Chapter 5

Testing on Discrete and Continuous Fuzzy Numbers

5.1. Fuzzy Data with Soft Computing

5.1.1. Membership Function

In the traditional statistical approach, we assume that each expert provides the exact estimate of the desired quantity x, so, after polling the experts, we get, for each possible value x_i of this quantity, the probability $p(x_i)$ that a randomly chosen expert selected this value as his/her estimate.

In the fuzzy case, we take into consideration that an expert often cannot provide a definite estimates. The simplest case is when an expert has in mind several possible values of the estimated quantity. In this case, instead of asking each expert for a definite estimate, we can ask each experts which of v values they consider possible and which not. As a result, for each value x_i, we can count the total number $N(x_i)$ of experts who consider this value possible.

Similar to the statistical approach, we do not want to ask all N experts, so we would like to get a representative sample of $n << N$ experts, and ask only these n experts. In doing this, we hope that for each x_i the portion $\mu(x_i) = n(x_i)/n$ of experts in this sample who consider x_i to be a possible estimate is approximately the same as the portion $N(x_i)/N$ corresponding to all N experts. Therefore, we must consider the ratios $\mu(x_i) = n(x_i)/n$.

Similarly to probabilities, these numbers belong to the interval [0, 1]. However, we can no longer call these values as probabilities, because they do not necessarily add up to 1. For example, if every expert considers all the values possible, then $\mu(x_i)=1$ for all i, so the sum of these values is no longer equal to 0. The values $\mu(x_i)$ are called *degrees of possibility*, and the function μ which maps each value x_i into the corresponding degree $\mu(x_i)$ is called a *membership function*, or a *fuzzy set.*

In considering the question related with fuzzy property, we consider the information itself has the uncertainty and fuzzy property. Hence, let's firstly give an easy and precise explanation about fuzzy numbers.

Definition 5.1.1. Fuzzy Number

Let U *denote a universal set,* $\{A_i\}_{i=1}^n$ *be a subset of discussion factors on* U *, and* $\Lambda(A_i)$ *be a level set of* A_i *for* $\boldsymbol{i=1,2,\cdots,n}$ *. The fuzzy number of a statement or a term* X *over* U *is defined as:*

$$\mu_U(X)=\sum_{i=1}^{n}\mu_i(X)I_{A_i}(X) \tag{2.1}$$

where $\{\mu_i(X), 0\le \mu_i(X)\le 1\}_{i=1}^n$ *are set of membership functions for corresponding factor in* $\{A_i\}_{i=1}^n$*, and* $I_A(x)=1$*, if* $x\in A$*;* $I_A(x)=0$*, if* $x\notin A$*. If the domain of the universal set is continuous, then the fuzzy number can be written as*

$$\mu_U(X)=\int_{A_i\subseteq A}\mu_i(X)I_{A_i}(X).$$

In the research of social science, the sampling survey is always used to evaluate and understand public opinion on certain issues. The traditional survey forces people to choose one answer from the survey, but it ignores the uncertainty of human thinking. For instance, when people need to choose the answer from a survey which lists five choices, including "very satisfactory," "satisfactory," "normal," "unsatisfactory," and "very unsatisfactory," the traditional survey method become quite impractical.

The advantages of evaluation with fuzzy number are listed as follows:

(i) Evaluation process becomes robust and consistent by reducing the degree of subjectivity of the evaluator.

(ii) Self-potentiality is highlighted by indicating individual distinctions.

(iii) Provide the evaluators with an encouraging, stimulating, self-reliant guide that emphasizes on individual characteristics. While the drawback is that the calculating process will be more complex than the conventional one.

5.2. Fuzzy Mean

In the traditional statistical approach, we start with a collection of real numbers, i.e., in more precise term, we used *number-valued* statistics. In an interval situation, we start with a collection of intervals instead of a collection of numbers. So, if we use statistical methods to process this collection, it is natural to call these statistical methods *interval-valued*.

Let us see what statistical characteristics we can naturally extract from this collection. In traditional statistical techniques, each expert presents a single number, and from this collection, we can extract the probabilities $p(x_i)$ and cumulative probabilities $F(x_i)$. In the interval case, each expert presents *two* numbers x^- and x^+. So instead of a single collection of numbers, we have two collections: a collection of the lower endpoints x^-, and the collection of upper endpoints x^+. It is therefore natural to apply the standard statistical procedure to each of these collections.

Definition 5.2.1. Fuzzy Mean (Data with Interval Values)

Let U *be the universe set, and* $\{FS_i = [a_i, b_i], a_i, b_i \in R, i = 1, 2, \cdots, n\}$ *be a sequence of random fuzzy sample on* U. *Then the fuzzy expected value is defined as*

$$\overline{FX} = [\frac{1}{n}\sum_{i=1}^{n} a_i, \frac{1}{n}\sum_{i=1}^{n} b_i].$$

Fuzzy techniques were developed to describe and analyze the situation when an expert is not sure about the value of the estimated quantity x, and may consider several different values to be possible.

We start with the set $X = \{x_1, \cdots, x_v\}$ of all possible values of x. In principle, the values which an expert considers possible can form an arbitrary subset of this set. Usually, however, not all subsets occur. Typically, if an expert believes that two values $x' < x''$ are possible, than all intermediate values x (i.e., values for which $x' < x < x''$) are possible as well. In such a situation, to describe the set of all the values x which an expert considers possible, it is sufficient to describe the smallest x^- and the largest x^+ of these values; then all the values between x^- and x^+ are possible as well. In other words, in such situation, the set of all values x_i which an expert considers possible forms an *interval* $[x^-, x^+] = \{x^- \le x \le x^+\}$.

Therefore, all we have to collect from the experts are these intervals, i.e., to be more precise, their endpoints. As a result, we have a collection of intervals.

Definition 5.2.2. Fuzzy Mean for Unbounded Sample

Let U *be the universe set, and* $\{FS_i = [a_i, \infty), a_i \in R, i = 1, 2, \cdots, n\}$ *be a* *sequence of random fuzzy sample on* U *. Then the fuzzy expected value is defined as*

$$\overline{FX} = [\frac{a_1 + a_2 + \cdots + a_n}{n}, \infty)$$

If $X_1 = (-\infty, b_1]$, …, $X_n = (-\infty, b_n]$, *then the fuzzy expected value is defined as*

$$\overline{FX} = (-\infty, \frac{b_1 + b_2 + \cdots + b_n}{n}].$$

Example 5.2.3.

In a survey with the starting salary for the new undergraduate students' salary, we find the following 5 data as follows:

[1000, 2000], [2000, 2500], [3000, 4000], [1500, 2000], [1000, 1500]

Then according to the definition 3.2.1, the fuzzy mean becomes:

$$\overline{FX} = [\frac{1+2+3+1.5+1}{5}, \frac{2+2.5+4+2+1.5}{5}] = [1.7, 2.4] \quad \text{(unit is thousand)}$$

Definition 5.2.4. Defuzzyfication for Discrete Fuzzy Data

Let D *be a fuzzy sample on universe domain* U *with ordered linguistic variable* $\{L_i : i = 1, 2, \cdots, k\}$. $\mu_D(L_i) = m_i$ *is the membership with respect to* L_i, $\sum_{i=1}^{n} \mu_D(L_i) = 1$. *We say* $D_f = \sum_{i=1}^{k} m_i L_i$ *is the defuzzyfication value for discrete fuzzy data* D.

Definition 5.2.5. Defuzzyfication for Interval Fuzzy Data

Let C *be a fuzzy sample on universe domain* U *with support on* $[a, b]$. *The membership of* $\mu_C(x) = f(x)$, $0 \le f(x) \le 1$ *if* $x \in [a, b]$, *is a convex function. Then*

$$C_f = \frac{\int_a^b xf(x)dx}{\int_a^b f(x)dx}$$

is called the defuzzyfication value for interval fuzzy data.

5.3. Some Properties and Soft Computing of Fuzzy Data

Soft computing is the most important tool on arithmetic manipulation of fuzzy numbers. After soft computing, we introduce some properties of fuzzy number including “fuzzy equal”, “fuzzy belongs”, “fuzzy index equal”, as well as “fuzzy distance”. We also give detailed discussion on “fuzzy distance” which includes four types. From theorem 5.3.1. to theorem 5.3.7., we give more detailed discussions on properties of distance.

5.3.1. Fuzzy Equal and Fuzzy Belongs for Fuzzy Data

We give the definition of the fuzzy equal for discrete data as the beginning of this section.

Definition 5.3.1. Fuzzy Equal for Discrete Data

Let U *be a universe domain and* $L = \{L_1, L_2, \cdots, L_k\}$ *be sequence of rank ordering of linguistic variables on* U. $\{X_i = \frac{m_{i1}}{L_1} + \frac{m_{i2}}{L_2} + \cdots + \frac{m_{ik}}{L_k}, i = 1, 2\}$, $\sum_{j=1}^{k} m_{ij} = 1$ *are two random samples from* U. *If* $m_{1j} = m_{2j}$ *for* $j = 1, 2, \cdots, k$. *Then we say that* X_1 *fuzzy equals to* X_2*, denoted by* $X_1 \approx_F X_2$.

Definition 5.3.2. Fuzzy Index Equal for Discrete Data

Let U *be a universe domain,* $L = \{L_1, L_2, \cdots, L_k\}$ *be sequence of rank* *ordering of linguistic variables on* U *, and* $\{X_i = \frac{m_{i1}}{L_1} + \frac{m_{i2}}{L_2} + ... + \frac{m_{ik}}{L_k}, i = 1, 2\}$,

$\sum_{j=1}^{k} m_{ij} = 1$ *are two random samples from* U *. Then the center of fuzzy number for discrete type is* $CX_i = \sum_{j=1}^{k} m_{ij} L_j$ *. If* $CX_1 = CX_2$ *, we say that* X_1 *fuzzy index equals to* X_2*, denoted by* $X_1 \approx_I X_2$.

Definition 5.3.3. Fuzzy Equal for Interval Data

Let A*,* B *be two fuzzy data with membership functions* $\mu_A(x) = f(x)$ *and* $\mu_B(y) = g(y)$ *where* $0 \le f(x) \le 1$ *if* $x \in [a, b]$ *and* $0 \le g(y) \le 1$ *if* $y \in [c, d]$ *. If* A*,* B *have the same support and* f *,* g *are all convex functions then we say* A *is fuzzy equal to* B *, written as* $A =_F B$ *, or briefly written as* $A =_F [a, b]$.

For left unbounded or right unbounded, the definitions are similar.

Definition 5.3.4. Fuzzy Belongs for Interval Data

Let A*,* B *be two fuzzy data with membership functions* $\mu_A(x) = f(x)$ *and* $\mu_B(y) = g(y)$ *where* $0 \le f(x) \le 1$ *if* $x \in [a, b]$ *and* $0 \le g(y) \le 1$ *if* $y \in [c, d]$ *. If the support of* A *is contained in the support* B *and* f *,* g *are all convex functions, then we say* A *is fuzzy belongs to* B*, written as* $A \in_F B$*, or briefly written.* $A \in_F [c, d]$.

In order to set up an appropriate testing hypothesis on the fuzzy data, it is necessary to give definitions about measurement of distance of fuzzy set. In the following, we set up firstly, the definition of fuzzy distance with fuzzy interval data. The definition is different from the traditional interval operations. Our consideration is concentrated on the statistical point of view.

Definition 5.3.5. Distance of Fuzzy Interval Set

Let A, B *be two fuzzy data with membership functions* $\mu_A(x) = f(x)$ *and* $\mu_B(y) = g(y)$ *where* $0 \le f(x) \le 1$ *if* $x \in [a,b]$ *and* $0 \le g(y) \le 1$ *if* $y \in [c,d]$. *We give four definitions of distance:*

$$d_1(A,B) = \inf\{|x-y| : x \in A, y \in B\}$$

$$d_2(A,B) = \sup\{|x-y| : x \in A, y \in B\}$$

$$d_3(A,B) = \inf\{\varepsilon_1, \varepsilon_2\}$$

$$d_4(A,B) = \sup\{\varepsilon_1, \varepsilon_2\}$$

where $\varepsilon_1 = \inf\{\varepsilon : [c,d] \subset [a-\varepsilon, b+\varepsilon]\}$ *and* $\varepsilon_2 = \inf\{\varepsilon : [a,b] \subset [c-\varepsilon, d+\varepsilon]\}$.

Example 5.3.6. Fuzzy Distance

Let A, B be two fuzzy data with support [1, 3], [2, 5]. Then it is easy to calculate $d_1(A,B)$ and $d_2(A,B)$ as follows:

$$d_1(A,B) = \inf\{|x-y| : x \in [1,3], y \in [2,5]\} = 0$$

$$d_2(A,B) = \sup\{|x-y| : x \in [1,3], y \in [2,5]\} = 4$$

To calculate $d_3(A,B)$ and $d_4(A,B)$, we should calculate ε_1 and ε_2 first as follows:

$$\varepsilon_1 = \inf\{\varepsilon : [2,5] \subset [1-\varepsilon, 3+\varepsilon]\} = 2$$

$$\varepsilon_2 = \inf\{\varepsilon : [1,3] \subset [2-\varepsilon, 5+\varepsilon]\} = 1$$

By Definition 5.3.5, the other two distant are as follows:

$$d_3(A,B) = \inf\{2,1\} = 1$$

$$d_4(A,B) = \sup\{2,1\} = 2.$$ ❑

5.3.2. Some Properties about Fuzzy Data

Theorem 5.3.1. Let A, B be two fuzzy data with membership functions

$\mu_A(x) = f(x)$ and $\mu_B(y) = g(y)$ where $0 \le f(x) \le 1$ if $x \in [a,b]$ and $0 \le g(y) \le 1$ *if* $y \in [c,d]$, the fuzzy equals implies fuzzy belongs.

Proof: Gives $A =_F [a,b]$, since $[a,b] = [a,b]$ implies $[a,b] \subseteq [a,b]$. Hence we conclude that $A \in_F [a,b]$. ❑

Theorem 5.3.2. For any fuzzy set C with support $[m,n]$ and has no intersection with the support of A and B.

(i) If $b < m$, $d < m$ and $d_1(A,C) = d_1(B,C)$, $d_2(A,C) = d_2(B,C)$, then $A =_F B$.

(ii) If $a > n$, $c > n$ and $d_1(A,C) = d_1(B,C)$, $d_2(A,C) = d_2(B,C)$, then $A =_F B$.

Proof: (i) By definition $d_1(A,C) = d_2(B,C)$ give that

$$\mathbf{inf\{| x - z |: x \in A, z \in C\} = inf\{| y - z |: y \in B, z \in C\}}$$

This imply that $m - b = m - d$, i.e. $b = d$. Similarly, by definition $d_2(A,C)$ = $d_2(B,C)$ give that

$$\mathbf{sup\{| x - z |: x \in A, z \in C\} = sup\{| y - z |: y \in B, z \in C\}}$$

This imply that $n - a = n - c$, i.e. $a = c$. Therefore $A =_F B$.

(ii) By using the similarly argument as part (i), it is easily to conduct the proof. ❑

Theorem 5.3.3. For any fuzzy set C with support $[m,n]$ and has no intersection with the support of A and B. If $A \in_F B$, then $d_1(A,C) \ge d_1(B,C)$ and $d_2(A,C) \le d_2(B,C)$.

Proof: We will only prove the case $d < m$, the proofs of other cases are similar. The fuzzy belongs $A \in_F B$ imply that

$$a \ge c \quad \text{and} \quad b \le d$$

Let m subtract by b and d; and let n subtract by a and c yield that

$$m - b \ge m - d \quad \text{and} \quad n - a \le n - c$$

Hence we have

$$d_1(A, C) \geq d_1(B, C) \quad \text{and} \quad d_2(A, C) \leq d_2(B, C). \qquad \square$$

We note that the inverse direction of theorem 3.3.3 is not true. For example, choose $[a, b] = [1, 3]$, $[c, d] = [15, 18]$, $[m, n] = [8, 10]$ then

$d_1(A, C) \geq d_1(B, C)$ and $d_2(A, C) \leq d_2(B, C)$ but $A \not\leq_F B$

Theorem 5.3.4. Suppose that $[a, b] \cap [c, d] = \phi$ and $d - c \geq b - a$. Then

$$d_3(A, B) = \begin{cases} c - a & \text{if } b < c \\ b - d & \text{if } a > d \end{cases}.$$

Proof: We will only consider the case $b < c$, the proof of case $a > d$ is similar. To obtain the fuzzy distance d_3, we firstly calculate ε_1 and ε_2 as follows:

$$\varepsilon_1 = \inf\{\varepsilon : [c, d] \subset [a - \varepsilon, b + \varepsilon]\} = d - b$$

and

$$\varepsilon_2 = \inf\{\varepsilon : [a, b] \subset [c - \varepsilon, d + \varepsilon]\} = c - a$$

By comparing these two values, we have

$$\varepsilon_1 - \varepsilon_2 = (d - c) - (b - a) \geq 0$$

Therefore

$$d_3(A, B) = \inf\{\varepsilon_1, \varepsilon_2\} = c - a \qquad \square$$

Theorem 5.3.5. Suppose that $[a, b] \cap [c, d] \neq \phi$, $[a, b] \not\subset [c, d]$, and $d - c \geq b - a$, then

$$d_3(A, B) = \begin{cases} c - a & \text{if } a < c \\ b - d & \text{if } b > d \end{cases}.$$

Proof: Since

$$d - c \geq b - a$$

we have

$$d - b \geq c - a$$

The other parts of proof are similar to that of Theorem 3.3.4. ❑

Theorem 5.3.6. If $[a,b]\subset[c,d]$ then $d_3(A,B)=0$.

Proof: First, we calculate ε_1 and ε_2 as follows:

$$\varepsilon_1=\inf\{\varepsilon:[c,d]\subset[a-\varepsilon,b+\varepsilon]\}>0$$

and

$$\varepsilon_2=\inf\{\varepsilon:[a,b]\subset[c-\varepsilon,d+\varepsilon]\}=0$$

Therefore $d_3(A,B)=0$. ❑

Theorem 5.3.7. Give a fuzzy set C with support $[m,n]$.

(i) If $m-n\geq d-c$ and $A\in_F B$ then $d_3(A,B)\leq d_3(B,C)$.

(ii) If $m-n\geq d-c$ and $A=_F B$ then $d_3(A,B)=d_3(B,C)$.

Proof: (i) The fuzzy belongs $A\in_F B$ imply that $a\geq c$ and $b\leq d$. Let consider the following cases:

Case 1: If $d<m$ then $d_3(A,C)=m-a\leq m-c=d_3(B,C)$.

Case 2: If $c>n$ then $d_3(A,C)=b-n\leq d-n=d_3(B,C)$.

Case 3: If $c<m$, $a<m$, and $d\geq m$ then

$$d_3(A,C)=m-a\leq m-c=d_3(B,C).$$

Case 4: If $c<m$ and $a\geq m$ then $d_3(A,C)=0<m-c=d_3(B,C)$.

Case 5: If $c\geq m$ and $d\leq n$ then $d_3(A,C)=0=d_3(B,C)$.

Case 6: If $c\leq n$, $b>n$, and $d>n$ then $d_3(A,C)=b-n\leq d-n=d_3(B,C)$.

Case 7: If $c\leq n$, $b\leq n$, and $d>n$ then $d_3(A,C)=0<d-n=d_3(B,C)$.

After considering all possible cases, we completed the proof of (i). Similarly, by considering all the possible cases, the proof of part (ii) can be easily obtained. ❑

5.4. Testing Hypothesis with Fuzzy Data

It is a new research topic about the hypothesis testing of fuzzy mean with interval values. First of all, we will give a definition about the defuzzification. Then under the fuzzy significant level δ, we make a one side or two side testing. The side of these methods are a little different from traditional significant level α. In order to get the robust properties, we will set up the rejection area level F_δ, according to the fuzzy population.

5.4.1. Testing Hypothesis for Fuzzy Equal

Let U be the universal set (a discussion domain), $L=\{L_1, L_2, \cdots, L_k\}$ a set of k-linguistic variables on U, and $A=\{A_1, A_2, \cdots, A_m\}$ and $B=\{B_1, B_2, \cdots, B_n\}$ be two sets drawn from categorical populations with numbers on U. For each sample in $\{A_j, B_j\}$, assign a linguistic variable L_j and a normalized membership m_{ij} where $\sum_{j=1}^{k} m_{ij} = 1$, and let $Fn_{ij} = \sum_{i \in A, B} Ln_{ij}$, $j = 1, 2, \cdots, k$ be the total memberships in the cell (i, j). The following statements are process for testing hypothesis

Testing hypothesis of fuzzy equal for discrete fuzzy mean

Consider a K-cell multinomial vector $\boldsymbol{m=\{n_1, n_2, ..., n_k\}}$ with $\sum_{i=1}^{k} n_i = n$.

The *Pearson chi-squared test* ($\chi^2 = \sum_i \sum_j \frac{(n_{ij}-e_{ij})^2}{e_{ij}}$) is a well known statistical test for investigating the significance of the differences between observed data arranged in K classes and the theoretically expected frequencies in the K classes. It is clear that the large discrepancies between the observed data and expected cell counts will result in larger values of χ^2.

However, a somewhat ambiguous question is whether (quantitative) discrete data can be considered categorical and use the traditional χ^2-test. For example, suppose a child is asked the following question: "*How much do you love your sister*?" If the responses is a fuzzy number (say, 70% of the time), it is certainly inappropriate to use the traditional χ^2-test for the analysis. We will present a χ^2-test for fuzzy data as follows:

Procedures for Testing Hypothesis of Fuzzy Equal for Discrete Fuzzy Mean

1. *Hypothesis*: Two populations have the same distribution ratio.
2. *Statistics*: $\chi^2 = \sum_{i\in A,B}\sum_{j=1}^{c}\frac{([Fn_{ij}]-e_{ij})^2}{e_{ij}}$.

   (In order to perform the chi-square test for fuzzy data, we transfer the decimal fractions of Fn_{ij} in each cell of fuzzy category into the integer $[Fn_{ij}]$ by counting 0.5 or higher fractions as 1 and discard the rest.)
3. *Decision rule*: Under significance level α, if $\chi^2 > \chi^2_\alpha(k-1)$, then we reject null hypothesis.

Testing Hypothesis of Fuzzy Index Equal for Discrete Fuzzy Mean

Let $\overline{FX}$ be the fuzzy sample mean, $\overline{X}_f$ be the defuzzyfication of $\overline{FX}$. Under the fuzzy significant level F_δ, and the corresponding critical value F_δ, we want to test

$$\boldsymbol{H_0: F\mu = \overline{FX}}$$

$$\boldsymbol{H_1: F\mu \neq \overline{FX}}$$

where $F\mu$ is the fuzzy mean of the underlying population. Let μ be the defuzzyfication value of $F\mu$, then the above hypothesis becomes $H_0: \mu = \mu_0$. The test procedure is listed as follows:

1. *Hypothesis*:

$$H_0: F\mu = F\mu_0$$

$$H_1: F\mu \neq F\mu_0$$

2. *Statistics*: Find the fuzzy mean $\overline{FX}$ from a random sample $\{S_i, i = 1, 2, ..., n\}$.

3. *Decision rule*: Under the fuzzy significant level F_δ, if $|\overline{X}_f - \mu_0| > \delta$, then reject H_0; otherwise do not reject H_0.

Note that the left side test

$$H_0: F\mu = F\mu_0$$

$$H_1: F\mu > F\mu_0$$

Under the fuzzy significant level F_δ, the rejection region is replaced by $\mu_0 - \overline{X}_f > \delta$. The right side test can be conducted by a similarly way.

Testing Hypothesis with Continuous Fuzzy Mean

1. *Hypothesis*:

$$H_0: F\mu =_F [a, b]$$

$$H_1: F\mu \neq_F [a, b]$$

2. *Statistics*: Find the fuzzy mean $\overline{FX} = [x_l, x_u]$ from a random sample $\{S_i, i = 1, 2, ..., n\}$.

3. *Decision rule*: Under the significant level F_δ, calculate the value k such that $k = \delta(b - a)$, if $|x_l - a| > k$ or $|x_u - b| > k$ then reject H_0; otherwise do

not reject H_0.

5.4.2. Testing Hypothesis for Fuzzy Belongs

The testing hypothesis for fuzzy belong can be performed in three ways, namely, fuzzy belong with bounded sample, fuzzy belong with bounded below sample, fuzzy belong with above sample. We will list three procedures as follows.

Testing of Fuzzy Belongs with Bounded Sample

1. *Hypothesis*:

   $H_0: F\mu \in_F [a, b]$

   $H_1: F\mu \notin_F [a, b]$

2. *Statistics*: Find the fuzzy mean $\overline{FX} = [x_l, x_u]$ from a random sample $\{S_i, i = 1, 2, ..., n\}$.

3. *Decision rule*: Under the significant level F_δ, calculate the value k such that $k = \delta(b-a)$, if $x_l < a-k$ or $x_u > b+k$ then reject H_0; otherwise do not reject H_0.

Testing of Fuzzy Belongs with Unbounded Below Sample

1. *Hypothesis*:

   $H_0: F\mu \in_F (-\infty, b]$

   $H_1: F\mu \notin_F (-\infty, b]$

2. *Statistics*: Find the fuzzy mean $\overline{FX} = (-\infty, x_u]$ from a random sample $\{S_i, i$

$= 1, 2, ..., n\}$.

3. *Decision rule*: Under the significant level F_δ, calculate the value k such that $k = \delta r$ where r is a constant, if $x_u > b + k$ then reject H_0; otherwise do not reject H_0.

Testing of Fuzzy Belongs to with Bounded Above Sample

1. *Hypothesis*:

$$H_0: F\mu \in_F [a, \infty)$$

$$H_1: F\mu \notin_F [a, \infty)$$

2. *Statistics*: Find the fuzzy mean $\overline{FX} = [x_l, \infty)$ from a random sample $\{S_i, i = 1, 2, ..., n\}$.

3. *Decision rule*: Under the significant level F_δ, calculate the value k such that $k = \delta r$ where r is a constant, if $x_l > a - k$ then reject H_0; otherwise do not reject H_0.

5.5. Empirical Studies

Example 5.5.1.

How do Chinese and English-speaking children's conditional reasoning and expressions develop over time? Is language different, such as English versus Chinese, related to children's understanding of conditionals? Among the testing stimuli, six conditional questions with different degrees of hypotheticality were asked based on a picture book to the two groups of children in their native language respectively. The questions are listed in the table 3.1.

Table 5.1: Types of Conditional Questions and Examples

Conditional Questions	**Examples**
1. Future open conditionals	If you ask your Mom whether she loves you, what will she say?
2. Present open conditionals	If somebody bites you, does it hurt?
3. Past open conditionals	There are lions in the zoo. If I have been to the zoo, would I see the lions?
4. Imaginative conditionals	Which animal would you like to be if you were the piglet? Why would you want to be a ______ ?
5. Present counterfactuals	The mother pig is afraid after the piglet becomes a lion because the lion might bite her with its sharp teeth. What if the lion didn't have sharp teeth?
6. Past counterfactuals	The piglet was a lion before. But he changed back to be a piglet again at the end. What if the piglet had not changed back to himself, what would the mommy pig have done then?

The test for fuzzy equals are shown in the table 3.2.

Table 5.2: A comparison of the traditional and fuzzy statistical analysis 6 conditional questions in table 5.1

Note: [a] Conditional response, [b] Indeterminate response, [c] Non-conditional response, [d] More than two cells numbers are ≤ 1, Chi-Square test is invalid. In order to perform the Chi-square test for fuzzy data, we transfer the decimal fractions in each cell of fuzzy category into the integer by counting 5 and higher fractions as 1 and discard the rest.

H_0: Language difference does not affect children's understanding of conditionals. i.e. $H_0 : F\mu_C =_F F\mu_E$		Con.[a]	Ind.[b]	Non. Con[c]	Chi-Square Test[d]
1. Future open conditionals	Chinese	29.8	21.2	5	$\chi^2 = 3.71$, $p = 0.16$ Accept H_0
	English	16.8	4.4	0.8	
2. Present open conditionals	Chinese	20.6	22.4	12.6	$\chi^2 = 4.68$, $p = 0.096$ Reject H_0
	English	14.2	5.6	2.2	
3. Past open conditionals	Chinese	22.8	23.2	10	$\chi^2 = 0.55$, $p = 0.76$ Accept H_0
	English	11.4	8	2.6	
4. Imaginative conditionals	Chinese	42.2	12.4	1.4	$\chi^2 = 0.41$ [d]
	English	17	4.8	0.2	
5. Present counterfactuals	Chinese	25	9.2	1.8	$\chi^2 = 0.34$, $p = 0.84$ Accept H_0
	English	14.4	6	1.6	
6. Past counterfactuals	Chinese	25.6	22.4	8	$\chi^2 = 2.76$, $p = 0.25$ Accept H_0
	English	10.8	5.2	6	

From Table 5.2 we find that there was no difference between these two groups. The fuzzy testing hypothesis of fuzzy equal for discrete fuzzy data, uses more differentiated categories and tends to reflect a more truthful picture of the data.

Example 5.5.2.

A farmer wants to adapt a new cooking style of fried chicken from traditional techniques. He invites 5 experts to join the evaluating experiment. After they tested the new fry chicken, they are asked to give a fuzzy grading with: very unsatisfactory (V. U.) = 1, unsatisfactory (U.) = 2, no difference (N. D.) = 3, satisfactory (S.) = 4, very satisfactory (V. S.) = 5. Table 5.3 shows the result of the 5 experts' evaluation.

Table 5.3: Evaluation Result for 5 Experts

Expert	V. U. Grade = 1	U. Grade = 2	N. D. Grade = 3	S. Grade = 4	V. S. Grade = 5
A	0	0	0	0.7	0.3
B	0	0	0	0	1.0
C	0	0.4	0.6	0	0
D	0	0	0	0.8	0.2
E	0.1	0.9	0	0	0

Let's set up the hypothesis testing for fuzzy index equal:

$$\boldsymbol{H_0} : \mathbf{F}_{\mu} = \mathbf{3}$$

$$\boldsymbol{H_1} : \mathbf{F}_{\mu} \neq \mathbf{3}$$

Under the significant level $\delta = 0.1$, since $\overline{X}_f = 2.4$ then we have

$$\mu_0 - \overline{X}_f = 3 - 2.4 = 0.6 > 0.1$$

Therefore we reject H_0. The manager conclude that fuzzy index $\overline{X}_f$ is less than 3, the manager will not apply this new cooking style.

Example 5.5.3.

A company administrator wants to control the time of turning on air-condition base on the energy saving reason. He feels that that the temperature over 28℃ will be hot and is the time to turn on. However, he wants to know how the other staff feels. So, he asks for five staffs at random to investigate and then gets five data as follows:

$$[27, \infty), [26, \infty), [29, \infty), [24, \infty), [26, \infty)$$

He need to do the following test

$$H_0: \mu = [28, \infty)$$

$$H_1: \mu \neq [28, \infty)$$

After simple computation, we have $\overline{FX}$ = [26.4, ∞). Under the significant level $\delta = 0.2$, since 28 – 26.4 > 0.2, we reject H_0 and suggest that turn on the air condition when the temperature is between 26℃ and 28℃.

Example 5.5.4.

The human resource department announced that 20 to 26 years old people request their salary between 20 thousands and 40 thousands with deviation 5 thousands. The manger asks the statistical department to check it up. Suppose they find 10 young man between 20 and 26 years old, survey their request salary, the sample data are as follows:

[3, 4], 1.8, [2, 3], [4, 6], [1.5, 2] [3, 4], 2, [2, 3], [3, 5], [2.5, 4]

where the unit is ten thousands. To test

$$H_0: F\mu \in_F [2, 4]$$

$$H_1: F\mu \notin_F [2, 4]$$

Here, w should treat 1.8 as [1.8, 1.8]. After simple computation, we get

$$\overline{FX} = [\frac{3+1.8+2+4+1.5+3+2+2+3+2.5}{10}, \frac{4+1.8+3+6+2+4+2+3+5+4}{10}]$$

= [2.38, 3.48]

Under the significant level $\delta = 0.5$, since 2.38 ＞ 2 – 0.5 and 3.48 ＜ 4 + 0.5 and 2.38 ＜ 2 + 0.5 but 3.48 ＜ 4 – 0.5, we do not accept what human resource department saying. We accept H_0, i.e. $\overline{FX} \in_F [2, 4]$.

Example5.5.5.

Suppose a salesman wonders how the living standard will influence the sales of volumn for two communities X and Y. They want to find out which has higher income level to make the sales strategy. He chooses 100 families at random from each community and gets data. He wants to test

H_0: Two communities have same income level

H_1: Two communities have different income level

After simple computation, we get $\overline{FX} = [4.3, 5.5]$ and $\overline{FY} = [6.7, 7.7]$ (in 10 thousands). Under the significant level $F_\delta = 1$, $\overline{FX}_\delta$ = 4.9, $\overline{FY}_\delta$ = 7.2. Since 7.2 – 4.9 ＞ 1, we conclude that community Y has higher income level than that of community X.