Entropy of bi-capacities

Ivan Kojadinovic
LINA CNRS FRE 2729
Site école polytechnique de l'univ. de Nantes
Rue Christian Pauc
44306 Nantes, France
`ivan.kojadinovic@univ-nantes.fr`

Jean-Luc Marichal
Applied Mathematics Unit
Univ. of Luxembourg
162A, avenue de la Faïencerie
L-1511 Luxembourg, G.D. Luxembourg
`jean-luc.marichal@uni.lu`

Abstract

The notion of entropy, recently generalized to capacities, is extended to bi-capacities and its main properties are studied.

Keywords: Multicriteria decision making, bi-capacity, Choquet integral, entropy.

1 Introduction

The well-known Shannon entropy [10] is a fundamental concept in probability theory and related fields. In a general non probabilistic setting, it is merely a measure of the uniformity (evenness) of a discrete probability distribution. In a probabilistic context, it can be naturally interpreted as a measure of unpredictability.

By relaxing the additivity property of probability measures, requiring only that they be monotone, one obtains Choquet capacities [1], also known as fuzzy measures [11], for which an extension of the Shannon entropy was recently defined [4, 5, 7, 8].

The concept of capacity can be further generalized. In the context of multicriteria decision making, *bi-capacities* have been recently introduced by Grabisch and Labreuche [2, 3] to model in a flexible way the preferences of a decision maker when the underlying scales are *bipolar*.

Since a bi-capacity can be regarded as a generalization of a capacity, the following natural question arises : how could one appraise the 'uniformity' or 'uncertainty' associated with a bi-capacity in the spirit of the Shannon entropy?

The main purpose of this paper is to propose a definition of an extension of the Shannon entropy to bi-capacities. The interpretation of this concept will be performed in the framework of multicriteria decision making based on the Choquet integral. Hence, we consider a set $N := \{1, \ldots, n\}$ of *criteria* and a set $\mathcal{A}$ of *alternatives* described according to these criteria, i.e., real-valued functions on N. Then, given an alternative $x \in \mathcal{A}$, for any $i \in N$, $x_i := x(i)$ is regarded as the *utility* of x w.r.t. to criterion i. The utilities are further considered to be *commensurate* and to lie either on a unipolar or on a *bipolar* scale. Compared to a unipolar scale, a biploar scale is characterized by the additional presence of a neutral value (usually 0) such that values above this neutral reference point are considered to be good by the decision maker, and values below it are considered to be bad. As in [2, 3], for simplicity reasons, we shall assume that the scale used for all utilities is $[0, 1]$ if the scale is unipolar, and $[-1, 1]$ with 0 as neutral value, if the scale is bipolar.

This paper is organized as follows. The second and third sections are devoted to a presentation of the notions of capacity, bi-capacity and Choquet integral in the framework of multicriteria decision making. In the last section, after recalling the definitions of the probabilistic Shannon entropy and of its extension to capacities, we propose a generalization of it to bi-capacities. We also give an interpretation of it in the context of multicriteria decision making and we study its main properties.

2 Capacities and bi-capacities

In the context of aggregation, capacities [1] and bi-capacities [2, 3] can be regarded as generalizations of weighting vectors involved in the calculations of weighted arithmetic means.

Let $\mathcal{P}(N)$ denote the power set of N and let $\mathcal{Q}(N) := \{(A,B) \in \mathcal{P}(N) \times \mathcal{P}(N) | A \cap B = \emptyset\}$.

Definition 2.1 *A function $\mu : \mathcal{P}(N) \to [0,1]$ is a capacity if it satisfies :*

(i) $\mu(\emptyset) = 0$, $\mu(N) = 1$,

(ii) for any $S, T \subseteq N$, $S \subseteq T \Rightarrow \mu(S) \leq \mu(T)$.

A capacity μ on N is said to be *additive* if $\mu(S \cup T) = \mu(S) + \mu(T)$ for all disjoint subsets $S, T \subseteq N$. A particular case of additive capacity is the *uniform* capacity on N. It is defined by

$$\mu^*(T) = |T|/n, \qquad \forall T \subseteq N.$$

The *dual* (or *conjugate*) of a capacity μ on N is a capacity $\bar{\mu}$ on N defined by $\bar{\mu}(A) = \mu(N) - \mu(N \setminus A)$, for all $A \subseteq N$.

Definition 2.2 *A function $v : \mathcal{Q}(N) \to \mathbb{R}$ is a bi-capacity if it satisfies :*

(i) $v(\emptyset,\emptyset) = 0$, $v(N,\emptyset) = 1$, $v(\emptyset, N) = -1$,

(ii) $A \subseteq B$ implies $v(A,\cdot) \leq v(B,\cdot)$ and $v(\cdot,A) \geq v(\cdot,B)$.

Furthermore, a bi-capacity v is said to be :

- of the *Cumulative Prospect Theory (CPT) type* [2, 3, 12] if there exist two capacities μ_1, μ_2 such that

  $$v(A,B) = \mu_1(A) - \mu_2(B), \qquad \forall (A,B) \in \mathcal{Q}(N).$$

  When $\mu_1 = \mu_2$ the bi-capacity is further said to be *symmetric*, and *asymmetric* when $\mu_2 = \bar{\mu}_1$

- *additive* if it is of the CPT type with μ_1, μ_2 additive, i.e. for any $(A,B) \in \mathcal{Q}(N)$

  $$v(A,B) = \sum_{i \in A} \mu_1(i) - \sum_{i \in B} \mu_2(i).$$

  Note that an additive bi-capacity with $\mu_1 = \mu_2$ is both symmetric and asymmetric since $\bar{\mu}_1 = \mu_1$.

As we continue, to indicate that a CPT type bi-capacity v is constructed from two capacities μ_1, μ_2, we shall denote it by v_{μ_1,μ_2}

Let us also consider a particular additive bi-capacity on N : the *uniform* bi-capacity. It is defined by

$$v^*(A,B) = \frac{|A| - |B|}{n}, \qquad \forall (A,B) \in \mathcal{Q}(N).$$

3 The Choquet integral

When utilities are considered to lie on a unipolar scale, the importance of the subsets of (interacting) criteria can be modeled by a capacity. A suitable aggregation operator that generalizes the weighted arithmetic mean is then the Choquet integral [6].

Definition 3.1 *The Choquet integral of a function $x : N \to \mathbb{R}^+$ represented by the profile $(x_1, \ldots, x_n)$ w.r.t a capacity μ on N is defined by*

$$C_\mu(x) := \sum_{i=1}^{n} x_{\sigma(i)} [\mu(A_{\sigma(i)}) - \mu(A_{\sigma(i+1)})],$$

where σ is a permutation on N such that $x_{\sigma(1)} \leq \cdots \leq x_{\sigma(n)}$, $A_{\sigma(i)} := \{\sigma(i), \ldots, \sigma(n)\}$, for all $i \in \{1, \ldots, n\}$, and $A_{\sigma(n+1)} := \emptyset$.

When the underlying utility scale is bipolar, Grabisch and Labreuche proposed to substitute a bi-capacity to the capacity and proposed a natural generalization of the Choquet integral [3].

Definition 3.2 *The Choquet integral of a function $x : N \to \mathbb{R}$ represented by the profile $(x_1, \ldots, x_n)$ w.r.t a bi-capacity v on N is defined by*

$$C_v(x) := C_{\nu^v_{N^+}}(|x|)$$

where $\nu^v_{N^+}$ is a game on N (i.e. a set function on N vanishing at the empty set) defined by

$$\nu^v_{N^+}(C) = v(C \cap N^+, C \cap N^-), \qquad \forall C \subseteq N,$$

and $N^+ := \{i \in N | x_i \geq 0\}$, $N^- := N \setminus N^+$.

As shown in [3], an equivalent expression of $C_v(x)$ is :

$$C_v(x) = \sum_{i \in N} |x_{\sigma(i)}| \Big[v(A_{\sigma(i)} \cap N^+, A_{\sigma(i)} \cap N^-) - v(A_{\sigma(i+1)} \cap N^+, A_{\sigma(i+1)} \cap N^-) \Big], \quad (1)$$

where $A_{\sigma(i)} := \{\sigma(i), \ldots, \sigma(n)\}$, $A_{\sigma(n+1)} := 0$, and σ is a permutation on N so that $|x_{\sigma(1)}| \leq \cdots \leq |x_{\sigma(n)}|$.

4 Entropy of a bi-capacity

4.1 The concept of probabilistic entropy

The fundamental concept of *entropy of a probability distribution* was initially proposed by Shannon [9, 10]. The Shannon entropy of a probability distribution p defined on a nonempty finite set $N := \{1, \ldots, n\}$ is defined by

$$H_S(p) := \sum_{i \in N} h[p(i)]$$

where

$$h(x) := \begin{cases} -x \ln x, & \text{if } x > 0, \\ 0, & \text{if } x = 0, \end{cases}$$

The quantity $H_S(p)$ is always non negative and zero if and only if p is a Dirac mass (*decisivity* property). As a function of p, H_S is strictly concave. Furthermore, it reaches its maximum value $(\ln n)$ if and only if p is uniform (*maximality* property).

In a general non probabilistic setting, $H_S(p)$ is nothing else than a measure of the uniformity of p. In a probabilistic context, it can be interpreted as a measure of the *information* contained in p.

4.2 Extension to capacities

Let μ be a capacity on N. The following entropy was proposed by Marichal [5, 7] (see also [8]) as an extension of the Shannon entropy to capacities :

$$H_M(\mu) := \sum_{i \in N} \sum_{S \subseteq N \setminus i} \gamma_s(n) h[\mu(S \cup i) - \mu(S)].$$

Regarded as a uniformity measure, H_M has been recently axiomatized by means of three axioms [4] : the symmetry property, a boundary condition for which H_M reduces to the Shannon entropy, and a generalized version of the well-known recursivity property.

A fundamental property of H_M is that it can be rewritten in terms of the maximal chains of the Hasse diagram of N [4], which is equivalent to :

$$H_M(\mu) = \frac{1}{n!} \sum_{\sigma \in \Pi_N} H_S(p_\sigma^\mu), \quad (2)$$

where Π_N denotes the set of permutations on N and, for any $\sigma \in \Pi_N$,

$$p_\sigma^\mu(i) := \mu(\{\sigma(i), \ldots, \sigma(n)\}) - \mu(\{\sigma(i+1), \ldots, \sigma(n)\}), \qquad \forall i \in N.$$

The quantity $H_M(\mu)$ can therefore simply be seen as an average over Π_N of the uniformity values of the probability distributions p_σ^μ calculated by means of the Shannon entropy. As shown in [4], in the context of aggregation by a Choquet integral w.r.t a capacity μ on N, $H_M(\mu)$ can be interpreted as a measure of the average value over all $x \in [0,1]^n$ of the degree to which the arguments $x_1, \ldots, x_n$ contribute to the calculation of the aggregated value $C_\mu(x)$.

To stress on the fact that H_M is an average of Shannon entropies, we shall equivalently denote it by $\overline{H}_S$ as we go on.

It has also been shown that $H_M = \overline{H}_S$ satisfies many properties that one would intuitively require from an entropy measure [4, 7]. The most important ones are :

1. **Boundary property for additive measures.** For any additive capacity μ on N, we have
$$\overline{H}_S(\mu) = H_S(p),$$
where p is the probability distribution on N defined by $p(i) = \mu(i)$ for all $i \in N$.

2. **Boundary property for cardinality-based measures.** For any cardinality-based capacity μ on N (i.e. such that, for any $T \subseteq N$, $\mu(T)$ depends only on $|T|$), we have
$$\overline{H}_S(\mu) = H_S(p^\mu),$$

where p^μ is the probability distribution on N defined by $p^\mu(i) = \mu(\{1,\dots,i\}) - \mu(\{1,\dots,i-1\})$ for all $i \in N$.

3. **Decisivity.** For any capacity μ on N,
$$\overline{H}_S(\mu) \geq 0.$$
Moreover, $\overline{H}_S(\mu) = 0$ if and only if μ is a binary-valued capacity, that is, such that $\mu(T) \in \{0,1\}$ for all $T \subseteq N$.

4. **Maximality.** For any capacity μ on N, we have
$$\overline{H}_S(\mu) \leq \ln n.$$
with equality if and only if μ is the uniform capacity μ^* on N.

5. **Increasing monotonicity toward $\boldsymbol{\mu^*}$.** Let μ be a capacity on N such that $\mu \neq \mu^*$ and, for any $\lambda \in [0,1]$, define the capacity μ_λ on N as $\mu_\lambda := \mu + \lambda(\mu^*_N - \mu)$. Then for any $0 \leq \lambda_1 < \lambda_2 \leq 1$ we have
$$\overline{H}_S(\mu_{\lambda_1}) < \overline{H}_S(\mu_{\lambda_2}).$$

6. **Strict concavity.** For any two capacities μ_1, μ_2 on N and any $\lambda \in]0,1[$, we have
$$\overline{H}_S(\lambda\,\mu_1 + (1-\lambda)\,\mu_2) > \lambda\,\overline{H}_S(\mu_1) + (1-\lambda)\,\overline{H}_S(\mu_2).$$

4.3 Generalization to bi-capacities

For any bi-capacity v on N and any $N^+ \subseteq N$, as in [3], we define the game $\nu^v_{N^+}$ on N by
$$\nu^v_{N^+}(C) := v(C \cap N^+, C \cap N^-), \qquad \forall C \subseteq N,$$
where $N^- := N \setminus N^+$.

Furthermore, for any $N^+ \subseteq N$, let p^v_{σ,N^+} be the probability distribution on N defined, for any $i \in N$, by
$$p^v_{\sigma,N^+}(i) := \frac{|\nu^v_{N^+}(A_{\sigma(i)}) - \nu^v_{N^+}(A_{\sigma(i+1)})|}{\sum_{j\in N} |\nu^v_{N^+}(A_{\sigma(j)}) - \nu^v_{N^+}(A_{\sigma(j+1)})|} \tag{3}$$
where $A_{\sigma(i)} := \{\sigma(i),\dots,\sigma(n)\}$, for all $i \in N$, and $A_{\sigma(n+1)} := \emptyset$

We then propose the following simple definition of the extension of the Shannon entropy to a bi-capacity v on N :
$$\overline{\overline{H}}_S(v) := \frac{1}{2^n} \sum_{N^+\subseteq N} \frac{1}{n!} \sum_{\sigma\in\Pi_N} H_S(p^v_{\sigma,N^+}) \tag{4}$$

As in the case of capacities, the extended Shannon entropy $\overline{\overline{H}}_S(v)$ is nothing else than an average of the uniformity values of the probability distributions p^v_{σ,N^+} calculated by means of H_S.

In the context of aggregation by a Choquet integral w.r.t a bi-capacity v on N, let us show that, as previously, $\overline{\overline{H}}_S(v)$ can be interpreted as a measure of the average value over all $x \in [-1,1]^n$ of the degree to which the arguments $x_1,\dots,x_n$ contribute to the calculation of the aggregated value $C_v(x)$.

In order to do so, consider an alternative $x \in [-1,1]^n$ and denote by $N^+ \subseteq N$ the subset of criteria for which $x \geq 0$. Then, from Eq. (1), we see that the Choquet integral of x w.r.t v is simply a weighted sum of $|x_{\sigma(1)}|,\dots,|x_{\sigma(n)}|$, where each $|x_{\sigma(i)}|$ is weighted by
$$\nu^v_{N^+}(A_{\sigma(i)}) - \nu^v_{N^+}(A_{\sigma(i+1)}).$$

Clearly, these weights are not always positive, nor do they sum up to one. From the monotonicity conditions of a bi-capacity, it follows that the weight corresponding to $|x_\sigma(i)|$ is positive if and only if $\sigma(i) \in N^+$.

Depending on the evenness of the distribution of the absolute values of the weights, the utilities $x_1,\dots,x_n$ will contribute more or less evenly in the calculation of $C_v(x)$.

A straightforward way to measure the evenness of the contribution of $x_1,\dots,x_n$ to $C_v(x)$ consists in measuring the uniformity of the probability distribution p^v_{σ,N^+} defined by Eq. (3). Note that p^v_{σ,N^+} is simply obtained by normalizing the distribution of the absolute values of the weights involved in the calculation of $C_v(x)$.

Clearly, the uniformity of p^v_{σ,N^+} can be measured by the Shannon entropy. Should $H_S(p^v_{\sigma,N^+})$ be close to $\ln n$, the distribution p^v_{σ,N^+} will be approximately uniform and all the partial evaluations $x_1,\dots,x_n$ will be involved almost equally in the calculation of $C_v(x)$. On the contrary, should $H_S(p^v_{\sigma,N^+})$ be close to zero, one $p^v_{\sigma,N^+}(i)$ will be very close to one and $C_v(x)$ will be almost proportional to the corresponding partial evaluation.

Let us now go back to the definition of the extended Shannon entropy. From Eq. (4), we clearly see that $\overline{\overline{\overline{H}}}_S(v)$ is nothing else than a measure of the average of the behavior we have just discussed, i.e. taking into account all the possibilities for σ and N^+ with uniform probability. More formally, for any $N^+ \subseteq N$, and any $\sigma \in \Pi_N$, define the set

$$\begin{aligned}\mathcal{O}_{\sigma,N^+} := \{x \in [-1,1]^n \mid \forall i \in N^+, x_i \in [0,1], \\ \forall i \in N^-, x_i \in [-1,0[, |x_{\sigma(1)}| \leq \cdots \leq |x_{\sigma(n)}|\}.\end{aligned}$$

We clearly have $\bigcup_{N^+ \subseteq N} \bigcup_{\sigma \in \Pi_N} \mathcal{O}_{\sigma,N^+} = [-1,1]^n$.

Let $x \in [-1,1]^n$ be fixed. Then there exist $N^+ \subseteq N$ and $\sigma \in \Pi_N$ such that $x \in \mathcal{O}_{\sigma,N^+}$ and hence $C_v(x)$ is proportional to $\sum_{i\in N} x_{\sigma(i)}\, p^v_{\sigma,N^+}(i)$.

Starting from Eq. (4) and using the fact that $\int_{x\in\mathcal{O}_{\sigma,N^+}} \mathrm{d}x = 1/n!$, the entropy $\overline{\overline{\overline{H}}}_S(v)$ can be rewritten as

$$\begin{aligned}H_M(\mu) &= \frac{1}{2^n} \sum_{N^+ \subseteq N} \sum_{\sigma \in \Pi_N} \int_{x\in\mathcal{O}_{\sigma,N^+}} H_S(p^v_{\sigma,N^+})\,\mathrm{d}x \\ &= \frac{1}{2^n} \int_{[-1,1]^n} H_S(p^v_{\sigma_x,N^+_x})\,\mathrm{d}x,\end{aligned}$$

where $N^+_x \subseteq N$ and $\sigma_x \in \Pi_N$ are defined such that $x \in \mathcal{O}_{\sigma_x,N^+_x}$.

We thus observe that $\overline{\overline{\overline{H}}}_S(v)$ measures the average value over all $x \in [-1,1]^n$ of the degree to which the arguments $x_1, \ldots, x_n$ contribute to the calculation of $C_v(x)$. In probabilistic terms, it corresponds to the expectation over all $x \in [-1,1]^n$, with uniform distribution, of the degree of contribution of arguments $x_1, \ldots, x_n$ in the calculation of $C_v(x)$.

4.4 Properties of $\overline{\overline{\overline{H}}}_S$

We first present two lemmas giving the form the probability distributions p^v_{σ,N^+} for CPT type bi-capacities.

Lemma 4.1 *For any bi-capacity v_{μ_1,μ_2} of the CPT type on N, any $N^+ \subseteq N$, and any $\sigma \in \Pi_N$, we have*

$$\begin{aligned}p^{v_{\mu_1,\mu_2}}_{\sigma,N^+}(i) = \Big[&\mu_1(A_{\sigma(i)} \cap N^+) - \mu_1(A_{\sigma(i+1)} \cap N^+) \\ &+\mu_2(A_{\sigma(i)} \cap N^-) - \mu_2(A_{\sigma(i+1)} \cap N^-)\Big] \\ &/ \left[\mu_1(N^+) + \mu_2(N^-)\right], \qquad \forall i \in N.\end{aligned}$$

Lemma 4.2 *For any CPT type asymmetric bi-capacity v_{μ_1,μ_2} on N, any $N^+ \subseteq N$, and any $\sigma \in \Pi_N$, we have*

$$\begin{aligned}p^{v_{\mu_1,\mu_2}}_{\sigma,N^+}(i) = \mu_1(A_{\sigma(i)} \cap N^+) - \mu_1(A_{\sigma(i+1)} \cap N^+) \\ + \bar{\mu}_1(A_{\sigma(i)} \cap N^-) - \bar{\mu}_1(A_{\sigma(i+1)} \cap N^-),\end{aligned}$$

for all $i \in N$.

We now state four important properties of $\overline{\overline{\overline{H}}}_S$.

Property 4.1 (Additive bi-capacity) *For any additive bi-capacity v_{μ_1,μ_2} on N, $\overline{\overline{\overline{H}}}_S(v_{\mu_1,\mu_2})$ equals*

$$\frac{1}{2^n} \sum_{N^+ \subseteq N} \sum_{i\in N} h\left[\frac{\mu_1(i \cap N^+) + \mu_2(i \cap N^-)}{\sum_{j\in N^+}\mu_1(j) + \sum_{j\in N^-}\mu_2(j)}\right]$$

Proof. Let v_{μ_1,μ_2} be an additive bi-capacity on N. Then, using Lemma 4.1, for any $N^+ \subseteq N$, any $\sigma \in \Pi_N$, any $i \in N$, we obtain that

$$\begin{aligned}|\nu^{v_{\mu_1,\mu_2}}_{N^+}(A_{\sigma(i)}) - \nu^{v_{\mu_1,\mu_2}}_{N^+}(A_{\sigma(i+1)})| \\ = \mu_1(\sigma(i) \cap N^+) + \mu_2(\sigma(i) \cap N^-).\end{aligned}$$

It follows that, for any $N^+ \subseteq N$,

$$H_S(p^{v_{\mu_1,\mu_2}}_{\sigma,N^+}) = \sum_{i\in N} h\left[\frac{\mu_1(i \cap N^+) + \mu_2(i \cap N^-)}{\sum_{j\in N^+}\mu_1(j) + \sum_{j\in N^-}\mu_2(j)}\right],$$

for all $\sigma \in \Pi_N$, from which we get the desired result. □

Property 4.2 (Add. sym./asym. bi-capacity) *For any additive asymmetric/symmetric bi-capacity v_{μ_1,μ_2} on N,*

$$\overline{\overline{\overline{H}}}_S(v_{\mu_1,\mu_2}) = H_S(p),$$

where p is the probability distribution on N defined by $p(i) := \mu_1(i)$ for all $i \in N$.

Proof. The result follows from Property 4.1. □

Property 4.3 (Decisivity) *For any bi-capacity v on N,*

$$\overline{\overline{\overline{H}}}_S(v) \geq 0.$$

Moreover, $\overline{\overline{\overline{H}}}_S(v) = 0$ if and only, for any $x \in [-1,1]^n$, only one partial evaluation is used in the calculation of $C_v(x)$.

Proof. From the decisivity property satisfied by the Shannon entropy, we have that, for any probability distribution p on N, $H_S(p) \geq 0$ with equality if and only if p is Dirac.

Let v be a bi-capacity on N. If follows that $\overline{\overline{H}}_S(v) \geq 0$ with equality if and only if, for any $N^+ \subseteq N$, any $\sigma \in \Pi_N$, p^v_{σ,N^+} is Dirac, which is clearly equivalent to having, for any $x \in [-1,1]^n$, only one partial evaluation contributing in the calculation of $C_v(x)$. □

Property 4.4 (Maximality) *For any bi-capacity v on N, we have*

$$\overline{\overline{H}}_S(v) \leq \ln n.$$

with equality if and only if v is the uniform capacity v^* on N.

Proof. From the maximality property satisfied by the Shannon entropy, we have that, for any probability distribution p on N, $H_S(p) \leq \ln n$ with equality if and only if p is uniform.

Let v be a bi-capacity on N. It follows that $\overline{\overline{H}}_S(v) \leq \ln n$ with equality if and only if, for any $N^+ \subseteq N$, and any $\sigma \in \Pi_N$, p^v_{σ,N^+} is uniform.

It is easy to see that if $v = v^*$, then $\overline{\overline{H}}_S(v) = \ln n$. Let us show that if $\overline{\overline{H}}_S(v) = \ln n$, then necessarily $v = v^*$.

To do so, consider first the case where $N^+ \in \{\emptyset, N\}$. From the normalization condition $v(N,\emptyset) = 1 = -v(\emptyset, N)$, it easy to verify that, for any $\sigma \in \Pi_N$,

$$\sum_{j \in N} |\nu^v_{N^+}(A_{\sigma(j)}) - \nu^v_{N^+}(A_{\sigma(j+1)})| = 1.$$

It follows that, if, for any $\sigma \in \Pi_N$, p^v_{σ,N^+} is uniform, then, for any $\sigma \in \Pi_N$,

$$|\nu^v_{N^+}(A_{\sigma(i)}) - \nu^v_{N^+}(A_{\sigma(i+1)})| = \frac{1}{n}, \qquad \forall i \in N.$$

This implies that,

$$v(i, \emptyset) = \frac{1}{n} = -v(\emptyset, i), \qquad \forall i \in N. \tag{5}$$

Consider now the case where $N^+ \in 2^N \setminus \{\emptyset, N\}$. If $\overline{\overline{H}}_S(v) = \ln n$, we know that, for any $\sigma \in \Pi_N$, p^v_{σ,N^+} is uniform. From Eq. (5), we have that, for any $\sigma \in \Pi_N$,

$$|\nu^v_{N^+}(A_{\sigma(n)}) - \nu^v_{N^+}(A_{\sigma(n+1)})| = \frac{1}{n}.$$

Since, for any $\sigma \in \Pi_N$, p^v_{σ,N^+} is uniform, we obtain that

$$|\nu^v_{N^+}(A_{\sigma(i)}) - \nu^v_{N^+}(A_{\sigma(i+1)})| = \frac{1}{n}, \qquad \forall i \in N.$$

□

References

[1] G. Choquet. Theory of capacities. *Annales de l'Institut Fourier*, 5:131–295, 1953.

[2] M. Grabisch and C. Labreuche. Bi-capacities I : definition, Möbius transform and interaction. *Fuzzy Sets and Systems*, 151:211–236, 2005.

[3] M. Grabisch and C. Labreuche. Bi-capacities II : the Choquet integral. *Fuzzy Sets and Systems*, 151:237–259, 2005.

[4] I. Kojadinovic, J.-L. Marichal, and M. Roubens. An axiomatic approach to the definition of the entropy of a discrete Choquet capacity. *Information Sciences*, 2005. In press.

[5] J.-L. Marichal. *Aggregation operators for multicriteria decison aid.* PhD thesis, University of Liège, Liège, Belgium, 1998.

[6] J.-L. Marichal. An axiomatic approach of the discrete Choquet integral as a tool to aggregate interacting criteria. *IEEE Transactions on Fuzzy Systems*, 8(6):800–807, 2000.

[7] J.-L. Marichal. Entropy of discrete Choquet capacities. *European Journal of Operational Research*, 3(137):612–624, 2002.

[8] J.-L. Marichal and M. Roubens. Entropy of discrete fuzzy measures. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 8(6):625–640, 2000.

[9] C. Shannon and W. Weaver. *A mathematical theory of communication.* University of Illinois, Urbana, 1949.

[10] C. E. Shannon. A mathematical theory of communication. *Bell Systems Technical Journal*, 27:379–623, 1948.

[11] M. Sugeno. *Theory of fuzzy integrals and its applications.* PhD thesis, Tokyo Institute of Technology, Tokyo, Japan, 1974.

[12] A. Tversky and D. Kahneman. Advances in prospect theory : cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5:297–323, 1992.