ENTROPY OF DISCRETE FUZZY MEASURES

J.-L. Marichal
M. Roubens

University of Liège
Belgium

Douz, April 2000

Suppose we have n states of nature.

How can we define the uncertainty related to these states ?

Classical view :

probability distribution :

$p(i) \geq 0$, $i \in \{1, \dots, n\} = N$

$$\sum_i p(i) = 1$$

Fuzzy measure (set function)

$v(S)$, $S \subset N$
$v(\emptyset) = 0$, $v(N) = 1$

$$S \subset T \Rightarrow v(S) \leq v(T)$$

Particular fuzzy measures

- additive (probabilistic)

  $v(S \cup T) = v(S) + v(T)$ when $S \cap T = \emptyset$

  $\Rightarrow \exists p(i)$ such that $v(S) = \sum_{i \in S} p(i)$

- cardinality-based

  $|S| = |T| \Rightarrow v(S) = v(T)$

  $\Rightarrow \exists : 0 = c_0 \leq c_1 \leq \cdots \leq c_n = 1$ with $v(S) = c_{|S|}$.

- binary

  $v(S) \in \{0, 1\}$

- Dirac

  $$v(S) = \begin{cases} 1 \text{ iff } S \ni i \\ 0 \text{ otherwise} \end{cases}$$

To summarize $v(S)$ in terms of a probabilistic measure, one can define the power indices called

Shapley values $Sh(i)$

$$Sh(i) = \sum_{T \subset n \backslash i} \gamma_t \delta_i v(T \cup i), \quad t = |t|$$

$\delta_i v(T \cup i) = v(T \cup i) - v(T)$: contribution of i when joining the coalition T

$$\gamma_t = \frac{(n-t-1)!t!}{n!}, \text{ Shapley coefficients}$$

The real power of i, $Sh(i)$ is such that

$$\sum_{i \in N} Sh(i) = 1$$

$v(S)$	$Sh(i)$
additive	$p(i)$
cardinality-based	$\frac{1}{n}$
Dirac	$\exists i : Sh(i) = 1$
	$\forall j \neq i : Sh(j) = 0$

Classical measure of uncertainty : Shannon entropy

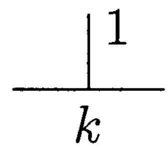
probabilistic measure : $\{p(i)\}_{i \in N}$

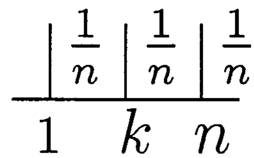
Shannon entropy : $H(p)$

$$H(p) = -\sum_i p(i) \ln p(i) = \sum_i h(p(i))$$

with $h(x) = -x \ln x$

$$H(p) = \begin{cases} 0 \text{ iff } p \text{ is deterministic} \\ \ln n \text{ iff } p \text{ is uniform} \end{cases}$$





$$0 \leq H(p) \leq \ln n$$

Extension to v : $H(v)$

Such that if v is probabilistic

$H(v)$ corresponds to the Shannon entropy

First proposal :

$$\begin{aligned} H_u(v) &= H(Sh), \qquad \text{Yager (1999)} \\ &= -\sum_i Sh(i) \ln Sh(i) \\ &= \sum_i h[Sh(i)] \quad (h(x) = -x \ln x) \\ &= \sum_i h\left[\sum_{T \subset N \setminus i} \gamma_t \delta_i v(T \cup i)\right] \end{aligned}$$

$$0 \leq H_u(v) \leq \ln n$$

We have the Jensen inequality

$$h\left(\sum_k \lambda_k t_k\right) \geq \sum_k \lambda_k h(t_k)$$

if h is strictly concave and $\sum_k \lambda_k = 1$.

$\Rightarrow$

$$\sum_i h[\sum_{T \subset N \setminus i} \gamma_t \delta_i v(T \cup i)] \geq \sum_i \sum_{T \subset N \setminus i} \gamma_t h[\delta_i v(T \cup i)]$$

$\uparrow$ $H_u(v)$ $\qquad$ $\uparrow$ $H_\ell(v)$ second proposal

$$H_\ell(v) = -\sum_{i\in N}\sum_{T\subset N\setminus i} \gamma_t \delta_i v(T\cup i) \ln \delta_i v(T\cup i)$$

$$0 \le H_\ell(v) \le \ln n$$

$$\underset{\text{lower entropy}}{H_\ell(v)} \quad \le \quad \underset{\text{upper entropy(Yager 99)}}{H_u(v)}$$

$H_\ell(v) = H_u(v)$ iff v is additive (probabilistic)

non-additivity $\Rightarrow H_\ell(u) < H_u(v)$

What to choose between H_ℓ and H_u ?

Depends on limit characterizations.

We remember that

$$\begin{array}{rcl} H(p) & = & 0 \quad \text{iff } p \text{ is deterministic} \\ H(p) & = & \ln n \quad \text{iff } p \text{ is uniform.} \end{array}$$

Let us extend these definitions

$$H(v) = 0 \text{ iff} \begin{cases} v \text{ is a binary fuzzy measure} \\ \Uparrow \\ v \text{ is a Dirac fuzzy measure} \end{cases}$$

$$H(v) = \ln n \text{ iff} \begin{cases} v(S) = \frac{|S|}{n} \\ \Downarrow \\ v \text{ is cardinality-based} \\ \Downarrow \\ Sh(i) = \frac{1}{n} (*) \end{cases}$$

(*) does not characterize the v's !

v's such that

$$Sh(i) = \frac{1}{n}$$

have been characterized by Marichal (1998)

$$v(S) = \frac{|S|}{n} + \sum_{\substack{T \subset N \\ |T| \geq 2}} \left[\sum_{j=1}^{|T \cap S|} \binom{|T \cap S|}{j} B_{|T|-j} \right] c(T)$$

where $c(T)$ are reals satisfying constraints

$$\frac{1}{n} + \sum_{\substack{T \ni i \\ |T| \geq 2}} \left[\sum_{j=0}^{|T \cap S|} \binom{|T \cap S|}{j} B_{|T|-j-1} \right] c(T) \geq 0$$

$\{B_n\}_{\in N_0}$ are the Bernoulli numbers :

$$B_0 = 1$$

$$\sum_{k=0}^{n} \binom{n+1}{k} B_k = 0, \quad n \in N_0$$

Cardinality-based fuzzy measure

$H_\ell(v) = \ln n$ $(\Rightarrow H_u(v) = \ln n)$	$\Leftrightarrow$	$v(S) = \frac{\lvert S\rvert}{n}$
		$\Downarrow$
$H_\ell(v)$: Shannon with $w_{n-i} = c_{i+1} - c_i$ and $(H_u(v) = \ln n)$	$\Leftarrow$	v : card.based fuzzy measure $(v(S) = c_{\lvert S\rvert})$
		$\Downarrow$
$H_u(v) = \ln n$	$\Leftrightarrow$	$Sh(i) = \frac{1}{n}$

About Choquet integrals

$$C(x_1, \ldots, x_u), \qquad x_{(1)} \leq x_{(2)} \leq \cdots \leq x_{(x)}$$

$$\begin{aligned} &= \sum_i x_{(i)}[v((i), \ldots, (n)) - v((i+1), \ldots, (n))] \\ &= \sum_i x_{(i)} \delta_i v((i), \ldots, (n)) \end{aligned}$$

OWA operator is a particular Choquet integral where

$v(S) = c_{|S|}$, cardinality-based fuzzy measure

Let $w_{n-i} = c_{i+1} - c_i$

$$OWA(x_1, \ldots, x_n) = w_1 x_{(i)} + \cdots + w_n x_{(n)}$$

$$\begin{cases} H_\ell(w) = H(w) = -\sum_i w_i \ln w_i \text{ (Shannon)} \\ H_u(w) = \ln n \end{cases}$$

H_ℓ : degree to which the aggregator uses the arguments.

$$C(x_1, \dots, x_n) \in \{x_1, \dots, x_n\}$$
$$\Leftrightarrow \quad v \text{ is binary} \Leftrightarrow H_\ell(v) = 0$$

Operators	H_ℓ	H_u (Yager 99)	
x_k	0	0	
$\frac{1}{n}\sum_i x_i$	$\ln n$	$\ln n$	add. measure
$\sum_i w_i x_i$	$-\sum_i w_i \ln w_i$	$-\sum_i w_i \ln w_i$	add. measure
$x_{(k)}$	0	$\ln n$	card. based
$\min_i x_i$	0	$\ln n$	card. based
$\max_i x_i$	0	$\ln n$	card. based
$\sum_i w_i x_{(i)}$	$\underbrace{-\sum_i w_i \ln w_i}$ as proposed by Yager in 1988	$\ln n$	card. based

Properties common to H_ℓ and H_u

Symmetry

$$H_\ell(\pi v) = H_\ell(v) \quad , \quad H_u(\pi v) = H_u(v)$$

π is a permutation of $N = \{1, \ldots, n\}$.

Expansibility

For Shannon,

$$H(p_1, \ldots, p_{n-1}, 0) = H(p_1, \ldots, p_{n-1}).$$

Let us consider a null element for $v : \{i\}$

$$v(T \cup i) = v(T) \text{ for all } T \subset N \setminus \{i\}$$

v_{-i} : restriction of v to $N \setminus \{i\}$.

We have

$$H_\ell(v) = H_\ell(v_{-i}) \quad , \quad H_u(v) = H_u(v_{-i})$$

Ordinal fuzzy measures and entropy

Suppose v is defined on an ordinal scale L

$$L : \{\ell_1, \dots, \ell_m\}$$

$$H_L(v) = \ell_{|R|-1}, \qquad R : \{v(S) \mid S \subset N\}$$

H_L is a measure of diversity of the coefficients of the fuzzy measure (extension of the ordinal entropy defined by Yager, 1999).

Properties

Symmetry : $H_L(v) = H_L(\pi v)$

Expansibility : $H_L(v) = H_L(v_{-i})$ if $\{i\}$ is a null element of v

$$\ell_1 \leq H_L(v) \leq \ell_K \qquad k = \min(2^n, m) - 1$$

$H_L(v) = \ell_1$ (min. index on the L scale)

iff $v(S) \in \{\ell_1, \ell_m\}$

$H_L(v) = \ell_m$ (max. index on the L scale)

iff $R = \{v(S), \mid S \subset N\} = L$

(If $m \geq 2^n$, all v's are distinct).

If one consider the Sugeno integral :

$$S(x_1, \dots, x_n) = \bigvee_{i=1}^{n} [x_{(i)} \wedge v(A_{(i)})]$$

$$A_i := \{(i), \dots, (n)\}$$

$$v(S) \in \{\ell_1, \ell_m\} \Leftrightarrow S(x_1, \dots, x_n) \in \{x_1, \dots, x_n\}$$

$H(v)$ is useful in MCDM

Consider the following MC problem :

	M	φ	L
a	18	16	10
b	10	12	18
c	14	15	15

- Students good in M and L should be favoured or φ and L

- M and φ give the same information about the profile of a student

- M and φ are more important than L

$$\Rightarrow \begin{cases} p(M) = p(\varphi) > p(L) \\ c > a > b \end{cases}$$

If weighted mean is used :

$$W(x) = x(M)p(M) + x(\varphi)p(\varphi) + x(L)p(L)$$

We end with a contradiction

$c > a$ and $(p(M) = p(\varphi))$

$$\Rightarrow \quad (14+15)p(M) + 15p(L) > (18+16)p(M) + 10p(L)$$
$$\Rightarrow \quad p(L) > p(M) \text{ !!!}$$

Use of Choquet integral can help

If

$$\begin{cases} v(M,\varphi) = .5 < v(M) + v(\varphi) = .9 & \text{: redondancy} \\ v(M,L) = v(\varphi,L) = .9 > v(M) + v(L) = .75 : & \text{: synergy} \\ v(M) = v(\varphi) = .45 & \\ v(L) = .3 & \end{cases}$$

	M	φ	L	Choquet	Weighted mean
a	18	16	10	13.9	15.25
b	10	12	18	13.6	12.75
c	14	15	15	14.6	14.625
				$\underline{c > a > b}$	$\underline{a > c > b}$

(with $p(M) = p(\varphi) = 3/8$,

$p(L) = 2/8$)

However :

$$\frac{v(M)}{v(L)} = \frac{p(M)}{p(L)} = \frac{3}{2} \quad !$$

$H_\ell(v) = .82$ v rather well distributed over the total capacity.

H_ℓ can be used as a (non-linear) objective to determine the v's.

v's ?

$$\max H_\ell(v)$$

under given constraints

$$\begin{cases} v(M,\varphi) > v(M) + v(\varphi) \\ v(M,L) < v(M) + v(L) \\ \quad\vdots \end{cases}$$

Re. paper by Marichal and Roubens

"Determination of weights of interacting criteria from a reference set"

EJOR, 00, to appear.