Axiomatic characterizations of (quasi-) L-statistics and S-statistics and the Producer Assessment Problem

Marek Gągolewski[*1,2] **Przemysław Grzegorzewski**[1,2]

[1]Systems Research Institute, Polish Academy of Sciences,
Newelska 6, 01-447 Warsaw, Poland
[2]Faculty of Mathematics and Information Science, Warsaw University of Technology,
Plac Politechniki 1, 00-661 Warsaw, Poland

Abstract

Two classes of aggregation functions: L-statistics and S-statistics and their generalizations called quasi-L-statistics and quasi-S-statistics are considered. Some interesting characterizations of these families of operators are given. The aforementioned functions are useful for various applications. In particular, they are very helpful for modeling the so-called Producer Assessment Problem.

Keywords: Aggregation operators, impact functions, scientometrics, L-statistics, S-statistics, OWA operators, OMA operators, OWMax operators, the Hirsch index.

1. Introduction

Aggregation operators, i.e. functions combining many numerical values into a single one, play an important role in many areas of everyday life including decision making, computer science, data summarization, statistics, economy, social sciences or operational research. Apart from particular applications, the theory of aggregation operators is a fascinating and rapidly developing mathematical domain where many problems are still open, there are plenty of facts to be discovered and there is a need for new ideas to be elaborated (we refer the reader to [10] for the recent state of the art monograph).

Operators projecting multidimensional state space into a single dimension are often called aggregation functions [10]. Typical examples widely used in the decision making are: the sample maximum and other quantiles, arithmetic mean, OWA [20] and OWMax [4] operators, etc.

In this paper we consider two useful classes of aggregation operators: L-statistics and S-statistics and their natural generalizations, called quasi-L-statistics and quasi-S-statistics, respectively. We discuss some of their properties, especially the characterization theorems.

It is worth mentioning that so well-known and popular classes of aggregation operators like OWA, OWMax or OMA operators appear as particular

cases of the families of the functions discussed in this paper.

This contribution extends our results given in [6] and recent results obtained by Mesiar and Mesiárova-Zemánková [15]. In particular, we analyze the aforementioned functions with respect to the so-called Producer Assessment Problem (which involves situations where given alternatives are rated not only with respect to the quality of delivered items but also to their productivity) discussed, e.g., in [6, 8]. The issue of fair assessment of scientists based on the number of citations gained by their papers is its the most representative instance. The h-index proposed by J.E. Hirsch [11], one of the most widely known tools used in this domain, is a special case of the discussed quasi-S-statistics.

The paper is organized as follows. In Sec. 2 we present the conventional notation used throughout the paper and recall the concept of impact functions which form a model for the Producer Assessment Problem. In Sec. 3 we describe the class of L-statistics and S-statistics, their natural generalizations and basic properties. Then in Sec. 4 we present a few characterization theorems for the aforementioned classes of functions. Finally, in Sec. 5 we discuss the problem of vector concatenation with respect to its influence on the impact function value.

2. Preliminaries

2.1. Basic notation

Let $\mathbb{I} = [a, b]$ denote any nonempty closed interval of extended real numbers $\bar{\mathbb{R}} = [-\infty, \infty]$. Unless stated otherwise, $n, m \in \mathbb{N}$. We assume that $\mathbb{N}_0 = \{0, 1, 2, \dots\}$ denotes the set of all nonnegative integers while $[n] = \{1, 2, \dots, n\}$.

Given any $\mathbf{x} = (x_1, \dots, x_n)$, $\mathbf{y} = (y_1, \dots, y_n)$ $\in \mathbb{I}^n$, we write $\mathbf{x} \leq \mathbf{y}$ if and only if $(\forall i \in [n])$ $x_i \leq y_i$. Similarly, we say that $\mathbf{x} < \mathbf{y}$ when $\mathbf{x} \leq \mathbf{y}$ and $\mathbf{x} \neq \mathbf{y}$. Additionally, $\mathbf{x} \cong \mathbf{y}$ if and only if there exists a permutation σ of $[n]$ such that $(x_1, \dots, x_n) = (y_{\sigma(1)}, \dots, y_{\sigma(n)})$.

Let $x_{(i)}$ denote the ith-smallest value of $\mathbf{x} = (x_1, \dots, x_n)$ while $(n * x)$ stands for the n-tuple

---
[*]Corresponding author; e-mail: gagolews@ibspan.waw.pl.

Table 1: The Producer Assessment Problem — typical examples.

	Producer	Products	Rating method	Discipline
A	Scientist	Scientific articles	Number of citations	Scientometrics
B	Scientific institute	Scientists	The h-index	Scientometrics
C	Web server	Web pages	Number of in-links	Webometrics
D	Artist	Paintings	Auction price	Auctions
E	Billboard company	Advertisements	Sale results	Marketing

$(x, x, \ldots, x) \in \mathbb{I}^n$.

For any $\mathbf{x} \in \mathbb{I}^n$, $\mathbf{y} \in \mathbb{I}^m$ and any function $\mathbf{g}$ defined on $\mathbb{I}^{n+m}$ the notation $\mathbf{g}(\mathbf{x}, \mathbf{y})$ is applied for $\mathbf{g}(x_1, \ldots, x_n, y_1, \ldots, y_m)$. The vector operators $+, \vee$ are understood component-wise.

Moreover, let $\mathbb{I}^{1,2,\cdots} = \bigcup_{n=1}^{\infty} \mathbb{I}^n$ denote the set of all vectors of arbitrary length with elements in $\mathbb{I}$. By $\mathcal{E}(\mathbb{I}) = \{\mathsf{F} : \mathbb{I}^{1,2,\cdots} \to \mathbb{I}\}$ we mean the set of all **aggregation operators** on $\mathbb{I}$.

Finally, $\mathbf{1}$ denotes the indicator function.

2.2. The Producer Assessment Problem

Consider a **producer** (e.g. a writer, scientist, artist, craftsman) and a nonempty set of his **products** (e.g. books, papers, works, goods). Suppose that each product is given a **rating** (of quality, popularity, etc.) which is a single number in $\mathbb{I} = [a, b]$, where a denotes the lowest admissible valuation. Some typical instances of such situation are listed in Table 1.

It is clear that each possible producer's state can be described by a point in $\mathbb{I}^{1,2,\cdots}$. The **Producer Assessment Problem** involves constructing and analyzing aggregation operators which can be used for rating producers (see [6, 8]). A family of such functions should take into account the two following aspects of producer's quality:

- the ability to output highly-rated products,
- overall productivity.

Clearly, the first component for fixed productivity can be properly described by a broad class of (extended) aggregation functions [10]. However, in practice we are also interested in distinguishing between entities of different productivity. Thus consider the following definition (its slightly modified version was given in [6]).

Definition 1. An **impact function** in $\mathbb{I}^{1,2,\cdots}$ is a function $\mathsf{J} \in \mathcal{E}(\mathbb{I})$ such that

(I1) J is nondecreasing in each variable, i.e.
$(\forall n)(\forall \mathbf{x}, \mathbf{y} \in \mathbb{I}^n) \ \mathbf{x} \le \mathbf{y} \Rightarrow \mathsf{J}(\mathbf{x}) \le \mathsf{J}(\mathbf{y})$;

(I2) J is arity-monotonic, i.e.
$(\forall n, m)(\forall \mathbf{x} \in \mathbb{I}^n)(\forall \mathbf{y} \in \mathbb{I}^m) \ \mathsf{J}(\mathbf{x}) \le \mathsf{J}(\mathbf{x}, \mathbf{y})$;

(I3) J is symmetric, i.e.
$(\forall n)(\forall \mathbf{x}, \mathbf{y} \in \mathbb{I}^n) \ \mathbf{x} \cong \mathbf{y} \Rightarrow \mathsf{J}(\mathbf{x}) = \mathsf{J}(\mathbf{y})$;

(I4) J fulfills the weak lower boundary condition, i.e. $\inf_{\mathbf{x} \in \mathbb{I}^{1,2,\cdots}} \mathsf{J}(\mathbf{x}) = a$.

Conditions (I1) and (I2) correspond to the principle called "the more the better", which is justified in many practical instances of the Producer Assessment Problem. According to (I3) each product is of equal importance and the overall rating is not affected by the presentation order of the products.

The family of all impact functions will be denoted by $\mathcal{E}_{\mathcal{I}}(\mathbb{I})$.

Such formal model given for the Producer Assessment Problem allows us to abstract from its context-dependent interpretation (to avoid any bias) and focus solely on the analysis of its mathematical properties.

It is worth noting that the set of requirements given in Def. 1 is similar to the axiomatization proposed by Woeginger [18, 19] for the so-called bibliometric impact indices (for other axiomatizations see [12, 13, 16, 17]).

Henceforth, let $P_{(\text{I1})}$ stand for the family of nondecreasing functions, $P_{(\text{I2})}$ denote the functions fulfilling (I2), and so on. Thus the family of all impact functions satisfies the following condition: $\mathcal{E}_{\mathcal{I}}(\mathbb{I}) = P_{(\text{I1})} \cap P_{(\text{I2})} \cap P_{(\text{I3})} \cap P_{(\text{I4})}$.

For the clarity of presentation, we also consider a subclass of impact functions that additionally satisfies the condition defined below (see e.g. [10, Def. 2.7]).

Definition 2. $\mathsf{F} \in \mathcal{E}_{\mathcal{I}}(\mathbb{I})$ fulfills the (strong) **lower boundary condition** (denoted $\mathsf{F} \in P_{(\text{lb})}$) if, for all n, $\inf_{\mathbf{x} \in \mathbb{I}^n} \mathsf{F}(\mathbf{x}) = a$.

One can easily see that $P_{(\text{lb})} \subset P_{(\text{I4})}$. Moreover, it can be shown that for any $\mathsf{F} \in P_{(\text{I1})}$ we have $\mathsf{F} \in P_{(\text{lb})}$ if and only if $(\forall n) \ \mathsf{F}(n * a) = a$. From now on, let $\mathcal{E}'_{\mathcal{I}}(\mathbb{I}) = P_{(\text{I1})} \cap P_{(\text{I2})} \cap P_{(\text{I3})} \cap P_{(\text{lb})}$.

3. (Quasi-)L- and S-statistics

3.1. Problem

In this section we consider two particular families of aggregation operators, called L-statistics and S-statistics and discuss their generalized versions, called quasi-L-statistics and quasi-S-statistics, respectively. We examine in which cases they also belong to the class of impact functions satisfying the lower boundary condition. Some corresponding results for usual L-statistics and S-statistics are given in [6].

Further on we assume that $\mathbb{I} = [0, \infty]$. The following two concepts will be also useful.

Definition 3. A **triangle of coefficients** (compare [14, 2]) is a sequence $\triangle = (c_{i,n} \in \mathbb{I} : i \in [n], n \in \mathbb{N})$.

Note that such triangle of coefficients can be represented graphically by

$$
\begin{array}{cccc}
c_{1,1} & & & \\
c_{1,2} & c_{2,2} & & \\
c_{1,3} & c_{2,3} & c_{3,3} & \\
\vdots & \vdots & \vdots & \ddots
\end{array}
$$

In a similar way one can define a triangle of functions.

Definition 4. A triangle of functions is a sequence $\triangle = (f_{i,n} \in \mathbb{I}^{\mathbb{I}} : i \in [n], n \in \mathbb{N})$.

3.2. L-statistics and quasi-L-statistics

Let us start from recalling the definition of L-statistic.

Definition 5. Given any triangle of coefficients $\triangle = (c_{i,n})_{i \in [n], n \in \mathbb{N}}$, the **L-statistic** associated with $\triangle$ is a function $\mathsf{L}_{\triangle} \in \mathcal{E}(\mathbb{I})$ such that

$$
\mathsf{L}_{\triangle}(\mathbf{x}) = \sum_{i=1}^{n} c_{i,n}\, x_{(n-i+1)}, \tag{1}
$$

for any $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{I}^n$.

Please note that the well-known OWA functions (*ordered weighted averaging functions*), introduced in [20], are just special cases of L-statistics. Actually, the triangle of coefficients for an OWA function $\triangle = (w_{i,n})_{i \in [n], n \in \mathbb{N}}$ satisfies the following conditions (for $a = 0$):

(a) $(\forall n)\ \sum_{j=1}^{n} w_{j,n} = 1$,
(b) $w_{i,n} \in [0, 1]$ for $i \in [n]$.

So defined triangle of coefficients is sometimes called a *weighting triangle* (see e.g. [2, 3] for an extension of OWA for input vectors of arbitrary length).

Now let us consider the following generalization of the usual L-statistic.

Definition 6. Given any triangle of functions $\triangle = (f_{i,n})_{i \in [n], n \in \mathbb{N}}$, the **quasi-L-statistic** associated with $\triangle$ is a function $\mathsf{qL}_{\triangle} \in \mathcal{E}(\mathbb{I})$ such that

$$
\mathsf{qL}_{\triangle}(\mathbf{x}) = \sum_{i=1}^{n} f_{i,n}(x_{(n-i+1)}) \tag{2}
$$

for any $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{I}^n$.

It is easily seen that quasi-L-statistics generalize L-statistics. We obtain an L-statistic by taking $f_{i,n}(x) = c_{i,n} x$. Moreover, it is worth noting that the OMA functions for $a = 0$ (*ordered modular averages*) are special cases of the suggested class of quasi-L-statistics for $\sum_{i=1}^{n} f_{i,n} = \mathsf{id}$, nondecreasing $f_{i,n}$ and $0 \preceq f_{i,n} \preceq b$ (see [15]).

One may be interested in which cases the (quasi-)L-statistics are impact functions satisfying the lower boundary condition. The proposition given below extends the result shown in [6].

Proposition 7. *Let* $\mathbb{I} = [0, \infty]$ *and let* $\triangle = (f_{i,n})_{n \in \mathbb{N}, i \in [n]}$ *denote any triangle of functions. Then* $\mathsf{qL}_{\triangle} \in \mathcal{E}'_{\mathcal{I}}(\mathbb{I})$ *if and only if* $(\forall n)(\forall i \in [n])$ *functions* $f_{i,n}$ *satisfy the following conditions*

- $f_{i,n}$ *is nondecreasing,*
- $f_{i,n}(0) = 0$,
- $\sum_{j=1}^{i} f_{j,n+1} \succeq \sum_{j=1}^{i} f_{j,n}$.

Proof. It is clear that $\mathsf{qL}_{\triangle} \in P_{(I1)} \cap P_{(I3)} \cap P_{(lb)}$ if and only if each $f_{i,n}$ is nondecreasing and $f_{i,n}(0) = 0$.

Such $\mathsf{qL}_{\triangle}$ is arity-monotonic if and only if $(\forall \mathbf{x} \in \mathbb{I}^n)\ \sum_{j=1}^{n} f_{j,n+1}(x_{(n-i+1)}) \geq \sum_{j=1}^{n} f_{j,n}(x_{(n-i+1)})$.

Fix any $\mathbf{x} \in \mathbb{I}^n$. We have $\mathsf{qL}_{\triangle}(x_{(n)}, n * 0) \geq \mathsf{qL}_{\triangle}(x_{(n)}, (n-1) * 0)$ iff $f_{1,n+1} \succeq f_{1,n}$. Then $\mathsf{qL}_{\triangle}(x_{(n)}, x_{(n-1)}, (n-1) * 0) \geq \mathsf{qL}_{\triangle}(x_{(n)}, x_{(n-1)}, (n-2)*0)$ iff additionally $f_{1,n+1} + f_{2,n+1} \succeq f_{1,n} + f_{2,n}$ and so on, QED. $\qquad\square$

3.3. S- and quasi-S-statistics

Now let us consider S-statistics and their extensions.

Definition 8. The **S-statistic** associated with a triangle of coefficients $\triangle = (c_{i,n})_{i \in [n], n \in \mathbb{N}}$ is a function $\mathsf{S}_{\triangle} \in \mathcal{E}(\mathbb{I})$ such that

$$
\mathsf{S}_{\triangle}(\mathbf{x}) = \bigvee_{i=1}^{n} c_{i,n} \wedge x_{(n-i+1)} \tag{3}
$$

for $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{I}^n$, where $\vee$ and $\wedge$ denote the supremum (and hence the name) and infimum operators, respectively.

S-statistics generalize the OWMax operators (ordered weighted maximum operator) for arbitrary number of arguments [4].

Basic statistical properties of S-statistics were examined in [7]. In particular, it turns out that for a random input data set (continuous i.i.d. random variables) the distribution of an S-statistic is asymptotically normal with the mean equal to the value of a well-defined characteristic of the underlying probability distribution. Moreover, there exists an S-statistic $\mathsf{S}_{\blacktriangle}$ that $\lfloor \mathsf{S}_{\blacktriangle} \rfloor = \mathsf{H}$, where H is the Hirsch h-index [11], a popular tool used in scientometrics.

Now, similarly as in the previous section, we propose a generalization of the S-statistic.

Definition 9. The **quasi-S-statistic** associated with a triangle of functions $\triangle = (f_{i,n})_{i \in [n], n \in \mathbb{N}}$ is a function $\mathsf{qS}_{\triangle} \in \mathcal{E}(\mathbb{I})$ such that

$$
\mathsf{qS}_{\triangle}(\mathbf{x}) = \bigvee_{i=1}^{n} f_{i,n}(x_{(n-i+1)}). \tag{4}
$$

Equation (4) is actually a generalization of S-statistic since by setting $f_{i,n}(x) = x \wedge c_{i,n}$ we obtain a usual S-statistic. Note that for $x, c_{i,n} \in \mathbb{I}$

$$
x \wedge c_{i,n} = \begin{cases} x & \text{if } x < c_{i,n}, \\ c_{i,n} & \text{if } x \geq c_{i,n}. \end{cases}
$$

Moreover, its right-sided derivative $\frac{d_+}{dx}(x \wedge c_{i,n}) = \mathbf{1}(x \leq c_{i,n})$.

Without loss of generality we assume further on that $(\forall n)$ $c_{1,n} \leq c_{2,n} \leq \cdots \leq c_{n,n}$ and $\mathsf{f}_{1,n} \preceq \mathsf{f}_{2,n} \preceq \cdots \preceq \mathsf{f}_{n,n}$.

We are going to show which triangles of coefficients generate a quasi-S-statistic that is an impact function fulfilling $P_{(\mathrm{lb})}$. Again, we generalize the result given in [6].

Proposition 10. *For $\mathbb{I} = [0, \infty]$ and any triangle of functions $\triangle = (\mathsf{f}_{i,n})_{n \in \mathbb{N}, i \in [n]}$ it holds $\mathsf{qS}_\triangle \in \mathcal{E}'_{\mathcal{I}}(\mathbb{I})$ if and only if $(\forall n)(\forall i \in [n])$*

- *$\mathsf{f}_{i,n}$ is nondecreasing,*
- *$\mathsf{f}_{n,n}(0) = 0$, and*
- *$\mathsf{f}_{i,n+1} \succeq \mathsf{f}_{i,n}$.*

Proof. It is obvious that $\mathsf{qS}_\triangle \in P_{(\mathrm{I}1)} \cap P_{(\mathrm{I}3)} \cap P_{(\mathrm{lb})}$ if and only if each $\mathsf{f}_{i,n}$ is nondecreasing and $\mathsf{f}_{1,n}(0) = \cdots = \mathsf{f}_{n,n}(0) = 0$.

Let $\mathbf{x} \in \mathbb{I}^n$. It suffices to show that $\mathsf{qS}_\triangle(\mathbf{x}, 0) \geq \mathsf{qS}_\triangle(\mathbf{x})$ if and only if $\mathsf{f}_{i,n+1} \succeq \mathsf{f}_{i,n}$ for all $i \in [n]$. We have $\mathsf{qS}_\triangle(x_{(n)}, n * 0) \geq \mathsf{qS}_\triangle(x_{(n)}, (n-1) * 0)$ if and only if $\mathsf{f}_{1,n+1} \succeq \mathsf{f}_{1,n}$. Then $\mathsf{qS}_\triangle(x_{(n)}, x_{(n-1)}, (n-1) * 0) \geq \mathsf{qS}_\triangle(x_{(n)}, x_{(n-1)}, (n-2) * 0)$ if and only if additionally $\mathsf{f}_{2,n+1} \succeq \mathsf{f}_{2,n}$ and so on. This completes the proof. $\qquad\square$

Let $\triangle = (c_{i,n})_{i \in [n], n \in \mathbb{N}}$. Mesiar and Mesiárova-Zemánková in [15] recently shown that $\mathsf{S}_\triangle = \mathsf{qL}_\triangledown$ with $\triangledown = (\mathsf{f}_{i,n})_{i \in [n], n \in \mathbb{N}}$ such that $\mathsf{f}_{i,n}(x) = 0 \vee ((x - c_{i-1,n}) \wedge (c_{i,n} - c_{i-1,n}))$, with convention that $c_{0,n} = 0$.

One may be also interested in the relationship between quasi-S-statistics and quasi-L-statistics. Next theorem defines the intersection of these two classes in the case of nondecreasing functions fulfilling the (strong) lower boundary condition.

Theorem 11. *Let $\mathbb{I} = [0, \infty]$ and $\triangle = (\mathsf{f}_{i,n})_{i \in [n], n \in \mathbb{N}}$ such that $\mathsf{qS}_\triangle \in P_{(\mathrm{I}1)} \cap P_{(\mathrm{lb})}$. Then there exists $\triangledown = (\mathsf{g}_{i,n})_{i \in [n], n \in \mathbb{N}}$ such that $\mathsf{qS}_\triangle = \mathsf{qL}_\triangledown$ if and only if there exist nondecreasing functions $\mathsf{w}_1, \mathsf{w}_2 \cdots \in \mathbb{I}^\mathbb{I}$, $\mathsf{w}_n(0) = 0$ for all n, and a triangle of coefficients $\blacktriangle = (c_{i,n})_{i \in [n], n \in \mathbb{N}}$ such that $\mathsf{f}_{i,n}(x) = \mathsf{w}_n(x) \wedge c_{i,n}$.*

Proof. Fix n. We have $(\forall i \in [n])$ $\mathsf{f}_{i,n}(0) = \mathsf{g}_{i,n}(0)$. We check for which $\triangle$ and $\triangledown$ $\bigvee_{i=1}^n \mathsf{f}_{i,n}(x_{(n-i+1)}) = \sum_{i=1}^n \mathsf{g}_{i,n}(x_{(n-i+1)})$ holds for all $\mathbf{x} \in \mathbb{I}^n$. We have $\mathsf{f}_{1,n} = \mathsf{g}_{1,n}$, because $\mathsf{qS}_\triangle(x_{(n)}, (n-1) * 0) = \mathsf{qL}_\triangledown(x_{(n)}, (n-1) * 0)$. Then, by considering the vector $(x_{(n)}, x_{(n-1)}, (n-2) * 0)$ we obtain

$$\mathsf{g}_{2,n}(x_{(n-1)}) = \begin{cases} 0 \text{ if } \mathsf{f}_{1,n}(x_{(n)}) - \mathsf{f}_{2,n}(x_{(n-1)}) \geq 0, \\ \mathsf{f}_{2,n}(x_{(n-1)}) - \mathsf{f}_{1,n}(x_{(n)}) \text{ otherwise.} \end{cases}$$

We have $\mathsf{f}_{1,n} \preceq \mathsf{f}_{2,n}$. However, the function $\mathsf{g}_{2,n}(x_{(n-1)})$ should not depend on $x_{(n)}$. Therefore $\mathsf{f}_{1,n}$ should be of the form

$$\mathsf{f}_{1,n}(x) = \begin{cases} \mathsf{f}_{2,n}(x) & \text{for } x \leq y, \\ c_{1,n} & \text{for } x > y, \end{cases}$$

for some $y, c_{1,n} \in \mathbb{I}$. This implies $\mathsf{f}_{1,n}(x) = \mathsf{f}_{2,n}(x) \wedge c_{1,n}$. By applying a similar reasoning for successive elements of $\mathbf{x}$, we obtain $\mathsf{w}_n = \mathsf{f}_{n,n}$, and the theorem holds. $\qquad\square$

4. Characterizations of the operators

In this section we characterize the subclasses of the mentioned functions, that is we discuss a certain set of properties that are fulfilled *if and only if* an aggregation operator is either a (quasi)-L- or S-statistic.

4.1. A characterization of L-statistics and quasi-L-statistics

A characterization of nondecreasing L- and quasi-L-statistics follows from [10] and [15], respectively. Here we recall these results for the sake of completeness.

Definition 12. *Let $\mathsf{F} \in \mathcal{E}(\mathbb{I})$. Then F is **symmetric additive** (denoted $\mathsf{F} \in P_{(\mathrm{sadd})}$), whenever $(\forall n)$ $(\forall \mathbf{x}, \mathbf{y} \in \mathbb{I}^n)$*

$$\mathsf{F}(\mathbf{x} \overset{S}{+} \mathbf{y}) = \mathsf{F}(\mathbf{x}) + \mathsf{F}(\mathbf{y}),$$

where $\mathbf{x} \overset{S}{+} \mathbf{y} = (x_{(1)} + y_{(1)}, \ldots, x_{(n)} + y_{(n)})$.

We see that $P_{(\mathrm{sadd})} \subseteq P_{(\mathrm{I}3)}$.

Theorem 13. *Let $\mathsf{F} \in P_{(\mathrm{I}1)}$. Then $\mathsf{F} \in P_{(\mathrm{sadd})}$ if and only if F is a nondecreasing L-statistic.*

The proof follows from [10, Prop. 2.116].

Definition 14. *Let $\mathsf{F} \in \mathcal{E}(\mathbb{I})$. Then F is **symmetric modular** (denoted $\mathsf{F} \in P_{(\mathrm{smod})}$), whenever $(\forall n)$ $(\forall \mathbf{x}, \mathbf{y} \in \mathbb{I}^n)$*

$$\mathsf{F}(\mathbf{x} \overset{S}{\vee} \mathbf{y}) + \mathsf{F}(\mathbf{x} \overset{S}{\wedge} \mathbf{y}) = \mathsf{F}(\mathbf{x}) + \mathsf{F}(\mathbf{y}),$$

where $\mathbf{x} \overset{S}{\vee} \mathbf{y} = (x_{(1)} \vee y_{(1)}, \ldots, x_{(n)} \vee y_{(n)})$ and $\mathbf{x} \overset{S}{\wedge} \mathbf{y} = (x_{(1)} \wedge y_{(1)}, \ldots, x_{(n)} \wedge y_{(n)})$.

We have $P_{(\mathrm{smod})} \subseteq P_{(\mathrm{I}1)} \cap P_{(\mathrm{I}3)}$.

Theorem 15. *$\mathsf{F} \in P_{(\mathrm{smod})}$ if and only if F is a nondecreasing quasi-L-statistic.*

The proof is omitted.

4.2. A characterization of S-statistics and quasi-S-statistics

Definition 16. *We say that $\mathsf{F} \in \mathcal{E}(\mathbb{I})$ has **regular linear increments** (denoted $\mathsf{F} \in P_{(\mathrm{rli})}$) iff, for any n, $\mathbf{x} \in \mathbb{I}^n$ and i, it holds*

$$\frac{\partial_+}{\partial x_i} \mathsf{F}(\mathbf{x}) = \begin{cases} 1 & \text{if } x_i \in [\mathsf{F}(\mathbf{x}); \mathsf{F}(\mathbf{x}'_i)), \\ 0 & \text{otherwise,} \end{cases}$$

where $\mathbf{x}'_i = (x_1, \ldots, x_{i-1}, \infty, x_{i+1}, \ldots, x_n)$ and $\frac{\partial_+}{\partial x}$ is the right-sided derivative.

Each $\mathsf{F} \in P_{(\mathrm{rli})}$ is also continuous and nondecreasing in each variable. Moreover, the above definition implies that $(\forall s \in \operatorname{ran} \mathsf{F} = [\mathsf{F}(n*a); \mathsf{F}(n*b)])$ $\mathsf{F}(n*s) = s$, i.e. it is a range-idempotent function, see def. 2.38 in [10].

We propose the following axiomatization of the class of S-statistics.

Theorem 17. *We have* $\mathsf{F} \in P_{(\mathrm{rli})} \cap P_{(\mathrm{I3})} \cap P_{(\mathrm{lb})}$ *if and only if* F *is an S-statistic.*

Proof. ($\Leftarrow$) Trivial.

($\Rightarrow$) Let $\mathsf{F} \in P_{(\mathrm{rli})} \cap P_{(\mathrm{I3})}$, $n \in \mathbb{N}$ and $\mathbf{x} \in \mathbb{I}^n$. Let $d_n = \mathsf{F}(n*0)$ and $c_{i,n} = \mathsf{F}(i*\infty, (n-i)*0) \geq d_n$. We have $\mathsf{F}(x_{(n)}, (n-1)*0) = (x_{(n)} \wedge c_{1,n}) \vee d_n$. If $x_{(n-1)} \leq (x_{(n)} \wedge c_{1,n}) \vee d_n$ then $\mathsf{F}(\mathbf{x}) = \mathsf{F}(x_{(n)}, (n-1)*0)$. Otherwise, $\mathsf{F}(x_{(n)}, x_{(n-1)}, (n-2)*0) = (x_{(n-1)} \wedge c_{2,n})$. That is, more generally,

$$\mathsf{F}(x_{(n)}, x_{(n-1)}, (n-2)*0) = $$
$$d_n \vee (x_{(n)} \wedge c_{1,n}) \vee (x_{(n-1)} \wedge c_{2,n}).$$

By considering successive values $x_{(n-2)}, \ldots, x_{(1)}$ in a similar manner we come to the equality

$$\mathsf{F}(\mathbf{x}) = d_n \vee \bigvee_{i=1}^n (x_{(n-i+1)} \wedge c_{i,n}).$$

Now, as F also fulfills the (strong) lower boundary condition, $d_n = 0$ and hence F is an S-statistic. $\square$

To characterize the quasi-S-statistics we use the symmetrized maxitivity property, see [10, def. 2.119].

Definition 18. Let $\mathsf{F} \in \mathcal{E}(\mathbb{I})$. Then F is **symmetric maxitive** (denoted $\mathsf{F} \in P_{(\mathrm{smax})}$), whenever $(\forall n)$ $(\forall \mathbf{x}, \mathbf{y} \in \mathbb{I}^n)$

$$\mathsf{F}(\mathbf{x} \overset{S}{\vee} \mathbf{y}) = \mathsf{F}(\mathbf{x}) \vee \mathsf{F}(\mathbf{y}).$$

It is worth pointing out that $P_{(\mathrm{smax})} \subset P_{(\mathrm{I3})}$. We are now ready to give the following axiomatization of the last class of functions.

Theorem 19. $\mathsf{F} \in P_{(\mathrm{smax})}$ *if and only if* F *is a nondecreasing quasi-S-statistic.*

We omit the simple proof.

5. Vector concatenation and impact function values

Consider the following problem. We are given an impact function $\mathsf{F} \in \mathcal{E}_{\mathcal{I}}(\mathbb{I})$ and two input vectors $\mathbf{x}, \mathbf{y} \in \mathbb{I}^{1,2,\cdots}$. Knowing both values of $\mathsf{F}(\mathbf{x})$ and $\mathsf{F}(\mathbf{y})$ we would like to estimate the value of $\mathsf{F}(\mathbf{x}, \mathbf{y})$, where $(\mathbf{x}, \mathbf{y})$ denotes vector concatenation (merging).

The problem is not only of theoretical manner. A real-life example of such situation was considered in the field of scientometrics where the Hirsch h-index and Egghe g-index was explored [5, 9]. It has some significant practical applications, e.g., when calculating the impact functions for a producer given his/her products' quality valuations obtained from disjoint data sources. Such a case is often encountered in bibliometrics, where citation data are gathered from diverse databases like the *Web of Science* and Elsevier *SciVerse Scopus* (see e.g. [5]).

Generally, properties $P_{(\mathrm{I1})}$, $P_{(\mathrm{I2})}$ and $P_{(\mathrm{I3})}$ imply

$$\mathsf{F}(\mathbf{x}, \mathbf{y}) \geq \mathsf{F}(\mathbf{x}) \vee \mathsf{F}(\mathbf{y}).$$

On the other hand, the only universal upper bound is given by $\mathsf{F}(\mathbf{x}, \mathbf{y}) \leq \infty = \sup \mathbb{I}$.

It is easily seen that for $(\mathbf{x}, \mathbf{y}) \in \mathbb{I}^n$ we have $\mathsf{S}_\triangle(\mathbf{x}, \mathbf{y}) \leq c_{n,n}$.

Below we consider two basic properties that may lead to exact value of the impact function. Furthermore, corresponding operators that satisfy these properties are characterized.

First we show when the above-mentioned lower bound is also an upper bound.

Definition 20. We say that $\mathsf{F} \in \mathcal{E}(\mathbb{I})$ is **merge-maxitive** (denoted $\mathsf{F} \in P_{(\mathrm{mmax})}$), if $(\forall \mathbf{x}, \mathbf{y} \in \mathbb{I}^{1,2,\cdots})$

$$\mathsf{F}(\mathbf{x}, \mathbf{y}) = \mathsf{F}(\mathbf{x}) \vee \mathsf{F}(\mathbf{y}).$$

Proposition 21. *Let* $\mathsf{F} \in \mathcal{E}_{\mathcal{I}}(\mathbb{I})$. *Then* $\mathsf{F} \in P_{(\mathrm{mmax})}$ *if and only if there exists a nondecreasing function* $\mathsf{g} : \mathbb{I} \to \mathbb{I}$, $\mathsf{g}(0) = 0$, *such that* $(\forall n)(\forall \mathbf{x} \in \mathbb{I}^n)$ $\mathsf{F}(\mathbf{x}) = \bigvee_{i=1}^n \mathsf{g}(x_i) = \mathsf{g}(x_{(n)})$.

Proof. ($\Rightarrow$) Let F be as assumed and $\mathbf{x} \in \mathbb{I}^n$. We have $\mathsf{F}(\mathbf{x}) = \mathsf{F}(x_{(1)}) \vee \cdots \vee \mathsf{F}(x_{(n)}) = \bigvee_{i=1}^n \mathsf{g}(x_{(i)})$. For F be nondecreasing in any variable, g must also be nondecreasing. $P_{(\mathrm{I4})}$ is satisfied by assuming $\mathsf{g}(0) = 0$. From that it follows $\mathsf{F}(\mathbf{x}) = \mathsf{g}(x_{(n)})$.

As ($\Leftarrow$) is trivial, the proof is complete. $\square$

Definition 22. We say that $\mathsf{F} \in \mathcal{E}(\mathbb{I})$ is **merge-additive** (denoted $\mathsf{F} \in P_{(\mathrm{madd})}$), if $(\forall \mathbf{x}, \mathbf{y} \in \mathbb{I}^{1,2,\cdots})$

$$\mathsf{F}(\mathbf{x}, \mathbf{y}) = \mathsf{F}(\mathbf{x}) + \mathsf{F}(\mathbf{y}).$$

Proposition 23. *Let* $\mathsf{F} \in \mathcal{E}_{\mathcal{I}}(\mathbb{I})$. *It holds* $\mathsf{F} \in P_{(\mathrm{madd})}$ *if and only if there exists a nondecreasing function* $\mathsf{g} : \mathbb{I} \to \mathbb{I}$, $\mathsf{g}(0) = 0$ *such that* $(\forall n)(\forall \mathbf{x} \in \mathbb{I}^n)$ $\mathsf{F}(\mathbf{x}) = \sum_{i=1}^n \mathsf{g}(x_i)$.

Proof. ($\Rightarrow$) By the symmetry property and merge-additivity we have $\mathsf{F}(\mathbf{x}) = \mathsf{F}(x_{\sigma(1)}) + \cdots + \mathsf{F}(x_{\sigma(n)}) = \sum_{i=1}^n \mathsf{g}(x_i)$ for some g. g must be nondecreasing and $\mathsf{g}(0) = 0$ to satisfy $P_{(\mathrm{I1})}$ and $P_{(\mathrm{I4})}$.

($\Leftarrow$) Trivial. $\square$

We see that the both properties are quite strong. The merge-maxitive impact function is a quasi-S-statistics for a triangle of functions $\triangle$ such that $\mathsf{f}_{i,n} = \mathsf{g}$ for all $n, i \in [n]$ (equivalent to quasi-L-statistics with $\triangle$ such that $\mathsf{f}_{1,n} = \mathsf{g}$ and $\mathsf{f}_{i,n} = 0$ for all $n, i \in \{2, 3, \ldots, n\}$). Merge-additive functions are equivalent to quasi-L-statistics with $\triangle$ such that $\mathsf{f}_{i,n} = \mathsf{g}$.

6. Conclusions

In this paper we have discussed some families of aggregation operators which may be regarded useful for modeling the Producer Assessment Problem. Especially, we have considered a particular subfamily of impact functions, called S-statistics (which generalize the well-known h-index and OWMax operators), their natural extensions called quasi-S-statistics and two other families of operators: L-statistics and quasi-L-statistics.

All tools proposed and discussed in this paper have interesting properties and hence may support the decision process and be useful in many areas, e.g. in scientometrics, marketing, management, etc.

References

[1] T. Calvo et al, Generation of weighting triangles associated with aggregation functions, **Intl. J. Uncertainty, Fuzziness and Knowledge-based Systems** 8(4):417–451, 2000.

[2] T. Calvo, A. Kolesarova, M. Komornikova and R. Mesiar, Aggregation operators: Properties, classes and construction methods. In T. Calvo, G. Mayor and R. Mesiar, editors, **Aggregation Operators. New Trends and Applications**, pages 3–104, Physica-Verlag, New York, 2002.

[3] T. Calvo and G. Mayor, Remarks on two types of extended aggregation functions, **Tatra Mountains Mathematical Publications** 16:235–253, 1999.

[4] D. Dubois, H. Prade and C. Testemale, Weighted fuzzy pattern matching, **Fuzzy Sets and Systems** 28: 313-331, 1998.

[5] L. Egghe, The influence of merging on h-type indices, **Journal of Informetrics** 2(3):252–262, 2008.

[6] M. Gagolewski and P. Grzegorzewski, Arity-monotonic extended aggregation operators, **CCIS** 80:693–702, 2010.

[7] M. Gagolewski and P. Grzegorzewski, S-statistics and their basic properties, In C. Borgelt et al, editors, **Combining Soft Computing and Statistical Methods in Data Analysis**, pages 281–288, Springer-Verlag, 2010.

[8] M. Gagolewski and P. Grzegorzewski, Possibilistic analysis of arity-monotonic aggregation operators and its relation to bibliometric impact assessment of individuals, **International Journal of Approximate Reasoning**, 2011, doi:10.1016/j.ijar.2011.01.010.

[9] W. Glänzel, H-index concatenation, **Scientometrics** 77(2):369–372, 2008.

[10] M. Grabisch, E. Pap, J.-L. Marichal and R. Mesiar. **Aggregation Functions**, Cambridge, 2009.

[11] J. E. Hirsch, An index to quantify individual's scientific research output, **PNAS** 102(46):16569–16572, 2005.

[12] T. Marchant, An axiomatic characterization of the ranking based on the h-index and some other bibliometric rankings of authors, **Scientometrics** 80(2):325–342, 2009.

[13] T. Marchant, Score-based bibliometric rankings of authors, **Journal of the American Society for Information Science and Technology** 60(6):1132–1137, 2009.

[14] G. Mayor and T. Calvo, On extended aggregation functions. In **Proc. IFSA 1997**, Vol. 1, pages 281–285, Academia, Prague, 1997.

[15] R. Mesiar and A. Mesiárova-Zemánková, The Ordered Modular Averages, **IEEE Transactions on Fuzzy Systems** 19(1):42–50, 2011.

[16] I. Palacios-Huerta and O. Volij, The measurement of intellectual influence, **Econometrica** 72(3):963–977, 2004.

[17] A. Quesada, More axiomatics for the Hirsch index, **Scientometrics** 82:413–418, 2010.

[18] G. J. Woeginger, An axiomatic analysis of Egghe's g-index, **Journal of Informetrics** 2(4):364–368, 2008.

[19] G. J. Woeginger, An axiomatic characterization of the Hirsch-index, **Mathematical Social Sciences** 56(2):224–232, 2008.

[20] R. R. Yager, On ordered weighted averaging aggregation operators in multictriteria decision making, **IEEE Trans. Systems, Man, and Cybernetics** 18(1):183–190, 1988.