Paper:

Identification of Fuzzy Measures with Distorted Probability Measures

Aoi Honda*,, and Yoshiaki Okazaki***

*Department of Systems Innovation and Informatics, Kyushu Institute of Technology
680-4 Kawazu, Iizuka 820-8502, Japan
E-mail: {aoi, okazaki}@ces.kyutech.ac.jp
**Department of Computer science and Mathematics, University of Paris I, Panthéon-Sorbonne
72, rue Regnault, 75013 Paris, France
[Received October 13, 2004; accepted February 4, 2005]

We consider the identification of fuzzy measures using a class of distorted probabilities – a scale transformation of probabilities. A fuzzy measure, which is a nonadditive set function with a high degree of freedom, enables us to express complicated interactions among evaluative items. Because of the high degree of freedom, however, it is difficult to identify all of the values $\mu(A)$ for every event A from known data $\mu(B), B \in \mathscr{A}$, where $\mathscr{A}$ is generally a small subclass of events. In this paper, we classify fuzzy measures by introducing "type," and propose an identifying fuzzy measures using a classified class of distorted probabilities.

Keywords: fuzzy measure, distortion, distorted probability measure, type, fuzzy measure identification

1. Introduction

Fuzzy measures are nonadditive set functions. Since they can be used to express interactions between items that are not expressible by regular measures, which have additivity, studies are being done on their application in fields such as the analysis of subjective evaluation problems, decision-making support, and pattern recognition [1–5].

When data is collected as fuzzy measures in the real world, fuzzy measures are rarely all given as known data, making it necessary to estimate the measures of whole events from partial known data. In applications, universal set X is a finite set consisting of a finite number of elements. If measures are additive, measures can be determined when the values of all singletons, that is, all one-point sets are found. Fuzzy measures are not additive, however, so the measures cannot be determined only from the values of singletons. Instead, the values of all subsets of X must be determined. This paper discusses the problem of identifying fuzzy measures from a subset of known data. To identify fuzzy measures, they can be restricted to a certain class, within which they are identified, and methods using λ-fuzzy measure and χ-fuzzy measure have been proposed [1, 6–8]. We propose identification using the class of distorted fuzzy measures, and verify its feasibility by comparing other methods of identification.

2. Preliminaries

Let X be n-point set $X = \{x_1, \ldots, x_n\}$. $\mathscr{P}(X)$ denote the power set of X. Brackets "$\{\cdot\}$" and commas denoting a set are omitted when this will not cause confusion, e.g., the two-point set $\{x_1, x_2\}$ is written "x_1x_2".

Definition 1 ([2, 9, 10]) A set function $g : \mathscr{P}(X) \to [0, 1]$ is a fuzzy measure if it satisfies the following conditions:
(i) $g(\emptyset) = 0$ and $g(X) = 1$.
(ii) For any $A, B \in \mathscr{P}(X)$, $g(A) \leq g(B)$ whenever $A \subset B$.

(i) expresses boundedness and (ii) monotonicity. Next, we introduce the concept of distortion.

Definition 2 (Chateauneuf [11], Honda and Okazaki [12]) Let g and h be two fuzzy measures on $\mathscr{P}(X)$. g is a distortion of h (distorted h) when a nondecreasing function $f : [0, 1] \to [0, 1]$ exists such that

$$g(A) = f \circ h(A) = f(h(A))$$

for any $A \in \mathscr{P}(X)$.

f is called a *scaling function*. Fuzzy measure g can be considered to be a distortion of fuzzy measure h by scaling function f. The magnitude relationship of h is propagated to g because f is a nondecreasing function. When h is a probability measure or some other readily-treatable measure, we can expect g to be treatable. Even when h is limited only to probability measures, a wide range of fuzzy measures can be covered because f is selectable at will. As examples of fuzzy measures, we introduce the possibility measure, λ-fuzzy measure and χ-fuzzy measure.

Definition 3 (Zadeh [13]) Fuzzy measure Π is called a possibility measure if there exists a function $\pi : X \to [0, 1]$ such that

$$\sup\{\pi(x) \mid x \in X\} = 1,$$

and

$$\Pi(A) = \sup\{\pi(x) \mid x \in A\}$$

holds for all $A \subset X$, where $\sup \pi(\emptyset) := 0$.

Definition 4 (Sugeno [2]) Fuzzy measure g_λ is called a λ-fuzzy measure if there exists a λ such that $-1 < \lambda < \infty$, and

$$g_\lambda(A \cup B) = g_\lambda(A) + g_\lambda(B) + \lambda g_\lambda(A) g_\lambda(B)$$

holds for any $A, B \in \mathscr{P}(X)$ such that $A \cap B = \emptyset$.

The above property of the λ-fuzzy measure is called λ-additivity.

Definition 5 (Okada and Tajima [8]) Fuzzy measure g_χ is called a χ-fuzzy measure if there exists a χ such that

$$\chi \geq 1 - \frac{\min_{x_i \in X} g(\{x_i\})}{\sum_{x_i \in X} g(\{x_i\})},$$

and

$$g_\chi(A) = \chi^{|A|-1}\left(\sum_{x_i \in A} g_\chi(\{x_i\})\right)$$

holds for any $A \in \mathscr{P}(X)$, where $|A|$ is the number of elements of A.

We next introduce the concept of "type" as a basis for fuzzy measure classification. Γ denotes the whole permutation of elements of $\mathscr{P}(X)$, and γ denotes an element of Γ. In other words,

$$\Gamma := \{\gamma = (A_1, \ldots, A_{2^n}) \mid A_i \in \mathscr{P}(X) \text{ for } i \in \{1, \ldots, 2^n\}, A_i \neq A_j \text{ for } i \neq j\}.$$

Note that $|\Gamma| = 2^n!$.

Definition 6 For $\gamma = (A_1, A_2, \ldots, A_{2^n}) \in \Gamma$, the set function $g : \mathscr{P}(X) \to [0,\ 1]$ is type γ ($\in \Gamma$) when

$$g(A_1) \leq g(A_2) \leq \cdots \leq g(A_{2^n}).$$

Note that the above definition also contains equalities, so a fuzzy measure may belong to multiple types simultaneously.

Definition 7 For $\gamma = (A_1, A_2, \ldots, A_{2^n}) \in \Gamma$, set function $g : \mathscr{P}(X) \to [0,\ 1]$ is strictly type γ ($\in \Gamma$) when

$$g(A_1) < g(A_2) < \cdots < g(A_{2^n}).$$

Theorem 8 Let $\mathscr{S}(\gamma)$ denote the set of all fuzzy measures of type γ. If a fuzzy measure g_0 is strictly type γ, any fuzzy measure $g \in \mathscr{S}(\gamma)$ is a distorted g_0. In particular, $\mathscr{S}(\gamma) = \{f \circ g_0 \mid f : [0,\ 1] \to [0,\ 1]$ is nondecreasing, $f(0) = 0$ and $f(1) = 1\}$.

Proof Since g_0 is strictly type γ and g is type γ, we have

$$g_0(A_1) < g_0(A_2) < \cdots < g_0(A_{2^n})$$

and

$$g(A_1) \leq g(A_2) \leq \cdots \leq g(A_{2^n}).$$

Thus, if we let $f(g_0(A_i)) = g(A_i), i = 1, \ldots, 2^n$ on point $g_0(A_i)$, and interpolate between points, say using straight lines, so they are nondecreasing, we obtain f as a nondecreasing function.

Theorem 8 yields that if class $\mathscr{S}(\gamma)$ contains a probability P which is strictly type γ, then by finding (i.e., identifying) this probability, any $g \in \mathscr{S}(\gamma)$ can be expressed as a distortion of P. This fact provides a starting point later in section 3 when we discuss the identification of fuzzy measures using distorted probability measures.

Theorem 9 Let g and g' be fuzzy measures. g is a distorted g' if
(C1) g and g_1 are type γ,
(C2) $g'(A) = g'(B)$ implies $g(A) = g(B)$ for any $A, B \in \mathscr{P}(X)$.

Proof By (C1), we have

$$\begin{aligned} &g(A_1) \leq g(A_2) \leq \cdots \leq g(A_{2^n}), \\ &g'(A_1) \leq g'(A_2) \leq \cdots \leq g'(A_{2^n}). \end{aligned} \quad \cdots\cdots (1)$$

Set the ranges of g and g' as

$$\begin{aligned} R(g) &:= \{g(A) | A \in \mathscr{P}(X)\} \subset [0,\ 1], \\ R(g') &:= \{g'(A) | A \in \mathscr{P}(X)\} \subset [0,\ 1]. \end{aligned}$$

For $f : R(g') \to R(g)$, we set $f(g'(A_i)) := g(A_i)$, $i = 1, \ldots, 2^n$. Thus, we can define f without contradiction by (C2), and f is a nondecreasing function on $R(g')$ by (1). We still must expand it so $f : [0,\ 1] \to [0,\ 1]$, which can be done by interpolating between points $(g'(A_i),\ g(A_i))$, $i = 1, 2, \ldots, 2^n$, using straight lines for instance, so the intervals are nondecreasing.

Theorem 10 (Honda and Okazaki [12]) Let g and g' be fuzzy measures. g is a distorted g' if
(C3) $g'(A) \leq g'(B)$ implies $g(A) \leq g(B)$ for any $A, B \in \mathscr{P}(X)$.

Proof As in the proof for Theorem 9, we set the ranges of g and g' as

$$\begin{aligned} R(g) &:= \{g(A) | A \in \mathscr{P}(X)\} \subset [0,\ 1], \\ R(g') &:= \{g'(A) | A \in \mathscr{P}(X)\} \subset [0,\ 1]. \end{aligned}$$

$g'(A) = g'(B)$ implies $g(A) = g(B)$ by (C3) so $f : R(g') \to R(g)$ can be defined without contradiction. To expand interpolate by, say, straight lines, just as in Theorem 9. f can be obtained as a nondecreasing function from (C3).

Corollary 11 (Honda and Okazaki [12]) Let g and g' be fuzzy measures. g is a distorted g' and the scaling function f can be obtained as a monotonically increasing function if $g'(A) \leq g'(B)$ if and only if $g(A) \leq g(B)$ for any $A, B \in \mathscr{P}(X)$.

Proof By the given condition $g'(A) = g'(B)$ if and only if $g(A) = g(B)$, and $g'(A) < g'(B)$ if and only if $g(A) < g(B)$. Thus, $f : R(g') \to R(g)$ is monotonically increasing from the proof of Theorem 9. By interpolating $f : R(g') \to R(g)$

using monotonically increasing functions (using straight lines piecewise fashion, for instance), we obtain monotonically increasing scaling function $f:[0,1]\to[0,1]$.

2.1. Distorted Probability

Distorted probability measures are measures obtained by distorting probabilities, which are linear measures, using scaling functions. A distorted probability measure is expressed by a combination of a probability and a scaling function. Since probabilities are additive, we can expect distorted probability measures to be relatively easy to treat.

Definition 12 A set function $P:\mathscr{P}(X)\to[0,1]$ is a probability measure if
(i) $P(\emptyset)=0,\ P(X)=1$,
(ii) for any $A,B\in\mathscr{P}(X)$, $P(A\cup B)=P(A)+P(B)$ whenever $A\cap B=\emptyset$.

From Theorem 10, a fuzzy measure is a distorted probability (Chateauneuf [11]) if $P(A)\le P(B)$ implies $g(A)\le g(B)$ for any $A,B\in\mathscr{P}(X)$. This means that the magnitude relationships of the measures of events A and B on probability P remain unchanged in the measures of g.

Probabilities have the following property:

$$P(A)\le P(B) \text{ implies } P(A\cup C)\le P(B\cup C)$$

for any $A,B,C\in\mathscr{P}(X)$ such that $A\cap C=B\cap C=\emptyset$ (Fine [14]).

This property is called consistency with direct sums. Since the magnitude relationship of events of distorted probability g is the same as that of probabilities, it follows similarly that

$$g(A)\le g(B) \text{ implies } g(A\cup C)\le g(B\cup C)$$

for any $A,B,C\in\mathscr{P}(X)$ such that $A\cap C=B\cap C=\emptyset$.

Note that the reverse does not hold. For instance, in the relation $g(x_1)<g(x_2)<g(x_3)=g(x_4)<g(x_1,x_2)<g(x_1,x_3)<g(x_1,x_4)<g(x_2,x_4)<g(x_2,x_3)<g(x_3,x_4)<g(x_1,x_2,x_3)<g(x_1,x_2,x_4)<g(x_1,x_3,x_4)<g(x_2,x_3,x_4)$, g is consistent with the direct sums, but it is not a distorted probability. (cf. also Narukawa and Torra [15]).

The discussion so far shows that, when discussing distorted probability measures, we must consider the magnitude relationship of the measures for events. In this regard, the concept of type introduced earlier becomes important.

Example 13 ([16]) All λ-fuzzy measures are distorted probabilities.

Proof This is evident by defining as the probability and scaling function

$$P(A):=\log_{(1+\lambda)}(1+\lambda g_\lambda(A))$$

for $A\in\mathscr{P}(X)$, and

$$f(x):=-\frac{1}{\lambda}+\frac{1}{\lambda}(1+\lambda)^x$$

for $x\in[0,1]$.

Example 14 All possibility measures on a finite set are distorted probabilities.

Proof For a finite set $X=\{x_1,x_2,\dots,x_n\}$, we can suppose $\Pi(\{x_1\})\le\Pi(\{x_2\})\le\cdots\le\Pi(\{x_n\})$ without loss of generality. Choose a probability such that

$$\sum_{i=1}^{k}P(\{x_i\})<P(\{x_{k+1}\})$$

and interpolate points $(P(\{x_i\}),\Pi(\{x_i\}))$ and $(\sum_{i=1}^{k}P(\{x_i\}),\Pi(\{x_k\}))$ using straight lines to obtain nondecreasing scaling function f. For instance, let $P(\{x_i\}):=2^{i-1}/(2^n-1)$.

Fact 15 Let X be a two-point set $X=\{x_1,x_2\}$, and $g(x_1)\le g(x_2)$. Then, all fuzzy measure is of the type:

$$\gamma_1=(\emptyset,\{x_1\},\{x_2\},X).$$

Moreover, any fuzzy measure is a distorted probability.

Fact 16 Let X be a three-point set $X=\{x_1,x_2,x_3\}$ and $g(x_1)\le g(x_2)\le g(x_3)$. Then, all fuzzy measure can be classified into one or more of the following eight types. Of these, those under types γ_1 and γ_2 are distorted probability measures.

$$\begin{aligned}
\gamma_1 &= (\emptyset,x_1,x_2,x_1x_2,x_3,x_1x_3,x_2x_3,X)\\
\gamma_2 &= (\emptyset,x_1,x_2,x_3,x_1x_2,x_1x_3,x_2x_3,X)\\
\gamma_3 &= (\emptyset,x_1,x_2,x_3,x_1x_3,x_1x_2,x_2x_3,X)\\
\gamma_4 &= (\emptyset,x_1,x_2,x_3,x_1x_3,x_2x_3,x_1x_2,X)\\
\gamma_5 &= (\emptyset,x_1,x_2,x_1x_2,x_3,x_2x_3,x_1x_3,X)\\
\gamma_6 &= (\emptyset,x_1,x_2,x_3,x_1x_2,x_2x_3,x_1x_3,X)\\
\gamma_7 &= (\emptyset,x_1,x_2,x_3,x_2x_3,x_1x_2,x_1x_3,X)\\
\gamma_8 &= (\emptyset,x_1,x_2,x_3,x_2x_3,x_1x_3,x_1x_2,X).
\end{aligned}$$

Fact 17 Let X be a four-point set $X=\{x_1,x_2,x_3,x_4\}$, and $g(x_1)\le g(x_2)\le g(x_3)\le g(x_4)$. All fuzzy measure can be classified into one of 70016 types. Of these, those under the following 14 types are distorted probability measures.

$$\begin{aligned}
\gamma_1 = (&\emptyset,x_1,x_2,x_3,x_4,x_1x_2,x_1x_3,x_1x_4,x_2x_3,x_2x_4,x_3x_4,\\
&x_1x_2x_3,x_1x_2x_4,x_1x_3x_4,x_2x_3x_4,X)\\
\gamma_2 = (&\emptyset,x_1,x_2,x_3,x_4,x_1x_2,x_1x_3,x_2x_3,x_1x_4,x_2x_4,x_3x_4,\\
&x_1x_2x_3,x_1x_2x_4,x_1x_3x_4,x_2x_3x_4,X)\\
\gamma_3 = (&\emptyset,x_1,x_2,x_3,x_1x_2,x_4,x_1x_3,x_1x_4,x_2x_3,x_2x_4,\\
&x_1x_2x_3,x_3x_4,x_1x_2x_4,x_1x_3x_4,x_2x_3x_4,X)\\
\gamma_4 = (&\emptyset,x_1,x_2,x_3,x_1x_2,x_4,x_1x_3,x_2x_3,x_1x_4,x_2x_4,\\
&x_1x_2x_3,x_3x_4,x_1x_2x_4,x_1x_3x_4,x_2x_3x_4,X)\\
\gamma_5 = (&\emptyset,x_1,x_2,x_3,x_1x_2,x_1x_3,x_4,x_1x_4,x_2x_3,x_1x_2x_3,\\
&x_2x_4,x_3x_4,x_1x_2x_4,x_1x_3x_4,x_2x_3x_4,X)
\end{aligned}$$

$\gamma_6 = (\emptyset, x_1, x_2, x_3, x_1x_2, x_1x_3, x_4, x_2x_3, x_1x_4, x_1x_2x_3, x_2x_4, x_3x_4, x_1x_2x_4, x_1x_3x_4, x_2x_3x_4, X)$

$\gamma_7 = (\emptyset, x_1, x_2, x_3, x_1x_2, x_1x_3, x_2x_3, x_4, x_1x_2x_3, x_1x_4, x_2x_4, x_3x_4, x_1x_2x_4, x_1x_3x_4, x_2x_3x_4, X)$

$\gamma_8 = (\emptyset, x_1, x_2, x_3, x_1x_2, x_1x_3, x_2x_3, x_1x_2x_3, x_4, x_1x_4, x_2x_4, x_3x_4, x_1x_2x_4, x_1x_3x_4, x_2x_3x_4, X)$

$\gamma_9 = (\emptyset, x_1, x_2, x_1x_2, x_3, x_4, x_1x_3, x_1x_4, x_2x_3, x_2x_4, x_1x_2x_3, x_1x_2x_4, x_3x_4, x_1x_3x_4, x_2x_3x_4, X)$

$\gamma_{10} = (\emptyset, x_1, x_2, x_1x_2, x_3, x_4, x_1x_3, x_2x_3, x_1x_4, x_2x_4, x_1x_2x_3, x_1x_2x_4, x_3x_4, x_1x_3x_4, x_2x_3x_4, X)$

$\gamma_{11} = (\emptyset, x_1, x_2, x_1x_2, x_3, x_1x_3, x_4, x_1x_4, x_2x_3, x_1x_2x_3, x_2x_4, x_1x_2x_4, x_3x_4, x_1x_3x_4, x_2x_3x_4, X)$

$\gamma_{12} = (\emptyset, x_1, x_2, x_1x_2, x_3, x_1x_3, x_4, x_2x_3, x_1x_4, x_1x_2x_3, x_2x_4, x_1x_2x_4, x_3x_4, x_1x_3x_4, x_2x_3x_4, X)$

$\gamma_{13} = (\emptyset, x_1, x_2, x_1x_2, x_3, x_1x_3, x_2x_3, x_4, x_1x_2x_3, x_1x_4, x_2x_4, x_1x_2x_4, x_3x_4, x_1x_3x_4, x_2x_3x_4, X)$

$\gamma_{14} = (\emptyset, x_1, x_2, x_1x_2, x_3, x_1x_3, x_2x_3, x_1x_2x_3, x_4, x_1x_4, x_2x_4, x_1x_2x_4, x_3x_4, x_1x_3x_4, x_2x_3x_4, X)$

When X is a three-point set, λ-fuzzy measures are type γ_1 or γ_2. Possibility measures are type γ_1. For a four-point set X, λ-fuzzy measures are type $\gamma_1, \ldots, \gamma_{13}$ or γ_{14}. Possibility measures are type γ_{14}.

Fact 18 Let X be a five-point set $X = \{x_1, x_2, x_3, x_4, x_5\}$ and $g(\{x_1\}) \leq g(\{x_2\}) \leq g(\{x_3\}) \leq g(\{x_4\}) \leq g(\{x_5\})$. Distorted probability measures are classified into 546 types, obtained by adding event sequences that include x_5 to the above 14 types (Fact 17) on $\{x_1, x_2, x_3, x_4\}$. No general algorithm has been established, however.

Above Facts 15-18 show that fuzzy measures have a much higher degree of freedom than probability measures.

3. Identification of Fuzzy Measures

Let X be an n-point set $X = \{x_1, x_2, \ldots, x_n\}$. Suppose that the fuzzy measures of partial events $\mathscr{A} \subset \mathscr{P}(X)$ are given, and that it has the value $g(A), A \in \mathscr{A}$. In this section, we determine fuzzy measures of all events in $\mathscr{P}(X) \setminus \mathscr{A}$, i.e., we discuss the identification problem of fuzzy measures. We wish to minimize squared errors of identification results relative to true values (least-squares estimation).

3.1. Simple Method

If the only information given consists of the fact that g is a fuzzy measure, then we know only that g is monotonic. The value of an unknown fuzzy measure $g(A), A \in \mathscr{P}(X) \setminus \mathscr{A}$ satisfies

$$m \leq g(A) \leq M$$

where $m := \max\{g(B) \mid A \supset B \in \mathscr{A}\}$ and $M := \min\{g(C) \mid A \subset C \in \mathscr{A}\}$.

Table 1. Probabilities assigned to $X = \{x_1, x_2, x_3, x_4\}$.

type	$P(x_1)$	$P(x_2)$	$P(x_3)$	$P(x_4)$
γ_1	5/29	7/29	8/29	9/29
γ_2	4/23	5/23	6/23	8/23
γ_3	3/33	8/33	10/33	12/33
γ_4	4/28	5/28	8/28	11/28
γ_5	2/25	6/25	7/25	10/25
γ_6	2/19	4/19	5/19	8/19
γ_7	2/21	4/21	5/21	10/21
γ_8	2/23	4/23	5/23	12/23
γ_9	2/25	4/25	9/25	10/25
γ_{10}	3/25	4/25	8/25	10/25
γ_{11}	2/29	6/29	9/29	12/29
γ_{12}	2/25	4/25	8/25	11/25
γ_{13}	2/27	4/27	8/27	13/27
γ_{14}	1/15	2/15	4/15	8/15

Although the estimated value can lie anywhere within the interval $[m, M]$, if we assume that the probability that the unknown data exists is the same throughout the range, the distribution of unknown data follows a uniform distribution $U[m, M]$. The estimated value that produces the smallest squared error is then given by $(m + M)/2$.

3.2. Proposed Method – Identification Using Distorted Probability Measures

If monotonicity is the only information available on data, then the best method is the simple one described above. In real-world situations, however, data is likely to have certain characteristics. When dealing with human subjective evaluations, for instance, it is natural to think that data will possess some property akin to additivity, even if additivity itself does not hold. We propose an estimation in which analysis is restricted to the class of distorted probabilities. Although distorted probability measures are not additive, magnitude relationships of probability measures are preserved. Estimating a fuzzy measure means, in this case, to determine probability P and scaling function f for g.

The technique consists of estimating the type of fuzzy measure from known data, then determining the probability of each singleton and nondecreasing function f. The type is selected from the magnitude relationship of known data. When several candidates exist for the type, the selection is narrowed based on the magnitude relationship assuming that additivity holds. If a suitable type cannot be found, known data can be modified after which the type is determined (see Appendix B). Probabilities can be assigned arbitrarily, in this paper, the probabilities are given for each type in advance (**Table 1**). Here, probabilities are given so they increase in equal increments when events are listed in ascending order of size. Thus, for type γ_{14}, the probability measures of $\{\emptyset, x_1, x_2, x_1x_2, x_3, x_1x_3, x_2x_3, x_1x_2x_3, x_4, x_1x_4, x_2x_4, x_1x_2x_4, x_3x_4, x_1x_3x_4, x_2x_3x_4, X\}$ are given as $\{0, \frac{1}{15}, \frac{2}{15}, \frac{3}{15}, \frac{4}{15}, \frac{5}{15},$

$\frac{6}{15}, \frac{7}{15}, \frac{8}{15}, \frac{9}{15}, \frac{10}{15}, \frac{11}{15}, \frac{12}{15}, \frac{13}{15}, \frac{14}{15}, 1\}$. Function f is determined as a piecewise linear function, as was done earlier in the proof of Theorem 9. The value of unknown data is obtained from $g(A) = f \circ P(A)$. When a measure of set A is given as known data, and $g(A) = \alpha$, the estimated value is $f(P(A)) = g(A) = \alpha$ as long as no modifications (as mentioned above) have been introduced, and no error will be introduced to the known data. In contrast, known data is changed and errors are introduced when λ- and χ-fuzzy measures are used for estimation.

When conducting piecewise linear approximation on nondecreasing function $f : [0,1] \to [0,1]$, where nodes are denoted $(x_i, f(x_i)), i = 1,2,\ldots,n$, a natural way to choose dividing points $\{x_i \mid i = 1,2,\ldots,n\}$ is to divide the interval into n equal subintervals. This is why we selected probability values so they are equally spaced in interval $[0,1]$ above. In actual practice, probabilities are given in the simplest fractional expressions possible so the probability distribution can be visualized – the given values are not necessarily the best. If some information is known at the outset about the shape of function f, then the probabilities assigned to the dividing points $\{x_i \mid i = 1,2,\ldots,n\}$ should be modified accordingly. If, for instance, f is expected to be a function with a "reverse S" shape, where $f(x)$ rises sharply in the vicinities of $x = 0$ and $x = 1$, it is better to assign probabilities at closer proximity in the vicinity of 0 – the probabilities of the complementary sets are necessarily assigned at closer proximity in the vicinity of 1.

3.3. Identification Using λ-Fuzzy Measures

We now identify fuzzy measures when they are limited to the class of λ-fuzzy measures [1]. Identifying fuzzy measures in this case means estimating parameter λ. Thus, the λ value that minimizes

$$\sum_{A \in \mathscr{A}} (g(A) - g_\lambda(A))^2$$

is obtained. This is expressed by a $(n-1)^2$-th degree equation with respect to λ when X is an n-point set. The estimated λ is used to determine unknown data. The known data is also recalculated, so an error is normally introduced between known data and recalculated values. If $\mathscr{A}$ does not contain all singletons in which case, estimation is still possible, λ is estimated and used to determine measures of unknown singletons, which are used along with the estimated value of λ for identification.

3.4. Identification Using χ-Fuzzy Measures

Here we identify fuzzy measures when they are limited to the class of χ-fuzzy measures ([7, 8]). χ-fuzzy measures also have a property similar to additivity, as shown in Definition 5. As in the case of λ-fuzzy measures, identifying a χ-fuzzy measure means to estimate a suitable value for parameter χ. In other words, the χ value that minimizes

$$\sum_{A \in \mathscr{A}} (g(A) - g_\chi(A))^2$$

is obtained. For an n-point set X, this is obtained by solving a $(n-1)^2$-th degree equation with respect to χ. Unknown data is determined from the estimated χ-value. As in the case of λ-fuzzy measures, known data is recalculated, which normally introduces errors with known data. If $\mathscr{A}$ does not contain all singletons, identification can still be done in the manner of λ-fuzzy measures above.

4. Case Study

4.1. Data Used

Data was collected using a questionnaire. Since most people find it difficult to give their subjective evaluations qualitatively, pair comparison is normally used to reduce the burden on the subject, although this requires a great many responses. We had subjects give their evaluations intuitively in the form of a number rating. Two questions were asked, framed to make it easier for subjects to answer. From the answers obtained from ten subjects, 20 data samples were collected.

1. The subject was asked to name some product s/he is interested in obtaining, along with four evaluation items s/he would consider when making a selection. The subject was then asked to rate the degree to which s/he would likely purchase the product when only one item, two items, three items, and all four items were satisfactorily provided. A personal computer, for example, may have the four evaluation items of expandability, CPU speed, stability, and functions. When the computer only has a high degree of expandability, it means that the other items are not provided satisfactorily.

2. The subject was asked to name four entertainers or television personalities who s/he generally favors (or at least does not dislike), and rate how much s/he would want to view a TV program in which each one alone, two together, three together, and all four appear.

Subjects were asked to give their ratings based on a 100-point scale.

The four items are denoted x_1, x_2, x_3 and x_4 (where $g(x_1) \leq g(x_2) \leq g(x_3) \leq g(x_4)$), and normalized so the measure of $X = \{x_1, x_2, x_3, x_4\}$ equals unity. It was assumed that $g(x_1), g(x_2), g(x_3), g(x_4), g(x_1x_2), g(x_3x_4), g(x_1x_2x_3)$ and $g(X)$ are known; these were used to identify remaining data values using the four methods above, and results were compared. The six unknown data consist of $g(x_1x_3), g(x_1x_4)$, $g(x_2x_3)$, $g(x_1x_2x_4), g(x_1x_3x_4)$ and $g(x_2x_3x_4)$. Steepest descent was used to estimate parameters λ and χ in methods 3.3 and 3.4. For identification using distorted probabilities, the type was first estimated from known data. There is no need to modify data if the order of the known data matches one of the

Table 2. Squared errors and dispersions of estimated values.

Estimation		simple method	distorted probability	λ-fuzzy measure	χ-fuzzy measure
Err_1	Squared error	0.1189	0.0390	0.0837	0.0871
	Variance	0.0081	0.0006	0.0023	0.0030
Err_2	Squared error	0.1189	0.0386	0.0531	0.0549
	Variance	0.0081	0.0006	0.0010	0.0012

following $k_1, \ldots, k_5$:

$$k_1 = \{\emptyset, x_1, x_2, x_3, x_4, x_1x_2, x_3x_4, x_1x_2x_3, X\}$$
$$k_2 = \{\emptyset, x_1, x_2, x_3, x_1x_2, x_4, x_1x_2x_3, x_3x_4, X\}$$
$$k_3 = \{\emptyset, x_1, x_2, x_3, x_1x_2, x_1x_2x_3, x_4, x_3x_4, X\}$$
$$k_4 = \{\emptyset, x_1, x_2, x_1x_2, x_3, x_4, x_1x_2x_3, x_3x_4, X\}$$
$$k_5 = \{\emptyset, x_1, x_2, x_1x_2, x_3, x_1x_2x_3, x_4, x_3x_4, X\}$$

For example, if known data matches k_1, the type is either γ_1 or γ_2 (see Theorem 10). If $g(x_1) + g(x_4) \leq g(x_2) + g(x_3)$, then the estimated type is γ_1, otherwise it is γ_2. If known data do not match any of k_1-k_5, then they are modified so they match one of the k's, after which the type is estimated. The method of modification is presented in Appendix B, and an example of identification in Appendix C.

4.2. Results and Discussion

Results of identification obtained from the four methods and the means of the sums of squared errors and their variances are presented in **Table 2**. Squared errors when known data was replaced by identification results (Err_1) and those when original known data was used (Err_2) are both presented. Because the proposed method uses the original known data when there is no need to modify it, it has an advantage over other methods in Err_1.

$$Err_1 = \frac{1}{20}\sum_{i=1}^{20}\sum_{A\in\mathscr{P}(X)}(g(A)-\bar{g}(A))^2$$
$$Err_2 = \frac{1}{20}\sum_{i=1}^{20}\sum_{A\in\mathscr{P}(X)\setminus\mathscr{A}}(g(A)-\bar{g}(A))^2,$$

where $\bar{g}$ is the estimated value of g.

While fairly satisfactory identification results were produced with all of the methods tested in the present case study, the method using distorted probabilities produced the best results when variances are taken into account. In terms of the number of best identification based on squared error Err_1 among the 20 data samples, 15 were by distorted probabilities, 3 by λ-fuzzy measures, and 1 by the simple method. When based on squared error Err_2, 13 best identification results were by distorted probabilities, 6 by χ-fuzzy measures, and 1 by the simple method.

Table 3. Squared errors E_2 for simple method and distorted probabilities.

Estimation	Simple method	distorted probability
Squared error	0.1182	0.0726
Variance	0.0071	0.0076

5. Human Subjective Evaluation

We now examine which of the fuzzy measure class best approaches subjective evaluations of the subjects in the present study.

5.1. Distorted Probability

All data satisfied monotonicity, so they are fuzzy measures. The type of the data is examined by comparing it with those of distorted probability measures. If there is no matching type, then the data is modified into each of the types of the distorted probability measures (see Appendix B), and the type that gives the smallest squared errors between the modified and original data is chosen. This squared error is considered as that between the data and the distorted probability measures. The mean and variance of squared errors between data and distorted probability measures were 0.0058 and 0.0000, respectively. If data matches two or more types, then the type with the smallest squared errors is chosen as the data type. Of the 20 data samples in the present study, there was one sample that was found to be distorted probabilities, requiring no modification. We next assumed that one of the data element was unknown and estimated its value using the simple method and the proposed one. Since X is a four-point set, each data set contains 16 values (elements), which is the number of subsets of X. Of the 14 data elements that remain after the empty and universal sets are excluded, one was assumed to be unknown. The squared error of the estimated value $\bar{g}$ against the true value g was obtained as follows:

$$E_2 = \frac{1}{20}\sum_{i=1}^{20}\sum_{A\in\mathscr{P}(X)}(g(A)-\bar{g}(A))^2$$

and the mean of the 20 squared errors for the 20 samples were obtained (see **Table 3**).

5.2. λ-Fuzzy Measures

λ values giving the best fit, i.e., with the smallest squared error, for each of the 20 data samples were com-

Table 4. Squared errors between data and fuzzy measures.

Estimation	Simple method	distorted probability	λ-fuzzy measure	χ-fuzzy measure
Squared error	0.0000	0.0058	0.0784	0.1323
Variance	0.0000	0.0000	0.0021	0.0630

puted, and the squared errors between the λ-fuzzy measures and the data were obtained as follows:

$$E_\lambda = \sum_{A\in\mathscr{P}(X)} (g(A) - g_\lambda(A))^2.$$

Then the mean of the 20 squared errors were obtained (**Table 4**). These λ values were used as the true λ values in section 4.2.

5.3. χ-Fuzzy Measures

χ values were obtained in a similar manner as for λ-fuzzy measures, and squared errors between χ-fuzzy measures and data were obtained as follows:

$$E_\chi = \sum_{A\in\mathscr{P}(X)} (g(A) - g_\chi(A))^2.$$

Then the mean of the 20 squared errors were obtained (**Table 4**). These χ values were used as the true χ values in section 4.2.

5.4. Results and Discussion

As **Table 3** shows, mean squared errors for distorted probabilities are smaller than those for the simple method, so we can state that data obtained for the present case are closer to distorted probabilities than they are to simple fuzzy measures. Such results are to be expected if we consider that distorted probabilities possess a high degree of freedom. In **Table 4** as well, the distorted probabilities produce the smallest mean squared errors. These results indicate that the evaluation scales of subjects in the present study are represented better as distorted probabilities rather than λ- or χ-fuzzy measures. However, this is not to suggest that the use of λ- or χ-fuzzy measures is inappropriate. In fact, results show that, on average, identification using λ-fuzzy measures produces sufficiently satisfactory results. The above analysis is based on the means and variance of squared errors. Of the 20 data samples, the smallest squared errors was obtained in 13 when estimated with distorted probabilities, 4 with χ-fuzzy measures, 2 with λ-fuzzy measures, and 1 with the simple method.

6. Issues for Future Investigation

Our proposal produced superior results as compared to previous conventional methods. We can state, furthermore, that subjective evaluation scales of the subjects in the present study are represented fairly well with distorted probabilities. It was found that distorted probabilities may better express subjective evaluation scales. We plan to implement further case studies in the future to verify this.

When there are more evaluation items (five, six or more) than in the present study, it is still necessary to assign probabilities to all types. This is possible, but not practical. In such cases, it will be necessary to select some suitable representative types.

Subjects were asked to give their intuitive ratings in the present case study. The prevailing view among subjects was that this placed a burden on them, and that pair comparison would have been easier to answer. Pair comparison requires $(2^n \times (2^n - 1))/2$ comparisons. Since $n = 4$ in the present case study, 105 pair comparisons are necessary, which is rather impractical. A more realistic method may be to obtain part of fuzzy measures by pair comparison, and estimate the remaining ones.

References:

[1] K. Asai, "Introduction to fuzzy management sciences," Ohmsha, 1992 (in Japanese).

[2] M. Sugeno, "Fuzzy measures and fuzzy integrals: a survey," in M. M. Gupta, G. N. Saridis, and B. R. Gains, editors, "Fuzzy automata and decision processes," pp. 89-102, North Holland, Amsterdam, 1977.

[3] T. Terano, K. Asai, and M. Sugeno (ed.), "Introduction to applied fuzzy systems," Ohmsha, 1989 (in Japanese).

[4] T. Terano, K. Asai, and M. Sugeno, "Fuzzy Systems Theory and its Applications," Academic Press, Sandiego, CA, 1992 (English translation of [3]).

[5] Z. Wang, and G. J. Klir, "Fuzzy Measure Theory," Plenum Press, New York and London, 1992.

[6] A. Honda, T. Nakano, and Y. Okazaki, "Subjective evaluation based on distorted probability," Joint Int. Conf. on Soft Computing and Intelligent Systems and 3rd Int. Symp. on Advanced Intelligent Systems, Tsukuba, Japan, October, 2002.

[7] I. Okada, and H. Tajima, "Analytic hierarchy processes with χ-fuzzy measures," The 12-th fuzzy system symposium, 1996, pp. 875-878 (in Japanese).

[8] I. Okada, and H. Tajima, "Mathematical properties and identifications of χ-fuzzy measures," Proceedings of the first workshop on evaluation of heart and mind, 1996, pp. 35-36 (in Japanese).

[9] M. Grabisch, T. Murofushi, and M. Sugeno, eds., "Fuzzy Measures and Integrals: Theory and Applications," Physica-Verlag, 2000.

[10] M. Sugeno, "Theory of fuzzy integrals and its applications," Doctoral Thesis, Tokyo Institute of Technology, 1974.

[11] A. Chateauneuf, "Decomposable capacities distorted to probabilities and concave capacities," Math. Social Sci., 31, pp. 19-37, 1996.

[12] A. Honda, and Y. Okazaki, "Invariance of fuzzy measure," Bull. of the Kyushu Inst. of Tech., pure and appl. Math., 46, pp. 1-8, 1999.

[13] L. A. Zadeh, "Fuzzy Sets as a Basis for a Theory of Possibility," Fuzzy Sets and Systems, 1, pp. 3-28, 1978.

[14] T. L. Fine, "Theories of Probability," Academic Press, New York, 1982.

[15] Y. Narukawa, and V. Torra, "Fuzzy measure and probability distributions: distorted probabilities," IEEE Trans. on Fuzzy Systems, to appear.

[16] R. Kruse, "A note on λ-additive fuzzy measures," Fuzzy Sets and Systems, Vol.8, pp. 219-222, 1982.

```
D(x[N], y[N]){
    i=0; /* point location of x*/
    j=0; /* point location of y*/
    k=0; /* counter for different elements */
    while i<N do {
        if list1 is empty then {
            / **** /
            while x[i] != y[j] do {
                add y[j] to end of list1[];
                j++;
                k++;
            }
            i++; j++;
            / **** /
        }
        else {
            /* find element that matches x[i] from list1[] */
            k++;
            if found then{
                delete that element from list1[], and carry over the
                other elements on the right by one position to the left;
                i++;
                if list1[] {\it is empty} then {
                    add k to list2[];
                    k=0;
                }
            }
            else /* not found */ {
                repeat steps in / **** / ` / **** /;
            }
        }
    }
    return the square root of sum of squares of elements of list2[];
}
```

Fig. 1. Function for returning similarity between lists.

Appendix A. Similarity of Types

We introduce the concept of affinity between types to quantify the similarity of two types.

Definition 19 Let $\theta = (m_1, \ldots, m_k)$ such that $\sum_{i=1}^{k} m_i = n$ and $m_i > 0$ be a partition of n elements. The partition of list $\gamma := (A_1, \ldots, A_n)$ by θ is given by

$$((A_1, \ldots, A_{m_1}), (A_{m_1+1}, \ldots, A_{m_1+m_2}), \ldots, \\ (A_{n-m_k+1}, \ldots, A_n)) =: (B_1, \ldots, B_k),$$

where $|B_i| = m_i$. Let Θ denote the set of all θ. Let γ and γ' be two lists consisting of the same n elements. The distance between γ and γ' is defined as follows:

$$D(\gamma, \gamma') = \inf_{\Theta} \left\{ \sqrt{\sum_{B_i \neq B_i'} |B_i|^2} \right\}.$$

$D(\gamma, \gamma')$ can be considered to be the difference between two types, so the smaller the value of D the more similar are the two types. The program code of a function that returns D is presented in **Fig.1**.

Example 20 For $\alpha_1 = (1,2,3,4,5,6,7,8,9,10), \alpha_2 = (1,3,2,4,5,8,7,6,9,10), \alpha_3 = (1,2,4,5,6,3,7,8,9,10)$, we have

$$D(\alpha_1, \alpha_2) = \sqrt{|\{2,3\}|^2 + |\{6,7,8\}|^2} = \sqrt{13},$$
$$D(\alpha_2, \alpha_3) = \sqrt{|\{3,2,4,5,8,7,6\}|^2} = 7.$$

Appendix B. Modification of Data

We describe how to modify fuzzy measures g of type γ to type γ'. Let θ be a partition of γ and γ', for which $D(\gamma, \gamma')$, and let $\gamma = \{B_1, \ldots, B_m\}$ and $\gamma' = \{B_1', \ldots, B_m'\}$ be the partitions of γ and γ', by θ. For the subset $B_i \neq B_i'$, replace the data with the mean of the data in B_i.

Example 21 To modify

$$g(C) \leq g(B) \leq g(A) \leq g(E) \leq g(D),$$

or type (C, B, A, E, D), to

$$g'(A) \leq g'(B) \leq g'(C) \leq g'(D) \leq g'(E),$$

or type (A, B, C, D, E), the partition is given by $(B_1, B_2) = ((C, B, A), (E, D))$, $(B_1', B_2') = ((A, B, C), (D, E))$, where

Table 6. Estimation results of methods (underlining denotes known data).

	$\underline{x_1}$	$\underline{x_2}$	$\underline{x_3}$	$\underline{x_4}$	$\underline{x_1,x_2}$	x_1,x_3	x_1,x_4	x_2,x_3
Original data	0.10	0.10	0.15	0.20	0.25	0.30	0.40	0.35
Simple method	0.1000	0.1000	0.1500	0.2000	0.2500	0.3750	0.6000	0.3750
Proposed method	0.1000	0.1000	0.1500	0.2250	0.2250	0.2250	0.4750	0.3500
λ-fuzzy measure	0.1000	0.1000	0.1500	0.2000	0.2359	0.3036	0.3715	0.3036
χ-fuzzy measure	0.1000	0.1000	0.1500	0.2000	0.2500	0.3125	0.3750	0.3125
	x_2,x_4	$\underline{x_3,x_4}$	$\underline{x_1,x_2,x_3}$	x_1,x_2,x_4	x_1,x_3,x_4	x_2,x_3,x_4	$\underline{X}$	
Original data	0.50	0.85	0.60	0.80	0.90	0.95	1.00	
Simple method	0.6000	0.8500	0.6000	0.6250	0.9250	0.9250	1.00	
Proposed method	0.7250	0.8500	0.6000	0.8750	0.9000	0.9500	1.000	
λ-fuzzy measure	0.3715	0.4573	0.5122	0.6044	0.7208	0.7208	0.9450	
χ-fuzzy measure	0.3750	0.4375	0.5468	0.6249	0.7030	0.7030	1.074	

Table 7. Squared errors of estimated values by various methods relative to true values.

Estimation	Simple method	Proposed method	λ-fuzzy measure	χ-fuzzy measure
Err_1	0.0881	0.0688	0.3074	0.3268
Err_2	0.0881	0.0619	0.1424	0.1483

Table 5. Response of one subject.

x_1	x_2	x_3	x_4
10	10	15	20
x_1,x_2	x_1,x_3	x_1,x_4	x_2,x_3
25	30	40	35
x_2,x_4	x_3,x_4	x_1,x_2,x_3	x_1,x_2,x_4
50	85	60	80
x_1,x_3,x_4	x_2,x_3,x_4	X	
90	95	100	

$\theta = (3,2)$. The modified values are:

$$g(C) = g(B) = g(A) = \frac{g(C)+g(B)+g(A)}{3},$$

$$g(E) = g(D) = \frac{g(E)+g(D)}{2}.$$

In sections 3.5 and 5.1, the types were modified to minimize the squared errors relative to unmodified data.

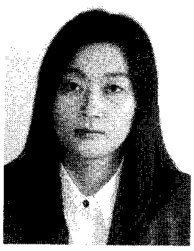
Appendix C. Estimation Example

The answer given by one of the subjects to question 1 is presented in **Table 5.**

The evaluation items for purchase of a notebook computer x_1 = screen size, x_2 = performance, x_3 = light weight, x_4 = stability

All ratings are divided by 100 to normalize to $g(X) = 1$. The fuzzy measures are of the type $\gamma = (\emptyset, x_1, x_2, x_3, x_4, x_1x_2, x_1x_3, x_2x_3, x_1x_4, x_2x_4, x_1x_2x_3, x_1x_2x_4, x_3x_4, x_1x_3x_4, x_2x_3x_4, X)$. Modification is necessary because the type does not match any of those for distorted probability measures. By making the modifications: $g'(x_1x_2) = g'(x_4) = (g(x_1x_2)+g(x_4))/2 = 0.225$ and $g'(x_3x_4) = g'(x_1x_2x_4) = (g(x_3x_4) + g(x_1x_2x_4))/2 = 0.825$, the type becomes γ_4 . This is also when the squared error is the smallest. Thus, γ_4 is the true type. ($D(\gamma,\gamma_4) = 2\sqrt{2}$) Assume now that the known data consist of $\{g(\emptyset), g(x_1), g(x_2), g(x_3), g(x_4), g(x_1,x2), g(x_3,x_4), g(x_1,x_2,x_3), g(X)\}$. In the proposed estimation method, the list of known data is given by $k = (\emptyset, \{x_1\}, \{x_2\}, \{x_3\}, \{x_1,x2\}, \{x_4\}, \{x_3,x_4\}, \{x_1,x_2,x_3\}, X)$. To choose the most similar from among k_1-k_5, we modify as $g'(x_1x_2) = g'(x_4) = (g(x_1x_2) + g(x_4))/2 = 0.225$, which results in k_2 ($D(k,k_2) = 2$). Since this also results in the smallest squared error (when compared before and after modification), we select k_2. The estimated type is either γ_5 or γ_6. Since $g(\{x_1\}) + g(\{x_4\}) \geq g(\{x_2\}) + g(\{x_3\})$, the estimated type is γ_6. Estimation using λ-fuzzy measures obtained $\lambda = 3.5758$, while that using χ-fuzzy measures obtained $\chi = 1.2499$. Estimation results of different methods are presented in **Table 6.**

The squared errors are presented in **Table 7.**



Name:
Aoi Honda

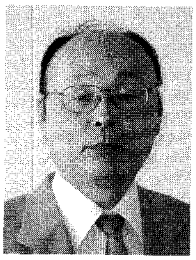
Affiliation:
Assistant, Department of Systems Innovation and Informatics, Kyushu Institute of Technology, and Visiting Scientist, Department of Computer science and Mathematics, University of Paris I, Panthéon-Sorbonne

Address:
680-4 Kawazu, Iizuka 820-8502, Japan, and 72, rue Regnault, 75013 Paris, France

Brief Biographical History:
Assistant, Department of Systems Innovation and Informatics, Kyushu Institute of Technology

Main Works:
- "Canonical fuzzy measure on (0, 1]," Fuzzy Sets and Systems, Vol.123, pp. 147-150, 2001.
- "Random translation of a product measure on R^∞," Proceedings of International Symposium on Banach and Function Spaces 2003, YokohamaTosho, 2004.

Membership in Learned Societies:
- The Mathematical Society of Japan
- The Institute of Electronics, Information and Communication Engineers
- Information Processing Society of Japan



Name:
Yoshiaki Okazaki

Affiliation:
Professor, Department of Systems Innovation and Informatics, Kyushu Institute of Technology

Address:
680-4 Kawazu, Iizuka 820-8502, Japan

Brief Biographical History:
1989- Assistant Professor, Department of Artificial Intelligence, Kyushu Institute of Technology
1993- Professor, Department of Systems Innovation and Informatics, Kyushu Institute of Technology

Main Works:
- "Subjective Evaluation Process by Multi-Dimensional Fuzzy Measure," SCIS & ISIS 2004, Keio Univ., Yokohama, Japan.

Membership in Learned Societies:
- The Mathematical Society of Japan
- The Institute of Electronics, Information and Communication Engineers
- Japan Society for Fuzzy Theory and Intelligent Informatics