CONSTRUCTION OF NORMAL FUZZY NUMBERS: A CASE STUDY WITH EARTHQUAKE WAVEFORM DATA

Dhruba Das  
Department of Statistics, Gauhati University, Guwahati-781014, Assam, India.  
Dhrubadas16@gmail.com

Anamika Dutta  
Department of Statistics, Gauhati University, Guwahati-781014, Assam, India.  
anamika.dut268@gmail.com

Supahi Mahanta  
Department of Statistics, Gauhati University, Guwahati-781014, Assam, India.  
supahi_mahanta@rediffmail.com

Hemanta K. Baruah  
Department of Statistics, Gauhati University, Guwahati-781014, Assam, India.  
hemanta_bh@yahoo.com

Abstract: This article demonstrates that a normal fuzzy number can be constructed from earthquake waveform data. According to the Randomness-Fuzziness Consistency Principle, two independent laws of randomness in $[\alpha, \beta]$ and $[\beta, \gamma]$ are necessary and sufficient to define a normal fuzzy number $[\alpha, \beta, \gamma]$. In this article, we have shown how to construct normal fuzzy numbers using data from earthquake waveform and have studied the pattern of the membership curve of the waveform.

Keywords: Superimposition of Sets, Distribution Function, Membership function.

1. Introduction

A fuzzy real number $[\alpha, \beta, \gamma]$ is an interval around the real number $\beta$ with the elements in the interval being partially present. Partial presence of an element in a fuzzy set is defined by the name membership function. Based on the Randomness- Fuzziness Consistency Principle (Baruah, 2010, 2011a, 2011b, 2011c, 2012), in this article we shall show how to construct normal fuzzy numbers using the data of minimum and maximum amplitudes of every individual oscillation of the waveform of an earthquake that had occurred in the city of Guwahati on May 25, 1998.

The basic problem in constructing normal fuzzy numbers was the lack of understanding as to how exactly to define partial presence of an element in an interval. Indeed, various explanations regarding the possible relationship between probability and fuzziness have come up, and no concrete conclusion could be arrived at. Partial presence of an element in a set is expressed in terms of the fuzzy membership function. But how exactly to construct the membership function of a fuzzy number mathematically remained a problem. Baruah (2010, 2011a, 2011b, 2011c, 2012) has shown that two laws of randomness are necessary as well as sufficient to define a normal law of fuzziness. In other words, trying to frame one single law of probability from a given law of fuzziness, as had been tried upon while formulating the existing probability-possibility consistency principles, was not mathematically meaningful an exercise, because we need two laws of randomness, probabilistic or otherwise, and not one single law of probability, to define a law of fuzziness. This has led to a proper measure theoretic explanation of partial presence, and construction of fuzzy numbers can therefore be based on that.

We need to understand that if a variable $X$ can assume values in an interval $[L, U]$ where $L$ follows a law of randomness in the interval $[\alpha, \beta]$ while $U$ follows another law of randomness in the interval $[\beta, \gamma]$, then we are in a situation defining fuzzy uncertainty, with randomness defined in the measure theoretic sense. In such a case, Baruah’s principle of...
consistency between randomness and fuzziness states that the distribution function of $L$, which is known as the left reference function also with reference to fuzziness, in the interval $[a, \beta]$ together with the complementary distribution function of $U$ which is known as the right reference function also in the interval $[\beta, \gamma]$, would give us the membership function of a normal fuzzy number $[a, \beta, \gamma]$. The two concerned laws of randomness may or may not be geared to laws of probability because measure theoretically speaking the notion of probability need not actually appear in the definition of randomness in the sense that a probabilistic variable is necessarily random while a random variable need not be probabilistic. It should be noted that the notion of probability does not enter into the definition of a random variable (Rohatgi and Saleh, 2001, pages 41 – 43). When a variable is probabilistic, it has to be random by definition, although when a variable is random, it need not be probabilistic. Accordingly, all results of the classical theory of probability are automatically applicable to a random variable defined in the measure theoretic sense.

In what follows, we are going to explain how exactly a fuzzy number originates. We are going to show how exactly to construct a fuzzy number. We shall not assume anything heuristic in principle. The data collected from the aforesaid earthquake waveform, we now proceed to construct a normal fuzzy number.

2. Methodology

We collected data of the waveform of an earthquake in the city of Guwahati, India. The earthquake occurred on May 25, 1998. To represent the waveform in the form of a normal fuzzy number, we have considered the part of the waveform that had appeared while the earthquake was at its maximum amplitude to zero amplitude. In fact, the waveform of an earthquake is indeed a combination of two independent normal fuzzy numbers, one represented by the waveform from amplitude zero to the maximum amplitude and the other represented by the waveform from the maximum amplitude to zero amplitude. Here our interest is to study the second case. The first case can also be studied in the same way.

We have collected 39 observations from the highest amplitude to zero amplitude of the waveform. We thus have collected the values taken by $L$ and $U$ on those 39 observations. These values are, say

\[(a_1, a_2, a_3, \ldots, a_{39})\]

and

\[(b_1, b_2, b_3, \ldots, b_{39})\]

respectively.

Now, using the operation of set superimposition defined by Baruah (2010, 2011a, 2011b, 2011c, 2012) we may proceed to construct normal fuzzy numbers as discussed in (Das et al, 2013), which would define the uncertainty associated with waveform variations.

Baruah defined the operation of superimposition of two real intervals $[a_1, b_1]$ and $[a_2, b_2]$ as

\[[a_1, b_1] \cap [a_2, b_2] = [a_1 \wedge a_2, b_1 \vee b_2] = [\min(a_1, a_2), \max(b_1, b_2)]\]

where for example $[a_2, b_1]$ represents the elements of $[a_2, b_1]$ with every element being present twice. A closer look would reveal that

\[[a_1, b_1] \cap [a_2, b_2] = [a_1 \wedge a_2, b_1 \vee b_2] = [a_1 \min(a_1, a_2), b_1 \max(b_1, b_2)]\]

where

\[a_1 \min(a_1, a_2), \quad b_1 \max(b_1, b_2)\]

Here it was assumed without loss of any generality that $[a_1, b_1] \cap [a_2, b_2]$ is not void, or in other words that $\max (a_i) \leq \min (b_1), i = 1, 2$.

The reader may observe the entry of ordered values in this expression. This would now lead us to the exact definition of partial presence of an element in a set, better known as fuzziness. Unless one looks into the matters through this lens, the way to construct a fuzzy number would not be clear. Double

http://japmnt.com/
representation as such may not be quite important. However, if the elements in the intervals \([a_1, b_1]\) and \([a_2, b_2]\) are presumed to be partially present with level of presence of all elements in the two intervals being equal to \(\frac{1}{2}\), then the identity would be

\[
[a_1, b_1]^{(1/2)} (S) [a_2, b_2]^{(1/2)} = [a_1, (a_2) \cup [a_2, b_1]^{(1/4)} U [b_1, b_2]^{(2/3)}]
\]

Here \([a_1, b_1]\) represents the interval \([a_1, b_1]\) where every element of the interval is partially present with level of presence of every element being equal to \(\frac{1}{2}\).

Now suppose

\[
a_1, a_2, \ldots, a_{39} \text{ arranged in increasing order of magnitude, and}
\]

\[
b_1, b_2, \ldots, b_{39} \text{ arranged in increasing order of magnitude.}
\]

are values of \(a_1, a_2, \ldots, a_{39}\) and \(b_1, b_2, \ldots, b_{39}\) respectively.

Now suppose

\[
a_1, a_2, a_3, \ldots, a_{39} \text{ are values of } a_1, a_2, \ldots, a_{39}\text{ arranged in}
\]

\[
b_1, b_2, b_3, \ldots, b_{39} \text{ arranged in increasing order of magnitude.}
\]

are values of \(b_1, b_2, \ldots, b_{39}\) and \(a_1, a_2, \ldots, a_{39}\) respectively.

Superimposing the intervals

\[
[a_1, b_1], [a_2, b_2], \ldots, [a_{39}, b_{39}]
\]

and thereafter normalizing in the aforesaid manner is equivalent to superimposing the intervals

\[
[a_1, b_1]^{(1/39)}, [a_2, b_2]^{(1/39)}, \ldots, [a_{39}, b_{39}]^{(1/39)}
\]

with constant level of partial presence \(1/39\) for every interval. We shall thus get, subject to the condition that \([a_1, b_1] \cap [a_2, b_2] \cap [a_3, b_3] \cap \ldots \cap [a_{39}, b_{39}]\) is not void,

\[
[a_1, b_1]^{(1/39)} (S) [a_2, b_2]^{(1/39)} (S) [a_3, b_3]^{(1/39)} (S) \ldots \ldots \ldots (S) [a_{39}, b_{39}]^{(1/39)} = [a_1, (a_2) \cup [a_2, a_3]^{(2/39)} U [a_3, a_4]^{(3/39)} \cup \ldots \ldots U [a_{39}, b_1]^{(1/4)} U [b_1, b_2]^{(38/39)} U [b_2, b_3]^{(37/39)} \cup \ldots \ldots U [b_{38}, b_39]^{(4/39)}]
\]

where for example, \([b_1, b_2]^{(38/39)}\) represents the interval \([b_1, b_2]\) with level of partial presence of every element being \(38/39\) for every element in the interval.

The actual diagram would thus be a simple function in the measure theoretic sense from \(a_1\) to \(a_{39}\), and another simple function from \(b_1\) to \(b_{39}\). The first of these two simple functions is non-decreasing while the second is non-increasing. Smoothing of these two functions will ultimately lead to the membership function of our normal fuzzy number representing one half of the waveform. We shall now proceed to construct the membership function from the data collected from an actual earthquake.

1. Construction of the membership curve for the earthquake waveform

The original picture of the waveform of the earthquake has been shown in fig. 1, which has been obtained from an earthquake measuring device at Gauhati University and the date of that particular waveform under study was 25-05-1998.

![Fig. 1: Waveform of the earthquake of May 25, 1998, in Guwahati.](http://japmnt.com/)

From the waveform, we determined data regarding the minimum and the maximum amplitudes for 39 observations. The minimum amplitudes in increasing order of magnitude (in cm.) were

-1.15, -1.14, -1.1, -1.05, -0.81, -0.8, -0.7, -0.6, -0.51, -0.5, -0.49, -0.48, -0.47, -0.46, -0.45, -0.44, -0.43, -0.42, -0.41, -0.37, -0.3, -0.29, -0.28, -0.27, -0.26, -0.25, -0.24, -0.23, -0.22, -0.21, -0.2, -0.19, -0.15, -0.14, -0.13, -0.12, -0.11, -0.1, 0

and similarly the maximum amplitudes in increasing order of magnitude (in cm.) were

0, 0.1, 0.11, 0.12, 0.13, 0.14, 0.15, 0.16, 0.17, 0.18, 0.19, 0.21, 0.22, 0.23, 0.24, 0.25, 0.26, 0.27, 0.28, 0.29, 0.3, 0.31, 0.33, 0.35, 0.41,
0.43, 0.44, 0.46, 0.47, 0.49, 0.5, 0.52, 0.55, 0.57, 0.6, 0.9, 1.05, 1.3, 1.35, 1.45.

We then superimposed the intervals making them equally fuzzy with constant level of partial presence equal to 1/39 in every case.

At this point, we would like to define what is known as an empirical distribution function in the statistical literature (Gibbons and Chakraborti, 1992, page-25). An empirical distribution function may be considered as an estimate of the cumulative distribution function defining the randomness concerned. For a sample of size n, this function $S_n(x)$, is defined as the proportion of values that do not exceed x. Accordingly, if $X_{(1)}$, $X_{(2)}$, ..., $X_{(n)}$ denote the order statistics of a random sample, its empirical distribution function would be given by

$$S_n(x) = 0, \text{ if } x < X_{(1)},$$
$$= k/n, \text{ if } X_{(k)} \leq x < X_{(k+1)},$$
$$k = 1, 2, \ldots, (n-1),$$
$$= 1, \text{ if } x \geq X_{(n)}.$$  

$x$ here being random, so would be $S_n(x)$. Writing

$$A_i(t) = 0, \text{ if } X_i > t,$$
$$= 1, \text{ otherwise},$$

we see that

$$S_n(x) = \sum A_i(x) / n.$$  

Therefore $nS_n(x)$ will have the law followed by the sum of n independent Bernoulli random variables $A_i(x)$. Indeed in such a case, we would have

$$\text{Prob} \left[ S_n(x) = k/n \right] = C_k \left[ F_X(x) \right]^k \left[ 1 - F_X(x) \right]^{n-k}$$

for $k = 0, 1, \ldots, n$. Hence the mathematical expectation of $S_n(x)$ would be given by

$$E \left[ S_n(x) \right] = F_X(x).$$  

Therefore, $S_n(x)$ converges uniformly to $F_X(x)$ almost surely. This leads to the Glivenko – Cantelli theorem that states that the limiting value of the supremum of the difference between $S_n(x)$ and $F_X(x)$, as $n$ becomes infinitely large, converges to zero almost surely.

After plotting the values (Fig. 2), we have seen that the minimum amplitudes of the waveform is an empirical distribution function of a random variable in the interval [-1.15, 0], while the maximum amplitudes of the waveform is a complementary empirical distribution function of another random variable in the interval [0, 1.45]. Now, according to Baruah’s randomness-fuzziness consistency principle, the theoretical distribution function of the minimum amplitudes in [-1.15, 0] and the theoretical complementary distribution function of the maximum amplitudes in [0, 1.45] define a normal fuzzy number [-1.15, 0, 1.45], which is clear from the diagram given below:

![Fig.2 Membership values of one half of the waveform of the earthquake](http://japmnt.com/)

Thus, the fuzzy membership function in this example can be approximated as

$$\mu_x(x) = \begin{cases} F(x) & \text{if } -1.15 \leq x \leq 0 \\ G(x) & \text{if } 0 \leq x \leq 1.45 \\ 0 & \text{if } x > 1.45 \end{cases}$$

Here $F(x)$ and $G(x)$ are the Dubois–Prade left and right reference functions. $F(x)$ is a non-decreasing continuous function from 0 to 1, and $G(x)$ is non-increasing from 1 to 0 for $-1.15 \leq x \leq 0$ and $0 \leq x \leq 1.45$ respectively.

From the diagram it is clear that the membership curve for the right reference function decreases nearly exponentially with the increase in the amplitudes of the waveform and that the left reference function increases nearly exponentially with the increase in the amplitudes of the waveform.
3. Fitting of the reference functions

Taking the different values of amplitudes as an independent variable X and the membership values as the dependent variable Y, we can fit the reference functions. As for the right reference function, let

\[ Y = a \cdot e^{bX} \]  \[ b < 0, \ X > 0, \]  

such that the maximum of Y will be unity for \( X = 0 \).

Using the method of least squares the estimates of a and b were computed. We have found that the estimated values of Y is 1.098383 for the dependent variable \( X=0 \). We therefore had to do a little bit of scaling in the sense that the estimated value of the parameter a had to be divided by 1.098383. The equation thus found was

\[ Y = 1.84118745 \cdot e^{-0.59969X} \]

The curve concerned has been depicted in Fig. 3 below.

![Figure 3: The right reference function](image)

The equation of the left reference function would similarly be

\[ Y = a \cdot e^{bX} \]  \[ b > 0, \ X < 0, \]  

on condition that the maximum of Y would be unity when \( X = 0 \). Once again, using the method of least squares we computed the estimates of a and b, and did a bit of scaling. The equation was found to be

\[ Y = 1.51248645 \cdot e^{0.70178X} \]

The curve has been shown in Fig. 4 below.

![Figure 4: The left reference function](image)

After combining the estimated left and right reference functions, we can construct the estimated membership curve (fig. 5), where the right reference function decreases nearly exponentially and the left reference function increases nearly exponentially.

![Figure 5: Estimated membership curve for the waveform of the earthquake](image)

4. Conclusions

Two laws of randomness can define a normal law of fuzziness. From one half of the waveform of an earthquake, starting from the maximum amplitude to zero amplitude, we have been able to show that the waveform is indeed a normal fuzzy number. This is how a normal fuzzy number has to be constructed.
References


