ICIC Express Letters
Volume **12**, Number **8**, August **2018**

ICIC International ©2018 ISSN 1881-803X
pp. **791–798**

TYPE-2 FUZZY CONTINGENCY TABLE AND SIMILARITY INDICES

Hiroaki Uesu

Practical Engineering Education Program
Kanazawa Institute of Technology
7-1 Ohgigaoka, Nonoichi, Ishikawa 921-8501, Japan
uesu@neptune.kanazawa-it.ac.jp

Received January 2018; accepted April 2018

Abstract. *The contingency table analysis provides the interrelation between two variables and can help find interactions between them. Previously, we have extended the contingency table by applying fuzzy theory and defined (type-1) fuzzy contingency table. Further, we extend type-1 fuzzy contingency table to type-2 in order to analyze inexact information. In this paper, we define the type-2 2×2 fuzzy contingency table, and apply it to similarity index calculation.*

Keywords: Contingency table, Type-2 fuzzy contingency table, Fuzzy numbers, T-norm, Similarity index

1. **Introduction.** Generally, we could efficiently analyze the inexact information and investigate the fuzzy relation by applying the fuzzy theory. Then, we would define the fuzzy contingency table to investigate the relational structure. We have investigated the relation of instructor structure and learning structure [6] and properties of similarity indices and connectivity indices [5] and so on. And then, we have defined "type-2" fuzzy contingency table and applied the questionnaire analysis [8]. In this paper, we would discuss about a fuzzy contingency table in Section 2 and discuss connectivity indices and similarity coefficients for type-1 fuzzy 2×2 contingency table in Section 3. Furthermore, we would propose "type-2" fuzzy contingency table and new "fuzzy" similarity indices by using the type-2 fuzzy contingency table in Section 4. Finally, some concluding remarks are given in Section 5.

2. **Fuzzy Contingency Table.**

Definition 2.1. *Cardinality of Fuzzy Set*

Consider a fuzzy set A in universe $U = \{x_i | i = 1, \ldots, N\}$. Cardinality $|A|$ of fuzzy set A is defined by

$$|A| = \sum_{k=1}^{N} \mu_A(x_k)$$

where $\mu_A(x)$ is a membership function of the fuzzy set A.

Definition 2.2. *Type-1 $m \times n$ Fuzzy Contingency Table*

For the fuzzy set A in universe $U = \{x_i | i = 1, \ldots, N\}$. Type-1 $m \times n$ fuzzy contingency table (Table 1) of fuzzy sets $A_1, \ldots, A_n$, $B_1, \ldots, B_m$ is defined in Table 1, where $f_{ij} = |A_j \cap B_i|$.

Assume

$$\sum_{i=1}^{n} \mu_{A_i}(x_k) = 1, \quad \sum_{i=1}^{m} \mu_{B_i}(x_k) = 1$$

DOI: 10.24507/icicel.12.08.791

and

$$\mu_{A_j \cap B_i}(x) = \mu_{A_j}(x) \cdot \mu_{B_i}(x)$$

then, we have for each $1 \le i \le m$ *and* $1 \le j \le n$

$$\sum_{i=1}^{m} f_{ij} = \sum_{k=1}^{N} \mu_{A_j}(x_k) \sum_{i=1}^{m} \mu_{B_i}(x_k) = |A_j|$$

$$\sum_{j=1}^{n} f_{ij} = \sum_{k=1}^{N} \mu_{B_i}(x_k) \sum_{j=1}^{n} \mu_{A_j}(x_k) = |B_i|.$$

TABLE 1. Type-1 fuzzy $m \times n$ contingency table

	A_1	$\cdots$	A_n	Sum
B_1	f_{11}	$\cdots$	f_{1n}	$\lvert B_i\rvert$
$\vdots$	$\vdots$		$\vdots$	$\vdots$
B_m	f_{m1}	$\cdots$	f_{mn}	$\lvert B_i\rvert$
Sum	$\lvert A_j\rvert$	$\cdots$	$\lvert A_j\rvert$	N

Here, we would expand Definition 2.2 and define a type-2 fuzzy contingency table. For the definition of type-2 fuzzy contingency table, we need a notion of mean value of fuzzy numbers which mean the product value of fuzzy numbers and the intersection of type-2 fuzzy sets. Then, we could clarify these definitions. Also, α^* is a fuzzy set that will be determined by the following membership function in the following definitions.

$$\mu_{\alpha^*}(x) = \alpha \quad (x \in \boldsymbol{R})$$

Definition 2.3. *Division of Summation of Fuzzy Numbers*
Let $x_1^*, x_2^*, x_3^*, \ldots, x_N^*$ *be fuzzy numbers with* α*-cuts*

$$C_\alpha(x_i^*) = [a_{\alpha,i}, b_{\alpha,i}] \quad (\alpha \in \mathbb{R},\ 0 \le \alpha \le 1)$$

and then the division of summation of fuzzy number into M *equals parts*

$$x^*/M = \bigcup_{\alpha \in (0,1]} (\alpha^* \cap C_\alpha(x^*/M))$$

where

$$C_\alpha(x^*/M) = \left[\frac{1}{M}\sum_{i=1}^{N} a_{\alpha,i},\ \frac{1}{M}\sum_{i=1}^{N} b_{\alpha,i}\right].$$

Definition 2.4. *Mean Value of Fuzzy Numbers*
Let $x_1^*, x_2^*, x_3^*, \ldots, x_N^*$ *be fuzzy numbers with* α*-cuts*

$$C_\alpha(x_i^*) = [a_{\alpha,i}, b_{\alpha,i}] \quad (\alpha \in \mathbb{R},\ 0 \le \alpha \le 1)$$

and then the mean value $\overline{x^*}$

$$\overline{x^*} = x^*/N.$$

Furthermore, we need some definitions [7] for type-2 fuzzy set.

Definition 2.5. *Product Value of Fuzzy Numbers*
Let u_1^*, u_2^* *be fuzzy numbers with* α*-cuts*

$$C_\alpha(u_i^*) = [a_{\alpha,i}, b_{\alpha,i}] \quad (\alpha \in \mathbb{R},\ 0 \le \alpha \le 1)$$

and then the product value $u_1^* \cdot u_2^*$ *is defined by*

$$u_1^* \cdot u_2^* = \bigcup_{\alpha \in (0,1]} (\alpha^* \cap C_\alpha(u_1^* \cdot u_2^*))$$

$$C_\alpha\left(u_1^* \cdot u_2^*\right) = \left[\min_{(x_1,x_2)\in C_\alpha(u_1^*)\times C_\alpha(u_2^*)} x_1 \cdot x_2, \max_{(x_1,x_2)\in C_\alpha(u_1^*)\times C_\alpha(u_2^*)} x_1 \cdot x_2\right].$$

Definition 2.6. *Intersection of Type-2 Fuzzy Sets*
Consider the type-2 fuzzy sets $\widetilde{A}$, $\widetilde{B}$ *in universe* $U = \{x_i | i = 1, \ldots, N\}$

$$\widetilde{A} = \{(x_i, u_i^*) | i = 1, \ldots, N\}, \ \widetilde{B} = \{(x_i, v_i^*) | i = 1, \ldots, N\}$$

where let u_i^*, v_i^* *be fuzzy numbers. Then the intersection* $\widetilde{A} \cap \widetilde{B}$ *is defined by*

$$\widetilde{A} \cap \widetilde{B} = \{(x_i, u_i^* \cdot v_i^*) | i = 1, \ldots, N\}.$$

Here, we would define the type-2 fuzzy contingency table by these definitions [8].

Definition 2.7. *Type-2 Fuzzy* $m \times n$ *Contingency Table*
For the type-2 fuzzy sets

$$\widetilde{A_1}, \ldots, \widetilde{A_n}, \ \widetilde{B_1}, \ldots, \widetilde{B_m}$$

in universe $U = \{x_i | i = 1, \ldots, N\}$, *put*

$$\widetilde{A_p} = \left\{\left(x_{i,p}, u_{i,p}^*\right) | i = 1, \ldots, N\right\},$$

$$\widetilde{B_q} = \left\{\left(x_{i,q}, u_{i,q}^*\right) | i = 1, \ldots, N\right\} \ (1 \le p \le n, 1 \le q \le m).$$

Here, let $u_{i,*}^*$ *be fuzzy numbers with* α*-cuts*

$$C_\alpha\left(u_{i,*}^*\right) = [a_{\alpha,i,*}, b_{\alpha,i,*}] \quad (\alpha \in \mathbb{R}, \ 0 \le \alpha \le 1), \quad [a_{\alpha,i,*}, b_{\alpha,i,*}] \subseteq [0, 1].$$

Type-2 fuzzy $m \times n$ *contingency table (Table 2) is defined by the following, where let* $\overline{f_{ij}}$ *be mean value* $\overline{u^* \cdot v^*}$ *of grades of intersection* $\widetilde{A_i} \cap \widetilde{B_j}$.

TABLE 2. Type-2 fuzzy $m \times n$ contingency table

	$\widetilde{A_1}$	$\cdots$	$\widetilde{A_n}$
$\widetilde{B_1}$	$\overline{f_{11}}$	$\cdots$	$\overline{f_{1n}}$
$\vdots$	$\vdots$		$\vdots$
$\widetilde{B_m}$	$\overline{f_{m1}}$	$\cdots$	$\overline{f_{mn}}$

3. Similarity Indices and Fuzzy 2 × 2 Contingency Table.

Definition 3.1. *Type-1 Fuzzy* 2×2 *Contingency Table*
Let $\boldsymbol{x}$ *and* $\boldsymbol{y}$ *be two vectors as follows:*

$$\boldsymbol{x} = (x_i), \ \boldsymbol{y} = (y_i), \ 0 \le x_i, y_i \le 1, \ 1 \le i \le N.$$

Then the type-1 fuzzy 2×2 *contingency table (Table 3) of* $\boldsymbol{x}$ *and* $\boldsymbol{y}$ *is defined accordingly.*

TABLE 3. Type-1 fuzzy 2×2 contingency table

$\boldsymbol{x}$ \ $\boldsymbol{y}$	1	0	Sum
1	a_{xy}	b_{xy}	$\sum_{k=1}^{N} x_k$
0	c_{xy}	d_{xy}	$N - \sum_{k=1}^{N} x_k$
Sum	$\sum_{k=1}^{N} y_k$	$N - \sum_{k=1}^{N} y_k$	N

Here, $T(p,q)$ is T-norm.

$$a_{xy}=\sum_{k=1}^{N}T(x_k,y_k),\quad b_{xy}=\sum_{k=1}^{N}x_k-a_{xy},$$

$$c_{xy}=\sum_{k=1}^{N}y_k-a_{xy},\quad d_{xy}=N-\sum_{k=1}^{N}y_k-b_{xy}=N-\sum_{k=1}^{N}x_k-c_{xy}.$$

From the fuzzy contingency table, we could compute any similarity coefficients. The typical similarity coefficients are

(1) Jaccard: $s_{xy}=\dfrac{a_{xy}}{a_{xy}+b_{xy}+c_{xy}}$

(2) Sorensen-Dice: $s_{xy}=\dfrac{2a_{xy}}{2a_{xy}+b_{xy}+c_{xy}}$

(3) Russell and Rao: $s_{xy}=\dfrac{a_{xy}}{a_{xy}+b_{xy}+c_{xy}+d_{xy}}$

(4) Simple Matching: $s_{xy}=\dfrac{a_{xy}+d_{xy}}{a_{xy}+b_{xy}+c_{xy}+d_{xy}}$

(5) Rogers-Tanimoto: $s_{xy}=\dfrac{a_{xy}+d_{xy}}{a_{xy}+2b_{xy}+2c_{xy}+d_{xy}}$

(6) Simpson: $s_{xy}=\dfrac{a_{xy}}{(a_{xy}+b_{xy})\wedge(a_{xy}+c_{xy})}$

(7) Ochiai: $s_{xy}=\dfrac{a_{xy}}{\sqrt{(a_{xy}+b_{xy})(a_{xy}+c_{xy})}}$

(8) Phi: $s_{xy}=\dfrac{a_{xy}d_{xy}-b_{xy}c_{xy}}{\sqrt{(a_{xy}+b_{xy})(a_{xy}+c_{xy})(b_{xy}+d_{xy})(c_{xy}+d_{xy})}}$

(9) Yule: $s_{xy}=\dfrac{a_{xy}d_{xy}-b_{xy}c_{xy}}{a_{xy}d_{xy}+b_{xy}c_{xy}}$

and so on.

Moreover, from the fuzzy contingency table, we could compute a connectivity index [5].

Definition 3.2. *Connectivity Index*

$$f_{xy}=\mu\frac{a_{xy}}{a_{xy}+c_{xy}}+(1-\mu)\frac{d_{xy}}{c_{xy}+d_{xy}}$$

where $a_{xy}+c_{xy}\neq 0$*,* $c_{xy}+d_{xy}\neq 0$*,* $0\leq\mu\leq 1$*.*

Here, considering a sample number, we define the following connectivity index.

Definition 3.3. *Tsuda's Connectivity Index*

$$t_{xy}=\frac{a_{xy}+d_{xy}}{(a_{xy}+c_{xy})+(c_{xy}+d_{xy})}\in[0,1]$$

where if $a_{xy}=c_{xy}=d_{xy}=0$*, then* $t_{ij}=1$*.*

In addition, applying T-norm, we could define the connectivity index.

Definition 3.4. *Connectivity Index by using T-norm*

$$f_{xy}=\frac{1}{N}\sum_{k=1}^{N}(1-T(x_k,1-y_k))$$

Here, $T(p,q)$ *is T-norm.*

4. **Similarity Indices and Type-2 Fuzzy 2 × 2 Contingency Table.** We extend the above definition of a fuzzy contingency table and similarity coefficients with a fuzzy 2 × 2 contingency table, and define type-2 fuzzy 2 × 2 contingency table and new fuzzy similarity coefficients. For these definitions, we need a notion of fuzzy negation, a notion of fuzzy complement of type-2 fuzzy set and a notion of a representation of fuzzy number. Then, we could clarify these definitions.

Definition 4.1. *Fuzzy Negation*
Let u^* be type-2 fuzzy grade value with α-cuts

$$C_\alpha(u^*) = [a_\alpha, b_\alpha] \quad (\alpha \in \mathbb{R},\ 0 \le \alpha \le 1), \quad [a_\alpha, b_\alpha] \subseteq [0,1].$$

Then the fuzzy negation of u^* is defined by the following

$$\neg u^* = \bigcup_{\alpha \in (0,1]} (\alpha^* \cap C_\alpha(1-u^*))$$

$$C_\alpha(1-u^*) = [1-b_\alpha, 1-a_\alpha].$$

Definition 4.2. *Fuzzy Complement of Type-2 Fuzzy Set*
Consider the type-2 fuzzy set $\widetilde{A}$ in universe $U = \{x_i | i = 1, \ldots, N\}$

$$\widetilde{A} = \{(x_i, u_i^*) | i = 1, \ldots, N\}$$

where let u_i^* be fuzzy numbers with α-cuts

$$C_\alpha(u_i^*) = [a_{\alpha,i,*}, b_{\alpha,i,*}] \quad (\alpha \in \mathbb{R},\ 0 \le \alpha \le 1), \quad [a_{\alpha,i,*}, b_{\alpha,i,*}] \subseteq [0,1].$$

Then fuzzy complement of $\widetilde{A}$ is defined by

$$\widetilde{A}^c = \{(x_i, \neg u_i^*) | i = 1, \ldots, N\}.$$

Definition 4.3. *Representation of Fuzzy Number*
Let u^* be fuzzy numbers with α-cuts

$$C_\alpha(u^*) = [a_\alpha, b_\alpha] \quad (\alpha \in \mathbb{R},\ 0 \le \alpha \le 1)$$

and then the representative value $Rep(u^*)$ is defined by

$$Rep(u^*) = \frac{a_1 + b_1}{2}.$$

Definition 4.4. *Type-2 Fuzzy* 2 × 2 *Contingency Table*
Let $\boldsymbol{x}$ and $\boldsymbol{y}$ be two fuzzy vectors as follows:

$$\boldsymbol{x} = (x_i^*),\ \boldsymbol{y} = (y_i^*),\ 1 \le i \le N.$$

Here, x_i^*, y_i^* are fuzzy numbers with α-cuts

$$C_\alpha(x_i^*) = [a_{\alpha,i}, b_{\alpha,i}] \quad (\alpha \in \mathbb{R},\ 0 \le \alpha \le 1),\ 0 \le C_1(x_i^*) \le 1,$$
$$C_\alpha(y_i^*) = [c_{\alpha,i}, d_{\alpha,i}] \quad (\alpha \in \mathbb{R},\ 0 \le \alpha \le 1),\ 0 \le C_1(y_i^*) \le 1.$$

For four type-2 fuzzy truth sets

$$\widetilde{1}_x,\ \widetilde{0}_x,\ \widetilde{1}_y,\ \widetilde{0}_y$$

in universe $U = \{1, \ldots, N\}$, put

$$\widetilde{1}_x = \{(i, x_i^*) | i = 1, \ldots, N\},\ \widetilde{0}_x = \widetilde{1}_x^{\,c} = \{(i, \neg x_i^*) | i = 1, \ldots, N\},$$
$$\widetilde{1}_y = \{(i, y_i^*) | i = 1, \ldots, N\},\ \widetilde{0}_y = \widetilde{1}_y^{\,c} = \{(i, \neg y_i^*) | i = 1, \ldots, N\}.$$

Then the type-2 fuzzy 2 × 2 *contingency table (Table 4) of $\boldsymbol{x}$ and $\boldsymbol{y}$ is defined accordingly, where let $\overline{f_{pq,M}}$ be the division of summation of fuzzy number into M equals parts of grades of intersection $\widetilde{p}_x \cap \widetilde{q}_y$ $(p, q = 0, 1)$. If the intersection $\widetilde{p}_x \cap \widetilde{q}_y$ is the following*

$$\widetilde{p}_x \cap \widetilde{q}_y = \left\{ \left(i, v_{pq,i}^*\right) | i = 1, \ldots, N \right\},$$

then

$$\widetilde{f_{pq,M}} = v^*_{pq}/M.$$

TABLE 4. Type-2 fuzzy 2×2 contingency table

$\boldsymbol{x} \backslash \boldsymbol{y}$	$\widetilde{1_y}$	$\widetilde{0_y}$
$\widetilde{1_x}$	$\widetilde{f_{11,M}}$	$\widetilde{f_{10,M}}$
$\widetilde{0_x}$	$\widetilde{f_{01,M}}$	$\widetilde{f_{00,M}}$

Example 4.1. *For four type-2 fuzzy truth sets*

$$\widetilde{1_x},\ \widetilde{0_x},\ \widetilde{1_y},\ \widetilde{0_y}$$

in universe $U = \{1, \ldots, N\}$, put

$$\widetilde{1_x} = \{(1, 0.3^*), (2, 0.4^*), (3, 0.9^*)\}$$
$$\widetilde{0_x} = \{(1, 0.7^*), (2, 0.6^*), (3, 0.1^*)\}$$
$$\widetilde{1_y} = \{(1, 0.8^*), (2, 0.1^*), (3, 0.2^*)\}$$
$$\widetilde{0_y} = \{(1, 0.2^*), (2, 0.9^*), (3, 0.8^*)\}.$$

Let a^* be fuzzy number with membership function $\mu_{a^*}(x)$:

$$\mu_{a^*}(x) = \max\left\{1 - \frac{1}{\min\{a, 1-a\}}|x - a|, 0\right\}.$$

From four type-2 fuzzy truth sets, we obtained the type-2 fuzzy 2×2 contingency table, and calculate "fuzzy" Jaccard similarity coefficient $\widetilde{s_{xy}}$. Jaccard similarity coefficient is defined by

$$s_{xy} = \frac{a_{xy}}{a_{xy} + b_{xy} + c_{xy}}.$$

To decide the denominator "$a_{xy} + b_{xy} + c_{xy}$", we define the division number M applying representation of fuzzy number.

$$M = Rep(v^*_{11}) + Rep(v^*_{10}) + Rep(v^*_{01})$$

To decide the division number M, we create "representative value" contingency table (Table 5).

TABLE 5. "Representative value" contingency table

$\boldsymbol{x} \backslash \boldsymbol{y}$	$\widetilde{1_y}$	$\widetilde{0_y}$
$\widetilde{1_x}$	0.46	1.14
$\widetilde{0_x}$	0.64	0.76

From "representative value" contingency table, we have obtained the division number M:

$$M = 0.46 + 1.14 + 0.64 = \underline{2.24}.$$

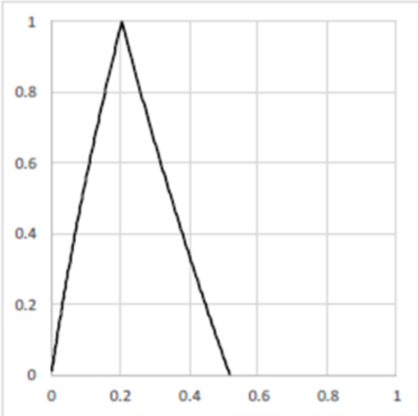
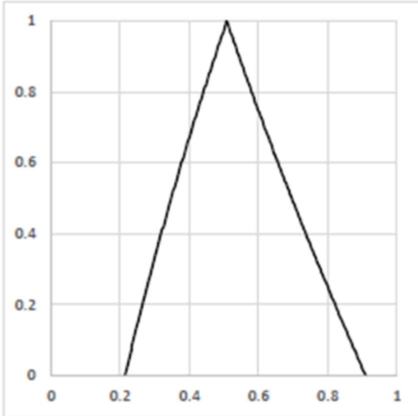
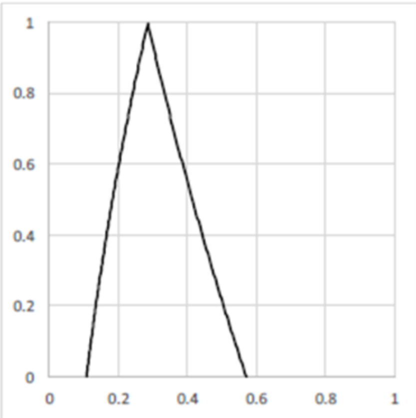
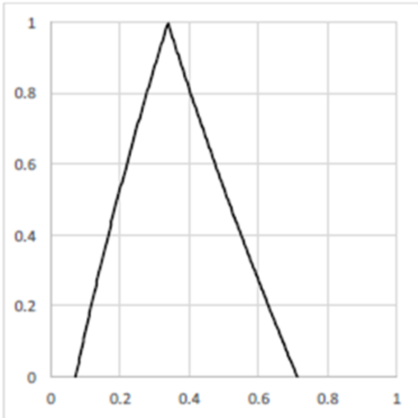
By the division number M, we have obtained type-2 fuzzy 2×2 contingency table (Table 6).

From this table, we have defined "fuzzy" Jaccard similarity coefficient $\widetilde{s_{xy}}$:

$$\widetilde{s_{xy}} = \overline{f_{11,M}}, \quad M = 2.24$$

Here, $\widetilde{s_{xy}}$ is fuzzy number with membership function as shown in Figure 1.

Table 6. Type-2 fuzzy 2×2 contingency table ($M = 2.24$)

$\boldsymbol{x}$ \ $\boldsymbol{y}$	$\widetilde{1}_y$	$\widetilde{0}_y$
$\widetilde{1}_x$		
$\widetilde{0}_x$		

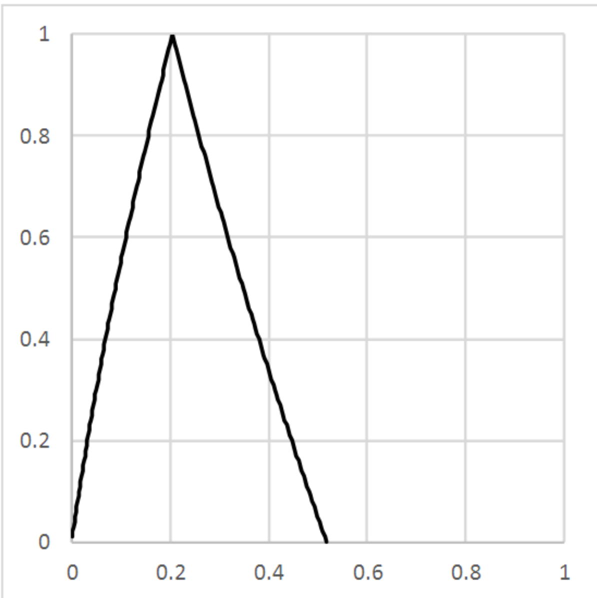


Figure 1. "Fuzzy" Jaccard similarity coefficient $\widetilde{s_{xy}}$

5. **Conclusions.** In this paper, we would extend fuzzy 2×2 contingency table to "type-2" fuzzy 2×2 contingency table. From this, we could calculate various "fuzzy" similarity coefficients. Further studies are needed in order to apply this method, and we consider various Type-2 fuzzy sets. This method is also available for the instruction/cognition analysis in education, the opinion poll in psychology and so on.

Acknowledgments. We thank the referees for useful comments.

REFERENCES

[1] C. E. Shannon, A mathematical theory of communication, *Bell System Tech. Journal*, vol.27, pp.379-423, 623-656, 1948.
[2] L. A. Zadeh, Open archive – Original research article, *Information and Control*, vol.8, no.3, pp.338-353, 1965.
[3] L. A. Zadeh, Fuzzy sets as a basis for a theory of possibility, *Fuzzy Sets and System*, vol.1, pp.3-28, 1978.
[4] A. Kaufmann, *Introduction to the Theory of Fuzzy Subsets*, Academic Press, New York, 1975.
[5] H. Uesu and H. Yamashita, Mathematical analysis of similarity index and connectivity index in fuzzy graph, *The 22nd Int'l Conference of the North American Fuzzy Information Processing Society*, pp.77-80, 2003.
[6] H. Uesu and H. Yamashita, Learning structure analysis system applying fuzzy theory, *The 4th IEEE International Conference on Advanced Learning Technologies*, pp.890-891, 2004.
[7] R. Viertl, *Statistical Methods for Fuzzy Data*, John Wiley & Sons, Ltd., 2011.
[8] H. Uesu, Students' needs analysis for media lectures applying Kano model, *Japan Society for Fuzzy Theory and Intelligent Informatics Soft Science Workshop*, pp.89-90, 2014 (in Japanese).
[9] H. Uesu and S. Kanagawa, Student needs analysis applying fuzzy contingency table, *The 27th Annual Conference of Biomedical Fuzzy System Association*, pp.45-46, 2014.
[10] H. Uesu, Type-2 fuzzy contingency table analysis and its application, *Kartografija I Geoinformacije*, vol.8, no.11, pp.176-178, 2015.
[11] H. Uesu, Student's needs analysis applying type-2 fuzzy contingency table for media lectures, *Proc. of the 28th Annual Conference of Biomedical Fuzzy Systems Association*, pp.293-296, 2015.
[12] H. Uesu and S. Takagi, Student needs analysis applying type-2 fuzzy contingency table and its application, *International Symposium on Information Theory and Its Applications*, p.567, 2016.
[13] H. Uesu, Contingency table analysis applying fuzzy number and its application, *The 8th International Joint Conference on Computational Intelligence*, pp.93-99, 2016.