

Measuring the Influence of the k th Largest Variable on Functions over the Unit Hypercube

Jean-Luc Marichal and Pierre Mathonet

Mathematics Research Unit, FSTC, University of Luxembourg,
6, rue Coudenhove-Kalergi, L-1359 Luxembourg,
Grand Duchy of Luxembourg.
`jean-luc.marichal[at]uni.lu, pierre.mathonet[at]uni.lu`

Abstract. By considering a least squares approximation of a given square integrable function $f: [0, 1]^n \rightarrow \mathbb{R}$ by a shifted L -statistic function (a shifted linear combination of order statistics), we define an index which measures the global influence of the k th largest variable on f . We show that this influence index has appealing properties and we interpret it as an average value of the difference quotient of f in the direction of the k th largest variable or, under certain natural conditions on f , as an average value of the derivative of f in the direction of the k th largest variable. We also discuss a few applications of this index in statistics and aggregation theory.

1 Introduction

Consider a real-valued function f of n variables x_1, \dots, x_n and suppose we want to measure a global influence degree of every variable x_i on f . A reasonable way to define such an influence degree consists in considering the coefficient of x_i in the best least squares approximation of f by affine functions of the form

$$g(x_1, \dots, x_n) = c_0 + \sum_{i=1}^n c_i x_i.$$

This approach was considered in [6, 10] for pseudo-Boolean functions $f: \{0, 1\}^n \rightarrow \mathbb{R}$ and in [9] for square integrable functions $f: [0, 1]^n \rightarrow \mathbb{R}$. It turns out that, in both cases, the influence index of x_i on f is given by an average “derivative” of f with respect to x_i .

Now, it is also natural to consider and measure a global influence degree of the smallest variable, or the largest variable, or even the k th largest variable for some $k \in \{1, \dots, n\}$. As an application, suppose we are to choose an appropriate aggregation function $f: [0, 1]^n \rightarrow \mathbb{R}$ to compute an average value of $[0, 1]$ -valued grades obtained by a student. If, for instance, we use the arithmetic mean function, we might expect that both the smallest and the largest variables are equally influent. However, if we use the geometric mean function, for which the value 0 (the left endpoint of the scale) is multiplicatively absorbent, we might anticipate that the smallest variable is more influent than the largest one.

Similarly to the previous problem, to define the influence of the k th largest variable on f it is natural to consider the coefficient of $x_{(k)}$ in the best least squares approximation of f by symmetric functions of the form

$$g(x_1, \dots, x_n) = a_0 + \sum_{i=1}^n a_i x_{(i)},$$

where $x_{(1)}, \dots, x_{(n)}$ are the order statistics obtained by rearranging the variables in ascending order of magnitude.

In this paper we solve this problem for square integrable functions $f: [0, 1]^n \rightarrow \mathbb{R}$. More precisely, we completely describe the least squares approximation problem above and derive an explicit expression for the corresponding influence index (§2). We also show that this index has several natural properties, such as linearity and continuity, and we give an interpretation of it as an average value of the difference quotient of f in the direction of the k th largest variable. Under certain natural conditions on f , we also interpret the index as an average value of the derivative of f in the direction of the k th largest variable (§3). We then provide some alternative formulas for the index to possibly simplify its computation (§4) and we consider some examples including the case when f is the Lovász extension of a pseudo-Boolean function (§5). Finally, we discuss a few applications of the index (§6).

We employ the following notation throughout the paper. Let \mathbb{I}^n denote the n -dimensional unit cube $[0, 1]^n$. We denote by $L^2(\mathbb{I}^n)$ the class of square integrable functions $f: \mathbb{I}^n \rightarrow \mathbb{R}$ modulo equality almost everywhere. For any $S \subseteq [n] = \{1, \dots, n\}$, we denote by $\mathbf{1}_S$ the characteristic vector of S in $\{0, 1\}^n$ (with the particular case $\mathbf{0} = \mathbf{1}_\emptyset$).

Recall that if the \mathbb{I} -valued variables x_1, \dots, x_n are rearranged in ascending order of magnitude $x_{(1)} \leq \dots \leq x_{(n)}$, then $x_{(k)}$ is called the *k th order statistic* and the function $\text{os}_k: \mathbb{I}^n \rightarrow \mathbb{R}$, defined as $\text{os}_k(\mathbf{x}) = x_{(k)}$, is the *k th order statistic function*. As a matter of convenience, we also formally define $\text{os}_0 \equiv 0$ and $\text{os}_{n+1} \equiv 1$. To stress on the arity of the function, we can replace the symbols $x_{(k)}$ and os_k with $x_{k:n}$ and $\text{os}_{k:n}$, respectively. For general background on order statistics, see for instance [1, 4].

Finally, we use the lattice notation \wedge and \vee to denote the minimum and maximum functions, respectively.

2 Influence Index for the k th Largest Variable

An *L -statistic* function is a linear combination of the functions $\text{os}_1, \dots, \text{os}_n$. A *shifted L -statistic* function is a constant plus an L -statistic function. Denote by V_L the set of shifted L -statistic functions. Clearly, V_L is spanned by the linearly independent set

$$B = \{\text{os}_1, \dots, \text{os}_n, \text{os}_{n+1}\} \tag{1}$$

and thus is a linear subspace of $L^2(\mathbb{I}^n)$ of dimension $n + 1$. For a given function $f \in L^2(\mathbb{I}^n)$, we define the *best shifted L -statistic approximation* of f as the

function $f_L \in V_L$ that minimizes the distance

$$\|f - g\|^2 = \int_{\mathbb{I}^n} (f(\mathbf{x}) - g(\mathbf{x}))^2 d\mathbf{x}$$

among all $g \in V_L$, where $\|\cdot\|$ is the norm in $L^2(\mathbb{I}^n)$ associated with the inner product $\langle f, g \rangle = \int_{\mathbb{I}^n} f(\mathbf{x})g(\mathbf{x}) d\mathbf{x}$. Using the general theory of Hilbert spaces, we immediately see that the solution of this approximation problem exists and is uniquely determined by the orthogonal projection of f onto V_L . This projection is given by

$$f_L = \sum_{j=1}^{n+1} a_j \text{os}_j, \quad (2)$$

where the coefficients a_j (for $j \in [n+1]$) are characterized by the conditions

$$\langle f - f_L, \text{os}_i \rangle = 0 \quad \text{for all } i \in [n+1]. \quad (3)$$

The coefficient matrix of this system is the square matrix M of order $n+1$ defined by $(M)_{ij} = \langle \text{os}_i, \text{os}_j \rangle$ for all $i, j \in [n+1]$.

Lemma 1. *For every $i, j \in [n+1]$, we have*

$$(M)_{ij} = \frac{\min(i, j)(\max(i, j) + 1)}{(n+1)(n+2)} \quad (4)$$

and

$$\frac{(M^{-1})_{ij}}{(n+1)(n+2)} = \begin{cases} 2, & \text{if } i = j < n+1, \\ \frac{n+1}{n+2}, & \text{if } i = j = n+1, \\ -1, & \text{if } |i - j| = 1, \\ 0, & \text{otherwise.} \end{cases} \quad (5)$$

Recall that the central second difference operator is defined for any real sequence $(z_k)_{k \geq 1}$ as $\delta_k^2 z_k = z_{k+1} - 2z_k + z_{k-1}$. For every $k \in [n]$, define the function $g_k \in L^2(\mathbb{I}^n)$ as

$$g_k = -(n+1)(n+2) \delta_k^2 \text{os}_k. \quad (6)$$

We immediately obtain the following explicit forms for the components of f_L in the basis (1).

Proposition 1. *The best shifted L -statistic approximation f_L of a function $f \in L^2(\mathbb{I}^n)$ is given by (2), where*

$$a_k = \begin{cases} \langle f, g_k \rangle, & \text{if } k \in [n], \\ (n+1)^2 \langle f, 1 \rangle - (n+1)(n+2) \langle f, \text{os}_n \rangle, & \text{if } k = n+1. \end{cases} \quad (7)$$

Now, to measure the global influence of the k th largest variable $x_{(k)}$ on an arbitrary function $f \in L^2(\mathbb{I}^n)$, we naturally define an index $I: L^2(\mathbb{I}^n) \times [n] \rightarrow \mathbb{R}$ as $I(f, k) = a_k$, where a_k is obtained from f by (7). We will see in the next section that this index indeed measures an influence degree.

Definition 1. Let $I: L^2(\mathbb{I}^n) \times [n] \rightarrow \mathbb{R}$ be defined as $I(f, k) = \langle f, g_k \rangle$, that is

$$I(f, k) = -(n+1)(n+2) \int_{\mathbb{I}^n} f(\mathbf{x}) \delta_k^2 x_{(k)} d\mathbf{x}. \quad (8)$$

3 Properties and Interpretations

In this section we present various properties and interpretations of the index $I(f, k)$. The first result follows immediately from Definition 1.

Proposition 2. For every $k \in [n]$, the mapping $f \mapsto I(f, k)$ is linear and continuous.

We now present an interpretation of $I(f, k)$ as a covariance. Considering the unit cube \mathbb{I}^n as a probability space with respect to the Lebesgue measure, we see that, for any $k \in [n]$, the index $I(f, k)$ is the covariance of the random variables f and g_k . Indeed, we have $I(f, k) = E(f g_k) = \text{cov}(f, g_k) + E(f) E(g_k)$, where $E(g_k) = \langle 1, g_k \rangle = I(1, k) = 0$. From the usual interpretation of the concept of covariance, we see that $I(f, k)$ is positive whenever the values of $f - E(f)$ and $g_k - E(g_k) = g_k$ have the same sign. Note that $g_k(\mathbf{x})$ is positive whenever $x_{(k)}$ is greater than $\frac{1}{2}(x_{(k+1)} + x_{(k-1)})$, which is the midpoint of the range of $x_{(k)}$ when the other order statistics are fixed at \mathbf{x} .

We now provide an interpretation of $I(f, k)$ as an expected value of the derivative of f in the direction of the k th largest variable (see Proposition 3).

Let S_n denote the symmetric group on $[n]$. Recall that the unit cube \mathbb{I}^n can be partitioned almost everywhere into the open standard simplexes

$$\mathbb{I}_\pi^n = \{\mathbf{x} \in \mathbb{I}^n : x_{\pi(1)} < \cdots < x_{\pi(n)}\} \quad (\pi \in S_n).$$

Definition 2. Given $k \in [n]$, let $f: \cup_{\pi \in S_n} \mathbb{I}_\pi^n \rightarrow \mathbb{R}$ be a function such that the partial derivative $D_{\pi(k)} f|_{\mathbb{I}_\pi^n}$ exists for every $\pi \in S_n$. The derivative of f in the direction (k) is the function $D_{(k)} f: \cup_{\pi \in S_n} \mathbb{I}_\pi^n \rightarrow \mathbb{R}$ defined as

$$D_{(k)} f(\mathbf{x}) = D_{\pi(k)} f(\mathbf{x}) \quad \text{for all } \mathbf{x} \in \mathbb{I}_\pi^n.$$

Now, for every $k \in [n]$, consider the function $h_k \in L^2(\mathbb{I}^n)$ defined as

$$h_k = (n+1)(n+2)(\text{os}_{k+1} - \text{os}_k)(\text{os}_k - \text{os}_{k-1}).$$

We easily see that h_k is a probability density function on \mathbb{I}^n . This fact can also be derived by choosing $f = \text{os}_k$ in the following result.

Proposition 3. For every $k \in [n]$ and every $f \in L^2(\mathbb{I}^n)$ such that $D_{(k)} f$ is continuous and integrable on $\cup_{\pi \in S_n} \mathbb{I}_\pi^n$, we have

$$I(f, k) = \int_{\mathbb{I}^n} h_k(\mathbf{x}) D_{(k)} f(\mathbf{x}) d\mathbf{x}. \quad (9)$$

We now give an alternative interpretation of $I(f, k)$ as an expected value, which does not require the additional assumptions of Proposition 3. In this more general framework, we naturally replace the derivative with a difference quotient. To this extent, we introduce some further notation. As usual, we denote by \mathbf{e}_i the i th vector of the standard basis for \mathbb{R}^n . For every $k \in [n]$ and every $h \in [0, 1]$, we define the (k) -difference (or *discrete (k) -derivative*) operator $\Delta_{(k),h}$ over the set of real functions on \mathbb{I}^n by

$$\Delta_{(k),h}f(\mathbf{x}) = f(\mathbf{x} + h\mathbf{e}_{\pi(k)}) - f(\mathbf{x})$$

for every $\mathbf{x} \in \mathbb{I}_\pi^n$ such that $\mathbf{x} + h\mathbf{e}_{\pi(k)} \in \mathbb{I}_\pi^n$. Thus defined, the value $\Delta_{(k),h}f(\mathbf{x})$ can be interpreted as the *marginal contribution* of $x_{(k)}$ on f at \mathbf{x} with respect to the increase h . For instance, we have $\Delta_{(k),h}x_{(k)} = h$.

Similarly, we define the (k) -difference quotient operator $Q_{(k),h}$ over the set of real functions on \mathbb{I}^n by $Q_{(k),h}f(\mathbf{x}) = \frac{1}{h}\Delta_{(k),h}f(\mathbf{x})$.

Theorem 1. *For every $k \in [n]$ and every $f \in L^2(\mathbb{I}^n)$, we have*

$$I(f, k) = (n+1)(n+2) \int_{\mathbb{I}^n} \int_{x_{(k)}}^{x_{(k+1)}} \Delta_{(k),y-x_{(k)}}f(\mathbf{x}) dy d\mathbf{x}. \quad (10)$$

As an immediate consequence of Theorem 1, we have the following interpretation of the index $I(f, k)$ as an expected value of a difference quotient with respect to some distribution.

Corollary 1. *For every $k \in [n]$ and every $f \in L^2(\mathbb{I}^n)$, we have*

$$I(f, k) = \int_{\mathbb{I}^n} \int_{x_{(k)}}^{x_{(k+1)}} p_k(\mathbf{x}, y) Q_{(k),y-x_{(k)}}f(\mathbf{x}) dy d\mathbf{x},$$

where $p_k(\mathbf{x}, y) = (n+1)(n+2)(y - x_{(k)})$ defines a probability density function on the set $\{(\mathbf{x}, y) : \mathbf{x} \in \mathbb{I}^n, y \in [x_{(k)}, x_{(k+1)}]\}$.

Another important feature of the index is its invariance under the action of permutations. Recall that a permutation $\pi \in S_n$ acts on a function $f: \mathbb{I}^n \rightarrow \mathbb{R}$ by $\pi(f)(x_1, \dots, x_n) = f(x_{\pi(1)}, \dots, x_{\pi(n)})$. By the change of variables theorem, we immediately see that every $\pi \in S_n$ is an isometry of $L^2(\mathbb{I}^n)$, that is, $\langle \pi(f), \pi(g) \rangle = \langle f, g \rangle$. From this fact, we derive the following result.

Proposition 4. *For every $f \in L^2(\mathbb{I}^n)$ and every $\pi \in S_n$, both functions f and $\pi(f)$ have the same best shifted L -statistic approximation f_L . Moreover, we have $\|\pi(f) - f_L\| = \|f - f_L\|$.*

With any function $f: \mathbb{I}^n \rightarrow \mathbb{R}$ we can associate the following symmetric function

$$\text{Sym}(f) = \frac{1}{n!} \sum_{\pi \in S_n} \pi(f).$$

It follows immediately from Propositions 2 and 4 that both functions f and $\text{Sym}(f)$ have the same best shifted L -statistic approximation f_L . Combining this observation with Proposition 4, we derive immediately the following corollary.

Corollary 2. For every $k \in [n]$, every $f \in L^2(\mathbb{I}^n)$, and every $\pi \in S_n$, we have $I(f, k) = I(\pi(f), k) = I(\text{Sym}(f), k)$.

Remark 1. Corollary 2 shows that, to compute $I(f, k)$, we can replace f with $\text{Sym}(f)$. For instance, if $f(\mathbf{x}) = x_i$ for some $i \in [n]$ then $\text{Sym}(f) = \frac{1}{n} \sum_{i=1}^n x_i = \frac{1}{n} \sum_{i=1}^n x_{(i)}$ and hence, using Proposition 3, we obtain $I(f, k) = \frac{1}{n}$.

Given $k \in [n]$, we say that the order statistic $x_{(k)}$ is *ineffective* almost everywhere for a function $f: \mathbb{I}^n \rightarrow \mathbb{R}$ if $\Delta_{(k), y - x_{(k)}} f(\mathbf{x}) = 0$ for almost all $\mathbf{x} \in \cup_{\pi \in S_n} \mathbb{I}_{\pi}^n$ and almost all $y \in]x_{(k-1)}, x_{(k+1)}[$. For instance, given unary functions $f_1, f_2 \in L^2(\mathbb{I})$, the order statistic $x_{(1)}$ is ineffective almost everywhere for the function $f: \mathbb{I}^2 \rightarrow \mathbb{R}$ such that

$$f(x_1, x_2) = \begin{cases} f_1(x_1), & \text{if } x_1 > x_2, \\ f_2(x_2), & \text{if } x_1 < x_2. \end{cases}$$

The following result immediately follows from Theorem 1.

Proposition 5. Let $k \in [n]$ and $f \in L^2(\mathbb{I}^n)$. If $x_{(k)}$ is ineffective almost everywhere for f , then $I(f, k) = 0$.

The *dual* of a function $f: \mathbb{I}^n \rightarrow \mathbb{R}$ is the function $f^d: \mathbb{I}^n \rightarrow \mathbb{R}$ defined by $f^d(\mathbf{x}) = 1 - f(\mathbf{1}_{[n]} - \mathbf{x})$. A function $f: \mathbb{I}^n \rightarrow \mathbb{R}$ is said to be *self-dual* if $f^d = f$. By using the change of variables theorem, we immediately derive the following result.

Proposition 6. For every $f \in L^2(\mathbb{I}^n)$ and every $k \in [n]$, we have $I(f^d, k) = I(f, n - k + 1)$. In particular, if f is self-dual, then $I(f, k) = I(f, n - k + 1)$.

4 Alternative Expressions for the Index

The computation of the index $I(f, k)$ by means of (8) or (9) might be not very convenient due to the presence of the order statistic functions. To make those integrals either more tractable or easier to evaluate numerically, we provide in this section some alternative expressions for the index $I(f, k)$ that do not involve any order statistic.

We first derive useful formulas for the computation of the integral $\langle f, \text{os}_k \rangle$ (Proposition 7). To this extent, we consider the following direct generalization of order statistic functions.

Definition 3. For every nonempty $S = \{i_1, \dots, i_s\} \subseteq [n]$, $s = |S|$, and every $k \in [s]$, we define the function $\text{os}_{k:S}: \mathbb{I}^n \rightarrow \mathbb{R}$ as $\text{os}_{k:S}(\mathbf{x}) = \text{os}_{k:s}(x_{i_1}, \dots, x_{i_s})$.

To simplify the notation, we will write $x_{k:S}$ for $\text{os}_{k:S}(\mathbf{x})$. Thus $x_{k:S}$ is the k th order statistic of the variables in S .

Lemma 2. For every $s \in [n]$ and every $k \in [s]$, we have

$$\sum_{\substack{S \subseteq [n] \\ |S|=s}} x_{k:S} = \sum_{j=k}^n \binom{j-1}{k-1} \binom{n-j}{s-k} x_{j:n}. \quad (11)$$

Lemma 3. For every $k \in [n]$, we have

$$x_{k:n} = \sum_{\substack{S \subseteq [n] \\ |S| \geq k}} (-1)^{|S|-k} \binom{|S|-1}{k-1} x_{|S|:S} \quad (12)$$

$$x_{k:n} = \sum_{\substack{S \subseteq [n] \\ |S| \geq n-k+1}} (-1)^{|S|-n+k-1} \binom{|S|-1}{n-k} x_{1:S} \quad (13)$$

Combining Lemma 3 with some classical results in measure theory, we compute several expressions of $\int_{\mathbb{I}^n} f(\mathbf{x}) x_{(k)} d\mathbf{x}$:

Proposition 7. For every function $f \in L^2(\mathbb{I}^n)$ and every $k \in [n]$, the integral $J_{k:n} = \int_{\mathbb{I}^n} f(\mathbf{x}) x_{(k)} d\mathbf{x}$ is given by each of the following expressions:

$$\int_{\mathbb{I}^n} f(\mathbf{x}) d\mathbf{x} - \sum_{S \subseteq [n]; |S| \geq k} (-1)^{|S|-k} \binom{|S|-1}{k-1} \int_0^1 \int_{[0,y]^S} \int_{[0,1]^{[n] \setminus S}} f(\mathbf{x}) d\mathbf{x} dy \quad (14)$$

$$\sum_{S \subseteq [n]; |S| \geq n-k+1} (-1)^{|S|-n+k-1} \binom{|S|-1}{n-k} \int_0^1 \int_{[y,1]^S} \int_{[0,1]^{[n] \setminus S}} f(\mathbf{x}) d\mathbf{x} dy \quad (15)$$

$$\int_{\mathbb{I}^n} f(\mathbf{x}) d\mathbf{x} - \sum_{S \subseteq [n]; |S| \geq k} \int_0^1 \int_{[0,y]^S} \int_{[y,1]^{[n] \setminus S}} f(\mathbf{x}) d\mathbf{x} dy \quad (16)$$

$$\sum_{S \subseteq [n]; |S| < k} \int_0^1 \int_{[0,y]^S} \int_{[y,1]^{[n] \setminus S}} f(\mathbf{x}) d\mathbf{x} dy \quad (17)$$

From Definition 1 and Proposition 7, we derive the following expressions for the quantity $\frac{I(f,k)}{(n+1)(n+2)}$:

$$\sum_{S \subseteq [n]; |S| \geq k-1} (-1)^{|S|+1-k} \binom{|S|+1}{k} \int_0^1 \int_{[0,y]^S} \int_{[0,1]^{[n] \setminus S}} f(\mathbf{x}) d\mathbf{x} dy \quad (18)$$

$$\sum_{S \subseteq [n]; |S| \geq n-k} (-1)^{|S|-n+k-1} \binom{|S|+1}{n-k+1} \int_0^1 \int_{[y,1]^S} \int_{[0,1]^{[n] \setminus S}} f(\mathbf{x}) d\mathbf{x} dy \quad (19)$$

$$\left(\sum_{S \subseteq [n]; |S|=k-1} - \sum_{S \subseteq [n]; |S|=k} \right) \int_0^1 \int_{[0,y]^S} \int_{[y,1]^{[n] \setminus S}} f(\mathbf{x}) d\mathbf{x} dy. \quad (20)$$

5 Some Examples

We now apply our results to two special classes of functions, namely the multiplicative functions and the Lovász extensions of pseudo-Boolean functions. The latter class includes the so-called discrete Choquet integrals, well-known in aggregation function theory.

5.1 Multiplicative functions

Consider the function $f(\mathbf{x}) = \prod_{i=1}^n \varphi_i(x_i)$, where $\varphi_i \in L^2(\mathbb{I})$, and set $\Phi_i(x) = \int_0^x \varphi_i(t) dt$ for $i = 1, \dots, n$. By using (18), we obtain

$$\frac{I(f, k)}{(n+1)(n+2)} = \sum_{\substack{S \subseteq [n] \\ |S| \geq k-1}} (-1)^{|S|+1-k} \binom{|S|+1}{k} \prod_{i \in [n] \setminus S} \Phi_i(1) \int_0^1 \prod_{i \in S} \Phi_i(y) dy \quad (21)$$

The following result gives a concise expression for $I(f, k)$ when f is symmetric.

Proposition 8. *Let $f: \mathbb{I}^n \rightarrow \mathbb{R}$ be given by $f(\mathbf{x}) = \prod_{i=1}^n \varphi(x_i)$, where $\varphi \in L^2(\mathbb{I})$, and let $\Phi(x) = \int_0^x \varphi(t) dt$. Then, for every $k \in [n]$, we have*

$$I(f, k) = \begin{cases} \Phi(1)^n \int_0^1 D_z h(z; k+1, n-k+2)|_{z=\Phi(y)/\Phi(1)} dy, & \text{if } \Phi(1) \neq 0, \\ (-1)^{n-k+1} (n+1) \frac{\Gamma(n+3)}{\Gamma(k+1)\Gamma(n-k+2)} \int_0^1 \Phi(y)^n dy, & \text{if } \Phi(1) = 0, \end{cases}$$

where $h(z; a, b) = z^{a-1}(1-z)^{b-1}/B(a, b)$ is the probability density function of the beta distribution with parameters a and b .

Example 1. Let $f: \mathbb{I}^n \rightarrow \mathbb{R}$ be given by $f(\mathbf{x}) = (\prod_{i=1}^n x_i)^c$, where $c > -\frac{1}{2}$. For instance, the product function corresponds to $c = 1$ and the geometric mean function to $c = 1/n$. We can calculate $I(f, k)$ by using Proposition 8 with $\varphi(x) = x^c$. Using the substitution $z = y^{c+1}$ and then integrating by parts, we obtain

$$I(f, k) = c \left(\frac{1}{c+1} \right)^{n+2} \frac{\Gamma(n+3) \Gamma(k-1 + \frac{1}{c+1})}{\Gamma(k+1) \Gamma(n+1 + \frac{1}{c+1})} = \frac{\Gamma(k-1 + \frac{1}{c+1})}{\Gamma(k+1) \Gamma(\frac{1}{c+1})} I(f, 1),$$

with

$$I(f, 1) = c \left(\frac{1}{c+1} \right)^{n+2} \frac{\Gamma(n+3) \Gamma(\frac{1}{c+1})}{\Gamma(n+1 + \frac{1}{c+1})}.$$

We observe that $I(f, k) \rightarrow I(f, 1)$ as $c \rightarrow -\frac{1}{2}$. Also, for $c > 0$, we have $I(f, k+1) < I(f, k)$ for every $k \in [n-1]$. As expected in this case, the smallest variables are more influent on f than the largest ones.

5.2 Lovász extensions

Recall that an n -place (lattice) term function $p: \mathbb{I}^n \rightarrow \mathbb{I}$ is a combination of projections $\mathbf{x} \mapsto x_i$ ($i \in [n]$) using the fundamental lattice operations \wedge and \vee ; see [2]. For instance,

$$p(x_1, x_2, x_3) = (x_1 \wedge x_2) \vee x_3$$

is a 3-place term function. Note that, since \mathbb{I} is a bounded chain, here the lattice operations \wedge and \vee reduce to the minimum and maximum functions, respectively.

Clearly, any shifted linear combination of n -place term functions

$$f(\mathbf{x}) = c_0 + \sum_{i=1}^m c_i p_i(\mathbf{x})$$

is a continuous function whose restriction to any standard simplex \mathbb{I}_π^n ($\pi \in S_n$) is a shifted linear function. According to Singer [11, §2], f is then the *Lovász extension* of the pseudo-Boolean function $f|_{\{0,1\}^n}$, that is, the continuous function $f: \mathbb{I}^n \rightarrow \mathbb{R}$ which is defined on each standard simplex \mathbb{I}_π^n as the unique affine function that coincides with $f|_{\{0,1\}^n}$ at the $n+1$ vertices of \mathbb{I}_π^n . Singer showed that a Lovász extension can always be written as

$$f(\mathbf{x}) = f_{n+1}^\pi + \sum_{i=1}^n (f_i^\pi - f_{i+1}^\pi) x_{\pi(i)} \quad (\mathbf{x} \in \mathbb{I}_\pi^n), \quad (22)$$

with $f_i^\pi = f(\mathbf{1}_{\{\pi(i), \dots, \pi(n)\}}) = v_f(\{\pi(i), \dots, \pi(n)\})$ for $i \in [n+1]$, where the set function $v_f: 2^{[n]} \rightarrow \mathbb{R}$ is defined as $v_f(S) = f(\mathbf{1}_S)$. In particular, $f_{n+1}^\pi = c_0 = f(\mathbf{0})$. Conversely, any continuous function $f: \mathbb{I}^n \rightarrow \mathbb{R}$ that reduces to an affine function on each standard simplex is a shifted linear combination of term functions:

$$f(\mathbf{x}) = \sum_{S \subseteq [n]} m_f(S) x_{1:S}, \quad (23)$$

where $m_f: 2^{[n]} \rightarrow \mathbb{R}$ is the *Möbius transform* of v_f , defined as

$$m_f(S) = \sum_{T \subseteq S} (-1)^{|S|-|T|} v_f(T).$$

Indeed, expression (23) reduces to an affine function on each standard simplex and agrees with $f(\mathbf{1}_S)$ at $\mathbf{1}_S$ for every $S \subseteq [n]$. Thus the class of shifted linear combinations of n -place term functions is precisely the class of n -place Lovász extensions.

Remark 2. A nondecreasing Lovász extension $f: \mathbb{I}^n \rightarrow \mathbb{R}$ such that $f(\mathbf{0}) = 0$ is also called a *discrete Choquet integral*. For general background, see for instance [5].

For every nonempty $S \subseteq [n]$ and every $k \in [|S|]$, the function $os_{k:S}$ is a Lovász extension and, from (12), we have

$$x_{k:S} = \sum_{\substack{T \subseteq S \\ |T| \geq k}} (-1)^{|T|-k} \binom{|T|-1}{k-1} x_{|T|:T}$$

The following proposition gives a concise expression for the index $I(os_{j:S}, k)$. We first compute the action of the symmetrizer Sym on such functions.

Lemma 4. For every nonempty $S \subseteq [n]$ and every $j \in [|S|]$, we have

$$\text{Sym}(\text{os}_{j:S}) = \frac{1}{\binom{n}{|S|}} \sum_{\substack{T \subseteq [n] \\ |T|=|S|}} \text{os}_{j:T}.$$

Proposition 9. For every nonempty $S \subseteq [n]$, every $j \in [|S|]$, and every $k \in [n]$, we have

$$I(\text{os}_{j:S}, k) = \frac{\binom{k-1}{j-1} \binom{n-k}{|S|-j}}{\binom{n}{|S|}} \quad (24)$$

if $0 \leq k - j \leq n - |S|$, and 0, otherwise.

The following proposition gives an explicit expression for the index $I(f, k)$ when f is a Lovász extension.

Proposition 10. If $f: \mathbb{I}^n \rightarrow \mathbb{R}$ is a Lovász extension, then

$$f(\mathbf{x}) = f(\mathbf{0}) + \sum_{i=1}^n x_{(i)} D_{(i)} f(\mathbf{x}). \quad (25)$$

Moreover, for every $k \in [n]$, we have

$$I(f, k) = \bar{v}_f(n - k + 1) - \bar{v}_f(n - k) = \sum_{s=1}^{n-k+1} \binom{n-k}{s-1} \bar{m}_f(s), \quad (26)$$

where $\bar{v}_f(s) = \binom{n}{s}^{-1} \sum_{S \subseteq [n]: |S|=s} v_f(S)$ and $\bar{m}_f(s) = \binom{n}{s}^{-1} \sum_{S \subseteq [n]: |S|=s} m_f(S)$.

We can readily see that the shifted L -statistic functions are precisely the symmetric Lovász extensions. From this observation we derive the following result.

Proposition 11. For any Lovász extension $f: \mathbb{I}^n \rightarrow \mathbb{R}$, we have $f_L = \text{Sym}(f)$ and

$$\text{Sym}(f) = f(\mathbf{0}) + \sum_{i=1}^n I(f, i) \text{os}_i.$$

6 Applications

We briefly discuss some applications of the influence index in aggregation theory and statistics.

6.1 Influence Index in Aggregation Theory

Several indexes (such as interaction, tolerance, and dispersion indexes) have been proposed and investigated in aggregation theory to better understand the general behavior of aggregation functions with respect to their variables; see [5, Chap. 10]. These indexes enable one to classify the aggregation functions according to their behavioral properties. The index $I(f, k)$ can also be very informative and thus contribute to such a classification. As an example, we have computed this index for the arithmetic mean and geometric mean functions (see Remark 1 and Example 1) and we can observe for instance that the smallest variable $x_{(1)}$ has a larger influence on the latter function.

Remark 3. Noteworthy aggregation functions are the so-called conjunctive aggregation functions, that is, nondecreasing functions $f: \mathbb{I}^n \rightarrow \mathbb{R}$ satisfying $0 \leq f(\mathbf{x}) \leq x_{(1)}$; see [5, Chap. 3]. Although these functions are bounded from above by $x_{(1)}$, the index $I(f, k)$ need not be maximum for $k = 1$. For instance, for the binary conjunctive aggregation function

$$f(x_1, x_2) = \begin{cases} 0, & \text{if } x_1 \vee x_2 < \frac{3}{4}, \\ x_1 \wedge x_2 \wedge \frac{1}{4}, & \text{otherwise,} \end{cases}$$

we have $I(f, 1) = \frac{17}{128}$ and $I(f, 2) = \frac{19}{64}$, and hence $I(f, 1) < I(f, 2)$.

In the framework of aggregation functions, it can be natural to consider and identify the functions $f \in L^2(\mathbb{I}^n)$ for which the order statistics are equally influent, that is, such that $I(f, k) = I(f, 1)$ for all $k \in [n]$. As far as the Lovász extensions are concerned, we have the following result, which can be easily derived from Proposition 10 and the immediate identities

$$\bar{v}_f(s) = \sum_{t=0}^s \binom{s}{t} \bar{m}_f(t) \quad \text{and} \quad \bar{m}_f(s) = \sum_{t=0}^s (-1)^{s-t} \binom{s}{t} \bar{v}_f(t).$$

Proposition 12. *If $f: \mathbb{I}^n \rightarrow \mathbb{R}$ is a Lovász extension, then the following are equivalent.*

- (a) *We have $I(f, k) = I(f, 1)$ for all $k \in [n]$.*
- (b) *The sequence $(\bar{v}_f(s))_{s=0}^n$ is in arithmetic progression.*
- (c) *We have $\bar{m}_f(s) = 0$ for $s = 2, \dots, n$.*

6.2 Influence index in statistics

It can be informative to assess the influence of every order statistic on a given statistic to measure, e.g., its behavior with respect to the extreme values. From this information we can also approximate the given statistic by a shifted L -statistic. Of course, for L -statistics (such as Winsorized means, trimmed means, linearly weighted means, quasi-ranges, Gini's mean difference; see [4, §6.3, §8.8, §9.4]), the computation of the influence indexes is immediate. However, for some other statistics such as the central moments, the indexes can be computed via (18)–(20).

Example 2. The closest shifted L -statistic to the variance $\sigma^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$ is given by

$$\sigma_L^2 = \frac{1 - n^2}{12n(n + 3)} + \sum_{k=1}^n I(\sigma^2, k) X_{(k)},$$

with $I(\sigma^2, k) = (n + 2)(2k - n - 1)/(n^2(n + 3))$, which can be computed from (20). We then immediately see that the smallest and largest variables are the most influent.

Acknowledgments

The authors wish to thank Samuel Nicolay for fruitful discussions. This research is supported by the internal research project F1R-MTH-PUL-09MRDO of the University of Luxembourg.

References

1. N. Balakrishnan and C. R. Rao, editors. *Order statistics: theory & methods*, volume 16 of *Handbook of Statist.* North-Holland, Amsterdam, 1998.
2. S. Burris and H. P. Sankappanavar. *A course in universal algebra*, volume 78 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1981.
3. F. N. David and N. L. Johnson. Statistical treatment of censored data. I. Fundamental formulae. *Biometrika*, 41:228–240, 1954.
4. H. David and H. Nagaraja. *Order statistics. 3rd ed.* Wiley Series in Probability and Statistics. Chichester: John Wiley & Sons., 2003.
5. M. Grabisch, J.-L. Marichal, R. Mesiar, and E. Pap. *Aggregation functions*, volume 127 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 2009.
6. P. Hammer and R. Holzman. Approximations of pseudo-Boolean functions; applications to game theory. *Z. Oper. Res.*, 36(1):3–21, 1992.
7. J. Kahn, G. Kalai, and N. Linial. The influence of variables on Boolean functions. In *Proc. 29th Annual Symposium on Foundations of Computational Science*, pages 68–80. Computer Society Press, 1988.
8. J.-L. Marichal. The influence of variables on pseudo-Boolean functions with applications to game theory and multicriteria decision making. *Discrete Appl. Math.*, 107(1-3):139–164, 2000.
9. J.-L. Marichal and P. Mathonet. Measuring the interactions among variables of functions over the unit hypercube. arXiv:0912.1547.
10. J.-L. Marichal and P. Mathonet. Weighted Banzhaf interaction index through weighted approximations of games. arXiv:1001.3052.
11. I. Singer. Extensions of functions of 0-1 variables and applications to combinatorial optimization. *Numer. Funct. Anal. Optimization*, 7:23–62, 1984.