BUBBLE CHART ANALYSIS FOR A LECTURE APPLYING TYPE-2 FUZZY CONTINGENCY TABLE

HIROAKI UESU

Mathematics and Science Education Research Center Kanazawa Institute of Technology 7-1 Ohgigaoka, Nonoichi, Ishikawa 921-8501, Japan uesu@neptune.kanazawa-it.ac.jp

Received November 2018; accepted February 2019

ABSTRACT. In general, a contingency table analysis is used to analyze the relationship between two categorical variables. However, in questionnaire surveys, the boundaries between categories are often ambiguous. For example, in a questionnaire using the Likert scale, the answer "I think so" and the answer "slightly agree" are considered to contain ambiguity at that boundary. Therefore, we have proposed the fuzzy $m \times n$ contingency table by applying the product operation of fuzzy set, and proposed a new analysis method of fuzzy data using fuzzy $m \times n$ contingency table, and researched the properties of the fuzzy contingency table and the various coefficients defined by it. In this paper, we propose the type-2 fuzzy $m \times n$ contingency table by extending the type-1 fuzzy $m \times n$ contingency table and the fuzzy coefficient of association. Furthermore, we propose a bubble chart analysis using fuzzy coefficient of association applying type-2 fuzzy contingency table, and we describe the effectiveness of this analysis by its application.

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

1. Introduction. In general, it is very useful to analyze inexact information efficiently by applying fuzzy theory [1,2]. In particular, we had defined 2×2 contingency table on fuzzy set and investigated fuzzy relationships. Then, we had demonstrated the effectiveness to analyze the relational structure by using it. For example, we have investigated the relation between instructor structure and learning structure [3], similarity index and connectivity index [4] by fuzzy 2×2 contingency table. Further, we generalized fuzzy 2×2 contingency table, we had defined (type-1) fuzzy contingency table to investigate fuzzy relationships [5,6].

Next, we had defined "type-2" fuzzy contingency table and we applied it to questionnaire analysis [7,8]. In some definition of fuzzy contingency table, one of them has been defined by Taheri et al. [9], but we propose another definition by fuzzy number. Our definition made it possible to perform various analyses from a different point of view. And then, we proposed to apply the type-2 fuzzy contingency table to a bubble chart analysis. "Bubble chart" is a diagram in which bubbles representing items are located on a diagram having two axes [10]. From the bubble chart, we could evaluate these items by three pieces of information – the size of bubble and the coordinate of bubble.

In this paper we propose our "type-2" fuzzy contingency table and show an application to bubble chart analysis using type-2 fuzzy contingency table.

2. Fuzzy Contingency Table. In fuzzy set operations, complementary law $(A \cup \overline{A} = U, A \cap \overline{A} = \phi)$ does not hold generally. However, in the fuzzy contingency table, we would like to define the total of row marginal total and column marginal total to be the cardinality of each fuzzy set. Therefore, we defined the cardinality of fuzzy set and the type-1 $m \times n$ fuzzy contingency table as follows.

DOI: 10.24507/icicel.13.06.483

Definition 2.1. Cardinality of fuzzy set

Consider a fuzzy set A in universe $U = \{x_i | i = 1, ..., N\}$. 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 $U = \{x_i | i = 1, ..., N\}$. Type-1 $m \times n$ fuzzy contingency table (Table 1) of fuzzy sets $A_1, ..., A_n, B_1, ..., B_m$ is defined by the following.

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

	A_1	•••	A_n
B_1	f_{11}	• • •	f_{1n}
:	:		•
B_m	f_{m1}	•••	f_{mn}

Here

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

and

$$f_{ij} = \frac{1}{N} |A_j \cap B_i| = \frac{1}{N} \sum_{k=1}^N \mu_{A_j \cap B_i}(x_k) = \frac{1}{N} \sum_{k=1}^N \mu_{A_j}(x_k) \cdot \mu_{B_i}(x_k)$$

Theorem 2.1. Row marginal total and column marginal total

For fuzzy sets A_1, \ldots, A_n , B_1, \ldots, B_m in universe $U = \{x_i | i = 1, \ldots, N\}$, in type-1 $m \times n$ fuzzy contingency table (Table 1), the following equation holds.

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

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.

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

We would like to define the mean value of fuzzy numbers. To define the mean value of fuzzy numbers, we have defined the division of summation of fuzzy numbers.

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;

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

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

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

where

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

Then, we would define the mean value of fuzzy numbers as follows.

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

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

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

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

Furthermore, we need some operations [7] for type-2 fuzzy set, and we defined as follows.

Definition 2.5. Product value of fuzzy numbers

Let u_1^* , u_2^* be fuzzy numbers with α -cuts

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

then the product value u_1^* , u_2^* is defined by

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

$$C_{\alpha}(u_{1}^{*} \cdot u_{2}^{*}) = \left[\min_{(x_{1}, x_{2}) \in C_{\alpha}(u_{1}^{*}) \times C_{\alpha}(u_{2}^{*})} x_{1} \cdot x_{2}, \max_{(x_{1}, x_{2}) \in C_{\alpha}(u_{1}^{*}) \times C_{\alpha}(u_{2}^{*})} x_{1} \cdot x_{2}\right].$$

Definition 2.6. Intersection of type-2 fuzzy sets

Consider the type-2 fuzzy sets \widetilde{A} , \widetilde{B} in universe $U = \{x_i | i = 1, ..., N\}$;

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

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

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

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

Definition 2.7. Representation of fuzzy number

Let u^* be fuzzy numbers with α -cuts

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

and if $C_1(u^*)$ is nonempty, then the representative value $\operatorname{Rep}(u^*)$ is defined by

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

And then, we have proposed the type-2 fuzzy $m \times n$ contingency table. In some definition of fuzzy contingency table, one of them has been defined by Taheri et al. [9], but we propose another definition by fuzzy number.

Definition 2.8. Type-2 fuzzy $m \times n$ contingency table

For the type-2 fuzzy sets

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

H. UESU

$$\widetilde{A_p} = \{ (x_{i,p}, u_{i,p}^*) | i = 1, \dots, N \}, \quad \widetilde{B_q} = \{ (x_{i,q}, u_{i,q}^*) | i = 1, \dots, N \}$$

(1 \le p \le n, \quad 1 \le q \le m)

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

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

Type-2 fuzzy $m \times n$ contingency table (Table 2) is defined by the following.

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

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

Here, let $\overline{f_{ij}}$ be mean value $\overline{u^* \cdot v^*}$ of grades of intersection $\widetilde{A_j} \cap \widetilde{B_i}$. **Theorem 2.2.** Let \boldsymbol{x} and \boldsymbol{y} be two fuzzy vectors as follows:

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

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

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

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

Above two fuzzy vectors \boldsymbol{x} , \boldsymbol{y} should satisfy the following condition:

$$a_{1,i} = b_{1,i} \text{ or } c_{1,i} = d_{1,i} \quad (1 \le i \le N).$$

Then the following equation holds.

$$\sum_{p=0}^{1} \sum_{q=0}^{1} \operatorname{Rep}\left(\overline{f_{pq,N}}\right) = 1$$

3. Fuzzy Coefficient of Association and Type-2 Fuzzy $m \times n$ Contingency Table. Here, we have proposed the fuzzy coefficient of association by fuzzifying Cramer's V. Cramer's V is an index of how strong a relationship between two variables is, and is used to calculate correlation in contingency table with more than 2×2 columns and rows. By fuzzifying Cramer's V, we could consider various applications.

Definition 3.1. Fuzzy coefficient association 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 $U = \{x_i | i = 1, ..., N\}$, we have the type-2 fuzzy $m \times n$ contingency table:

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

486

Here, fuzzy coefficient association V^* is defined by the following steps.

Step 1: Calculating the table of fuzzy expected frequencies E_{ij}

From the type-2 fuzzy $m \times n$ contingency table, we get the following E_{ij} of fuzzy expected frequencies:

$$E_{ij} = \left(e_{ij}^*\right)$$

Here, e_{ij}^* is the fuzzy expected frequencies with membership function $\mu_{e_{ij}^*}(u)$:

$$\mu_{e_{ij}^*}(u) = \max\left\{1 - \frac{1}{\min\{e_{ij}, 1 - e_{ij}\}} |u - e_{ij}|, 0\right\}, \quad e_{ij}^* = Rep\left(\sum_{p=1}^m \overline{f_{pj}}\right) \cdot Rep\left(\sum_{p=1}^n \overline{f_{ip}}\right)$$

Step 2: Calculating the fuzzy chi square χ^{2*} Fuzzy chi square χ^{2*} is defined by the following:

$$\chi^{2*} = \sum_{p=1}^{m} \sum_{q=1}^{n} \left\{ \left(\overline{f_{pq}} - e_{pq}^* \right) \cdot \left(\overline{f_{pq}} - e_{pq}^* \right) \right\} / Rep(e_{pq}^*)$$

Step 3: Calculating the fuzzy coefficient association V^* Fuzzy coefficient association V^* is defined by the following:

$$V^* = \sqrt{\chi^{2*} / \min(m - 1, n - 1)}$$

Here, let u^* be fuzzy numbers with α -cuts

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

and then $\sqrt{u^*}$ is defined by

$$\sqrt{u^*} = \bigcup_{\alpha \in (0,1]} \left(\alpha^* \cap C_\alpha \left(\sqrt{u^*} \right) \right)$$
$$C_\alpha \left(\sqrt{u^*} \right) = \left[\min_{x_1 \in C_\alpha(u^*)} \operatorname{sgn}(x_1) \cdot \sqrt{|x_1|}, \max_{x_2 \in C_\alpha(u^*)} \sqrt{x_2} \right].$$

4. **Bubble Chart Analysis.** We consider items with two dimension fuzzy vector (these elements are fuzzy number). From this fuzzy vector, we have created the bubble chart. In this bubble chart, bubbles are located on a diagram having two axes. From the bubble chart, we could consider priority of improvement items. The priority can be measured for each quadrant as shown in Figure 1.



FIGURE 1. Priority for each quadrant

H. UESU

Here, the positioning of the four quadrants is as follows

1st Quadrant: Strong item

The influence and the satisfaction are also high, so it is necessary to keep the item satisfaction from continuing to fall.

2nd Quadrant: Maintenance items

The satisfaction is high, but the influence on overall evaluation is small. Since the satisfaction is high, it is necessary to maintain.

3rd Quadrant: Alert items

The satisfaction and the influence on overall evaluation are low. It is an item that needs improvement after the priority improvement item.

4th Quadrant: Important improvement items

Items located here are items of low importance but high satisfaction. In other words, it must be improved with the highest priority in order to raise the overall satisfaction level.

And then, the size of the bubble represents the fuzziness. The bigger the bubble is, the fuzzier the value is.

5. Bubble Chart Analysis Applying Type-2 Fuzzy 5×5 Contingency Table. We create a survey questionnaire based on "Frederick Herzberg's motivation-hygiene theory". Then, we execute the questionnaire to the first year students of the university as shown in Figure 2. The lecture name is "mathematics for engineering", and the number of students is 56.

Q1. As for the class of 'Mathematics for engineering', in each of the follow these statements you most agree with or is most applicable to you.	ing	choo	se fi	rom	a sc	ale o	of 1-	5 wł	nich of
1. I think that the composition of this lecture and time allocation were appropriate.	1	-	2	-	3	-	4	-	5
2. I think that the items and contents handled in this lecture are appropriate.	1	-	2	-	3	-	4	-	5
3 . I think the grade evaluation of this lecture is appropriate.	1	-	2	-	3	-	4	-	5
4. I think that teacher's way of speaking (explanation, pause, etc.) is appropriate.	1	-	2	-	3	-	4	-	5
5. I think that the teacher was progressing the lecture while confirming the understanding degree of the students.	1	-	2	-	3	-	4	-	5
6. I think that the teacher was considering the students to express their questions and opinions.	1	-	2	-	3	-	4	-	5
I think that the teacher was taking appropriate measures to students who are talking in whispers at this lecture.	1	-	2	-	3	-	4	-	5

FIGURE 2. Questionnaire

From this questionnaire, we obtain the response table $R = (r_{ik})$.

Since each question belongs to the major factors "policy", "evaluation", "teacher", "environment", "human relationship", "achievement", "total evaluation", by averaging the answer values of the corresponding questions we obtain the category value table $T = (t_{ij}), 1 \le i \le n, j = 7.$

From the category value table $T = (t_{ij})$, the membership function is defined by the following equation for each type-2 fuzzy set E_p , $1 \le p \le 5$.

$$\mu_q(u) = \left\{ 1 - \left| \frac{1}{(1-q) \wedge q} (u-q) \right| \right\} \lor 0, \quad 0 \le q \le 1$$

$$\mu_{E_p}(u) = \begin{cases} \mu_{p-t_{ij+1}}(u), & [t_{ij}] = p \\ \mu_{t_{ij}-p+1}(u), & [t_{ij}] = p-1, & 0 \le u \le 1 \\ 0, & \text{otherwise} \end{cases}$$

Here, the correspondence between the value of p and the type-2 fuzzy set is as shown in Table 3.

TABLE 3 .	Value of	p and type-2	fuzzy set
-------------	----------	--------------	-----------

Category value	1	2	3	4	5
Category	Strongly disagree	Disagree	Neither	Agree	Strongly agree
Type-2 fuzzy set	E_1	E_2	E_3	E_4	E_5

We calculate the fuzzy coefficient association for each category vector and the total evaluation vector. Then, we create the fuzzy influence table as shown in Table 4.



TABLE 4. Fuzzy influence table

From the category value table $T = (t_{ij})$, we calculate the average value $\overline{t_{*j}}$, $1 \le j \le 6$, we fuzzify these value, and we create a fuzzy satisfaction table for each category. At this fuzzification, the membership function is defined as follows:

$$\mu_{\overline{t_{*j}}}(u) = \left\{ 1 - \left| \frac{1}{\left(1 - \overline{t_{*j}} + [\overline{t_{*j}}]\right) \wedge \left(\overline{t_{*j}} - [\overline{t_{*j}}]\right)} (u - \overline{t_{*j}}) \right| \right\} \lor 0$$

Standardizing the fuzzy coefficient association and the fuzzy satisfaction, we obtain a standard fuzzy coefficient association table and a standard fuzzy satisfaction table.

Then, we create a bubble chart. The diameter of each bubble is the width of the α -level set ($\alpha = 0.5$) of each standard fuzzy influence, and it is illustrated on this chart as a relative value.

From the bubble chart, we could consider priority to improve quality of lectures. Since the vertical axis is the average values of the standard fuzzy influence and the horizontal axis is the standard fuzzy satisfaction, the priority can be measured for each quadrant as shown in Figure 1.

H. UESU

	Policy	Evaluation	Teacher				
Fuzzy satisfaction							
	Environment	Human relationship	Achievement				
Fuzzy satisfaction							

TABLE 5. Fuzzy satisfaction table

TABLE 6. Standard fuzzy coefficient association table



TABLE 7. Standard fuzzy satisfaction table



As shown in Figure 3, since items in the first quadrant are both high influence and satisfaction level, "teacher" and "evaluation" are considered to be an item to maintain and reinforce as the strength of the subject "mathematics for engineering". If we fail to maintain this level of these items it may largely reduce total satisfaction and should be actively maintained. Also, regarding the items in the fourth quadrant, since the influence is high but the satisfaction is low, it is considered that "achievement" is items to be



FIGURE 3. Bubble chart

improved preferentially. By improving this item preferentially, total satisfaction can be effectively increased. In this area, the size of "achievement" is big, but there is only this item in the improvement area, we would not consider about fuzziness in this survey.

6. Conclusion. In this paper, we have presented the type-2 fuzzy $m \times n$ contingency table, and we proposed the fuzzy bubble chart analysis as an application. By extending the conventional method, we were able to analyze more fuzzy information.

In the future, we would like to proceed with the research on type-2 fuzzy contingency table and propose various applications. Also, we would like to consider about its generalization.

REFERENCES

- [1] L. A. Zadeh, Fuzzy sets, Inform and Control, vol.8, pp.338-353, 1965.
- [2] L. A. Zadeh, Fuzzy sets as a basis for a theory of possibility, *Fuzzy Sets and System*, vol.1, pp.3-28, 1978.
- [3] 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.
- [4] H. Uesu and H. Yamashita, Mathematical analysis of similarity index and connectivity index in fuzzy graph, The 22nd International Conference of the North American Fuzzy Information Processing Society, pp.77-80, 2003.
- [5] 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.
- [6] 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.
- [7] H. Uesu, Type-2 fuzzy contingency table analysis and its application, The 7th International Conference on Dynamic Systems and Applications, 2015.
- [8] 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.
- S. M. Taheri, G. Hesamian and R. Viertl, Contingency tables with fuzzy information, Communications in Statistics – Theory and Methods, vol.45, pp.5906-5917, 2016.
- [10] H. Uesu, Contingency table analysis applying fuzzy number and its application, The 8th International Joint Conference on Computational Intelligence, pp.93-99, 2016.
- [11] R. Viertl, Statistical Methods for Fuzzy Data, John Wiley & Sons, Ltd., 2011.
- [12] H. Uesu, Type-2 fuzzy contingency table and similarity indices, *ICIC Express Letters*, vol.12, no.8, pp.791-798, 2018.