'Useful' fuzzy measures of information, integrated ambiguity and directed divergence
D. S. Hooda a; Rakesh K. Bajaj b
a Jaypee University of Engineering and Technology, Guna, India b Jaypee University of Information Technology, Waknaghat, India

First published on: 15 June 2010

To cite this Article Hooda, D. S. and Bajaj, Rakesh K.(2010) 'Useful' fuzzy measures of information, integrated ambiguity and directed divergence', International Journal of General Systems, 39: 6, 647 — 658, First published on: 15 June 2010 (iFirst)
To link to this Article DOI: 10.1080/03081079.2010.495467
URL: http://dx.doi.org/10.1080/03081079.2010.495467

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: http://www.informaworld.com/terms-and-conditions-of-access.pdf
This article may be used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.
‘Useful’ fuzzy measures of information, integrated ambiguity and directed divergence

D.S. Hooda\textsuperscript{a} and Rakesh K. Bajaj\textsuperscript{b}\textsuperscript{*}

\textsuperscript{a}Jaypee University of Engineering and Technology, Guna, India; \textsuperscript{b}Jaypee University of Information Technology, Waknaghat, India

(Received 7 July 2008; final version received 8 April 2010)

In the present paper, a new concept of ‘useful’ fuzzy information, based on utility, is introduced by considering the uncertainties of fuzziness and probabilities of random events. A ‘useful’ fuzzy measure of integrated ambiguity is obtained by integration of fuzzy and probabilistic uncertainties with utility. A new ‘useful’ measure of fuzzy-directed divergence of a fuzzy set from another fuzzy set is proposed and its validity proved. Finally, the constrained optimisation of ‘useful’ fuzzy entropy and ‘useful’ fuzzy-directed divergence is studied.

Keywords: fuzzy entropy; ‘useful’ fuzzy information; total fuzzy information and constrained optimisation

AMS Subject Classification: 94D05; 94A15

1. Introduction

Fuzzy set theory (Zadeh 1965) makes use of entropy to measure the degree of fuzziness in a fuzzy set, which is also called fuzzy entropy (Ebanks 1983, Pal and Bezdek 1994). Fuzzy entropy is the measurement of fuzziness in a fuzzy set, and thus has an important position in fuzzy systems, such as fuzzy pattern recognition systems, fuzzy neural network systems, fuzzy knowledge base systems, fuzzy decision making systems, fuzzy control systems and fuzzy management information systems.

Let $S$ be any set. A fuzzy subset $A$ in $S$ is characterised by a membership function $\mu_A: S \rightarrow [0, 1]$. The value $\mu_A(x)$ represents the grade of membership of $x \in S$ in $A$. The membership function may be described as follows:

$$
\mu_A(x) = \begin{cases} 
0, & \text{if } x \notin A \text{ and there is no ambiguity,} \\
1, & \text{if } x \in A \text{ and there is no ambiguity,} \\
0.5, & \text{if there is maximum ambiguity whether } x \in A \text{ or } x \notin A.
\end{cases}
$$

Let $X$ be a discrete random variable with its realisation in the set $X = (x_1, x_2, \ldots, x_n)$ with probabilities $P = (p_1, p_2, \ldots, p_n)$; $0 < p_i \leq 1$, in an experiment. We call $(X, P)$ a discrete probabilistic framework. The information contained in this experiment is given by the

*Corresponding author. Email: rakesh.bajaj@gmail.com

ISSN 0308-1079 print/ISSN 1563-5104 online
© 2010 Taylor & Francis
DOI: 10.1080/03081079.2010.495467
http://www.informaworld.com
well-known Shannon entropy (1948):

\[ H(P) = -\sum_{i=1}^{n} p_i \log p_i. \]  

(1)

It may be noted that the meaning of fuzzy entropy is quite different from the classical Shannon entropy because no probabilistic concept is needed in order to define it. Fuzzy entropy deals with vagueness and ambiguous uncertainties, whereas Shannon entropy deals with probabilistic uncertainties. De Luca and Termini (1972) characterised the fuzzy entropy and introduced a set of following properties (P1–P4) for which fuzzy entropy should satisfy them:

- **Sharpness** (P1): \( H(A) \) is minimum iff \( A \) is a crisp set, i.e. \( \mu_A(x) = 0 \) or \( 1 \); for all \( x \).
- **Maximality** (P2): \( H(A) \) is maximum iff \( A \) is most fuzzy set, i.e. \( \mu_A(x) = 0.5 \); for all \( x \).
- **Resolution** (P3): \( H(A) \geq H(A^*) \), where \( A^* \) is a sharpened version of \( A \), i.e.
  
  (i) if \( \mu_A(x_i) \leq 0.5 \) then \( \mu_A^*(x_i) \leq \mu_A(x_i) \) and
  
  (ii) if \( \mu_A(x_i) \geq 0.5 \) then \( \mu_A^*(x_i) \geq \mu_A(x_i) \).
- **Symmetry** (P4): \( H(A) = H(\overline{A}) \), where \( \overline{A} \) is the complement of \( A \), i.e. \( \mu_{\overline{A}}(x_i) = 1 - \mu_A(x_i) \); for all \( i \).

The properties of fuzzy entropy are widely accepted and have become a criterion for defining any new fuzzy entropy.

Corresponding to entropy due to Shannon (1948), De Luca and Termini (1972) suggested the following measure of fuzzy entropy:

\[ H(A) = -\sum_{i=1}^{n} [\mu_A(x_i) \log \mu_A(x_i) + (1 - \mu_A(x_i)) \log (1 - \mu_A(x_i))]. \]  

(2)

As (2) satisfies all four properties (P1) to (P4), it is a valid measure of fuzzy entropy. Two fuzzy sets \( A \) and \( B \) are said to be fuzzy-equivalent if \( \mu_B(x_i) = \) either \( \mu_A(x_i) \) or \( 1 - \mu_A(x_i) \) for each value of \( i \). It is clear that fuzzy-equivalent sets have the same entropy but two sets may have the same fuzzy entropy without being fuzzy equivalent. From the fuzziness point of view, there is no essential difference between fuzzy equivalent sets, i.e. \( \mu_A(x) \) and \( 1 - \mu_A(x) \) give the same degree of fuzziness. A **standard fuzzy set** is that member of the class of fuzzy equivalent sets all of whose membership values are less than or equal to 0.5.

Later on Bhandari and Pal (1993) undertook a survey on information measures on fuzzy sets and gave some measures of fuzzy entropy. Corresponding to Renyi’s (1961) entropy, they suggested the following measure:

\[ H_\alpha(A) = \frac{1}{1 - \alpha} \sum_{i=1}^{n} \log [\mu_A^\alpha(x_i) + (1 - \mu_A(x_i))^{\alpha}]; \quad \alpha \neq 1, \alpha > 0. \]  

(3)

Pal and Pal’s (1989) exponential entropy introduced

\[ H_e(A) = \frac{1}{n \sqrt{e}} - \frac{1}{1} \sum_{i=1}^{n} \log [\mu_A(x_i) e^{1 - \mu_A(x_i)} + (1 - \mu_A(x_i)) e^{\mu_A(x_i)} - 1]. \]  

(4)

It may be noted that the Shannon entropy does not take into account the effectiveness or importance of the events, whereas in some practical situations of probabilistic nature, subjective considerations also play their own role. Belis and Guiasu (1968) considered the
qualitative aspect of information and attached a utility distribution \( U = (u_1, u_2, \ldots, u_n) \) with \( u_i > 0 \) for each \( i \), where \( u_i \) is utility or importance of an event \( x_i \) whose probability of occurrence is \( p_i \). In general, \( u_i \) is independent of \( p_i \). They suggested that the occurrence of an event removes two types of uncertainties: the quantitative type related to its probability of occurrence, and the qualitative type related to its utility (importance) for the fulfillment of some goal set by the experimenter. Bhaker and Hooda (1993) gave the generalised mean value characterisation of the useful information measures for incomplete probability distributions:

\[
H(P; U) = - \frac{\sum_{i=1}^{n} u_i p_i \log p_i}{\sum_{i=1}^{n} u_i}, \quad u_i > 0, \tag{5}
\]

and

\[
H_a(P; U) = \frac{1}{1 - \alpha} \log \frac{\sum_{i=1}^{n} u_i p_i^\alpha}{\sum_{i=1}^{n} u_i p_i^{\alpha}}; \quad \alpha \neq 1, \alpha > 0. \tag{6}
\]

The first attempt to quantify the uncertainty associated with a fuzzy event in the context of a discrete probabilistic framework appears to have been made by Zadeh (1968), who defined the (weighted) entropy of \( A \) with respect to \( (X, P) \) as

\[
H(A, P) = - \sum_{i=1}^{n} \mu_A(x_i) p_i \log p_i, \tag{7}
\]

where \( \mu_A \) is the membership function of \( A \) and \( p_i \) is the probability of occurrence of \( x_i \). It may be noted that the situation contains different types of uncertainties, e.g. randomness, ambiguity and vagueness; i.e. randomness and fuzziness. This measure does not satisfy properties (P1) to (P4), nor was it meant to. \( H(A, P) \) of a fuzzy event with respect to \( P \) is less than Shannon’s entropy which is of \( P \) alone.

In Section 2 of the present paper, we introduce a new concept of ‘useful’ fuzzy information measure, based on utility, by incorporating the uncertainties of fuzziness and probabilities of randomness having utilities. In Section 3, a new measure of total ‘useful’ fuzzy information, by considering the usefulness of an event along with fuzzy uncertainties and random uncertainties, is defined. In Section 4, we define a ‘useful’ measure of fuzzy-directed divergence of fuzzy set \( A \) from fuzzy set \( B \) and also prove its validity. Constrained optimisation of ‘useful’ fuzzy entropy and ‘useful’ fuzzy-directed divergence is discussed in Sections 5 and 6, respectively.

2. ‘Useful’ fuzzy information measures

Let \( X = \{x_1, x_2, \ldots, x_n\} \) be a universe of discourse with probability distribution \( P = \{(p_1, p_2, \ldots, p_n) \) with \( 0 < p_i \leq 1 \) for all \( i = 1, 2, \ldots, n \) \) and utility distribution \( U = \{(u_1, u_2, \ldots, u_n) \) with \( u_i > 0 \) for all \( i = 1, 2, \ldots, n \). Let \( A \) be a fuzzy set defined on \( X \) having membership values \( \mu_A(x_i), i = 1, 2, \ldots, n \). Then, \( \mu_A(x_1), \mu_A(x_2), \ldots, \mu_A(x_n) \) are ambiguities or uncertainties which lie between 0 and 1, but these are not probabilities because their sum is not unity.

As discussed earlier, \( \mu_A(x) \) and \( 1 - \mu_A(x) \) give the same degree of fuzziness, therefore incorporating altogether the uncertainties of fuzziness and probabilities of randomness having utilities. We suggest the following measure of fuzziness of fuzzy set analogous to
De Luca and Termini’s fuzzy entropy (2):

\[
H(A; P; U) = - \frac{\sum_{i=1}^{n} u_ip_i[\mu_A(x_i)\log \mu_A(x_i) + (1 - \mu_A(x_i))\log(1 - \mu_A(x_i))]}{\sum_{i=1}^{n} u_i p_i},
\]

\[
u_i > 0 \ \forall \ i,
\]

which can be called ‘useful’ fuzzy entropy of fuzzy set \(A\).

Theorem 1. The measure (8) satisfies the four properties (P1) to (P4) which are necessary for the validity of a measure of fuzziness in fuzzy set \(A\).

Proof. (P1) Sharpness:
If \(H(A; P; U) = 0\), then

\[
\sum_{i=1}^{n} u_ip_i[\mu_A(x_i)\log \mu_A(x_i) + (1 - \mu_A(x_i))\log(1 - \mu_A(x_i))] = 0,
\]

i.e. \(\mu_A(x_i)\log \mu_A(x_i) + (1 - \mu_A(x_i))\log(1 - \mu_A(x_i)) = 0\) for all \(i\),

i.e. \(\mu_A(x_i) = 0\) or \(1\) for all \(i = 1, 2, \ldots, n\),

i.e. \(A\) is a crisp set.

Conversely, let \(A\) be a crisp set, then either \(\mu_A(x_i) = 0\) or \(\mu_A(x_i) = 1\) for all \(i\).

i.e. \(\mu_A(x_i)\log \mu_A(x_i) + (1 - \mu_A(x_i))\log(1 - \mu_A(x_i)) = 0\) for all \(i\)

i.e. \(\sum_{i=1}^{n} u_ip_i[\mu_A(x_i)\log \mu_A(x_i) + (1 - \mu_A(x_i))\log(1 - \mu_A(x_i))] = 0, [u_i, p_i > 0 \ for \ all \ i]\)

i.e. \(H(A; P; U) = 0\).

Hence, \(H(A; P; U) = 0\) if and only if \(A\) is a non-fuzzy set or crisp set.

(P2) Maximality:
Differentiating \(H(A; P; U)\) with respect to \(\mu_A(x_i)\), we have

\[
\frac{\partial H(A; P; U)}{\partial \mu_A(x_i)} = \sum_{i=1}^{n} u_ip_i \log \frac{1 - \mu_A(x_i)}{\mu_A(x_i)}.
\]

Case 1. \(0 < \mu_A(x_i) < 0.5\).
In this case, \(\log(1 - \mu_A(x_i))/\mu_A(x_i)) > 0\) which gives \((\partial H(A; P; U))/\partial \mu_A(x_i)) > 0\). Therefore, \(H(A; P; U)\) is an increasing function of \(\mu_A(x_i)\) satisfying \(0 < \mu_A(x_i) < 0.5\).

Case 2. \(0.5 < \mu_A(x_i) < 1\).
In this case, \(\log(1 - \mu_A(x_i))/\mu_A(x_i)) < 0\) which gives \((\partial H(A; P; U))/\partial \mu_A(x_i)) < 0\). Therefore, \(H(A; P; U)\) is a decreasing function of \(\mu_A(x_i)\) satisfying \(0.5 < \mu_A(x_i) < 1\).
Also from (9), it may be noted that
\[
\frac{\partial H(A; P; U)}{\partial \mu_A(x_i)} = 0; \text{ when } \mu_A(x_i) = 0.5.
\] (10)

Hence, \( H(A; P; U) \) is a concave function and it has a global maximum at \( \mu_A(x_i) = 0.5 \) which shows that \( H(A; P; U) \) is maximum if and only if \( A \) is the most fuzzy set.

\textbf{(P3) Resolution:}

Since \( H(A; P; U) \) is an increasing function of \( \mu_A(x_i) \) in \([0, 0.5)\) and decreasing function in \((0.5, 1]\), therefore

if \( \mu_A^*(x_i) \leq \mu_A(x_i) \) then \( H(A^*; P; U) \leq H(A; P; U) \) in \([0, 0.5]\),

and

if \( \mu_A^*(x_i) \geq \mu_A(x_i) \) then \( H(A^*; P; U) \leq H(A; P; U) \) in \((0.5, 1]\). (12)

Taking (11) and (12) together, we get \( H(A^*; P; U) \leq H(A; P; U) \).

\textbf{(P4) Symmetry:}

Clearly, from the definition of \( H(A; P; U) \) and with \( \mu_A(x_i) = 1 - \mu_A(x_i) \), we conclude that \( H(\bar{A}; P; U) = H(A; P; U) \). Hence, \( H(A; P; U) \) satisfies all the properties of fuzzy entropy and is therefore a valid ‘useful’ measure of fuzzy entropy.

The measure (8) can be further generalised parametrically as given below:

\[
H_a(A; P; U) = \frac{1}{1 - \alpha} \sum_{i=1}^{n} u_i p_i \log\left[ \mu_A^\alpha(x_i) + (1 - \mu_A(x_i))^\alpha \right]; \alpha(\neq 1) > 0,
\] (13)

and

\[
H_e(A; P; U) = \frac{1}{n^{\sqrt{e}} - 1} \sum_{i=1}^{n} u_i p_i \log\left[ \mu_A(x_i)e^{1-\mu_A(x_i)} + (1 - \mu_A(x_i))e^{\mu_A(x_i)} - 1 \right].
\] (14)

which are generalised ‘useful’ fuzzy information measures corresponding to (3) and (4) respectively.

\[\square\]

3. ‘Useful’ fuzzy measure of integrated ambiguity

There have been several attempts to combine probabilistic and fuzzy uncertainties when \((X, P)\) is a discrete probability framework. The entropy given by (7) is a measure of uncertainty associated with a fuzzy event, and was the first composite measure of probabilistic and fuzzy uncertainty.

The total ‘useful’ information measure, which is integration of fuzzy and probabilistic uncertainties with utilities, was introduced and studied by De Luca and Termini (1972). This has been used by some authors as a measure of uncertainties of fuzziness and randomness of events in an experiment as follows:
(1) The uncertainty deduced from the ‘random’ nature of the experiment. The average of this uncertainty is computed by Shannon’s entropy, i.e. $H(P)$ given by

$$H(P) = -\sum_{i=1}^{n} p_i \log p_i,$$

(15)

(2) The uncertainty that arises from the fuzziness of the fuzzy set relative to the ordinary set. This amount is given by the fuzzy entropy (2).

(3) The statistical average of the ambiguity of the whole set is given by

$$S(A, P) = -\sum_{i=1}^{n} p_i [\mu_A(x_i) \log \mu_A(x_i) + (1 - \mu_A(x_i)) \log (1 - \mu_A(x_i))].$$

(16)

(4) The total fuzzy information measure is obtained by adding two kinds of uncertainties (15) and (16).


Furthermore, we combine probabilistic and fuzzy uncertainties with utilities when $(A; P; U)$ is a discrete probabilistic framework with utilities and fuzzy set $A$ which is characterised by membership function $\mu_A$. If we also consider importance or usefulness of an event, then the total ‘useful’ information measure is a measure of fuzzy uncertainties, random uncertainties and utilities of events, and is obtained as follows:

(1) If we consider the uncertainties with ‘usefulness’ from ‘random’ nature of an experiment, then average measure of these uncertainties was introduced and characterised by Bhaker and Hooda (1993):

$$H(P; U) = -\sum_{i=1}^{n} u_i p_i \log p_i; \quad u_i > 0 \quad \text{for all } i.$$  

(17)

(2) If we consider the uncertainty that arises from fuzziness of the fuzzy set, then we can compute the amount of ambiguity by taking the above proposed ‘useful’ fuzzy entropy which is average as suggested by De Luca and Termini (1972):

$$H(A; P; U) = -\sum_{i=1}^{n} u_i p_i H_i(A),$$

(18)

where $H_i(A) = [\mu_A(x_i) \log \mu_A(x_i) + (1 - \mu_A(x_i)) \log (1 - \mu_A(x_i))].$

(3) ‘Useful’ measure of total fuzzy information of fuzzy set $A$ in a random experiment is given by the sum of (17) and (18):

$$H_{Total}(A; P : U) = H(P; U) + H(A; P; U).$$

(19)

4. ‘Useful’ fuzzy-directed divergence measures

Let $A$ and $B$ be two standard fuzzy sets with same supporting points $x_1, x_2, \ldots, x_n$ and with fuzzy vectors $\mu_A(x_1), \mu_A(x_2), \ldots, \mu_A(x_n)$ and $\mu_B(x_1), \mu_B(x_2), \ldots, \mu_B(x_n)$. The simplest measure of fuzzy-directed divergence as suggested by Bhandari and Pal (1993) is

$$I(A, B) = \sum_{i=1}^{n} \left[ \mu_A(x_i) \log \frac{\mu_A(x_i)}{\mu_B(x_i)} + (1 - \mu_A(x_i)) \ln \frac{1 - \mu_A(x_i)}{1 - \mu_B(x_i)} \right].$$

(20)
and it is a valid measure if it is always non-negative and zero if \( \mu_A(x_i) = \mu_B(x_i) \) for each \( i = 1, 2, \ldots, n \).

Next, considering altogether the uncertainties of fuzziness of fuzzy set \( A \) from fuzzy set \( B \) and probabilities of random events having utilities, corresponding to (20) we define the following ‘useful’ measure of fuzzy-directed divergence of fuzzy set \( A \) from fuzzy set \( B \):

\[
I(A; B; P; U) = \sum_{i=1}^{n} u_i p_i \left[ \frac{\mu_A(x_i) \log(\mu_A(x_i))}{\mu_B(x_i)} + (1 - \mu_A(x_i)) \log((1 - \mu_A(x_i)))/(1 - \mu_B(x_i)) \right],
\]

\( u_i > 0 \) for all \( i \),

(21)

and ‘useful’ fuzzy symmetric divergence measure as

\[
J(A; B; P; U) = I(A; B; P; U) + I(B; A; P; U).
\]

(22)

**Theorem 2.** \( I(A; B; P; U) \) is a valid measure, i.e. \( I(A; B; P; U) \geq 0 \) with equality if \( \mu_A(x_i) = \mu_B(x_i) \) for each \( i = 1, 2, \ldots, n \).

**Proof.**

Let

\[
\sum_{i=1}^{n} \mu_A(x_i) = s, \sum_{i=1}^{n} \mu_B(x_i) = t \text{ and } \sum_{i=1}^{n} u_i p_i = u,
\]

then

\[
\sum_{i=1}^{n} u_i p_i \left( \frac{\mu_A(x_i) \log(\mu_A(x_i))}{\mu_B(x_i)} \right) \geq us \log \frac{s}{t}.
\]

(23)

Similarly, we can prove that

\[
\sum_{i=1}^{n} u_i p_i \left( 1 - \frac{\mu_A(x_i)}{1 - \mu_B(x_i)} \right) \geq u(n - s) \log \frac{n - s}{n - t}.
\]

(24)

Adding (23) and (24), we get

\[
\sum_{i=1}^{n} u_i p_i \left( \frac{\mu_A(x_i) \log(\mu_A(x_i))}{\mu_B(x_i)} + (1 - \mu_A(x_i)) \log((1 - \mu_A(x_i)))/(1 - \mu_B(x_i)) \right)
\]

\[
\geq u \left[ s \log \frac{s}{t} + (n - s) \log \frac{n - s}{n - t} \right],
\]

(25)
Let

\[ f(s) = s \log \frac{s}{t} + (n - s) \log \frac{n - s}{n - t}, \text{ then} \]

\[ f'(s) = \left( \log \frac{s}{t} - \log \frac{n - s}{n - t} \right) \text{ and} \]

\[ f''(s) = \frac{1}{s} + \frac{1}{n - s} = \frac{n}{s(n - s)} > 0. \]

Thus \( f''(s) > 0 \), which shows that \( f(s) \) is a convex function of \( s \) and has its minimum value when \( s/t = (n - s)/(n - t) = (n/n) = 1 \). Now, if \( A = B \) (i.e. \( s = t \)), then \( f(s) = 0 \). Hence, \( f(s) > 0 \) and vanishes only when \( s = t \). As \( \sum_{i=1}^{n} u_i p_i > 0 \), therefore, \( I(A, B; P; U) \geq 0 \) and vanishes only when \( A = B \). Thus, \( I(A, B; P; U) \) is a valid measure of ‘useful’ fuzzy-directed divergence of fuzzy sets \( A \) and \( B \); and consequently, \( J(A, B; P; U) \) is a valid ‘useful’ measure of symmetric divergence.

It may be noted that if \( B = A_F \), the most fuzzy set, i.e. \( \mu_B(x_i) = 0.5 \) for all \( x_i \), then \( I(A, A_F; P; U) = n \log 2 - H(A; P; U) \).

The measure (21) can be further generalised parametrically. Corresponding to the following measure of fuzzy-directed divergence as defined by Bajaj and Hooda (2010):

\[ I_\alpha(A, B) = \frac{1}{\alpha - 1} \sum_{i=1}^{n} \log \left[ \mu_A^\alpha(x_i) \mu_B^{1-\alpha}(x_i) + (1 - \mu_A(x_i))^\alpha (1 - \mu_B(x_i))^{1-\alpha} \right], \quad (26) \]

we suggest the following ‘useful’ measure of fuzzy-directed divergence of fuzzy set \( A \) from fuzzy set \( B \) with \( u_i > 0 \) for all \( i \):

\[ I_\alpha(A, B; P; U) = \frac{1}{\alpha - 1} \sum_{i=1}^{n} u_i p_i \log \left[ \mu_A^\alpha(x_i) \mu_B^{1-\alpha}(x_i) + (1 - \mu_A(x_i))^\alpha (1 - \mu_B(x_i))^{1-\alpha} \right] / \sum_{i=1}^{n} u_i p_i \]

\[ (27) \]

and ‘useful’ measure of fuzzy symmetric divergence as

\[ J_\alpha(A, B; P; U) = I_\alpha(A, B; P; U) + I_\alpha(B, A; P; U). \]

Furthermore, it can be proved that \( I_\alpha(A, B; P; U) \) is a valid measure by showing \( I_\alpha(A, B; P; U) \) with equality if \( \mu_A(x_i) = \mu_B(x_i) \) for each \( i = 1, 2, \ldots, n \).

Particular cases:

- \( \lim_{\alpha \to 1} I_\alpha(A, B) = I(A, B) \), i.e. as \( \alpha \to 1 \), (26) \( \to \) (20).
- \( \lim_{\alpha \to 1} I_\alpha(A, B; P; U) = I(A, B; P; U) \), i.e. as \( \alpha \to 1 \), (27) \( \to \) (21).
- If \( B = A_F \), the most fuzzy set, i.e. \( \mu_B(x_i) = 0.5 \) for all \( x_i \), then it can be easily shown that

\[ I_\alpha(A, A_F; P; U) = n \log 2 - H_\alpha(A; P; U). \]

5. Constrained optimisation of ‘useful’ fuzzy information measure

On considering the ‘useful’ fuzzy information measure given by (8) subject to \( \sum_{i=1}^{n} \mu_A(x_i) = a \) and using Lagrange’s multiplier method, we have
\[ u_i p_i \log \frac{\mu_A(x_i)}{1 - \mu_A(x_i)} = m, \quad (29) \]

i.e., \[ \mu_A(x_i) = \frac{e^{(m/u_i p_i)}}{1 + e^{(m/u_i p_i)}} = \frac{1}{1 + e^{-(m/u_i p_i)}}, \quad (30) \]

where \( m \) is determined by

\[ \phi(m) = \sum_{i=1}^{n} \frac{e^{(m/u_i p_i)}}{1 + e^{(m/u_i p_i)}} - a = 0, \quad (31) \]

and

\[ \phi'(m) = \sum_{i=1}^{n} \frac{1/(u_i p_i)e^{(m/u_i p_i)}}{(1 + e^{(m/u_i p_i)})^2} * \geq 0. \quad (32) \]

Note that

\[ \phi(-\infty) = -a, \phi(0) = \frac{n}{2} - a, \phi(\infty) = n - a. \quad (33) \]

Using (33), we see that Equation (31) has a unique root which is negative if \( a < (n/2) \) and is positive if \( a > (n/2) \). It may be observed that

- If \( m \) is negative, i.e., if \( a < (n/2) \), then \( \mu_A(x_i) \leq 1/2 \) for all \( i \).
- If \( m \) is positive, i.e., if \( a > (n/2) \), then \( \mu_A(x_i) \geq 1/2 \) for all \( i \).
- If \( m = 0 \), i.e., if \( a = (n/2) \), then \( \mu_A(x_i) = (1/2) \) for all \( i \).

Thus for maximisation, \( \mu_A(x_i) \)’s are either all less than or equal to 1/2 or all greater than or equal to 1/2 or all are equal to 1/2.

Now if \( \mu_A(x_i) \) is replaced by \( 1 - \mu_A(x_i) \), the fuzziness of the \( i \)th element does not change, but its contribution to the sum \( \sum_{i=1}^{n} \mu_A(x_i) \) changes. As such, it will be better to take a fuzzy set in standard form, i.e., \( \mu_A(x_i) \leq 1/2 \) for all \( i \).

Now, consider (8) where \( \mu_A(x_i) \) is given by (30). Differentiating (8) with respect to \( \mu_A(x_i) \), we get

\[ \frac{dH}{d\mu_A(x_i)} = -\sum_{i=1}^{n} u_i p_i \log \frac{\mu_A(x_i)}{1 - \mu_A(x_i)} > 0 \text{ because } \mu_A(x_i) \leq \frac{1}{2}. \quad (34) \]

Differentiating (30) with respect to \( m \), we get

\[ \frac{d\mu_A(x_i)}{dm} = \frac{1/(u_i p_i)e^{-(m/u_i p_i)}}{(1 + e^{-m/u_i p_i})^2} > 0. \quad (35) \]

Differentiating (31) with respect to \( m \), we get

\[ \frac{da}{dm} = \sum_{i=1}^{n} \frac{1/(u_i p_i)e^{-(m/u_i p_i)}}{(1 + e^{-m/u_i p_i})^2} > 0. \quad (36) \]

Combining (34), (35) and (36), we conclude that \( H(A; P; U) \) is an increasing function of \( a \). Hence, \( H(A; P; U) \) is maximum when \( a = (n/2) \) and then \( m = 0, \mu_A(x_i) = 1/2 \) for all \( i \).
Thus $H_{\max}(A; P; U) = \log 2$ and $H_{\max}$ increases from 0 to log 2 as $a$ increases from 0 to $n/2$.

Now, minimum value of entropy will occur when as many of the $\mu_A(x_i)$’s are 0 or 1 as possible subject to, of course, the constraint being satisfied. Also without loss of generality, we assume that $u_1p_1 \leq u_2p_2 \leq \cdots \leq u_np_n$.

- When $a = 0$, all $\mu_A(x_i)$’s have to be zero which gives $H_{\min}(A; P; U) = 0$.
- When $a = 1/2$, one of the $\mu_A(x_i)$’s can be 1/2 and rest have to be 0, so that
  \[ H_{\min}(A; P; U) = \frac{u_1p_1 \log 2}{\sum_{i=1}^{n} u_ip_i}. \]

- Similarly, when $a = 1$, two of the $\mu_A(x_i)$’s can be 1/2 and the rest have to be 0, so that
  \[ H_{\min}(A; P; U) = \frac{(u_1p_1 + u_2p_2) \log 2}{\sum_{i=1}^{n} u_ip_i}. \]

- When $a = n/2$, $H_{\min}(A; P; U) = \log 2$.
- When $a$ lies between $(m - 1)/2$ and $m/2$, $m - 1$ of $\mu_A(x_i)$’s can be 1/2, one can be a fraction $\eta$ and the rest can be 0, so that
  \[ H_{\min}(A; P; U) = \frac{u_1p_1 + u_2p_2 + \cdots + u_{m-1}p_{m-1}) \log 2 - u_np_m(\eta \log \eta + (1 - \eta)\log(1 - \eta))}{\sum_{i=1}^{n} u_ip_i}. \]

Thus as $a$ varies from $(m - 1)/2$ to $m/2$, $H_{\min}$ varies from $((u_1p_1 + u_2p_2 + \cdots + u_{m-1}p_{m-1}) \log 2)/\sum_{i=1}^{n} u_ip_i)$ to log 2. Therefore, $H_{\max}$ increases from 0 to log 2 continuously whereas $H_{\min}$ also increases from 0 to the same value but in a piecewise continuous manner.

### 6. Constrained optimisation of ‘useful’ fuzzy-directed divergence measure

The maximum value of the ‘useful’ fuzzy-directed divergence measure is obtained when as many possible values of $\mu_A(x_i)$’s are 0 or 1.

- When $a = 0$, the maximum value is $-\sum_{i=1}^{n} u_ip_i \log (1 - \mu_B(x_i))$
- When $a = 1/2$, the maximum value is $u_np_n \log \frac{1 - \mu_B(x_n)}{\mu_B(x_n)} - \sum_{i=1}^{n} u_ip_i \log (1 - \mu_B(x_i))$
  provided $u_1p_1 \leq u_2p_2 \leq \cdots \leq u_np_n$.
- Similarly, when $a = 1$, the maximum value is given by
  \[ u_np_n \log \frac{1 - \mu_B(x_n)}{\mu_B(x_n)} + u_{n-1}p_{n-1} \log \frac{1 - \mu_B(x_{n-1})}{\mu_B(x_{n-1})} - \sum_{i=1}^{n} u_ip_i \log (1 - \mu_B(x_i)). \]

- Finally, when $a = n/2$, the maximum value is $-\sum_{i=1}^{n} u_ip_i \log \mu_B(x_i)$.

Note that, the maximum value of ‘useful’ fuzzy-directed divergence measure is a piecewise continuous function which increases from $-\sum_{i=1}^{n} u_ip_i \log (1 - \mu_B(x_i))$ to $-\sum_{i=1}^{n} u_ip_i \log \mu_B(x_i)$. 
Next, for minimisation of the ‘useful’ fuzzy-directed divergence measure of a standard fuzzy set $A$ from a standard fuzzy set $B$, consider (21) subject to $\sum_{i=1}^{n} \mu_A(x_i) = a \leq n/2$. Again, using the method of Lagrange’s multiplier, we get

$$u_i p_i \log \left( \frac{\mu_A(x_i)(1 - \mu_B(x_i))}{(1 - \mu_A(x_i))\mu_B(x_i)} \right) = m,$$

(37)

$$\Rightarrow \frac{\mu_A(x_i)}{1 - \mu_A(x_i)} = e^{(m/u_i p_i)} \frac{\mu_B(x_i)}{1 - \mu_B(x_i)} = e^{(m/u_i p_i)} \beta_i,$$

(38)

where $m$ is determined by

$$\psi(m) = \sum_{i=1}^{n} \frac{e^{(m/u_i p_i)} \beta_i}{1 + e^{(m/u_i p_i)}} - a = 0,$$

(39)

and

$$\psi'(m) = \sum_{i=1}^{n} \frac{e^{(m/u_i p_i)} (\beta_i/u_i p_i)}{(1 + e^{(m/u_i p_i)})^2} \geq 0.$$

(40)

Note that

$$\psi(-\infty) = -a, \psi(0) = \sum_{i=1}^{n} \mu_B(x_i) - \sum_{i=1}^{n} \mu_A(x_i), \psi(\infty) = n - a.$$

(41)

Now using (41), we see that Equation (39) has a unique root which is

- positive if $\sum_{i=1}^{n} \mu_B(x_i) < \sum_{i=1}^{n} \mu_A(x_i)$,
- negative if $\sum_{i=1}^{n} \mu_B(x_i) > \sum_{i=1}^{n} \mu_A(x_i)$,
- zero if $\sum_{i=1}^{n} \mu_B(x_i) = \sum_{i=1}^{n} \mu_A(x_i)$.

If $m \geq 0$, then $\mu_A(x_i) \equiv \mu_B(x_i)$ so that each minimising value is less than the corresponding given value. Thus, in every case, either $\mu_A(x_i) < \mu_B(x_i)$ or in every case $\mu_A(x_i) > \mu_B(x_i)$ or in every case $\mu_A(x_i) = \mu_B(x_i)$.

7. Conclusion

The ‘useful’ fuzzy measure of information and directed divergence studied in the present paper can further be generalised parametrically. The optimisation of these measures under constraints has been studied; however, their illustration by numerical examples and applications would be studied and reported in due course.

Acknowledgements

The authors are thankful to the reviewers for their valuable suggestions and comments.
Notes on contributors

D.S. Hooda received his MA (Maths) from Delhi University in 1969 and MPhil from Meerut University, Meerut, in 1976. He joined CCS Haryana Agricultural University, Hisar as a Lecturer in 1970 and was appointed Asst Professor of mathematics in 1972. He received his PhD from KU Kurukshetra in 1981. His field of specialization is information theory and its applications and his research interests are information measures, source coding, entropy optimization principles and their applications in statistics, finance mathematics, survival analysis, and bounds on probabilities of error, pattern recognition and fuzzy information. Currently, Professor Hooda is working as Dean (Research) and Head, Department of Mathematics, Jaypee University of Engineering and Technology, Guna. He is engaged in research in information theory and its applications in engineering, applied and biological sciences.

Rakesh K Bajaj received his BSc degree with honours in mathematics from Banaras Hindu University, Varanasi and the MSc from the Indian Institute of Technology, Kanpur in 2000 and 2002, respectively. He received his MTech (Comp Sc) from RVD University, Udaipur and PhD (Mathematics) from Jaypee University of Information Technology (JUIT), Waknaghat in 2009. He has worked as Sr Lecturer in the Department of Mathematics, JUIT, Waknaghat since 2003. His interests include fuzzy information measures, pattern recognition, fuzzy clustering, fuzzy statistics and fuzzy mathematics in image processing.

References


