Soft Comput (2007) 11: 317–321
DOI 10.1007/s00500-006-0083-5

Xi-Zhao Wang · Su-Fang Zhang · Jun-Hai Zhai

A nonlinear integral defined on partition and its application to decision trees

Published online: 20 April 2006

Abstract Nonlinear integrals play an important role in information fusion. So far, all existing nonlinear integrals of a function with respect to a set function are defined on a subset of a space. In many of the problems with information fusion, such as decision tree generation in inductive learning, we often need to deal with the function defined on a partition of the space. Motivated by minimizing the classification information entropy of a partition while generating decision trees, this paper proposes a nonlinear integral of a function with respect to a nonnegative set function on a partition, and provides the conclusion that the sum of the weighted entropy of the union of several subsets is not less than the sum of the weighted entropy of a single subset. It is shown that selecting the entropy of a single attribute is better than selecting the entropy of the union of several attributes in generating rules by decision trees.

Keywords Decision tree · Nonlinear integral · Information fusion · Partition · Information entropy

1 Introduction

Information fusion means to extract useful information from many different information sources and then combine to achieve a result. As a tool of information fusion, the integral plays an important role in many fields such as data mining, pattern recognition, object classification and decision-making. Aggregation with different backgrounds in information fusion requires different integrals. A traditional aggregation tool used in information fusion is the weighted average method. It is essentially the Lebesgue-like integral with respect to a classical additive measure [1] and is valid for many linear models. Using a linear model needs a basic assumption that there is no integration among information sources for their contribution to an objective attribute. However, in many real situations such a linear model fails due to the inherent interaction among attributes. Hence, many nonlinear integrals with respect to nonadditive set function have been introduced. In information fusion or data mining set function can be explained as an efficiency function or an importance measure [2–4]. The nonadditivity of set function indicates the inherent interaction among information sources or the attributes to the objective attribute. So, nonlinear integrals can efficiently deal with many nonlinear models [5–10]. By now, there have already been developed many kinds of nonlinear integrals [5–15], the main ones being the Choquet integral [11–13], Sugeno integral [5,14], pan-integral [16, 17] and Wang integral [18]. These nonlinear integrals are all defined on a subset of a space. In many fields such as data mining and machine learning, in order to efficiently deal with the data and extract the useful rules, we need to divide a data set into several disjoint data subsets according to one criteria, and different criterion often lead to different partitions of the set. For example, in a decision tree, an attribute corresponds to a partition of a sample set, and a different attribute usually corresponds to a different partition. So information fusion can also be influenced by the partitions of the set. In this paper, we propose a nonlinear integral defined on a partition of the set. If we define the entropy of the set as the set function and let the integrand be equal to 1 in the nonlinear integral, then the value of the integral can be achieved at the most refined partition of the set, which is consistent with the maximum information gain of an attribute in generating a decision tree. Therefore, this paper to some extent, and from the viewpoint of minimum entropy, confirms that using a single attribute (attributes with equal entropy can be seen as one attribute) in ID3 algorithm of generating decision trees is more efficient compared with using several attributes merging at a node for selecting expanded attributes.

This paper is organized as follows. Section 2 gives some basic concepts, the new type of integral and some of its fundamental properties. Section 3 derives the formula for computing the new integral and Sect. 4 concludes this paper.

X.-Z. Wang(✉) · S.-F. Zhang · J.-H. Zhai
Department of Mathematics and Computer Science,
Hebei University, Baoding City,
Hebei Province, 071002, China
E-mail: xizhaowang@ieee.org

2 Some basic concepts and the new nonlinear integral

For convenience, suppose X be a nonempty set, $\boldsymbol{F}$ be an σ-algebra of subset of X. We call $(X, \boldsymbol{F})$ a measurable space. Let $\boldsymbol{G}$ be the set of all nonnegative functions defined on X. Here, X is not necessarily finite.

2.1 Some basic concepts and the definition of the new nonlinear integral

Definition 2.1 *A function μ defined on $\boldsymbol{F}$ is called a* set function.

(1) The set function μ is said to be additive *if $\mu(A \cup B) = \mu(A)+\mu(B)$, whenever $A \in \boldsymbol{F}$, $B \in \boldsymbol{F}$ with $A \cap B = \phi$. Otherwise, μ is called* nonadditive.

(2) The set function μ is said to be super-additive *if $\mu(A \cup B) \geq \mu(A) + \mu(B)$, whenever $A \in \boldsymbol{F}$, $B \in \boldsymbol{F}$ with $A \cap B = \phi$.*

(3) The set function μ is said to be sub-additive *if $\mu(A \cup B) \leq \mu(A)+\mu(B)$, whenever $A \in \boldsymbol{F}$, $B \in \boldsymbol{F}$, with $A \cap B = \phi$.*

Definition 2.2 *The* information entropy *is a measure of uncertainty of information at the statistical (incomplete) description of a system S with the use of a distribution of probabilities $p = \{p_i\}$ $(0 \leq p_i \leq 1,\ i = 1, 2, \ldots, n)$, e.g. The Boltzmann-Shannon entropy:*

$$E(S) = \sum_{i=1}^{n} p_i \ln p_i. \tag{1}$$

In many 2-class classification problems, given a collection S, containing positive and negative examples of some target concept, the entropy of S relative to this Boolean classification is $E(S) = -p_+ \log_2 p_+ - p_- \log_2 p_-$, where p_+ is the proportion of positive examples in S and p_- is the proportion of negative examples in S.

Definition 2.3 *Let $(X, \boldsymbol{F})$ be a measurable space, $\tau = \{A_1, A_2, \ldots, A_n\}$ is called a* partition of X *if and only if $\overset{n}{\underset{i=1}{\cup}} A_i = X$ and $A_i \cap A_j = \phi$ $(i \neq j)$.*

Definition 2.4 *Suppose τ_1 and τ_2 are two partitions of the set X and $\tau_1 = \{A_i\}_1^n$, $\tau_2 = \left\{B_j\right\}_1^m$, the partition τ_1 is* a refinement of the partition τ_2, *in symbol $\tau_1 < \tau_2$, if B_j is the union of some A_i, i.e. $B_j = \underset{k \in E}{\cup} A_k$, where $E \subset \{1, 2, \ldots, n\}$. We use $\tau_1 \leq \tau_2$ to denote $\tau_1 < \tau_2$ or $\tau_1 = \tau_2$.*

For example, given a set $X = \{x_1, x_2, x_3, x_4\}$, let $\tau_1 = \{x_1, x_2\} \cup \{x_3\} \cup \{x_4\}$, $\tau_2 = \{x_1, x_2\} \cup \{x_3, x_4\}$ then $\tau_1 < \tau_2$, i.e. the partition τ_1 is a refinement of the partition τ_2.

In what follows, we will give the definition of the new nonlinear integral defined on a partition of a set.

Definition 2.5 *Let $\upsilon \colon \boldsymbol{F} \to [0, \infty)$ be the Lebesgue measure, $\mu : \boldsymbol{F} \to [0, \infty)$ be a nonnegative set function, and $f : X \to [0, \infty)$ a nonnegative function. Given a partition of X, τ, then the integral of the function f with respect to μ on the partition τ, denoted by $\int_\tau f \mathrm{d}\mu$, is defined as:*

$$\int\limits_{\tau} f \mathrm{d}\mu = \inf_{\tau \leq \sigma} \sum_{i=1}^{n} \left(\frac{\int_{A_i} f \mathrm{d}\upsilon}{\int_X f \mathrm{d}\upsilon} \right) \mu(A_i), \tag{2}$$

where $\sigma = \{A_i\}_{i=1}^n$ is an arbitrary partition of X.

2.2 Fundamental properties

Without loss of generality, let $\mu : \boldsymbol{F} \to [0, \infty)$ be monotonic and satisfying $\mu(\phi) = 0$, and F be a nonnegative function defined on X. The integral introduced above possesses the following basic properties.

Proposition *Let $f, g \in \boldsymbol{G}$, $A, B \in \boldsymbol{F}$, $\int_X f \mathrm{d}\upsilon \neq 0$ and $c \in R_+$, then we have the following:*

(1) If $\mu(X) = 0$, then $\int_\tau f \mathrm{d}\mu = 0$.

(2) If $\tau_1 \leq \tau_2$, then $\int_{\tau_1} f \mathrm{d}\mu \leq \int_{\tau_2} f \mathrm{d}\mu$.

(3) If μ is super-additive, then $\int_\tau c \mathrm{d}\mu \leq \mu(X)$.

(4) $\int_\tau cf d\mu = \int_\tau f \mathrm{d}\mu$.

(5) For any constant $c_1 > 0, c_2 > 0$, if $f \neq 0$ and $g \neq 0$, then:

$$\int\limits_{\tau} (c_1 f + c_2 g) \mathrm{d}\mu \leq \int\limits_{\tau} f \mathrm{d}\mu + \int\limits_{\tau} g \mathrm{d}\mu.$$

Proof We only need to prove (3), (4) and (5); the remaining properties can be obtained directly from the definition. Suppose τ is a given partition of X.

For (3), we have

$$\begin{aligned}\int\limits_{\tau} c \mathrm{d}\mu &= \inf_{\tau \leq \sigma} \sum_{i=1}^{n} \frac{\int_{A_i} c \mathrm{d}\upsilon}{\int_X c \mathrm{d}\upsilon} \mu(A_i) \\ &= \inf_{\tau \leq \sigma} \sum_{i=1}^{n} \frac{c \int_{A_i} \mathrm{d}\upsilon}{c \int_X \mathrm{d}\upsilon} \mu(A_i) \\ &\leq \inf_{\tau \leq \sigma} \sum_{i=1}^{n} \mu(A_i) \leq \sum_{i=1}^{n} \mu\,(A_i) \leq \mu(X),\end{aligned} \tag{3}$$

where $\sigma = \{A_i\}_{i=1}^n$ is an arbitrary partition of X.

For (4), we have:

$$\begin{aligned}\int\limits_{\tau} cf \mathrm{d}\mu &= \inf_{\tau \leq \sigma} \sum_{i=1}^{n} \frac{\int_{A_i} cf \mathrm{d}\upsilon}{\int_X cf \mathrm{d}\upsilon} \mu(A_i) \\ &= \inf_{\tau \leq \sigma} \sum_{i=1}^{n} \frac{c \int_{A_i} f \mathrm{d}\upsilon}{c \int_X f \mathrm{d}\upsilon} \mu(A_i) \\ &= \inf_{\tau \leq \sigma} \sum_{i=1}^{n} \frac{\int_{A_i} f \mathrm{d}\upsilon}{\int_X f \mathrm{d}\upsilon} \mu(A_i) = \int\limits_{\tau} f \mathrm{d}\mu,\end{aligned} \tag{4}$$

where $\sigma = \{A_i\}_{i=1}^n$ is an arbitrary partition of X.

For (5), we have:

$$\begin{aligned}
&\int\limits_{\tau} (c_1 f + c_2 g)\,\mathrm{d}\mu \\
&= \inf_{\tau \le \sigma} \left\{ \sum_{i=1}^{n} \left(\frac{\int_{A_i} (c_1 f + c_2 g)\,\mathrm{d}\upsilon}{\int_X (c_1 f + c_2 g)\,\mathrm{d}\upsilon} \right) \mu(A_i) \right\} \\
&= \inf_{\tau \le \sigma} \left\{ \sum_{i=1}^{n} \left(\frac{\int_{A_i} c_1 f \mathrm{d}\upsilon}{\int_X (c_1 f + c_2 g)\,\mathrm{d}\upsilon} \right) \mu(A_i) \right\} \\
&\quad + \inf_{\tau \le \sigma} \left\{ \sum_{i=1}^{n} \left(\frac{\int_{A_i} c_2 g \mathrm{d}\upsilon}{\int_X (c_1 f + c_2 g)\,\mathrm{d}\upsilon} \right) \mu(A_i) \right\} \\
&\le \inf_{\tau \le \sigma} \left\{ \sum_{i=1}^{n} \left(\frac{\int_{A_i} c_1 f \mathrm{d}\upsilon}{\int_X c_1 f \mathrm{d}\upsilon} \right) \mu(A_i) \right\} \\
&\quad + \inf_{\tau \le \sigma} \left\{ \sum_{i=1}^{n} \left(\frac{\int_{A_i} c_2 g \mathrm{d}\upsilon}{\int_X c_2 g \mathrm{d}\upsilon} \right) \mu(A_i) \right\} \\
&= \inf_{\tau \le \sigma} \left\{ \sum_{i=1}^{n} \left(\frac{\int_{A_i} f \mathrm{d}\upsilon}{\int_X f \mathrm{d}\upsilon} \right) \mu(A_i) \right\} \\
&\quad + \inf_{\tau \le \sigma} \left\{ \sum_{i=1}^{n} \left(\frac{\int_{A_i} g \mathrm{d}\upsilon}{\int_X g \mathrm{d}\upsilon} \right) \mu(A_i) \right\} \\
&= \int\limits_{\tau} f \mathrm{d}\mu + \int\limits_{\tau} g \mathrm{d}\mu, \qquad (5)
\end{aligned}$$

where $\sigma = \{A_i\}_{i=1}^{n}$ is an arbitrary partition of X. □

The following example shows that the equality in (3) of the Proposition may not hold.

Example 1 Let $X = \{a, b\}$, $\boldsymbol{F} = P(X)$, $\upsilon(A) = |A|$ where $|A|$ is the number of the elements in A, $f(x) = 1$ when $x \in X$. $\tau = \{a\} \cup \{b\}$ is a partition of X, and:

$$\mu(E) = \begin{cases} 0 & \text{if } E = \phi \\ 0.5 & \text{if } E = \{a\} \\ 0.6 & \text{if } E = \{b\} \\ 1 & \text{if } E = X \end{cases} \qquad (6)$$

then

$$\int\limits_{\tau} f \mathrm{d}\mu = \frac{1}{2}\mu(\{a\}) + \frac{1}{2}\mu(\{b\}) = \frac{0.5 + 0.6}{2} = 0.55 \qquad (7)$$

but $\mu(\{X\}) = 1$.

3 The calculation of the nonlinear integral in special cases

In general, the value of the nonlinear integral is not necessarily achieved at the most refine partition τ. For example, suppose $A = \{x_1, x_2, x_3\}$, $f(x_i) = 1$, $(i = 1, 2, 3)$.

Let $\mu(x_1) = 1$, $\mu(x_2) = 2$, $\mu(x_3) = 3$, $\mu(x_1, x_2) = 1$, $\mu(x_2, x_3) = 2$, $\mu(x_1, x_3) = 3$.

$\tau = \{x_1\} \cup \{x_2\} \cup \{x_3\}$, $\upsilon(A) = |A|$, where $|A|$ is the number of elements in set A. Then $\int_\tau d\mu = (5/3)$, which is achieved at the partition $\tau_1 = \{x_1, x_2\} \cup \{x_3\}$, not at the partition τ, and $\tau < \tau_1$. But in special cases, it can be held.

Theorem *Suppose $X = \{x_1, x_2, \ldots, x_n\}$, $f(x_i) = 1$, $\tau = \{S_i\}_{i=1}^{l}$ be a given partition of X, and the existing Boolean function is denoted by $g : X \to \{-1, 1\}$. Let $\mu(A_i) = E(A_i)$, $\upsilon(A_i) = |A_i|$, where $|A_i|$ is the number of elements in set A_i, then:*

$$\begin{aligned}
\int\limits_{\tau} f \mathrm{d}\mu &= \inf_{\tau \le \sigma} \sum_{i=1}^{m} \left(\frac{\int_{A_i} f \mathrm{d}\upsilon}{\int_X f \mathrm{d}\upsilon} \right) \mu(A_i) \\
&= \sum_{i=1}^{l} \left(\frac{\int_{S_i} f \mathrm{d}\upsilon}{\int_X f \mathrm{d}\upsilon} \right) \mu(S_i) = \sum_{\mathrm{i}=1}^{l} \frac{|S_i|}{n} \mu(S_i), \qquad (8)
\end{aligned}$$

where $\sigma = \{A_i\}_{i=1}^{m}$ is an arbitrary partition of X.

Proof We first prove that Eq. (8) is held while $l = 2$. Then according to mathematical induction, we can easily prove Eq. (8) in any case.

Suppose $S_1 = \left(m_1^+, n_1^-\right)$, $S_2 = \left(m_2^+, n_2^-\right)$, $S_i \subset X$ $(i = 1, 2)$ and $S_1 \cup S_2 = X$, where $m_i^+ (i = 1, 2)$ denotes that there are $m_i (i = 1, 2)$ elements x_{ki} $(k = 1, 2, \ldots, m_i)$ satisfying $g\left(x_{ki}\right) = 1$ in $S_i (i = 1, 2)$. Similarly, $n_i^- (i = 1, 2)$ denotes that there are $n_i (i = 1, 2)$ elements y_{ji} $(j = 1, 2, \ldots, n_i)$ satisfying $g\left(y_{ji}\right) = -1$ in $S_i (i = 1, 2)$, and $m_1 + n_1 + m_2 + n_2 = n$. We can prove the following is held, that is:

$$\frac{|S_1|}{|S_1 \cup S_2|} E\left(S_1\right) + \frac{|S_2|}{|S_1 \cup S_2|} E\left(S_2\right) \le E\left(S_1 \cup S_2\right). \qquad (9)$$

First we prove

$$\begin{aligned}
&x \log_2(x + c) + y \log_2(y + d) \\
&\quad -(x + y) \log_2(x + y + c + d) \\
&\le x \log_2 x + y \log_2 y - (x + y) \log_2(x + y), \qquad (10)
\end{aligned}$$

where $x > 0$, $y > 0$, $c > 0$, $d > 0$.

Let

$$\begin{aligned}
u(c, d) &= x \log_2(x + c) + y \log_2(y + d) \\
&\quad -(x + y) \log_2(x + y + c + d). \qquad (11)
\end{aligned}$$

Then

$$u_c'(c, d) = \frac{1}{\ln 2} \left(\frac{x}{x + c} - \frac{x + y}{x + y + c + d} \right) \qquad (12)$$

Let $u_c'(c, d) = 0$, then $x/y = c/d$.

While $x/y = c/d$, we have

$$\begin{aligned}
u_c'' &= -\frac{1}{\ln 2} \left(\frac{x}{(x + c)^2} - \frac{x + y}{(x + y + c + d)^2} \right) \\
&= -\frac{1}{\ln 2} \frac{xy}{(x + c)^2 (x + y)} < 0. \qquad (13)
\end{aligned}$$

Similarly, by the following

$$u'_d(c,d) = \frac{1}{\ln 2}\left(\frac{y}{y+d} - \frac{x+y}{x+y+c+d}\right) = 0 \qquad (14)$$

we can also obtain $x/y = c/d$.

While $x/y = c/d$, we have

$$u''_d = -\frac{1}{\ln 2}\left(\frac{y}{(y+d)^2} - \frac{x+y}{(x+y+c+d)^2}\right)$$
$$= -\frac{1}{\ln 2}\frac{x^3}{y(x+c)^2(x+y)} < 0. \qquad (15)$$

So, while $x/y = c/d$, Eq. (11) can achieve its maximum, that is:

$$u_{\max}(c,d) = x\log_2 x + y\log_2 y - (x+y)\log_2(x+y). \qquad (16)$$

Let $x = m_1, y = n_1, c = m_2, d = n_2$ in Eq. (10), we can obtain:

$$m_1\log_2 m_1 + n_1\log_2 n_1 - (m_1+n_1)\log_2(m_1+n_1)$$
$$\geq m_1\log_2(m_1+m_2) + n_1\log_2(n_1+n_2)$$
$$-(m_1+n_1)\log_2 n. \qquad (17)$$

Similarly, let $x = m_2, y = n_2, c = m_1, d = n_1$ in Eq. (10), and we can obtain:

$$m_2\log_2 m_2 + n_2\log_2 n_2 - (m_2+n_2)\log_2(m_2+n_2)$$
$$\geq m_2\log_2(m_1+m_2) + n_2\log_2(n_1+n_2)$$
$$-(m_2+n_2)\log_2 n. \qquad (18)$$

By Eqs. (17)+(18), we can obtain the following:

$$m_1\log_2\frac{m_1}{m_1+n_1} + n_1\log_2\frac{n_1}{m_1+n_1} + m_2\log_2\frac{m_2}{m_2+n_2}$$
$$+n_2\log_2\frac{m_2}{m_2+n_2}$$
$$\geq (m_1+m_2)\log_2\frac{m_1+m_2}{n} + (n_1+n_2)\log_2\frac{n_1+n_2}{n}. \qquad (19)$$

So we can conclude

$$\frac{m_1+n_1}{n}$$
$$\times\left(\frac{m_1}{m_1+n_1}\log_2\frac{m_1}{m_1+n_1} + \frac{n_1}{m_1+n_1}\log_2\frac{n_1}{m_1+n_1}\right)$$
$$+\frac{m_2+n_2}{n}$$
$$\times\left(\frac{m_2}{m_2+n_2}\log_2\frac{m_2}{m_2+n_2} + \frac{n_2}{m_2+n_2}\log_2\frac{n_2}{m_2+n_2}\right)$$
$$\geq \frac{m_1+m_2}{n}\log_2\frac{m_1+m_2}{n} + \frac{n_1+n_2}{n}\log_2\frac{n_1+n_2}{n}. \qquad (20)$$

Therefore, Eq. (9) is held. By Eq. (9) we can obtain the following ($S_1 \cup S_2 \subset X$):

$$\frac{|S_1|}{|X|}E(S_1) + \frac{|S_2|}{|X|}E(S_2) \leq \frac{|S_1 \cup S_2|}{|X|}E(S_1 \cup S_2). \qquad (21)$$

Using Eq. (21), the validation of Eq. (8) is given by the mathematical induction. Then, we conclude that the sum of the weighted entropy of the union of several subsets is not less than the sum of the weighted entropy of a single subset. It is shown that using a single attribute (attributes with equal entropy can be seen as one attribute) in ID3 algorithm of generating decision trees is more efficient compared with using several attributes merging at a node for selecting expanded attributes. □

4 Conclusions

In this paper, we proposed a new nonlinear integral based on a partition of a set. The formula for computing this integral shows that the minimum entropy is attained at the most refined partition. Paying attention to the fact that the ID3 algorithm for generating decision trees selects the expanded attributes according to the integral given in this paper, we conclude that using the single attribute at a node in ID3 algorithm is more efficient compared with using several attributes (branches) merging from the viewpoint of entropy minimization.

Acknowledgements The research is supported by the National Natural Science Foundation of China (60473045, 60573069) and the Natural Science Foundation of Hebei Province (603137).

References

1. Halmos PR (1967) Measure theory. Van Nostrand, New York
2. Wang Z, Klir GJ (1992) Fuzzy measure theory. Plenum, New York
3. Wang Z, Leung KS, Wang J (1999) A genetic algorithm for determining non-additive set function in information fusion. Fuzzy Sets Syst 102:463–469
4. Wang Z, Leung KS, Xu K (1998) A new nonlinear fusion: an application of nonlinear integrals and genetic algorithms. Proceedings of FUSION ’98, Las Vegas, pp 299–306
5. Sugeno M (1974) Theory of Fuzzy integral and its applications. PhD Dissertation, Tokyo Institute of Technology
6. Grabisch M, Nguyen HT, Walker EA (1995) Fundamentals of uncertainty calculi, with applications to fuzzy inference. Kluwer, Dordrecht
7. Inoue K, Anzai T (1991) A study on the industrial design evaluation based upon non-additive measures. In: 7th Fuzzy System Sym., pp 521–524, Nagoya, Japan, June 1991. In Japan
8. Tanaka K, Sugeno M (1991) A study on subjective evaluation of color printing images. Int J Approx Reason 5:213–222
9. Washio T, Takahashi H, Kitamura M (1992) A method for supporting decision making on plant operation based on human reliability analysis by fuzzy integral. In: 2nd Int conf on fuzzy logic and neural networks. Iizuka, Japan, July 1992, pp 841–845
10. Yager RR (1988) On ordered weighted averaging aggregation operators in multi-criteria decision making. IEEE Trans Syst Man Cybern 18:183–190
11. de Campos LM, Bola∼nos MJ (1992) Characterization and comparison of Sugeno and Choquet integrals. Fuzzy Sets Syst 52:61–67
12. Choquet G (1954) Theory of capacities. Anna Inst Fourier 5:131–295
13. Murofushi T, Sugeno M (1989) An interpretation of fuzzy measure and the Choquet integral as an integral with respect to a fuzzy measure. Fuzzy Sets Syst 29:201–227

14. Ralescu D, Adams G (1980) The fuzzy integrals. J Math Anal 75:562–575
15. Zhang D, Guo C (1995) Generalized fuzzy integrals of set-valued functions. Fuzzy Sets Syst 76
16. Wang Z, Klir GJ (1997) PFB-integrals and PFA-integrals with respect to monotone set functions. Int J Uncertain Fuzz Knowl Based Syst 5(2):163–175
17. Wang Z, Wang W, Klir GJ (1996) Pan-integrals with respect to imprecise probabilities. Int J Gen Syst 25:229–243
18. Wang Z, Leung KS, Wong ML, Fang J (2000) A new type of nonlinear integrals and the computational algorithm. Fuzzy Sets Syst 112:223–231