

On the cardinality index of fuzzy measures and the signatures of coherent systems

Jean-Luc Marichal¹ and Pierre Mathonet²

¹ University of Luxembourg, Mathematics Research Unit
6, rue Coudenhove-Kalergi, L-1359 Luxembourg, Luxembourg
jean-luc.marichal@uni.lu

² University of Liège, Department of Mathematics
Grande Traverse, 12, B-4000 Sart-Tilman, Belgium
p.mathonet@ulg.ac.be

Boolean and pseudo-Boolean functions play a central role in various areas of applied mathematics. We will focus here on their use in decision making, cooperative game theory, and engineering reliability theory.

A *discrete fuzzy measure* on the finite set $X = \{1, \dots, n\}$ is a nondecreasing set function $\mu: 2^X \rightarrow [0, 1]$ satisfying the boundary conditions $\mu(\emptyset) = 0$ and $\mu(X) = 1$. For any subset $S \subseteq X$, the number $\mu(S)$ can be interpreted as the certitude that we have that a variable will take on its value in the set $S \subseteq X$.

A *cooperative game* on a finite set of players $N = \{1, \dots, n\}$ is a set function $v: 2^N \rightarrow \mathbb{R}$ which assigns to each coalition S of players a real number $v(S)$. This number represents the *worth* of S . (Even though the condition $v(\emptyset) = 0$ is often required for v to define a game, here we do not need this restriction.)

A *system* is defined by a finite set of components $C = \{1, \dots, n\}$ that are interconnected according to a certain structure. The components are either in function or in a failed state, and the same holds for the whole system. It is common to associate the Boolean value 0 with a failed state and the value 1 with a component that is in function. Therefore the *structure function* of a system is the function ϕ from 2^C to $\mathbb{B} = \{0, 1\}$ which associates with any set A of components that are in function the corresponding state of the system. The system is semicoherent if the structure function is nondecreasing and satisfies the conditions $\phi(\emptyset) = 0$ and $\phi(C) = 1$. It is coherent if in addition all the components are essential.

We identify any subset S of $\{1, \dots, n\}$ with its characteristic vector $\mathbf{1}_S \in \{0, 1\}^n$ (defined by $(\mathbf{1}_S)_k = 1$ if and only if k is in S). This identification allows us to identify set functions and pseudo-Boolean functions, i.e., functions from \mathbb{B}^n to \mathbb{R} . Therefore discrete fuzzy measures, cooperative games, and structure functions of coherent systems are all described by pseudo-Boolean functions.

The use of discrete fuzzy measures allows us to model real situations where additivity is not suitable since the set of such measures is richer than the set of classical additive measures. In the same way, cooperative games allow us to take into account possible interactions between the players and need not be additive. Finally, the set of all increasing pseudo-Boolean functions is necessary to describe all the possible semicoherent systems. However, the variousness of this set of functions also has the drawback that a general set function (fuzzy measure, cooperative game, or structure function of a system) might be difficult to interpret or analyze.

Various kinds of power indexes, or values, are used in cooperative game theory to overcome this problem. They measure the influence that a given player has on the outcome of the game or define a way of sharing the benefits of the game among the players. The best known values, due to Shapley [10] and Banzhaf [1], are defined in the following way. The Shapley value of player k in a game v on the set of players $[n] = \{1, \dots, n\}$ is defined by

$$\phi_{\text{Sh}}(v, k) = \sum_{S \subseteq [n] \setminus \{k\}} \frac{(n-s-1)!s!}{n!} (v(S \cup \{k\}) - v(S)), \quad (1)$$

while the Banzhaf value is given by

$$\phi_B(v, k) = \frac{1}{2^{n-1}} \sum_{S \subseteq [n] \setminus \{k\}} (v(S \cup \{k\}) - v(S)) = \frac{1}{2^{n-1}} \sum_{S \ni k} v(S) - \frac{1}{2^{n-1}} \sum_{S \not\ni k} v(S). \quad (2)$$

There are several axiomatic characterizations these values. They are also used to analyze fuzzy measures and were generalized by the concepts of Shapley or Banzhaf *interaction indexes*; see, e.g., [5].

In reliability theory of coherent systems, the importance of component k for system S can also be measured in various ways. Assuming that the components of the system have continuous i.i.d. lifetimes T_1, \dots, T_n , Barlow and Proschan [2] introduced in 1975 the n -tuple \mathbf{I}_{BP} (the Barlow-Proschan index) whose k th coordinate ($k \in [n]$) is the probability that the failure of component k causes the system to fail; that is,

$$I_{\text{BP}}^{(k)} = \Pr(T_S = T_k),$$

where T_S denotes the system lifetime. It turns out that for continuous i.i.d. component lifetimes, this index reduces to the Shapley value of the system structure function.

In this note we consider slightly different importance indexes that do not measure the influence of a given variable over a function but rather the influence of adding a variable to a given subset of variables. These indexes are the *cardinality index* introduced in 2002 by Yager [11] in the context of fuzzy measures and the *signature* of coherent systems introduced in 1985 by Samaniego [8, 9].

The cardinality index associated with a fuzzy measure μ on $X = \{1, \dots, n\}$ is the n -tuple (C_0, \dots, C_{n-1}) , where C_k is the average gain in certitude that we obtain by adding an arbitrary element to an arbitrary k -element subset, that is,

$$C_k = \frac{1}{(n-k) \binom{n}{k}} \sum_{|S|=k} \sum_{x \notin S} (\mu(S \cup \{x\}) - \mu(S)).$$

We observe that this expression, which resembles the Banzhaf value (2), could be used in cooperative game theory to measure the marginal contribution of an additional player to a k -element coalition. It is also clear that this index can be written as

$$C_k = \frac{1}{\binom{n}{k+1}} \sum_{|S|=k+1} \mu(S) - \frac{1}{\binom{n}{k}} \sum_{|S|=k} \mu(S). \quad (3)$$

The signature of a *system* consisting of n interconnected components having continuous and i.i.d. lifetimes T_1, \dots, T_n is defined as the n -tuple $(s_1, \dots, s_n) \in [0, 1]^n$ with $s_k = \Pr(T_S = T_{(k)})$, where T_S denotes the system lifetime and $T_{(k)}$ is the k th order statistic derived from T_1, \dots, T_n , i.e., the k th smallest lifetime. Thus, s_k is the probability that the k th failure causes the system to fail. It was proved [3] that

$$s_k = \frac{1}{\binom{n}{n-k+1}} \sum_{|\mathbf{x}|=n-k+1} \phi(\mathbf{x}) - \frac{1}{\binom{n}{n-k}} \sum_{|\mathbf{x}|=n-k} \phi(\mathbf{x}), \quad (4)$$

where $\phi: \mathbb{B}^n \rightarrow \mathbb{B}$ is the structure function of the system.

Clearly, Equations (3) and (4) show that for a given pseudo-Boolean function, the cardinality index and the signature are related by the formula $s_k = C_{n-k}$ for $1 \leq k \leq n$.

We now show in detail some properties of the cardinality index and signatures that are similar to properties of the Banzhaf and Shapley values. First we show how this index can be computed from the multilinear extension of the pseudo-Boolean function (or set function) under consideration.

Recall that any set function $f: 2^{[n]} \rightarrow \mathbb{R}$ can be represented in a unique way as a multilinear polynomial. This representation is given by

$$f(x_1, \dots, x_n) = \sum_{S \subseteq [n]} f(S) \prod_{i \in S} x_i \prod_{i \notin S} (1 - x_i), \quad x_i \in \mathbb{B}.$$

The multilinear extension \hat{f} of f is then given by the same polynomial expression but for variables in $[0, 1]$. For a function $g: [0, 1]^n \rightarrow \mathbb{R}$, we simply denote by $g(x)$ the polynomial function $g(x, \dots, x)$. The Banzhaf and Shapley indexes for f can be computed easily from the multilinear extension of f . In fact, the Shapley index of player k in f is given by

$$\phi_{\text{Sh}}(f, k) = \int_0^1 \left(\frac{\partial}{\partial x_k} \hat{f} \right) (x) dx.$$

The Banzhaf index of player k in f is given by $\phi_{\text{B}}(f, k) = \left(\frac{\partial}{\partial x_k} \hat{f} \right) \left(\frac{1}{2} \right)$.

It is possible to obtain a similar formula for the *tail signature*, or the *cumulative cardinality index* that we now introduce. These are the $(n+1)$ -tuples $\bar{\mathbf{S}} = (\bar{S}_0, \dots, \bar{S}_n)$ and $\bar{\mathbf{C}} = (\bar{C}_0, \dots, \bar{C}_n)$, respectively, defined by (see (4) and (3))

$$\bar{S}_k = \sum_{i=k+1}^n s_i = \frac{1}{\binom{n}{n-k}} \sum_{|A|=n-k} \phi(A) = \bar{C}_{n-k}. \quad (5)$$

With any pseudo-Boolean function f on $\{0, 1\}^n$, we associate the polynomial function p_f defined by $p_f(x) = x^n \hat{f}(1/x)$.

Theorem 1. *The cumulative cardinality index and the tail signature are obtain from p_f by*

$$p_f(x) = \sum_{k=0}^n \binom{n}{k} \bar{C}_{n-k} (x-1)^k = \sum_{k=0}^n \binom{n}{k} \bar{S}_k (x-1)^k.$$

The Banzhaf and Shapley values can be obtained via (weighted) least squares approximations of the pseudo-Boolean function by a pseudo-Boolean function of degree 1; see, e.g., [4, 6]. A similar property holds for the cardinality index and signature : they can be obtained via least squares approximations of the given pseudo-Boolean function by a symmetric pseudo-Boolean function. Recall that for $k \in \{1, \dots, n\}$ the k th order statistic function is the function $os_k: \mathbb{B}^n \rightarrow \mathbb{B}$, defined by the condition $os_k(\mathbf{x}) = 1$, if $|\mathbf{x}| = \sum_{i=1}^n x_i \geq n - k + 1$, and 0, otherwise. We also formally define $os_{n+1} \equiv 1$. Then it can be showed that the space of symmetric pseudo-Boolean functions is spanned by the order statistic functions. The *best symmetric approximation* of a pseudo-Boolean function f is the unique symmetric pseudo-Boolean function f_S that minimizes the weighted squared distance

$$\|f - g\|^2 = \sum_{\mathbf{x} \in \mathbb{B}^n} \frac{1}{\binom{n}{|\mathbf{x}|}} (f(\mathbf{x}) - g(\mathbf{x}))^2$$

from among all symmetric functions pseudo-Boolean functions g .

Theorem 2. *The best symmetric approximation of a pseudo-Boolean function f such that $f(0, \dots, 0) = 0$ is given by*

$$f_S = \sum_{k=1}^n s_k os_k = \sum_{k=1}^n C_{n-k} os_k \quad (6)$$

where s_k and C_k are derived from f by using (4) and (3), respectively.

References

1. J. Banzhaf. Weighted voting doesn't work : A mathematical analysis. *Rutgers Law Review*, 19:317–343, 1965.
2. R.E. Barlow and F. Proschan. Importance of system components and fault tree events. *Stochastic Processes and Their Applications*, 3:153–172, 1975.
3. Ph. J. Boland. Signatures of indirect majority systems. *J. Appl. Prob.*, 38:597–603, 2001.
4. A. Charnes, B. Golany, M. Keane, and J. Rousseau. Extremal principle solutions of games in characteristic function form: core, Chebychev and Shapley value generalizations. In *Economics of planning and efficiency*, volume 11 of *Adv. Stud. Theoret. Appl. Econometrics*, pages 123–133. Kluwer Acad. Publ., Dordrecht, 1988.
5. M. Grabisch, J.-L. Marichal, and M. Roubens. Equivalent representations of set functions. *Math. Oper. Res.*, 25(2):157–178, 2000.
6. P. L. Hammer and R. Holzman. Approximations of pseudo-Boolean functions; applications to game theory. *Z. Oper. Res.*, 36(1):3–21, 1992.
7. G. Owen. Multilinear extensions of games. *Management Sci.*, 18:P64–P79, 1971/72.
8. F.J. Samaniego. On closure of the IFR class under formation of coherent systems. *IEEE Trans. Reliability*, 34:69–72, 1985.
9. F.J. Samaniego. *System signatures and their applications in engineering reliability*. Int. Series in Operations Research & Management Science, 110. New York: Springer, 2007.
10. L. S. Shapley. A value for n -person games. In *Contributions to the theory of games, vol. 2*, Annals of Mathematics Studies, no. 28, pages 307–317. Princeton University Press, Princeton, N. J., 1953.
11. R. Yager. On the cardinality index and attitudinal character of fuzzy measures *Int. J. Gen. Syst.*, 31(3):303–329, 2002.