an ordered weighted averaging operator), which means that the information sources (for instance, criteria or experts' opinions) have the same importance. However, in many practical applications (for instance, in many multiattribute decision making problems) it is necessary to consider that the information sources have different importance. For this reason, in this paper we propose a natural generalization of the Winsorized means so that the sources of information can be weighted differently. The new functions, which we will call Winsorized weighted means, are a specific case of the Choquet integral and they are analyzed through several indices for which we give closed-form expressions: the orness degree, k-conjunctiveness and k-disjunctiveness indices, veto and favor indices, Shapley values and interaction indices. We also provide a closed-form expression for the Möbius transform and we show how we can aggregate data so that each information source has the desired weighting and outliers have no influence in the aggregated value.

Keywords Winsorized weighted means · Winsorized means · Choquet integral · Shapley values · SUOWA operators

1 Introduction

The aggregation of information is a fundamental process in many fields of science. Many times, this process is carried out through functions that, from a given data set, return a single value. Obviously, the choice of the function is of crucial importance in this process (see, for instance, Zhang and Xu 2014). For instance, in the field of statistics, it is well known that the mean is quite sensitive to extreme values and, consequently, the sample mean is not a robust estimator of the population mean. So,

✉ Bonifacio Llamazares
boni@eco.uva.es

1   Departamento de Economía Aplicada, Instituto de Matemáticas (IMUVA), Universidad de Valladolid, Avda. Valle de Esgueva 6, 47011 Valladolid, Spain

the mean value may not be very representative in data where outliers may occur. For overcoming this shortcoming, other estimators have been proposed in the literature (see, for instance, Huber and Ronchetti 2009). Two of the most popular are the trimmed means and the Winsorized means. In the trimmed means, the lowest and the highest values are removed before calculating the mean whereas in the Winsorized means they are replaced with the less extreme adjacent values.

Trimmed and Winsorized means can also be seen as symmetric aggregation functions. In fact, they are specific cases of OWA operators (Yager 1988): Trimmed means are OWA operators associated with weighting vectors of the form $(0, \ldots, 0, 1/m, \ldots, 1/m, 0, \ldots, 0)$ (where we are supposing that $n - m$ values have been removed) while Winsorized means are OWA operators associated with weighting vectors of the form $(0, \ldots, 0, (r+1)/n, 1/n, \ldots, 1/n, (s+1)/n, 0, \ldots, 0)$ (where we are supposing that the r lowest values and the s highest values have been replaced). The use of symmetric functions means that all information sources (for instance, criteria or experts' opinions) are treated equally. However, in many practical applications (for instance, in many multiattribute decision making (MADM) problems) it is necessary to consider that the information sources have different importance. But, as what we have previously reported for the mean value, the use of the weighted mean, which is known in the MADM literature as the simple additive weighting (SAW) method, may not be adequate in presence of outliers. For this reason, several proposals have been suggested in the literature with the purpose of introducing different weights for the information sources in an OWA-type aggregation [see, among others, Torra (1997), Llamazares (2013, 2015a), and Beliakov and Dujmović (2016)].

It is interesting to note that the three families of functions introduced by the above-mentioned authors are specific cases of the Choquet integral (Choquet 1953). Since its appearance, the Choquet integral has received increased attention from the scientific community, due mainly to its simplicity, versatility and good properties (for instance, it is well known that the Choquet integral is continuous, monotonic, idempotent, compensative, and ratio scale invariant; see Grabisch et al 2009). For these reasons, the Choquet integral has been a widely used tool in economics to deal with problems related to decision making under risk and uncertainty, finance, insurance, social welfare and quality of life (see, for instance, Heilpern 2002). Likewise, the integral of Choquet has also received much attention in the MADM literature because it allows to take into account the interaction that usually exists between the information sources, which is very useful in this field (among the vast literature that exists on this topic, see, for instance, Grabisch 1995; Grabisch and Labreuche 2010, 2016).

The aim of this paper is to present a natural generalization of the Winsorized means that allows us to consider weighting vectors for the information sources. The new functions, which we will call Winsorized weighted means, are also a specific case of the Choquet integral whose capacities are known. The analysis of these functions is carried out through the study of several indices that provide us with a more precise knowledge about their behavior in the aggregation processes. In this sense, we give closed-form expressions for the following indices: the orness degree (Marichal 2004), k-conjunctiveness and k-disjunctiveness indices (Marichal 2007), veto and favor indices (Marichal 2004, 2007), Shapley values (Shapley 1953), and interaction indices (Owen 1972; Murofushi and Sugeno 1993); and also for their Möbius trans-

form (Rota 1964). Of particular interest are the closed-form expressions obtained for the Shapley values because they reflects the global importance of each information source (as it is well known, in the case of a Choquet integral with respect to a capacity μ, the importance of the ith information source is no given by $\mu(\{i\})$ but by importance indices, and, among them, Shapley values are usually the most used). From them, we can determine the weights that allow us to get Shapley values previously fixed, and, in this way, we can aggregate data so that each information source has the desired weighting and outliers have no influence in the aggregated value (the procedure used will be illustrated with an example in Sect. 5).

The remainder of the paper is organized as follows. In Sect. 2 we recall some basic concepts on Choquet integrals. Section 3 is devoted to introduce the Winsorized weighted means. In Sect. 4 we show the main results of the paper: the Möbius transform and several indices of the Winsorized weighted means are given in closed-form expressions. In Sect. 5 we show by an example how Winsorized weighted means can be applied. Finally, some concluding remarks are provided in Sect. 6. All proofs are given in the Appendix.

2 Choquet Integral

The following notation will be used throughout the paper: $N = \{1, \ldots, n\}$; given $T \subseteq N$, $|T|$ will denote the cardinality of T; and vectors will be denoted in bold. Given a vector $\boldsymbol{x} \in \mathbb{R}^n$, $(\cdot)$ and $[\cdot]$ will denote permutations such that $x_{(1)} \leq \cdots \leq x_{(n)}$ and $x_{[1]} \geq \cdots \geq x_{[n]}$.

Choquet integral was introduced by Choquet (1953) by using the concept of capacity (Choquet 1953), which was also introduced in an independent way by Sugeno (1974) under the name of fuzzy measure.

Definition 1 A capacity (or fuzzy measure) μ on N is a set function, $\mu : 2^N \longrightarrow [0, \infty)$ satisfying $\mu(\varnothing) = 0$ and $\mu(A) \leq \mu(B)$ whenever $A \subseteq B$. A capacity μ is said to be normalized if $\mu(N) = 1$.

The Choquet integral is a functional that generalizes the Lebesgue integral (see, for instance, Choquet 1953, Murofushi and Sugeno 1991, and Denneberg 1994). Nevertheless, in the discrete case, it can be seen as an aggregation function over $\mathbb{R}^n$ (Grabisch et al 2009, p. 181). Notice that we define the Choquet integral for all vectors of $\mathbb{R}^n$ instead of nonnegative vectors given that we are actually considering the asymmetric Choquet integral with respect to μ (on this, see again Grabisch et al 2009, p. 182).

Definition 2 Given a capacity μ on N, the Choquet integral with respect to μ is the function $\mathcal{C}_\mu : \mathbb{R}^n \longrightarrow \mathbb{R}$ defined by

$$\mathcal{C}_\mu(\boldsymbol{x}) = \sum_{i=1}^n \mu(A_{(i)})\big(x_{(i)} - x_{(i-1)}\big),$$

where $A_{(i)} = \{(i), \ldots, (n)\}$, and we adopt the convention that $x_{(0)} = 0$.

Weighted means and OWA operators (Yager 1988) are two of best well-known specific cases of Choquet integrals. Both are defined through weighting vectors; that is, non-negative vectors whose components sum to one.[1]

Definition 3 1. Given a weighting vector $\boldsymbol{p}$, the weighted mean associated with $\boldsymbol{p}$ is the function $M_{\boldsymbol{p}} : \mathbb{R}^n \longrightarrow \mathbb{R}$ defined by

$$M_{\boldsymbol{p}}(\boldsymbol{x}) = \sum_{i=1}^{n} p_i x_i.$$

2. Given a weighting vector $\boldsymbol{w}$, the OWA operator associated with $\boldsymbol{w}$ is the function $O_{\boldsymbol{w}} : \mathbb{R}^n \longrightarrow \mathbb{R}$ defined by

$$O_{\boldsymbol{w}}(\boldsymbol{x}) = \sum_{i=1}^{n} w_i x_{[i]},$$

where as we have said, $[\cdot]$ is a permutation on N such that $x_{[1]} \geq \cdots \geq x_{[n]}$.

3 Winsorized Weighted Means

Winsorized means received this name in honor of Winsor, who suggested replacing the magnitude of extreme observations by the magnitude of the next largest (or smallest) observation (see Dixon 1960; Wainer 1976). Although it is common to define the Winsorized means through the proportion of the replaced values, we will use, for convenience, the number of replaced values (see, for instance, Barnett and Lewis 1994).

Definition 4 Given an integer r, with $0 \leq r < n/2$, the r-fold Winsorized mean is defined by

$$M^r(\boldsymbol{x}) = \frac{1}{n} \left(r x_{(r+1)} + \sum_{i=r+1}^{n-r} x_{(i)} + r x_{(n-r)} \right)$$

Notice that the arithmetic mean (when $r = 0$) and the median (when r is the largest integer less than $n/2$) are specific cases of the Winsorized means. The above definition can be easily extended to consider asymmetrical Winsorized means. For this, we will use the following notation:

$$\mathcal{R} = \{(r, s) \in \{0, 1, \ldots, n-1\}^2 \mid r + s < n\}.$$

---

[1] It is worth noting that the choice of the weight distribution has generated a large literature [in the case of OWA operators, see, for instance, Llamazares (2007), Liu (2011), Bai et al (2017) and Lenormand (2018)].

Definition 5 Let $(r, s) \in \mathcal{R}$. The (r, s)-fold Winsorized mean is defined by

$$M^{(r,s)}(\boldsymbol{x}) = \frac{1}{n} \left(r x_{(r+1)} + \sum_{i=r+1}^{n-s} x_{(i)} + s x_{(n-s)} \right).$$

Note that the kth order statistic $(OS_k(\boldsymbol{x}) = x_{(k)})$ is obtained with the pair $(k - 1, n - k) \in \mathcal{R}$; and the extreme cases of these pairs allow us to get the minimum and the maximum. Obviously, asymmetrical Winsorized means are specific cases of OWA operators.

Winsorized means can also be easily extended to include a weighting vector $\boldsymbol{p}$: Given $\boldsymbol{x}$, it is sufficient to replace the r lowest values and the s highest values by $x_{(r+1)}$ and $x_{(n-s)}$, respectively, and to consider the weighted mean associated with $\boldsymbol{p}$.

Definition 6 Let $\boldsymbol{p}$ be a weighting vector and $(r, s) \in \mathcal{R}$. The (r, s)-fold Winsorized weighted mean is defined by

$$M_{\boldsymbol{p}}^{(r,s)}(\boldsymbol{x}) = \left(\sum_{i=1}^{r} p_{(i)} \right) x_{(r+1)} + \sum_{i=r+1}^{n-s} p_{(i)} x_{(i)} + \left(\sum_{i=n-s+1}^{n} p_{(i)} \right) x_{(n-s)}.$$

Notice that $M_{\boldsymbol{p}}^{(0,0)} = M_{\boldsymbol{p}}$ for any weighting vector $\boldsymbol{p}$, and $M_{\boldsymbol{p}}^{(k-1,n-k)} = OS_k$ for any $k \in N$ and any weighting vector $\boldsymbol{p}$. Moreover, it is worth emphasizing that Winsorized weighted means have been obtained as specific cases of SUOWA operators (Llamazares 2018b); which are in turn particular cases of the Choquet integral [on SUOWA operators see Llamazares (2015a, b, 2016a, b, 2018a, b, 2019a, b)]. It is easy to see (Llamazares 2018b) that the capacity associated with the (r, s)-fold Winsorized weighted mean is

$$\mu_{\boldsymbol{p}}^{(r,s)}(T) = \begin{cases} 0, & \text{if } |T| \leq s, \\ \sum_{i \in T} p_i, & \text{if } s < |T| < n - r, \\ 1, & \text{if } |T| \geq n - r, \end{cases} \tag{1}$$

where $T \subseteq N$. In the specific case of order statistics, where $r + s = n - 1$, the capacity is

$$\mu_{\boldsymbol{p}}^{(r,s)}(T) = \begin{cases} 0, & \text{if } |T| \leq s, \\ 1, & \text{if } |T| \geq n - r = s + 1. \end{cases} \tag{2}$$

4 The Results

There exist in the literature several indices such as the orness and andness degrees, the tolerance indices, the importance and interaction indices, etc., that allow us to know the behavior of the functions used in the aggregation processes. The aim of this section is to analyze the Winsorized weighted means by showing closed-form expressions of the

following indices: the orness degree, k-conjunctiveness and k-disjunctiveness indices, veto and favor indices, the Shapley values, and interaction indices. Moreover, we also provide the expression of the Möbius transforms of the capacities associated with the Winsorized weighted means.

4.1 The Orness Degree

The notion of orness was introduced by Yager (1988) in the field of OWA operators to measure the degree to which the aggregation is disjunctive (i.e., it is like an *or* operation). Subsequently, and by using the notion of average value, Marichal (1998, 2004) generalized it to the case of Choquet integrals and gave an expression in terms of the capacity.

Remark 1 Let μ be a normalized capacity on N. Then

$$\text{orness}(\mathcal{C}_\mu) = \frac{1}{n-1} \sum_{t=1}^{n-1} \frac{1}{\binom{n}{t}} \sum_{\substack{T \subseteq N \\ |T|=t}} \mu(T).$$

The orness degree of the functions $M_p^{(r,s)}$ is shown in the next proposition.

Proposition 1 Let p be a weighting vector and $(r,s) \in \mathcal{R}$. Then,

$$\text{orness}\left(M_p^{(r,s)}\right) = \frac{1}{2} + \frac{r(r+1) - s(s+1)}{2n(n-1)}.$$

It is worth noting that the orness degree of $M_p^{(r,s)}$ does not depend on the weighting vector p; that is, $\text{orness}\left(M_p^{(r,s)}\right) = \text{orness}\left(M^{(r,s)}\right)$ for any weighting vector p. Other immediate consequences of the previous proposition are gathered in the following corollary.

Corollary 1 Let $(r,s) \in \mathcal{R}$. Then:

1. If $r > s$ then $\text{orness}\left(M^{(r,s)}\right) > 0.5$; if $r = s$ then $\text{orness}\left(M^{(r,s)}\right) = 0.5$; and if $r < s$ then $\text{orness}\left(M^{(r,s)}\right) < 0.5$.
2. $\text{orness}\left(M^{(r,s)}\right) = 1 - \text{orness}\left(M^{(s,r)}\right)$.
3. If $r + s \leq n - 2$, then

$$\text{orness}\left(M^{(r+1,s)}\right) = \text{orness}\left(M^{(r,s)}\right) + \frac{r+1}{n(n-1)},$$

$$\text{orness}\left(M^{(r,s+1)}\right) = \text{orness}\left(M^{(r,s)}\right) - \frac{s+1}{n(n-1)}.$$

4. If $r + s \leq n - 3$, then

$$\text{orness}\left(M^{(r+1,s+1)}\right) = \text{orness}\left(M^{(r,s)}\right) + \frac{r-s}{n(n-1)}.$$

Table 1 Orness degree of the functions $M^{(r,s)}$ when $n = 5$

r	s				
	4	3	2	1	0
0	0	0.2	0.35	0.45	0.5
1		0.25	0.4	0.5	0.55
2			0.5	0.6	0.65
3				0.75	0.8
4					1

The properties given in the previous corollary can be easily observed in Table 1, where we show the orness degree of the functions $M^{(r,s)}$ when $n = 5$.

Notice that Table 1 has a triangular structure (there are no values below the main diagonal of the table) so that the values increase as we move down or to the right. Moreover, in the main diagonal appear the orness degree of the order statistics (from the minimum to the maximum), and in the upper right corner appears the orness degree of the weighted mean.

It is also interesting to note that the properties given in Corollary 1 are reflected in the structure of Table 1:

1. The values below the secondary diagonal are greater than 0.5; the values in the secondary diagonal are 0.5; and the values above the secondary diagonal are less than 0.5.
2. The symmetric values with respect to the secondary diagonal sum to one.
3. The value of a cell in the row labeled r increases the amount $\frac{r+1}{n(n-1)}$ when we move a position down; the value of a cell in the column labeled s decreases the amount $\frac{s+1}{n(n-1)}$ when we move a position to the left (alternatively, the value of a cell in the column labeled s increases the amount $\frac{s}{n(n-1)}$ when we move a position to the right).
4. The value of a cell labeled (r, s) varies the amount $\frac{r-s}{n(n-1)}$ when we move down to the left (alternatively, the value of a cell labeled (r, s) increases the amount $\frac{r+s+1}{n(n-1)}$ when we move down to the right).

4.2 *k*-Conjunctiveness and *k*-Disjunctiveness Indices

The notions of k-conjunctive and k-disjunctive Choquet integrals (which were originally called *at most k-intolerant* and *at most k-tolerant* Choquet integrals) were introduced by Marichal (2007) to determine the conjunctive/disjunctive character of aggregation (see also Komorníková and Mesiar 2011).

Definition 7 Let $k \in N$ and let μ be a normalized capacity on N.

1. $\mathcal{C}_\mu$ is k-conjunctive if $\mathcal{C}_\mu \leq \mathrm{OS}_k$; i.e., $\mathcal{C}_\mu(\boldsymbol{x}) \leq x_{(k)}$ for any $\boldsymbol{x} \in \mathbb{R}^n$.
2. $\mathcal{C}_\mu$ is k-disjunctive if $\mathcal{C}_\mu \geq \mathrm{OS}_{n-k+1}$; i.e., $\mathcal{C}_\mu(\boldsymbol{x}) \geq x_{(n-k+1)}$ for any $\boldsymbol{x} \in \mathbb{R}^n$.

Since k-conjunctive and k-disjunctive Choquet integrals are infrequent in practice, Marichal (2007) suggested two indices for measuring the degree to which a Choquet integral is k-conjunctive or k-disjunctive.

Definition 8 Let $k \in N \backslash \{n\}$ and let μ be a normalized capacity on N. The k-conjunctiveness and k-disjunctiveness indices for $\mathcal{C}_\mu$ are defined by

$$\mathrm{conj}_k(\mathcal{C}_\mu) = 1 - \frac{1}{n-k} \sum_{t=1}^{n-k} \frac{1}{\binom{n}{t}} \sum_{\substack{T \subseteq N \\ |T|=t}} \mu(T),$$

$$\mathrm{disj}_k(\mathcal{C}_\mu) = \frac{1}{n-k} \sum_{t=k}^{n} \frac{1}{\binom{n}{t}} \sum_{\substack{T \subseteq N \\ |T|=t}} \mu(T) - \frac{1}{n-k} = \frac{1}{n-k} \sum_{t=k}^{n-1} \frac{1}{\binom{n}{t}} \sum_{\substack{T \subseteq N \\ |T|=t}} \mu(T).$$

In the following propositions we show the k-conjunctiveness and k-disjunctiveness indices for the functions $M_p^{(r,s)}$.

Proposition 2 *Let p be a weighting vector, $(r, s) \in \mathcal{R}$, and $k \in N \backslash \{n\}$. Then,*

$$\mathrm{conj}_k\big(M_p^{(r,s)}\big) = \begin{cases} 1 & \text{if } k \geq n - s, \\ \frac{n(n-1)+s(s+1)-k(k-1)}{2n(n-k)} & \text{if } r < k < n - s, \\ \frac{n(n-1)+s(s+1)-r(r+1)}{2n(n-k)} & \text{if } k \leq r. \end{cases}$$

Proposition 3 *Let p be a weighting vector, $(r, s) \in \mathcal{R}$, and $k \in N \backslash \{n\}$. Then,*

$$\mathrm{disj}_k\big(M_p^{(r,s)}\big) = \begin{cases} 1 & \text{if } k \geq n - r, \\ \frac{n(n-1)+r(r+1)-k(k-1)}{2n(n-k)} & \text{if } s < k < n - r, \\ \frac{n(n-1)+r(r+1)-s(s+1)}{2n(n-k)} & \text{if } k \leq s. \end{cases}$$

As in the case of the orness degree, k-conjunctiveness and k-disjunctiveness indices do not depend on the weighting vector p; that is, $\mathrm{conj}_k\big(M_p^{(r,s)}\big) = \mathrm{conj}_k\big(M^{(r,s)}\big)$ and $\mathrm{disj}_k\big(M_p^{(r,s)}\big) = \mathrm{disj}_k\big(M^{(r,s)}\big)$ for any weighting vector p. Notice also that $\mathrm{conj}_k\big(M^{(r,s)}\big) = \mathrm{disj}_k\big(M^{(s,r)}\big)$.

4.3 Veto and Favor Indices

The notions of veto and favor (originally called dictator) were suggested by Dubois and Koning (1991) in the framework of the social choice functions.

Definition 9 Let $j \in N$ and let μ be a normalized capacity on N.

1. j is a veto for $\mathcal{C}_\mu$ if $\mathcal{C}_\mu(x) \leq x_j$ for any $x \in \mathbb{R}^n$.
2. j is a favor for $\mathcal{C}_\mu$ if $\mathcal{C}_\mu(x) \geq x_j$ for any $x \in \mathbb{R}^n$.

Since veto and favor criteria are rather unusual, Marichal (2004, 2007) proposed two indices to measure the degree with which the behavior of a criterion is like a veto or a favor.

Definition 10 Let $j \in N$ and let μ be a normalized capacity on N. The veto and favor indices of criterion j with respect to μ are defined by

$$\text{veto}(\mathcal{C}_\mu, j) = 1 - \frac{1}{n-1} \sum_{t=1}^{n-1} \frac{1}{\binom{n-1}{t}} \sum_{\substack{T \subseteq N \setminus \{j\} \\ |T|=t}} \mu(T),$$

$$\text{favor}(\mathcal{C}_\mu, j) = \frac{1}{n-1} \sum_{t=0}^{n-1} \frac{1}{\binom{n-1}{t}} \sum_{\substack{T \subseteq N \setminus \{j\} \\ |T|=t}} \mu(T \cup \{j\}) - \frac{1}{n-1}.$$

In the following propositions we show the veto and favor indices of criteria with respect to the capacities $\mu_{\boldsymbol{p}}^{(r,s)}$.

Proposition 4 *Let $\boldsymbol{p}$ be a weighting vector, $(r, s) \in \mathcal{R}$ and $j \in N$. Then,*

$$\text{veto}(M_{\boldsymbol{p}}^{(r,s)}, j) = 1 - \frac{r}{n-1} - \frac{(1 - p_j)\big((n-r)(n-r-1) - s(s+1)\big)}{2(n-1)^2}.$$

Proposition 5 *Let $\boldsymbol{p}$ be a weighting vector, $(r, s) \in \mathcal{R}$ and $j \in N$. Then,*

$$\text{favor}(M_{\boldsymbol{p}}^{(r,s)}, j) = (1 - p_j)\frac{(n-1)(n-2) + r(r+1) - s(s-1)}{2(n-1)^2}$$
$$+ p_j \left(1 - \frac{s}{n-1}\right).$$

4.4 The Shapley Values

The Shapley values were introduced in the cooperative game theory as a solution to the problem of distributing the value $\mu(N)$ among the players. They can be interpreted as a type of average of the contribution of player j alone in all coalitions (see Shapley 1953; Marichal 2007).

Definition 11 Let $j \in N$ and let μ be a normalized capacity on N. The Shapley value of criterion j with respect to μ is defined by

$$\phi_j(\mu) = \frac{1}{n} \sum_{t=0}^{n-1} \frac{1}{\binom{n-1}{t}} \sum_{\substack{T \subseteq N \setminus \{j\} \\ |T|=t}} \Big(\mu(T \cup \{j\}) - \mu(T)\Big).$$

The Shapley values of the capacities $\mu_{\boldsymbol{p}}^{(r,s)}$ are shown in the following proposition.

Proposition 6 *Let $\boldsymbol{p}$ be a weighting vector, $(r, s) \in \mathcal{R}$ and $j \in N$. Then,*

$$\phi_j(\mu_{\boldsymbol{p}}^{(r,s)}) = \frac{r+s}{n-1} \frac{1}{n} + \left(1 - \frac{r+s}{n-1}\right) p_j. \tag{3}$$

It is worth mentioning that the Shapley value $\phi_j\big(\mu_p^{(r,s)}\big)$ is a convex combination between $1/n$ and p_j, which are the Shapley values of the capacities of any OWA operator (for instance, $\mu_p^{(r,n-1-r)}$ or $\mu_p^{(n-1-s,s)}$) and $\mu_p^{(0,0)}$, respectively. Moreover, given $j \in N$, the values $\phi_j\big(\mu_p^{(r,s)}\big)$ are the same for all capacities $\mu_p^{(r,s)}$ having the same value of $r + s$. Other immediate consequences of the previous proposition are gathered in the following corollary.

Corollary 2 *Let p be a weighting vector and $j \in N$.*

1. *If $r + s = n - 1$ then $\phi_j\big(\mu_p^{(r,s)}\big) = 1/n$.*
2. *If $r + s < n - 1$ then*

$$p_j > \frac{1}{n} \implies \phi_j\big(\mu_p^{(r,s)}\big) > \frac{1}{n},$$

$$p_j = \frac{1}{n} \implies \phi_j\big(\mu_p^{(r,s)}\big) = \frac{1}{n},$$

$$p_j < \frac{1}{n} \implies \phi_j\big(\mu_p^{(r,s)}\big) < \frac{1}{n}.$$

Since the Shapley value reflects the global importance of each criterion, it seems very interesting to be able to determine the weights that allow us to obtain Shapley values previously fixed. Notice that when $r + s < n - 1$ we can express the weight p_j in terms of $\phi_j(\mu_p^{(r,s)})$:

$$
\begin{aligned}
p_j &= \frac{n-1}{n-1-(r+s)}\left(\phi_j\big(\mu_p^{(r,s)}\big) - \frac{r+s}{n-1}\frac{1}{n}\right)\\
&= \frac{n-1}{n-1-(r+s)}\left(\phi_j\big(\mu_p^{(r,s)}\big) - \frac{1}{n} + \frac{1}{n}\left(1 - \frac{r+s}{n-1}\right)\right)\\
&= \frac{1}{n} + \frac{n-1}{n-1-(r+s)}\left(\phi_j\big(\mu_p^{(r,s)}\big) - \frac{1}{n}\right).
\end{aligned}
$$

From the previous expressions it is easy to check that $p_j \geq 0$ if and only if $\phi_j(\mu_p^{(r,s)}) \geq \frac{r+s}{n(n-1)}$, and $\sum_{j=1}^{n} p_j = 1$. Therefore we have the following corollary.

Corollary 3 *Let $(\phi_1, \ldots, \phi_n)$ be a weighting vector. Given $(r, s) \in \mathcal{R}$ such that $r+s < n - 1$, the following conditions are equivalent:*

1. $\min\limits_{j \in N} \phi_j \geq \dfrac{r+s}{n(n-1)}.$
2. *The vector p defined by*

$$p_j = \frac{1}{n} + \frac{n-1}{n-1-(r+s)}\left(\phi_j - \frac{1}{n}\right), \quad j = 1, \ldots, n,$$

is a weighting vector such that $\phi_j(\mu_p^{(r,s)}) = \phi_j$ for any $j \in N$.

4.5 Interaction Indices

Interaction indices allow us to measure the interaction degree between two elements of N. This concept was initially proposed by Owen (1972) under the name of covalue, and, in an independent way, by Murofushi and Soneda (1993). Afterwards, it has been generalized by Grabisch (1997) to subsets of N with any number of elements.

Definition 12 Let μ be a normalized capacity on N. The interaction index of elements $j, k \in N$ is defined by

$$
\begin{aligned}
I_{jk}(\mu) = &\frac{1}{n-1} \sum_{t=0}^{n-2} \frac{1}{\binom{n-2}{t}} \\
&\cdot \sum_{\substack{T \subseteq N \setminus \{j,k\} \\ |T|=t}} \Big(\mu(T \cup \{j,k\}) - \mu(T \cup \{j\}) - \mu(T \cup \{k\}) + \mu(T) \Big).
\end{aligned}
$$

We next give the interaction indices of the capacities $\mu_{\boldsymbol{p}}^{(r,s)}$.

Proposition 7 *Let $\boldsymbol{p}$ be a weighting vector, $(r, s) \in \mathcal{R}$ and $j, k \in N$. Then,*

$$
I_{jk}\big(\mu_{\boldsymbol{p}}^{(r,s)}\big) = \begin{cases}
\frac{1}{n-1} & \text{if } r = 0 \text{ and } s = n-1, \\
\frac{1}{n-2}\left(p_j + p_k - \frac{1}{n-1}\right) & \text{if } r = 0 \text{ and } 0 < s < n-1 \\
-\frac{1}{n-2}\left(p_j + p_k - \frac{1}{n-1}\right) & \text{if } s = 0 \text{ and } 0 < r < n-1 \\
-\frac{1}{n-1} & \text{if } s = 0 \text{ and } r = n-1, \\
0 & \text{otherwise.}
\end{cases}
$$

4.6 The Möbius Transform

The Möbius transform (Rota 1964) is a relevant concept in several fields such as combinatorics and cooperative game theory. In this last field it is known as *Harsanyi dividends* (Harsanyi 1959) and it can be interpreted as the contribution of each coalition by itself, without considering its parts.

Definition 13 Let μ be a normalized capacity on N. The Möbius transform of μ is the set function m^μ on N defined by

$$
m^\mu(A) = \sum_{B \subseteq A} (-1)^{|A \setminus B|} \mu(B) \quad (A \subseteq N).
$$

It is worth mentioning that given m^μ, it is possible to recover μ by the expression

$$
\mu(A) = \sum_{B \subseteq A} m^\mu(B) \quad (A \subseteq N),
$$

which it is known as the Zeta transform. Moreover, the Shapley values are related to the Möbius transform through the following expression:

$$\phi_j(\mu) = \sum_{j \in A \subseteq N} \frac{m^\mu(A)}{|A|}.$$

In the following proposition we give the Möbius transforms of the capacities $\mu_p^{(r,s)}$.

Proposition 8 *Let p be a weighting vector and $(r,s) \in \mathcal{R}$. Then,*

$$m^{\mu_p^{(r,s)}}(A) = \begin{cases} 0 & \text{if } |A| \le s, \\ (-1)^{|A|-s-1}\binom{|A|-2}{s-1}\left(\sum_{i \in A} p_i\right) & \text{if } s < |A| < n-r, \\ \left(\sum_{i \in A} p_i\right)\left((-1)^{|A|-s-1}\binom{|A|-2}{s-1}\right. & \\ \left. -(-1)^{|A|-n+r}\binom{|A|-2}{n-r-2}\right) & \\ + (-1)^{|A|-n+r}\binom{|A|-1}{n-r-1} & \text{if } |A| \ge n-r. \end{cases}$$

5 Example

In this section we show the use of Winsorized weighted means from a practical point of view. For this, we will use an example taken from Llamazares (2019b). Suppose that the Department of Mathematics in a Faculty of Economics offers a research assistantship for the students accepted into the M. Sc. in Economics. Applicants are evaluated with respect to seven subjects: Mathematics I (MatI), Mathematics II (MatII), Mathematics III (MatIII), Statistics I (StaI), Statistics II (StaII), Econometrics I (EcoI), and Econometrics II (EcoII), and the members of the committee would like to take into account the following aspects:

1. Each one of the first three subjects is considered twice as important as each one of the remaining four.
2. Outliers should be discarded.[2]

It is worth to notice that the above requirements fit perfectly into a Winsorized weighted mean-type aggregation where, in principle, the weighting vector p is $(0.2, 0.2, 0.2, 0.1, 0.1, 0.1, 0.1)$. Table 2 collects the marks obtained by three students (marks are given on a scale from 0 to 10). Note that, in Statistics I, student A gets its highest mark whereas students B and C get their lowest grades. Furthermore, these marks are also very different from those obtained in the other subjects and, as we will see later, they could be considered outliers.

---

[2] In this framework, outliers may be due to the fact that the same subject may have been taught by different teachers, students may have been ill, or they may have copied answers, etc.

Table 2 Marks of three students in the different subjects

Student	MatI	MatII	MatIII	StaI	StaII	EcoI	EcoII
A	7.9	7.8	7.7	9.8	7.5	7.6	7.4
B	7.7	7.8	7.9	5.2	8.3	8.4	8.5
C	8.2	8.4	8.5	5.2	7.7	7.8	7.9

Table 3 Ranking of the students when the functions $M_p^{(r,s)}$ are used

r	s 6	5	4	3	2	1	0
0	A≻B∼C	A≻B∼C	A≻C≻B	A≻C≻B	C≻A∼B	C≻A≻B	A≻C≻B
1		B∼C≻A	C≻B≻A	C≻B≻A	C≻B≻A	C≻B≻A	C≻B≻A
2			B∼C≻A	C≻B≻A	C≻B≻A	C≻B≻A	C≻B≻A
3				B∼C≻A	C≻B≻A	C≻B≻A	C≻B≻A
4					B≻C≻A	B≻C≻A	B≻C≻A
5						B∼C≻A	C≻B≻A
6							A≻B∼C

Table 4 Shapley values of the Winsorized weighted means

$r+s$	ϕ_1, ϕ_2, ϕ_3	$\phi_4, \phi_5, \phi_6, \phi_7$
0	0.2	0.1
1	$0.1\overline{90476}$	$0.10\overline{714285}$
2	$0.1\overline{809523}$	$0.1\overline{142857}$
3	$0.1\overline{714285}$	$0.12\overline{142857}$
4	$0.1\overline{619047}$	$0.1\overline{285714}$
5	$0.1\overline{523809}$	$0.13\overline{571428}$
6	$0.\overline{142857}$	$0.\overline{142857}$

Table 3 gathers the ranking of the three students when the Winsorized weighted means are used. As one can observe, student A wins when the weighted mean, the minimum or the maximum are used. Moreover, he/she also wins with the pairs (0, 5), (0, 4), and (0, 3). In the case of student B, he/she is the winner with the pairs (4, 2), (4, 1), and (4, 0); and he/she ties with C in the first position with the order statistics OS_2, OS_3, OS_4 and OS_6. In the remaining cases, the winner is student C.

In Table 4 we collect the Shapley values of the Winsorized weighted means obtained by using expression (3). Notice that when $r+s = 0$ (that is, in the case of the weighted mean) we get the desired values. However, as the value of $r+s$ increases, the Shapley values are approaching each other until they finally coincide, what happens when $r + s = 6$ (in this case, the Winsorized weighted means are order statistics and their Shapley values are 1/7).

Given that the Shapley values reflect the global importance of each criterion, the weighting vector p should be chosen so that the Shapley values are 0.2 (for the subjects MatI, MatII and MatIII) and 0.1 (for the remaining subjects). This can be done by using

Table 5 Weights for which the Shapley values of the subjects are 0.2 and 0.1

$r+s$	p_1, p_2, p_3	p_4, p_5, p_6, p_7
0	0.2	0.1
1	$0.21\overline{142857}$	$0.09\overline{142857}$
2	$0.2\overline{285714}$	$0.07\overline{857142}$
3	$0.2\overline{571428}$	$0.0\overline{571428}$
4	$0.3\overline{142857}$	$0.0\overline{142857}$

Table 6 Ranking of the students when the new weights are used

r	s				
	4	3	2	1	0
0	C≻B≻A	C≻B≻A	C≻B≻A	C≻A≻B	A≻C≻B
1		C≻B≻A	C≻B≻A	C≻B≻A	C≻B≻A
2			C≻B≻A	C≻B≻A	C≻B≻A
3				C≻B≻A	C≻B≻A
4					C≻B≻A

Corollary 3. Notice that, since $\min_{j \in N} \phi_j = 0.1$, the weights, which are shown in Table 5, can only be obtained when $r + s \leq 4$.

Table 6 shows the ranking of the students obtained through $M_p^{(r,s)}$ when the new weighting vectors are considered. It is interesting to note that now, in all cases, except for the weighted mean, the winner is student C. In fact, in all cases, except for the pairs (0, 0) and (0, 1), the ranking is C≻B≻A.

Notice that, from a practical point of view, it is necessary to choose a Winsorized weighted mean to rate the students. For that, we can use the following strategies:

1. To use a pair (r, s) previously fixed (some typical percentages are 10%, 15%, 20% or 25%; although smaller percentages are also used; see Hoitash and Hoitash 2009). In this case the pair (r, s) is chosen independently of the outliers present in the data. For instance, suppose we choose a percentage of 20% at both ends. Then, in our example, $0.2 \cdot 7 = 1.4$, and when we round down 1.4 to the nearest integer we have $r = s = 1$. As we can see in Table 6, the winner is student C.

2. To choose a pair (r, s) so that the outliers of all students are removed. For that, first of all we have to choose a method to detect outliers.[3] Usual procedures to detect outliers in the case of one-dimensional data are the boxplot rule (Tukey 1977) and the MAD–median rule (see, for instance, Iglewicz and Hoaglin 1993, Wilcox 2012, and Leys et al 2013). In our example, with both methods we get the following outliers:

   (a) For the student A: 9.8.
   (b) For the student B: 5.2.

---

[3] There is an abundant literature on this subject; see, for instance, Iglewicz and Hoaglin (1993), Barnett and Lewis (1994), Wilcox and Keselman (2003), Seo (2006) and Aggarwal (2017).

(c) For the student C: 5.2. Therefore, to remove the outliers of the three students we should take the pair $(1, 1)$,[4] and the winner is student C (see again Table 6).

6 Conclusion

In this paper we have generalized Winsorized means to include weighting vectors. For this, we have introduced, in a natural way, a new family of functions called Winsorized weighted means. These functions are a specific case of Choquet integrals (in fact, they are a specific case of SUOWA operators) and their capacities are relatively simple. This fact allows that several indices such as the orness degree, the k-conjunctiveness and k-disjunctiveness indices, the veto and favor indices, etc., can be given in closed-form expressions. Of particular interest are the closed-form expressions obtained for the Shapley values because they reflects the global importance of each information source. From them, we can determine the weights that allow us to get Shapley values previously fixed, and, in this way, we can aggregate data so that each information source has the desired weighting and outliers have no influence in the aggregation. It is important to note that other families of functions built in the framework of the Choquet integral do not exhibit this behavior (see Beliakov 2018).

A Proofs

We first recall the definition and some properties of binomial coefficients (see, for instance, Riordan (1968, pp. 1–3) and Grabisch (2016, p. 3).

Remark 2 Let $m \in \mathbb{N}$ and $k \in \mathbb{Z}$. Then:

1. $\binom{m}{k} = \begin{cases} \dfrac{m!}{k!(m-k)!} & \text{if } 0 \le k \le m, \\ 0 & \text{otherwise.} \end{cases}$

2. $\binom{m}{k} = \binom{m}{m-k}.$

3. If $0 \le k \le m$, then $\displaystyle\sum_{j=0}^{k} (-1)^j \binom{m}{j} = (-1)^k \binom{m-1}{k}.$

The following simple remarks on summation will be useful in some of the proofs.

---

[4] In general, the pair (r, s) is obtained by taking $r = \max r_i$ and $s = \max s_i$, where (r_i, s_i) is the pair used for removing the outlier of the ith alternative.

Remark 3 Let $p, q \in \mathbb{N}$, with $p \le q + 1$.[5] Then:

$$\sum_{t=p}^{q} t = \sum_{t=1}^{q} t - \sum_{t=1}^{p-1} t = \frac{q(q+1) - p(p-1)}{2}.$$

Remark 4 Let $\boldsymbol{p}$ be a weighting vector. If $\varnothing \subsetneq A \subseteq N$ and $t \ge 1$, then

$$\sum_{\substack{T \subseteq A \\ |T|=t}} \sum_{i \in T} p_i = \binom{|A| - 1}{t - 1} \sum_{i \in A} p_i.$$

In particular, when $A = N$ we have

$$\sum_{\substack{T \subseteq N \\ |T|=t}} \sum_{i \in T} p_i = \binom{n-1}{t-1} \sum_{i=1}^{n} p_i = \binom{n-1}{t-1} = \binom{n}{t} \frac{t}{n}.$$

Remark 5 Let $\boldsymbol{p}$ be a weighting vector and $j \in N$. If $t \ge 1$, then

$$\sum_{\substack{T \subseteq N \setminus \{j\} \\ |T|=t}} \sum_{i \in T} p_i = \binom{n-2}{t-1} \sum_{\substack{i=1 \\ i \ne j}}^{n} p_i = \binom{n-2}{t-1}(1 - p_j) = \binom{n-1}{t} \frac{t(1 - p_j)}{n-1}.$$

Remark 6 Let $\boldsymbol{p}$ be a weighting vector and $j, k \in N$. If $t \ge 1$, then

$$\sum_{\substack{T \subseteq N \setminus \{j,k\} \\ |T|=t}} \sum_{i \in T} p_i = \binom{n-3}{t-1} \sum_{\substack{i=1 \\ i \ne j,k}}^{n} p_i = \binom{n-3}{t-1}(1 - p_j - p_k)$$

$$= \binom{n-2}{t} \frac{t(1 - p_j - p_k)}{n-2}.$$

Proof of Proposition 1 Let $\boldsymbol{p}$ be a weighting vector and $(r, s) \in \mathcal{R}$. Since

$$\text{orness}\left(M_{\boldsymbol{p}}^{(r,s)}\right) = \frac{1}{n-1} \sum_{t=1}^{n-1} \frac{1}{\binom{n}{t}} \sum_{\substack{T \subseteq N \\ |T|=t}} \mu_{\boldsymbol{p}}^{(r,s)}(T),$$

and $\mu_{\boldsymbol{p}}^{(r,s)}$ is given by expression (1) (or expression (2) when $r + s = n - 1$), we distinguish two cases:

---

[5] Notice that we use the convention $\sum_{t=p}^{q} f(t) = 0$ when $p > q$.

1. If $r + s = n - 1$, then

$$\text{orness}\left(M_{\boldsymbol{p}}^{(r,s)}\right) = \frac{1}{n-1} \sum_{t=s+1}^{n-1} 1 = \frac{n-s-1}{n-1} = \frac{r}{n-1}.$$

2. If $r + s < n - 1$, then, by Remarks 4 and 3, we have

$$\text{orness}\left(M_{\boldsymbol{p}}^{(r,s)}\right)$$

$$= \frac{1}{n-1} \left(\sum_{t=s+1}^{n-r-1} \frac{1}{\binom{n}{t}} \sum_{\substack{T \subseteq N \\ |T|=t}} \sum_{i \in T} p_i + r \right)$$

$$= \frac{1}{n-1} \left(\frac{1}{n} \sum_{t=s+1}^{n-r-1} t + r \right)$$

$$= \frac{1}{n-1} \left(\frac{(n-r)(n-r-1) - s(s+1)}{2n} + r \right)$$

$$= \frac{1}{n-1} \frac{n(n-1) + r(r+1) - s(s+1)}{2n}$$

$$= \frac{1}{2} + \frac{r(r+1) - s(s+1)}{2n(n-1)}.$$

Notice that, when $r + s = n - 1$, the previous expression returns $r/(n-1)$. So, it is also valid in the case $r + s = n - 1$.

$\square$

Proof of Proposition 2 Let $\boldsymbol{p}$ be a weighting vector, $(r, s) \in \mathcal{R}$, and $k \in N \setminus \{n\}$. Since

$$\text{conj}_k\left(M_{\boldsymbol{p}}^{(r,s)}\right) = 1 - \frac{1}{n-k} \sum_{t=1}^{n-k} \frac{1}{\binom{n}{t}} \sum_{\substack{T \subseteq N \\ |T|=t}} \mu_{\boldsymbol{p}}^{(r,s)}(T),$$

and $\mu_{\boldsymbol{p}}^{(r,s)}$ is given by expression (1) (or expression (2) when $r + s = n - 1$), we distinguish the following cases:

1. If $n - k \leq s$ (or, equivalently, $k \geq n - s$), then

$$\text{conj}_k\left(M_{\boldsymbol{p}}^{(r,s)}\right) = 1.$$

2. If $s < n - k < n - r$ (or, equivalently, $r < k < n - s$), by Remarks 4 and 3 we have

$$\text{conj}_k\left(M_{\boldsymbol{p}}^{(r,s)}\right) = 1 - \frac{1}{n-k} \sum_{t=s+1}^{n-k} \frac{1}{\binom{n}{t}} \sum_{\substack{T \subseteq N \\ |T|=t}} \sum_{i \in T} p_i = 1 - \frac{1}{n-k} \sum_{t=s+1}^{n-k} \frac{t}{n}$$

$$= 1 - \frac{1}{n-k} \frac{(n-k)(n-k+1) - s(s+1)}{2n}$$

$$= 1 - \frac{n-k+1}{2n} + \frac{s(s+1)}{2n(n-k)}$$

$$= \frac{n+k-1}{2n} + \frac{s(s+1)}{2n(n-k)}$$

$$= \frac{n(n-1) + s(s+1) - k(k-1)}{2n(n-k)}.$$

3. If $n - k \geq n - r$ (or, equivalently, $k \leq r$), we distinguish two cases:

   (a) If $r + s = n - 1$, then

   $$\mathrm{conj}_k\big(M_{\boldsymbol{p}}^{(r,s)}\big) = 1 - \frac{1}{n-k} \sum_{t=n-r}^{n-k} 1 = 1 - \frac{r-k+1}{n-k} = \frac{n-(r+1)}{n-k}.$$

   (b) If $r + s < n - 1$, then

   $$\mathrm{conj}_k\big(M_{\boldsymbol{p}}^{(r,s)}\big)$$

   $$= 1 - \frac{1}{n-k} \left(\sum_{t=s+1}^{n-r-1} \frac{1}{\binom{n}{t}} \sum_{\substack{T \subseteq N \\ |T|=t}} \sum_{i \in T} p_i + r - k + 1 \right)$$

   $$= 1 - \frac{1}{n-k} \left(\sum_{t=s+1}^{n-r-1} \frac{t}{n} + r - k + 1 \right)$$

   $$= 1 - \frac{1}{n-k} \left(\frac{(n-r)(n-r-1) - s(s+1)}{2n} + r - k + 1 \right)$$

   $$= 1 - \frac{1}{n-k} \left(\frac{n+1-2k}{2} + \frac{r(r+1) - s(s+1)}{2n} \right)$$

   $$= \frac{n(n-1) + s(s+1) - r(r+1)}{2n(n-k)}.$$

   Notice also that, when $r + s = n - 1$, the previous expression returns $(n - r - 1)/(n - k)$. So, it is also valid in the case $r + s = n - 1$.

   $\square$

Proof of Proposition 3 Let $\boldsymbol{p}$ be a weighting vector, $(r, s) \in \mathcal{R}$, and $k \in N \backslash \{n\}$. Since

$$\mathrm{disj}_k\big(M_{\boldsymbol{p}}^{(r,s)}\big) = \frac{1}{n-k} \sum_{t=k}^{n-1} \frac{1}{\binom{n}{t}} \sum_{\substack{T \subseteq N \\ |T|=t}} \mu_{\boldsymbol{p}}^{(r,s)}(T),$$

and $\mu_p^{(r,s)}$ is given by expression (1) (or expression (2) when $r + s = n - 1$), we distinguish the following cases:

1. If $k \geq n - r$, then

$$\mathrm{disj}_k\left(M_p^{(r,s)}\right) = \frac{1}{n-k} \sum_{t=k}^{n-1} 1 = 1.$$

2. If $s < k < n - r$, by Remarks 4 and 3 we have

$$
\begin{aligned}
\mathrm{disj}_k\left(M_p^{(r,s)}\right) &= \frac{1}{n-k}\left(\sum_{t=k}^{n-r-1} \frac{1}{\binom{n}{t}} \sum_{\substack{T \subseteq N \\ |T|=t}} \sum_{i \in T} p_i + r\right) \\
&= \frac{1}{n-k}\left(\sum_{t=k}^{n-r-1} \frac{t}{n} + r\right) \\
&= \frac{1}{n-k}\left(\frac{(n-r)(n-r-1) - k(k-1)}{2n} + r\right) \\
&= \frac{n(n-1) + r(r+1) - k(k-1)}{2n(n-k)}.
\end{aligned}
$$

3. If $k \leq s$, we distinguish two cases:

   (a) If $r + s = n - 1$, then

   $$\mathrm{disj}_k\left(M_p^{(r,s)}\right) = \frac{1}{n-k} \sum_{t=n-r}^{n-1} 1 = \frac{r}{n-k}.$$

   (b) If $r + s < n - 1$, then

   $$\mathrm{disj}_k\left(M_p^{(r,s)}\right) = \frac{1}{n-k}\left(\sum_{t=s+1}^{n-r-1} \frac{1}{\binom{n}{t}} \sum_{\substack{T \subseteq N \\ |T|=t}} \sum_{i \in T} p_i + r\right).$$

   Notice that the above expression coincides with that of the second item when $k = s + 1$. Therefore,

   $$\mathrm{disj}_k\left(M_p^{(r,s)}\right) = \frac{n(n-1) + r(r+1) - s(s+1)}{2n(n-k)}.$$

   Notice also that, when $r + s = n - 1$, the previous expression returns $r/(n-k)$. So, it is also valid in the case $r + s = n - 1$.

   $\square$

Proof of Proposition 4 Let $\boldsymbol{p}$ be a weighting vector, $(r,s) \in \mathcal{R}$ and $j \in N$. Since

$$\text{veto}\big(M_{\boldsymbol{p}}^{(r,s)}, j\big) = 1 - \frac{1}{n-1} \sum_{t=1}^{n-1} \frac{1}{\binom{n-1}{t}} \sum_{\substack{T \subseteq N \setminus \{j\} \\ |T|=t}} \mu_{\boldsymbol{p}}^{(r,s)}(T),$$

and $\mu_{\boldsymbol{p}}^{(r,s)}$ is given by expression (1) (or expression (2) when $r+s = n-1$), we distinguish two cases:

1. If $r+s = n-1$, then

$$\text{veto}\big(M_{\boldsymbol{p}}^{(r,s)}, j\big) = 1 - \frac{1}{n-1} \sum_{t=s+1}^{n-1} 1 = 1 - \frac{r}{n-1} = \frac{s}{n-1}.$$

2. If $r+s < n-1$, then, by Remarks 5 and 3, we get

$$\text{veto}\big(M_{\boldsymbol{p}}^{(r,s)}, j\big) = 1 - \frac{1}{n-1} \left(\sum_{t=s+1}^{n-r-1} \frac{1}{\binom{n-1}{t}} \sum_{\substack{T \subseteq N \setminus \{j\} \\ |T|=t}} \sum_{i \in T} p_i + r \right)$$

$$= 1 - \frac{1}{n-1} \left(\frac{1-p_j}{n-1} \sum_{t=s+1}^{n-r-1} t + r \right)$$

$$= 1 - \frac{r}{n-1} - \frac{(1-p_j)\big((n-r)(n-r-1) - s(s+1)\big)}{2(n-1)^2}.$$

Notice that, when $r+s = n-1$, the previous expression returns $s/(n-1)$. So, it is also valid in the case $r+s = n-1$.

$\square$

Proof of Proposition 5 Let $\boldsymbol{p}$ be a weighting vector, $(r,s) \in \mathcal{R}$ and $j \in N$. Since

$$\text{favor}\big(M_{\boldsymbol{p}}^{(r,s)}, j\big) = \frac{1}{n-1} \sum_{t=0}^{n-1} \frac{1}{\binom{n-1}{t}} \sum_{\substack{T \subseteq N \setminus \{j\} \\ |T|=t}} \mu_{\boldsymbol{p}}^{(r,s)}(T \cup \{j\}) - \frac{1}{n-1},$$

and $\mu_{\boldsymbol{p}}^{(r,s)}$ is given by expression (1) (or expression (2) when $r+s = n-1$), we distinguish two cases:

1. If $r+s = n-1$, then

$$\text{favor}\big(M_{\boldsymbol{p}}^{(r,s)}, j\big) = \frac{1}{n-1} \sum_{t=s}^{n-1} 1 - \frac{1}{n-1} = \frac{n-s-1}{n-1} = \frac{r}{n-1}.$$

2. If $r + s < n - 1$, then, by Remarks 5 and 3, we get

$$\text{favor}\left(M_{\boldsymbol{p}}^{(r,s)}, j\right)$$

$$= \frac{1}{n-1}\left(\sum_{t=s}^{n-r-2}\frac{1}{\binom{n-1}{t}}\sum_{\substack{T\subseteq N\setminus\{j\}\\|T|=t}}\left(\sum_{i\in T}p_i + p_j\right) + r + 1\right) - \frac{1}{n-1}$$

$$= \frac{1}{n-1}\left(\frac{1-p_j}{n-1}\sum_{t=s}^{n-r-2}t + p_j(n-r-s-1) + r\right)$$

$$= \frac{(1-p_j)\big((n-r-2)(n-r-1) - s(s-1) + 2r(n-1)\big)}{2(n-1)^2}$$

$$+ \frac{p_j(n-s-1)}{n-1}$$

$$= (1-p_j)\frac{(n-1)(n-2) + r(r+1) - s(s-1)}{2(n-1)^2} + p_j\left(1 - \frac{s}{n-1}\right).$$

Notice that, when $r + s = n - 1$, the previous expression returns $r/(n-1)$. So, it is also valid in the case $r + s = n - 1$.

$\square$

Proof of Proposition 6 Let $\boldsymbol{p}$ be a weighting vector, $(r, s) \in \mathcal{R}$ and $j \in N$. Since

$$\phi_j\left(\mu_{\boldsymbol{p}}^{(r,s)}\right) = \frac{1}{n}\sum_{t=0}^{n-1}\frac{1}{\binom{n-1}{t}}\sum_{\substack{T\subseteq N\setminus\{j\}\\|T|=t}}\mu_{\boldsymbol{p}}^{(r,s)}(T \cup \{j\})$$

$$- \frac{1}{n}\sum_{t=0}^{n-1}\frac{1}{\binom{n-1}{t}}\sum_{\substack{T\subseteq N\setminus\{j\}\\|T|=t}}\mu_{\boldsymbol{p}}^{(r,s)}(T).$$

and $\mu_{\boldsymbol{p}}^{(r,s)}$ is given by expression (1) (or expression (2) when $r + s = n - 1$), we distinguish two cases:

1. If $r + s = n - 1$, then

$$\phi_j\left(\mu_{\boldsymbol{p}}^{(r,s)}\right) = \frac{1}{n}\sum_{t=s}^{n-1}1 - \frac{1}{n}\sum_{t=s+1}^{n-1}1 = \frac{1}{n}.$$

2. If $r + s < n - 1$, then, by Remark 5, we have

$$\phi_j\left(\mu_{\boldsymbol{p}}^{(r,s)}\right) = \frac{1}{n}\left(\sum_{t=s}^{n-r-2}\frac{1}{\binom{n-1}{t}}\sum_{\substack{T\subseteq N\setminus\{j\}\\|T|=t}}\left(p_j + \sum_{i\in T}p_i\right) + r + 1\right)$$

$$-\frac{1}{n}\left(\sum_{t=s+1}^{n-r-1}\frac{1}{\binom{n-1}{t}}\sum_{\substack{T\subseteq N\setminus\{j\}\\|T|=t}}\sum_{i\in T}p_i+r\right)$$

$$=\frac{1}{n}\left((n-r-s-1)p_j+\frac{1-p_j}{n-1}(s-(n-r-1))+1\right)$$

$$=\frac{1}{n}\left(\frac{s+r+n(n-r-s-1)p_j}{n-1}\right)$$

$$=\frac{r+s}{n-1}\frac{1}{n}+\frac{n-1-r-s}{n-1}p_j$$

$$=\frac{r+s}{n-1}\frac{1}{n}+\left(1-\frac{r+s}{n-1}\right)p_j.$$

Notice that, when $r+s=n-1$, the previous expression returns $1/n$. So, it is also valid in the case $r+s=n-1$.

$\square$

Proof of Proposition 7 Let p be a weighting vector, $(r,s)\in\mathcal{R}$ and $j,k\in N$. Notice that $I_{jk}\left(\mu_p^{(r,s)}\right)$ can be written as

$$I_{jk}\left(\mu_p^{(r,s)}\right)=\frac{1}{n-1}\left(I_{\{j,k\}}-I_{\{j\}}-I_{\{k\}}+I_\varnothing\right),$$

where I_K, $K\subseteq N$, is defined by

$$I_K=\sum_{t=0}^{n-2}\frac{1}{\binom{n-2}{t}}\sum_{\substack{T\subseteq N\setminus\{j,k\}\\|T|=t}}\mu_p^{(r,s)}(T\cup K).$$

Since $\mu_p^{(r,s)}$ is given by expression (1) (or expression (2) when $r+s=n-1$), we distinguish two cases:

1. If $r+s=n-1$, then it is easy to check that $I_{\{j,k\}}$, $I_{\{j\}}$, $I_{\{k\}}$, and $I_\varnothing$ take the following values:

$$I_{\{j,k\}}=\begin{cases}\sum_{t=0}^{n-2}1=n-1 & \text{if }s=0,\\ \sum_{t=s-1}^{n-2}1=n-s & \text{otherwise,}\end{cases}$$

$$I_{\{j\}}=I_{\{k\}}=\sum_{t=s}^{n-2}1=n-1-s,$$

$$I_\varnothing=\begin{cases}\sum_{t=s+1}^{n-2}1=n-2-s & \text{if }s<n-1,\\ 0 & \text{otherwise.}\end{cases}$$

We now calculate $I_{jk}\left(\mu_p^{(r,s)}\right)$ taking into account the different values of $I_{\{j,k\}}$, $I_{\{j\}}$, $I_{\{k\}}$, and $I_\varnothing$. We distinguish three cases:

(a) If $s = 0$, which is equivalent to $r = n - 1$, we get

$$I_{jk}\left(\mu_p^{(r,s)}\right) = \frac{1}{n-1}(n - 1 - 2(n - 1) + n - 2) = -\frac{1}{n-1}.$$

(b) If $0 < s < n - 1$ we have

$$I_{jk}\left(\mu_p^{(r,s)}\right) = \frac{1}{n-1}(n - s - 2(n - 1 - s) + n - 2 - s) = 0.$$

(c) If $s = n - 1$, which is equivalent to $r = 0$, we obtain

$$I_{jk}\left(\mu_p^{(r,s)}\right) = \frac{1}{n-1}.$$

Therefore,

$$I_{jk}\left(\mu_p^{(r,s)}\right) = \begin{cases} \frac{1}{n-1} & \text{if } r = 0, \\ -\frac{1}{n-1} & \text{if } r = n - 1, \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

2. If $r + s < n - 1$, then, by Remark 6, we can see that $I_{\{j,k\}}$, $I_{\{j\}}$, $I_{\{k\}}$, and $I_\varnothing$ take the following values:

$$I_{\{j,k\}} = \begin{cases} \sum_{t=0}^{n-2} 1 = n - 1 & \text{if } s = 0 \text{ and } r = n - 2, \\ \sum_{t=0}^{n-r-3}\left(p_j + p_k + \frac{t(1-p_j-p_k)}{n-2}\right) + r + 1 & \text{if } s = 0 \text{ and } r < n - 2, \\ \sum_{t=s-1}^{n-r-3}\left(p_j + p_k + \frac{t(1-p_j-p_k)}{n-2}\right) + r + 1 & \text{otherwise,} \end{cases}$$

$$I_{\{j\}} = \sum_{t=s}^{n-r-2}\left(p_j + \frac{t(1 - p_j - p_k)}{n - 2}\right) + r,$$

$$I_{\{k\}} = \sum_{t=s}^{n-r-2}\left(p_k + \frac{t(1 - p_j - p_k)}{n - 2}\right) + r,$$

$$I_\varnothing = \begin{cases} 0 & \text{if } r = 0 \text{ and } s = n - 2, \\ \sum_{t=s+1}^{n-2} \frac{t(1-p_j-p_k)}{n-2} & \text{if } r = 0 \text{ and } s < n - 2, \\ \sum_{t=s+1}^{n-r-1} \frac{t(1-p_j-p_k)}{n-2} + r - 1 & \text{otherwise,} \end{cases}$$

We now calculate $I_{jk}\left(\mu_p^{(r,s)}\right)$ taking into account the different values of $I_{\{j,k\}}$, $I_{\{j\}}$, $I_{\{k\}}$, and $I_\varnothing$. For instance, when $s = 0$ and $r = n - 2$ we get

$$I_{jk}\left(\mu_{\boldsymbol{p}}^{(r,s)}\right)$$

$$= \frac{1}{n-1}\left(n-1-p_j-(n-2)-p_k-(n-2)+\frac{1-p_j-p_k}{n-2}+n-3\right)$$

$$= \frac{1}{n-1}\frac{1-(n-1)(p_j+p_k)}{n-2} = -\frac{1}{n-2}\left(p_j+p_k-\frac{1}{n-1}\right).$$

Once all the cases have been analyzed (to avoid tedious calculations, the remaining cases are left to the reader), we have

$$I_{jk}\left(\mu_{\boldsymbol{p}}^{(r,s)}\right) = \begin{cases} \frac{1}{n-2}\left(p_j+p_k-\frac{1}{n-1}\right) & \text{if } r=0 \text{ and } 0<s<n-1, \\ -\frac{1}{n-2}\left(p_j+p_k-\frac{1}{n-1}\right) & \text{if } s=0 \text{ and } 0<r<n-1, \quad (5) \\ 0 & \text{otherwise.} \end{cases}$$

Expressions (4) and (5) together establish the truth of Proposition 7. □

Proof of Proposition 8 Let $\boldsymbol{p}$ be a weighting vector and $(r,s) \in \mathcal{R}$. Given $A \subseteq N$, since

$$m^{\mu_{\boldsymbol{p}}^{(r,s)}}(A) = \sum_{t=1}^{|A|}(-1)^{|A|-t}\sum_{\substack{T\subseteq A \\ |T|=t}}\mu_{\boldsymbol{p}}^{(r,s)}(T),$$

and $\mu_{\boldsymbol{p}}^{(r,s)}$ is given by expression (1) (or expression (2) when $r+s=n-1$), we distinguish the following cases:

1. If $|A| \le s$, then

$$m^{\mu_{\boldsymbol{p}}^{(r,s)}}(A) = 0.$$

2. If $s < |A| < n-r$, by Remark 4 and the third item of Remark 2, we get

$$m^{\mu_{\boldsymbol{p}}^{(r,s)}}(A)$$

$$= \sum_{t=s+1}^{|A|}(-1)^{|A|-t}\sum_{\substack{T\subseteq A \\ |T|=t}}\sum_{i\in T}p_i = \sum_{t=s+1}^{|A|}(-1)^{|A|-t}\binom{|A|-1}{t-1}\sum_{i\in A}p_i$$

$$= \left(\sum_{i\in A}p_i\right)\sum_{t=s+1}^{|A|}(-1)^{|A|-t}\binom{|A|-1}{|A|-t} = \left(\sum_{i\in A}p_i\right)\sum_{j=0}^{|A|-s-1}(-1)^j\binom{|A|-1}{j}$$

$$= \left(\sum_{i\in A}p_i\right)(-1)^{|A|-s-1}\binom{|A|-2}{|A|-s-1} = (-1)^{|A|-s-1}\binom{|A|-2}{s-1}\left(\sum_{i\in A}p_i\right).$$

Notice that when $s=0$ we have $m^{\mu_{\boldsymbol{p}}^{(r,s)}}(A) = 0$.

3 If $|A| \geq n - r$, we distinguish two cases:

(a) If $r + s = n - 1$, then, by the third item of Remark 2 we have

$$
m^{\mu_p^{(r,s)}}(A) = \sum_{t=s+1}^{|A|} (-1)^{|A|-t} \binom{|A|}{t} = \sum_{t=s+1}^{|A|} (-1)^{|A|-t} \binom{|A|}{|A|-t}
$$

$$
= \sum_{j=0}^{|A|-s-1} (-1)^j \binom{|A|}{j} = (-1)^{|A|-s-1} \binom{|A|-1}{|A|-s-1}
$$

$$
= (-1)^{|A|-s-1} \binom{|A|-1}{s}.
$$

(b) If $r + s < n - 1$, by Remark 4 and the third item of Remark 2, we have

$$
m^{\mu_p^{(r,s)}}(A)
$$

$$
= \sum_{t=s+1}^{n-r-1} (-1)^{|A|-t} \binom{|A|-1}{t-1} \sum_{i \in A} p_i + \sum_{t=n-r}^{|A|} (-1)^{|A|-t} \binom{|A|}{t}
$$

$$
= \left(\sum_{i \in A} p_i\right) \sum_{t=s+1}^{n-r-1} (-1)^{|A|-t} \binom{|A|-1}{|A|-t} + \sum_{t=n-r}^{|A|} (-1)^{|A|-t} \binom{|A|}{|A|-t}
$$

$$
= \left(\sum_{i \in A} p_i\right) \sum_{j=|A|-n+r+1}^{|A|-s-1} (-1)^j \binom{|A|-1}{j} + \sum_{j=0}^{|A|-n+r} (-1)^j \binom{|A|}{j}
$$

$$
= \left(\sum_{i \in A} p_i\right) \left((-1)^{|A|-s-1} \binom{|A|-2}{|A|-s-1}\right.
$$

$$
\left. -(-1)^{|A|-n+r} \binom{|A|-2}{|A|-n+r}\right) + (-1)^{|A|-n+r} \binom{|A|-1}{|A|-n+r}
$$

$$
= \left(\sum_{i \in A} p_i\right) \left((-1)^{|A|-s-1} \binom{|A|-2}{s-1} - (-1)^{|A|-n+r} \binom{|A|-2}{n-r-2}\right)
$$

$$
+ (-1)^{|A|-n+r} \binom{|A|-1}{n-r-1}.
$$

Notice also that, when $r + s = n - 1$, the previous expression returns $(-1)^{|A|-s-1} \binom{|A|-1}{s}$. So, it is also valid in the case $r + s = n - 1$.

$\square$

References

Aggarwal CC (2017) Outlier analysis, 2nd edn. Springer, Cham

Bai C, Zhang R, Song C, Wu Y (2017) A new ordered weighted averaging operator to obtain the associated weights based on the principle of least mean square errors. Int J Intell Syst 32(3):213–226

Barnett V, Lewis T (1994) Outliers in statistical data, 3rd edn. Wiley, Chichester

Beliakov G (2018) Comparing apples and oranges: the weighted OWA function. Int J Intell Syst 33(5):1089–1108

Beliakov G, Dujmović J (2016) Extension of bivariate means to weighted means of several arguments by using binary trees. Inf Sci 331:137–147

Choquet G (1953) Theory of capacities. Ann Inst Fourier 5:131–295

Denneberg D (1994) Non-additive measures and integral. Kluwer Academic Publisher, Dordrecht

Dixon WJ (1960) Simplified estimation from censored normal samples. Ann Math Stat 31(2):385–391

Dubois D, Koning JL (1991) Social choice axioms for fuzzy set aggregation. Fuzzy Sets Syst 43(3):257–274

Grabisch M (1995) Fuzzy integral in multicriteria decision making. Fuzzy Sets Syst 69(3):279–298

Grabisch M (1997) k-order additive discrete fuzzy measures and their representation. Fuzzy Sets Syst 92(2):167–189

Grabisch M (2016) Set functions, games and capacities in decision making, theory and decision library, series C, vol 46. Springer, Berlin

Grabisch M, Labreuche C (2010) A decade of application of the Choquet and Sugeno integrals in multi-criteria decision aid. Ann Oper Res 175(1):247–286

Grabisch M, Labreuche C (2016) Fuzzy measures and integrals in MCDA. In: Greco S, Ehrgott M, Figueira RJ (eds) Multiple criteria decision analysis: state of the art surveys, international series in operations research and management science, vol 233, 2nd edn. Springer, New York, pp 553–603

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

Harsanyi JC (1959) A bargaining model for cooperative n-person games. In: Tucker AW, Luce RD (eds) Contributions to the theory of games. Vol. 4, Annals of Mathematics Studies, vol 40. Princeton University Press, Princeton, pp 325–355

Heilpern S (2002) Using Choquet integral in economics. Stat Pap 43(1):53–73

Hoitash U, Hoitash R (2009) Conflicting objectives within the board: evidence from overlapping audit and compensation committee members. Group Decis Negot 18(1):57–73

Huber PJ, Ronchetti EM (2009) Robust statistics, 2nd edn. Wiley, Hoboken

Iglewicz B, Hoaglin D (1993) How to detect and handle outliers. ASQC Quality Press, Milwaukee

Komorníková M, Mesiar R (2011) Aggregation functions on bounded partially ordered sets and their classification. Fuzzy Sets Syst 175(1):48–56

Lenormand M (2018) Generating OWA weights using truncated distributions. Int J Intell Syst 33(4):791–801

Leys C, Ley C, Klein O, Bernard P, Licata L (2013) Detecting outliers: do not use standard deviation around the mean, use absolute deviation around the median. J Exp Soc Psychol 49(4):764–766

Liu X (2011) A review of the OWA determination methods: classification and some extensions. In: Yager RR, Kacprzyk J, Beliakov G (eds) Recent developments in the ordered weighted averaging operators: theory and practice. Springer, Berlin, pp 49–90

Llamazares B (2007) Choosing OWA operator weights in the field of social choice. Inf Sci 177(21):4745–4756

Llamazares B (2013) An analysis of some functions that generalizes weighted means and OWA operators. Int J Intell Syst 28(4):380–393

Llamazares B (2015a) Constructing Choquet integral-based operators that generalize weighted means and OWA operators. Inf Fus 23:131–138

Llamazares B (2015b) A study of SUOWA operators in two dimensions. Math Probl Eng 2015: Article ID 271,491

Llamazares B (2016a) A behavioral analysis of WOWA and SUOWA operators. Int J Intell Syst 31(8):827–851

Llamazares B (2016b) SUOWA operators: constructing semi-uninorms and analyzing specific cases. Fuzzy Sets Syst 287:119–136

Llamazares B (2018a) Closed-form expressions for some indices of SUOWA operators. Inf Fus 41:80–90

Llamazares B (2018b) Construction of Choquet integrals through unimodal weighting vectors. Int J Intell Syst 33(4):771–790

Llamazares B (2019a) SUOWA operators: a review of the state of the art. Int J Intell Syst 34(5):790–818

Llamazares B (2019b) SUOWA operators: an analysis of their conjunctive/disjunctive character. Fuzzy Sets Syst 357:117–134

Marichal JL (1998) Aggregation operators for multicriteria decision aid. Ph.D. thesis, University of Liège

Marichal JL (2004) Tolerant or intolerant character of interacting criteria in aggregation by the Choquet integral. Eur J Oper Res 155(3):771–791

Marichal JL (2007) k-intolerant capacities and Choquet integrals. Eur J Oper Res 177(3):1453–1468

Murofushi T, Soneda S (1993) Techniques for reading fuzzy measures (iii): interaction index. In: Proceedings of the 9th fuzzy systems symposium, Sapporo (Japan), pp 693–696

Murofushi T, Sugeno M (1991) A theory of fuzzy measures. Representation, the Choquet integral and null sets. J Math Anal Appl 159(2):532–549

Murofushi T, Sugeno M (1993) Some quantities represented by the Choquet integral. Fuzzy Sets Syst 56(2):229–235

Owen G (1972) Multilinear extensions of games. Manag Sci 18(5–part–2):64–79

Riordan J (1968) Combinatorial identities. Wiley, New York

Rota GC (1964) On the foundations of combinatorial theory I. Theory of Möbius functions. Z Wahrschein-lichkeitstheorie Verwandte Geb 2(4):340–368

Seo S (2006) A review and comparison of methods for detecting outliers in univariate data sets. Master's thesis, University of Pittsburgh

Shapley LS (1953) A value for n-person games. In: Kuhn H, Tucker AW (eds) Contributions to the theory of games, vol 2. Princeton University Press, Princeton, pp 307–317

Sugeno M (1974) Theory of fuzzy integrals and its applications. Ph.D. thesis, Tokyo Institute of Technology

Torra V (1997) The weighted OWA operator. Int J Intell Syst 12(2):153–166

Tukey JW (1977) Exploratory data analysis. Addison-Wesley, Reading

Wainer H (1976) Robust statistics: a survey and some prescriptions. J Educ Stat 1(4):285–312

Wilcox RR (2012) Modern statistics for the social and behavioral sciences: a practical introduction. CRC Press, Boca Raton

Wilcox RR, Keselman HJ (2003) Modern robust data analysis methods: measures of central tendency. Psychol Methods 8(3):254–274

Yager RR (1988) On ordered weighted averaging operators in multicriteria decision making. IEEE Trans Syst Man Cybern 18(1):183–190

Zhang Z, Xu Z (2014) Analysis on aggregation function spaces. Int J Uncertain Fuzziness Knowl Based Syst 22(05):737–747