2 Methodologies

2.1 Membership function

First of all, we would like to focus attention on indicator function first. The indicator function I_A of crisp set A can be define by

$$I_A(x) = \begin{cases} 1 & \text{for} \quad x \in A \\ 0 & \text{for} \quad x \notin A \end{cases} \tag{2.1}$$

$I_{\{x_k\}}$ and $I_{[a,b]}$ are two special indicator functions. Many authors view the corresponding membership function of a fuzzy set as the generalization of indicator function in crisp sets. The value of membership function represents the degree of the truth as an extension of valuation.

People often confused probabilities with degrees of membership. Membership functions are used to describe the degree of truth in all states not probability distributions. Fuzziness and randomness are two different concepts. Roughly speaking, randomness means the occurrences of events; but fuzziness means the vagueness of meaning. For example, the outcome of rolling a dice is a probability event and the meaning of an old man is a fuzzy event.

Definition 2.1 *For the universe U, a membership function on U is a real function from U to [0, 1]. The membership function which represents a fuzzy set* $\tilde{A}$ *is usually denoted by* $u_{\tilde{A}}$. *For* $x \in U$, *the value* $u_{\tilde{A}}(x)$ *is called the membership degree of x in the fuzzy set* $\tilde{A}$. *The membership degree* $u_{\tilde{A}}(x)$ *quantifies the grade of membership of the element x to the fuzzy set* $\tilde{A}$.

Example 2.1 $\tilde{A}$ = *"A man whose age closes to 20".*

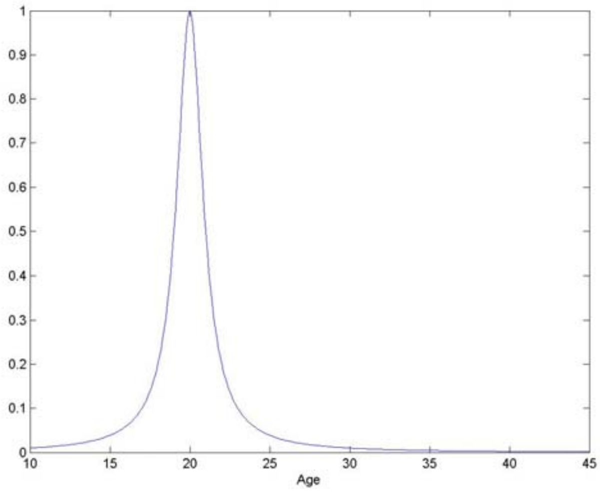

Figure 2 Membership function for a man whose age closes to 20

Fig. 2 is one probably membership function of a man whose age closes to 20. And it indicates the degree of membership increased when the x coordinate comes close to 20. On the contrary, the degree of membership decreased when the x coordinate goes away from 20.

2.2 Fuzzy numbers

Many measurements are contained with a remarkable amount of uncertainty such as the definition of an old man. Using fuzzy data to interpret the results of measurements should be better. This uncertainty is not the same with measurement errors and is called fuzziness. The concepts of fuzziness help us to preserve the original information of the data more completely.

Generally speaking, we divided fuzzy data into two kinds. One is discrete type; the other is continuous type. The former called discrete fuzzy data and the latter is called continuous fuzzy data. The following are the two definitions of discrete fuzzy data and continuous fuzzy data, respectively.

Definition 2.2 (Fuzzy Number, Nguyen and Wu (2006)) *Let* U *be a universal set,* $L=\{L_1, L_2, \cdots L_n\}$ *be a set of n-linguistic variables in* U *. For any term or statement* x *on* U *, its membership corresponding to* $\{L_1, L_2, \cdots L_n\}$ *is* $\{u_1(x), u_2(x), \cdots u_n(x)\}$ *, where* $u \rightarrow [0,1]$ *is a real function. x is called discrete fuzzy number if its corresponding membership function can be written as* $u_U(x)=\frac{u_1(x)}{L_1}+\frac{u_2(x)}{L_2}+\cdots+\frac{u_n(x)}{L_n}$ *, where "+" stands for "or", and "$\div$" stands for the membership* $u_i(x)$ *on* L_i *.*

If the domain of the universal set is continuous, x is called continuous fuzzy number. And its corresponding membership function can be written as $u_U(x)=\int_{L_i \subseteq L} \frac{u_i(x)}{L_i}$

Example 2.2 *"$\tilde{A}$ =How many hours do you watch TV for a day".*

The fuzzy number for watching TV per day may be written as

$$u_{\tilde{A}}(x)=\frac{0.2}{0}+\frac{0.1}{1}+\frac{0.4}{2}+\frac{0.1}{3}+\frac{0.2}{4}$$

Example 2.3 $\tilde{A}$ *="real numbers close to 10".*

$\tilde{A}=\{(x, u_{\tilde{A}}(x)) | u_{\tilde{A}}(x)=(1+(x-10)^2)^{-1}\}$ *, and the graph looks like Fig 3.*

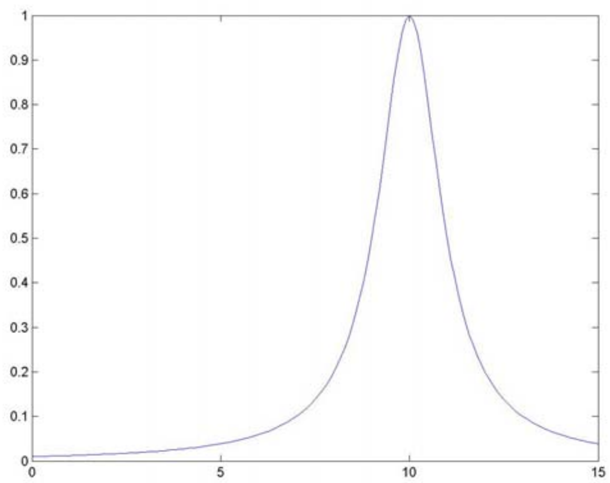

Figure 3 Real numbers close to 10

Continuous fuzzy data can be classified into several types, Such as interval-valued numbers, triangular numbers, trapezoid numbers, and exponential numbers etc. Most fuzzy numbers get these names from the sharp of membership function. Even though there are varies types of fuzzy numbers, but here we limit the discussion to three usual types: interval-valued numbers, triangular numbers and trapezoid numbers. The definitions of the three types of fuzzy data are given as follows.

Definition 2.3 $\tilde{A} = [a, b, c, d]$ *is called trapezoid fuzzy number, if its corresponding membership function can be defined as*

$$u_{\tilde{A}}(x) = \begin{cases} u_{\tilde{A}}^{L}(x) & \text{for} \quad a \le x \le b \\ 1 & \text{for} \quad b \le x \le c \\ u_{\tilde{B}}^{R}(x) & \text{for} \quad c \le x \le d \\ 0 & \text{for} \quad \text{otherwise} \end{cases}$$

where $a < b < c < d$ *and* $u_{\tilde{A}}^{L}(x)$ *is a monotone increasing real function and* $u_{\tilde{A}}^{R}(x)$ *is a monotone decreasing real function..*

Note that when $b = c$, $\tilde{A}$ *is called triangular fuzzy number;*

when $a = b$ and $c = d$, $\tilde{A}$ *is called interval-valued fuzzy number.*

The following examples are these kinds of fuzzy data.

Example 2.4 *Fig. 4 is an example of triangular fuzzy number, and its corresponding membership function is as follows.*

$$u_{\tilde{A}}(x) = \begin{cases} \dfrac{x-3}{4-3} & \text{for} \quad 3 \le x \le 4 \\ 1 & \text{for} \quad x = 4 \\ \dfrac{5-x}{5-4} & \text{for} \quad 4 \le x \le 5 \end{cases}$$

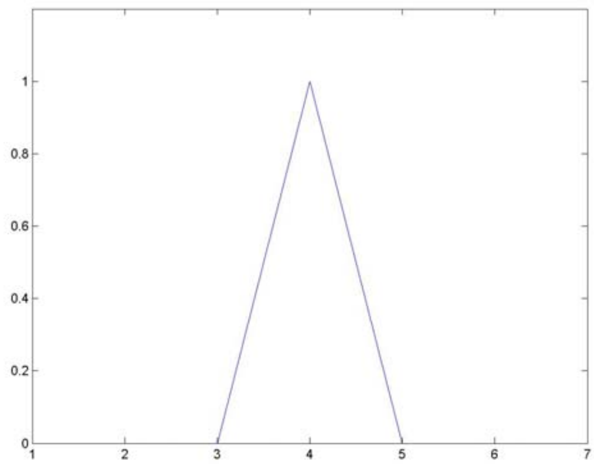

Figure 4   A triangular fuzzy number

Example 2.5 *Fig. 5 is an example of trapezoid fuzzy number, and its corresponding membership function is as follows.*

$$u_{\tilde{A}}(x)=\begin{cases}\dfrac{x-3}{4-3} & \text{for} \quad 3\le x\le 4\\ 1 & \text{for} \quad 4\le x\le 5\\ \dfrac{5-x}{5-4} & \text{for} \quad 5\le x\le 6\end{cases}$$

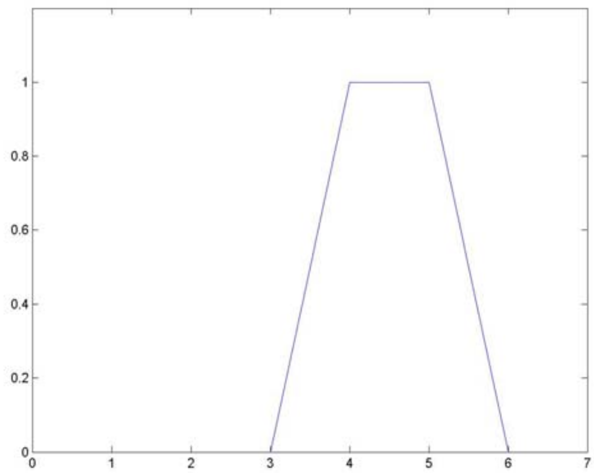

Figure 5 A trapezoid fuzzy number

Example 2.6 *Fig. 6 is an example of interval-valued fuzzy number, and its corresponding membership function is as follows.*

$$u_{\tilde{A}}(x) = \begin{cases} 1 & \text{for} \quad 3 \le x \le 4 \\ 0 & \quad\quad \text{otherwise} \end{cases}$$

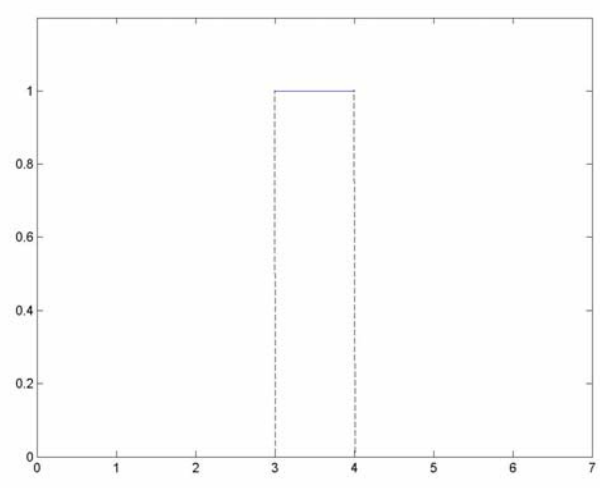

Figure 6 A interval-valued fuzzy number

2.3 Getting a continuous fuzzy data

Respondents choose one single answer or certain range of the answer in traditional sampling survey. But traditional method is not able to reflect the complex thoughts of each respondent sufficiently. If people can express the degree of their feelings by using membership functions, the answer presented will be closer to real human thoughts. But unfortunately scholars disagree in opinion about the construction of continuous fuzzy data. Many studies use continuous fuzzy without describing the construction method. The core of all the questions is fuzzy data is determined by its membership function, but the construction of membership function is quite subjective. To success this, we try to determine the membership function by the respondents themselves based on GSP.

Fig.7 is the image of fuzzy questionnaire which is about the prime time for marriage. Before answering the fuzzy questionnaire, respondents could click the three buttons to realize the meaning of each section and points. One example about the prime time for marriage is showed as Fig. 8. $\overline{AB}$ represents the desire for marriage grows continuously for 2 years from 26. $\overline{BC}$ represents the desire for marriage remains at the highest level between 28~30. $\overline{CD}$ represents the desire for marriage fall continuously from 30 until it reached 35.

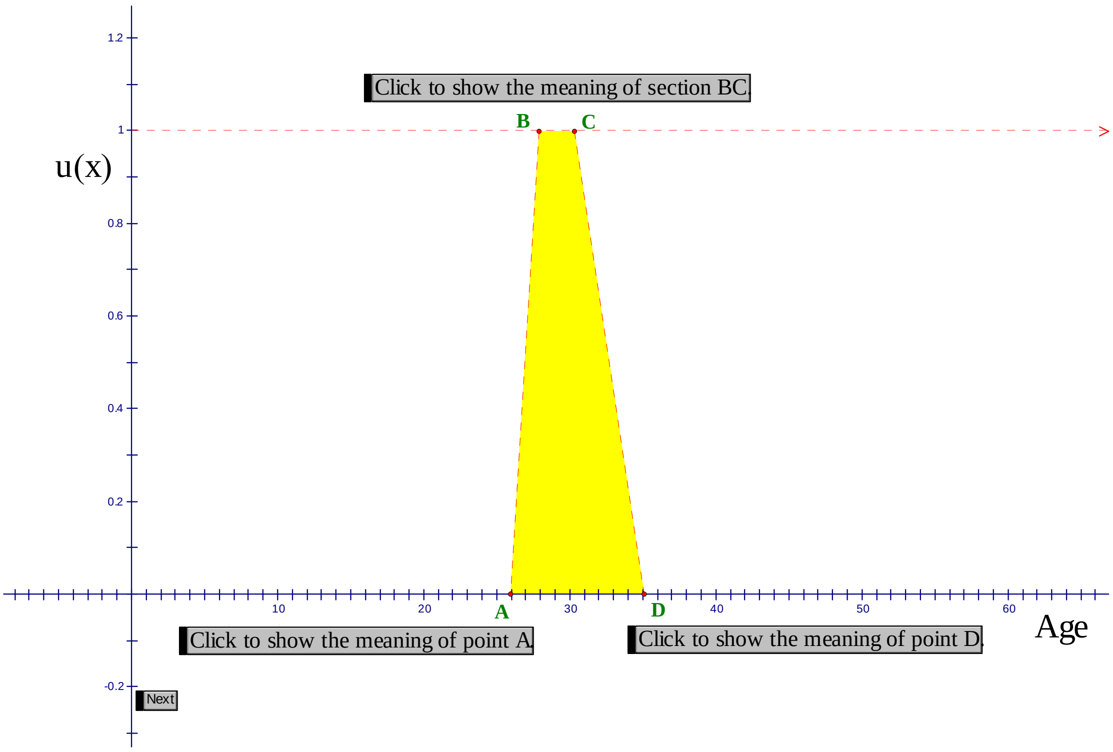

Figure 7 Fuzzy questionnaires about the prime time for marriage

Respondents can decide their own membership function of the prime time for marriage by moving the four points *A*, *B*, *C*, and *D*. By moving the four points, the age corresponding to the points will be change automatically. There are probably three types of fuzzy data: The first is trapezoid; the second is triangular (Fig. 9); the third is interval-valued type (Fig. 10).

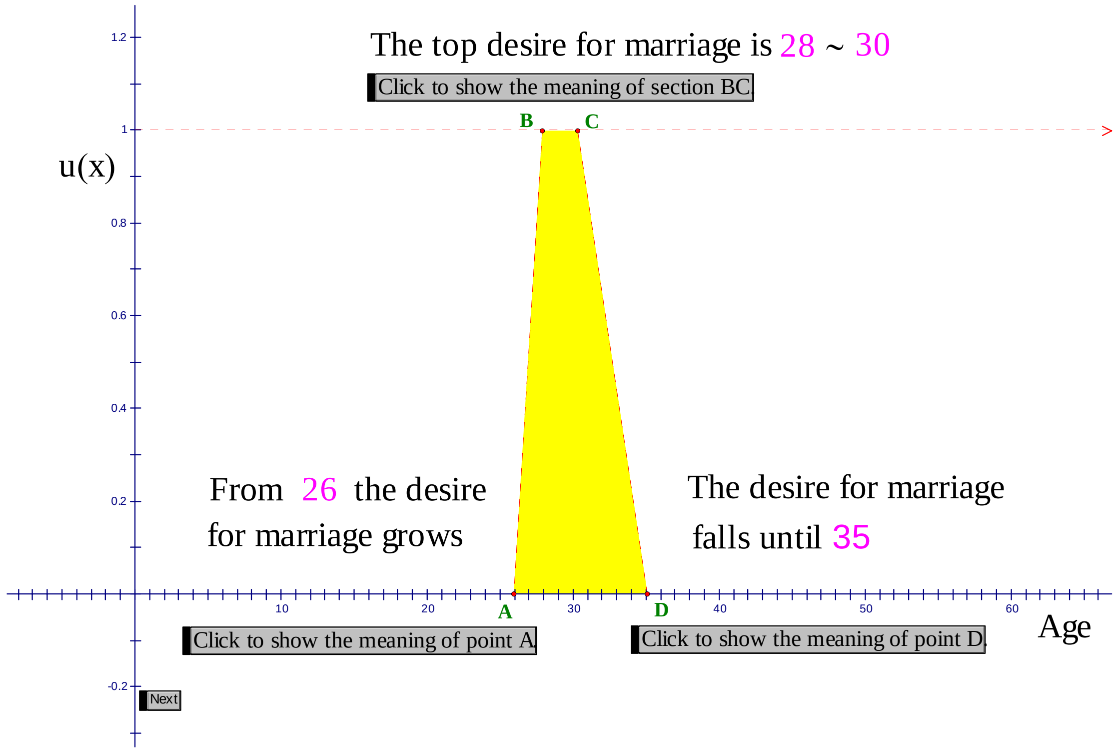

Figure 8 Sample for marriage

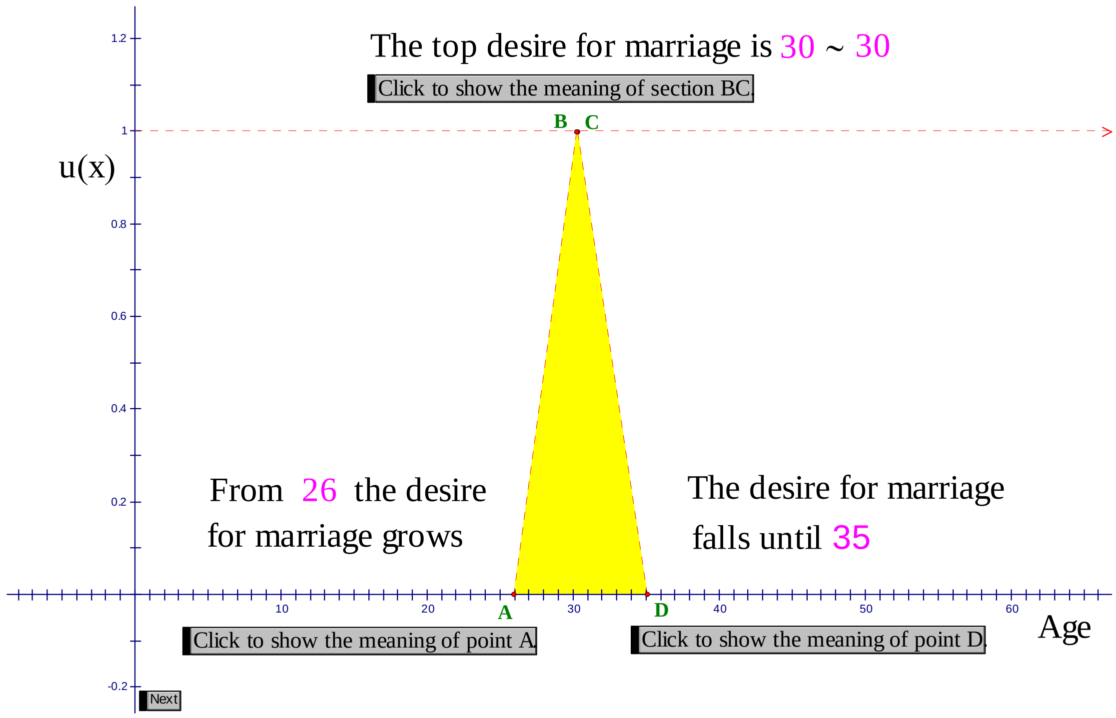

Figure 9 Sample for marriage

Fig. 9 is a special case of trapezoid when point *B* equals to point *C*. It represents the prime time for marriage is only 30. Fig. 10 shows the prime time for marriage is 28~30; moreover, the other ages are not concerned.

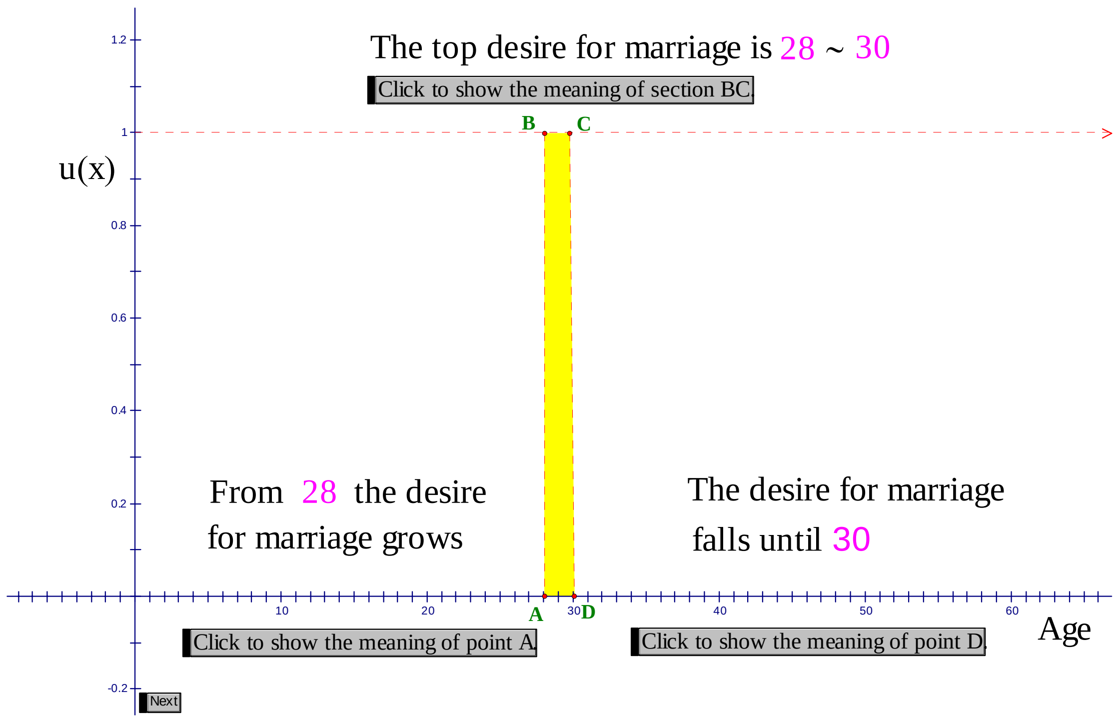

Figure 10 Sample for marriage

The following procedures illustrate how to get a fuzzy sample.

Step 1. Click the three buttons.

Step 2. Move points A, B, C, and D.

2.4 Ranking fuzzy data

Ranking method for continuous fuzzy numbers has so far not been defined uniquely semantically, and probably never will. Now we put our focus on two common methods. One uses the center point of each fuzzy data to rank interval-valued data. The other is called distance method proposed by Cheng [1]. It ranks fuzzy data by using the distance between geometric point and original point. The first method limit fuzzy data to interval-valued ones and Cheng's distance method is relatively complex in computing. In order to solve the problems we proposed a method is not only relatively easy in computing but also has fewer limitations to fuzzy data types.

The idea of our method is to generalize the ranking method for discrete types to continuous ones. Most scholars have no disagreement about the ranking method for discrete fuzzy numbers. The most common approach is as following.

Definition 2.4 *Suppose* $L'_1, L'_2, \cdots L'_n$ *are real values corresponding to n-linguistic variables. Let* $\tilde{A} = \frac{u_1(x)}{L'_1} + \frac{u_2(x)}{L'_2} + \cdots + \frac{u_n(x)}{L'_n}$, $\tilde{B} = \frac{u_1^{'}(x)}{L'_1} + \frac{u_2^{'}(x)}{L'_2} + \cdots + \frac{u_n^{'}(x)}{L'_n}$ *be two discrete fuzzy numbers and* $R(\tilde{A}) = \sum_{i=1}^{n} L'_i u_i(x)$, $R(\tilde{B}) = \sum_{i=1}^{n} L'_i u'_i(x)$. *We said*

$$\begin{array}{lll} \tilde{A} > \tilde{B} & \text{if} & R(\tilde{A}) > R(\tilde{B}), \\ \tilde{A} < \tilde{B} & \text{if} & R(\tilde{A}) < R(\tilde{B}), \\ \tilde{A} = \tilde{B} & \text{if} & R(\tilde{A}) = R(\tilde{B}). \end{array}$$

Continuous fuzzy numbers can generally be described by corresponding membership functions in the following form.

$$u_{\tilde{A}}(x) = \begin{cases} u_{\tilde{A}}^{L}(x) & \text{for} & a \le x \le b \\ 1 & \text{for} & b \le x \le c \\ u_{\tilde{B}}^{R}(x) & \text{for} & c \le x \le d \\ 0 & \text{for} & \text{otherwise} \end{cases} \qquad (2.2)$$

where $a \le b \le c \le d$. $u_{\tilde{A}}^{L}(x)$ is a monotone increasing real function and $u_{\tilde{A}}^{R}(x)$ is a monotone decreasing real function.

Property 2.1 *Let* $\tilde{A} = [a, b, c, d]$ *be a fuzzy number and* $u_{\tilde{A}}(x)$ *is its corresponding membership function. If* $u_{\tilde{A}}(x)$ *is bounded and* $\tilde{A}$ *can be covered by a finite interval then* $Q(\tilde{A}) = \frac{\int x u_{\tilde{A}}(x) dx}{\int u_{\tilde{A}}(x) dx}$

Proof

Let $x_0 = a,\ x_n = b$ and $x_i < x_{i+1}$ for $i = 0, 1, \cdots, n-1$ and then normalize each $u_{\tilde{A}}(x_i)$ by dividing $\sum_{i=1}^{n} u_{\tilde{A}}(x_i)$. We can find

$$Q(\tilde{A})=\frac{\sum_{i=1}^{n}x_i u_{\tilde{A}}(x_i)}{\sum_{i=1}^{n}u_{\tilde{A}}(x_i)}$$

Because $u_{\tilde{A}}(x)$ is bounded and $\tilde{A}$ is covered by a finite interval, hence as n approaches to infinity

$$Q(\tilde{A})=\lim_{n\to\infty}\frac{\sum_{i=1}^{n}x_i u_{\tilde{A}}(x_i)}{\sum_{i=1}^{n}u_{\tilde{A}}(x_i)}=\frac{\int x u_{\tilde{A}}(x)dx}{\int u_{\tilde{A}}(x)dx}.$$

For convenience,

$$\begin{aligned} Q(\tilde{A}) &= \frac{-(a+b)^2+(c+d)^2+(ab-cd)}{-3((a+b)-(c+d))}, && \text{if } \tilde{A} \text{ is trapezoid;} \\ Q(\tilde{A}) &= \frac{a+b+d}{3}, && \text{if } \tilde{A} \text{ is triangular;} \\ Q(\tilde{A}) &= \frac{b+c}{2}, && \text{if } \tilde{A} \text{ is interval-valued.} \end{aligned} \tag{2.3}$$

Additionally, let $S=\{\tilde{A}_1,\tilde{A}_2,\cdots,\tilde{A}_n\}$ be a convex fuzzy set. For any $\tilde{A}_i,\tilde{A}_j\in S$, we can rank

$$\begin{aligned} \tilde{A}_i &> \tilde{A}_j && \text{if } Q(\tilde{A}_i)>Q(\tilde{A}_j), \\ \tilde{A}_i &< \tilde{A}_j && \text{if } Q(\tilde{A}_i)<Q(\tilde{A}_j), \\ \tilde{A}_i &= \tilde{A}_j && \text{if } Q(\tilde{A}_i)=Q(\tilde{A}_j). \end{aligned} \tag{2.4}$$

We summarize the three methods mentioned above in Table 1.

Table 1 Compare with three different ranking methods

Method	Data Restriction	Ranking Rules
Center point	Only interval-value	Center point of each fuzzy data
Distance method	No	$\tilde{x}(i)=\frac{\int xu(x)dx}{\int u(x)dx}$, $\tilde{y}(i)=\frac{\int_0^1 yg(y)dy}{\int_0^1 g(y)dy}$ $Q(\tilde{A}_i)=\sqrt{(\tilde{x})^2+(\tilde{y})^2}$
Lin	No	$Q(\tilde{A}_i)=\frac{\int xu(x)dx}{\int u(x)dx}$

Example 2.7 *In Fig. 11 three fuzzy numbers* $\tilde{A}, \tilde{B}, \tilde{C}$ *are shown. The following are their membership functions, respectively.*

$$u_{\tilde{A}}(x) = \begin{cases} \dfrac{x-3}{3} & \text{for} \quad 3 \le x \le 6 \\ 1 & \text{for} \quad x = 6 \\ \dfrac{7-x}{1} & \text{for} \quad 6 \le x \le 7 \end{cases} \qquad u_{\tilde{B}}(x) = \begin{cases} \dfrac{x-3}{1} & \text{for} \quad 3 \le x \le 4 \\ 1 & \text{for} \quad 4 \le x \le 5 \\ \dfrac{7-x}{2} & \text{for} \quad 5 \le x \le 7 \end{cases}$$

$$u_{\tilde{C}}(x) = \begin{cases} \dfrac{x-3}{1.5} & \text{for} \quad 3 \le x \le 4.5 \\ 1 & \text{for} \quad x = 4.5 \\ \dfrac{7-x}{2.5} & \text{for} \quad 4.5 \le x \le 7 \end{cases}$$

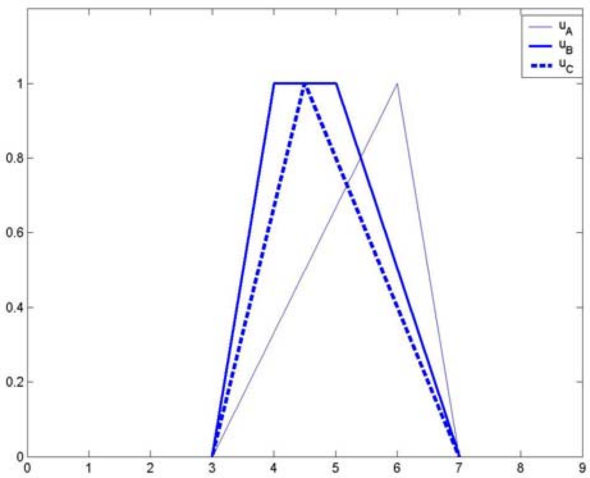

Figure 11 Shapes with three fuzzy numbers

By using Equ. (2.3) and Equ. (2.4)

$$Q(\tilde{A}) = \frac{3+6+7}{3} = 5.333$$

$$Q(\tilde{B}) = \frac{-(3+4)^2 + (5+7)^2 + (12-35)}{-3((3+4)-(5+7))} = 4.8$$

$$Q(\tilde{C}) = \frac{3+4.5+7}{3} = 4.833$$

$$\because Q(\tilde{B}) < Q(\tilde{C}) < Q(\tilde{A}) \quad \therefore \tilde{B} < \tilde{C} < \tilde{A}$$

Definition 2.5 *Let* $\tilde{A}$, $\tilde{B}$ *be the fuzzy two fuzzy numbers,* $d(\tilde{A},\tilde{B})$ *is called the distance between two fuzzy numbers if* $d(\tilde{A},\tilde{B}) = \left|Q(\tilde{A}) - Q(\tilde{B})\right|$.

Example 2.8 *In Fig. 12 two fuzzy numbers* $\tilde{A},\tilde{B}$ *are shown. The following are their membership functions, respectively.*

$$u_{\tilde{A}}(x) = \begin{cases} \dfrac{x-3}{3} & \text{for} \quad 3 \le x \le 6 \\ 1 & \text{for} \quad x = 6 \\ \dfrac{7-x}{1} & \text{for} \quad 6 \le x \le 7 \end{cases} \qquad u_{\tilde{B}}(x) = \begin{cases} \dfrac{x-6}{2} & \text{for} \quad 6 \le x \le 8 \\ 1 & \text{for} \quad x = 8 \\ \dfrac{8-x}{3} & \text{for} \quad 8 \le x \le 11 \end{cases}$$

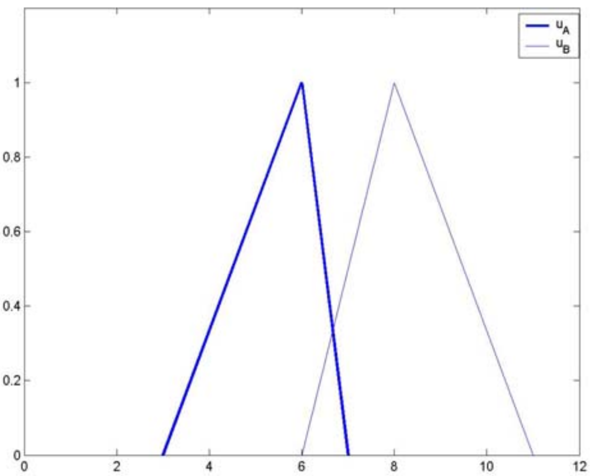

Figure 12 Shapes with two fuzzy numbers

By using Equ. (2.3)

$$Q(\tilde{A}) = \frac{3+6+7}{3} = \frac{16}{3}, \quad Q(\tilde{B}) = \frac{6+8+11}{3} = \frac{25}{3}$$

By Def. 2.5 given above, $d(\tilde{A},\tilde{B}) = \left|Q(\tilde{A}) - Q(\tilde{B})\right| = \left|\frac{16}{3} - \frac{25}{3}\right| = 3$.

Hence the distance between $\tilde{A}$ and $\tilde{B}$ is 3.