A CONNECTION BETWEEN FUZZY NUMBERS AND RANDOM INTERVALS

by

Maria Angeles Gil

Memorandum No. UCB/ERL M90/74

22 August 1990

A CONNECTION BETWEEN FUZZY NUMBERS AND RANDOM INTERVALS

by

Maria Angeles Gil

Memorandum No. UCB/ERL M90/74

22 August 1990

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

TITLE PAGE

A CONNECTION BETWEEN FUZZY NUMBERS AND RANDOM INTERVALS

by

Maria Angeles Gil

Memorandum No. UCB/ERL M90/74

22 August 1990

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

ARTICLE TITLE : A CONNECTION BETWEEN FUZZY NUMBERS AND RANDOM INTERVALS *

AUTHOR'S NAME : María Angeles GIL **.

AFFILIATION : Department of Electrical Engineering and Computer Sciences
University of California
Berkeley, CA 94720

* This research is supported in part by NASA Grant NCC 2-275, California MICRO Grant 88-094, LLNL Grant 442433-26705, NSF PYI Grant DMC-84511622 and a Grant from the spanish DGICYT. Their financial support is gratefully acknowledged.

** Visiting from the Department of Mathematics, University of Oviedo, 33071 Oviedo, SPAIN

A CONNECTION BETWEEN FUZZY NUMBERS AND RANDOM INTERVALS

MARIA ANGELES GIL

Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, CA 94720

ABSTRACT

Some connections between Fuzzy Set Theory and Probability Theory have been developed in the literature (cf. Goodman, Nguyen, Wang & Sanchez, Wang, and others). In this way, Goodman proved that any random set determines a naturally corresponding fuzzy set, although the converse of this result is not immediate.

In this paper we will analyze the converse result when dealing with fuzzy numbers. Thus, we will verify that for most of fuzzy numbers we can easily and intuitively associate random intervals such that the the membership degree of any element to a given fuzzy number coincides numerically with the probability of that element belonging to the associated random interval. We will finally discuss some of the advantages and inconveniences of that special connection.

1. INTRODUCTION

The problem of connecting exhaustively Fuzzy Sets and Probability Theories has been studied for many authors. In particular, Goodman (1976, 1982), Nguyen (1979), Wang (1982), Wang and Sanchez (1982), and others, have tried to establish such a connection in terms of the concepts of fuzzy sets and random sets, where if $\mathbb{X}$ is a given (referential) space, then

DEFINITION 1.1. A *fuzzy set* $\mathcal{A}$ of $\mathbb{X}$ is characterized by a *membership function*

$$\mu_{\mathcal{A}}: \mathbb{X} \to [0,1]$$

where $\mu_{\mathcal{A}}(x)$ is the degree to which x belongs to $\mathcal{A}$ (or degree to which x agrees or is compatible with $\mathcal{A}$).

If $(\Omega,\mathcal{C},P)$ is a probability space, and $\mathbb{X}$ is a subset of the euclidean space $\mathbb{R}$, then

DEFINITION 1.2. A *random set* $\mathbb{S}$ of $\mathbb{X}$ associated with $(\Omega,\mathcal{C},P)$ is a function

$$\mathbb{S}: \Omega \to \mathcal{F}(\mathbb{X})$$

where the image of $\mathbb{S}$ is contained in $\mathcal{K}(\mathbb{X})$ = collection of all nonempty compact subsets of $\mathbb{X}$, and satisfying the following measurability condition: $G_{\mathbb{S}}$ = graph of $\mathbb{S} = \{(\omega,x) \in \Omega \times \mathbb{X} \mid x \in \mathbb{S}(\omega)\} \in \mathcal{C} \times \mathcal{B}_{\mathbb{X}}$ (where $\mathcal{B}_{\mathbb{X}}$ is the smallest Borel σ-field on $\mathbb{X}$).

In some of his studies, Goodman proved that for any random set there is an immediate fuzzy set satisfying that

$$P(x \in \mathbb{S}_{\mathcal{A}}) = \mu_{\mathcal{A}}(x), \text{ for all } x \in \mathbb{X} \qquad (1)$$

Goodman also formulated the converse result: Given a fuzzy set $\mathcal{A}$ of $\mathbb{X}$, does there exist a random set $\mathbb{S}_{\mathcal{A}}$ such that (1) is true?. Goodman answered his question in affirmative. For each of the fuzzy sets he constructed the class of all random sets associated with it through (1). Nevertheless, Goodman's answer is general and, due to this generality, it is not immediate and easy to interpret and discuss the relationships between each fuzzy set and its associated random sets. For this reason, the aim of this paper is to analyze a slightly less general result, but more intuitive, immediate and easier to interpret and discuss.

The basic idea for this analysis can be illustrated by means of the example below:

Example 1.1. Let $\mathcal{A}$ be the class of *"tall"* people. This class if often regarded as a fuzzy set of the space $\mathbb{X} = \mathbb{R}$ characterized by a membership function $\mu_{\mathcal{A}}$, say that in Figure 1

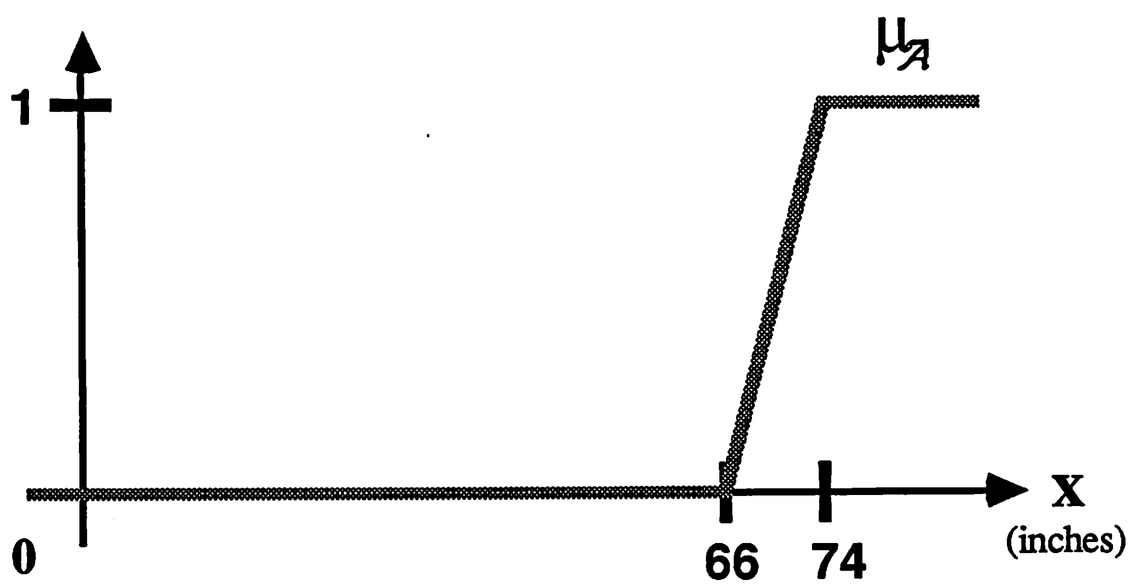

Fig. 1. Membership function of the fuzzy set $\mathcal{A}$ = class of *"tall"* people.

However, this class could alternatively be identified with the random set $\mathbb{S}_{\mathcal{A}}$ of $\mathbb{X}$, $\mathbb{S}_{\mathcal{A}} = [\mathbf{Z}, +\infty)$, where $\mathbf{Z}$ is a random variable uniformly distributed on $[66,74]$. That is, the event *"being tall"* could be described by means of an interval, $[\mathbf{Z}, +\infty)$ (what means that a person is considered as *"being tall"* whenever his height is higher or equal to $\mathbf{Z}$), but where $\mathbf{Z}$ is not a fixed value, but a random one, and having a uniform distribution on $[66,74]$. In this case, the

$$P(\mathbb{S}_{\mathcal{A}} \text{ contains } x) = P(x \in \mathbb{S}_{\mathcal{A}}) = P(\mathbf{Z} \le x) = \mu_{\mathcal{A}}(x), \text{ for all } x \in \mathbb{X}.$$

2. PARTICULARIZING GOODMAN'S QUESTION

The purpose of this paper is to formalize this last idea by particularizing Goodman's question to the case in which we deal with fuzzy numbers and random intervals. Thus, according to a general definition, we have

DEFINITION 2.1. A *fuzzy number* $\mathcal{A}$ is a fuzzy set of $\mathbb{X} = \mathbb{R}$ characterized by a *membership function* $\mu_{\mathcal{A}}: \mathbb{R} \to [0,1]$ such that

i) $\mu_{\mathcal{A}}$ is piecewise continuous.

ii) There exist $a_{\mathcal{A}}, b_{\mathcal{A}}, c_{\mathcal{A}}, d_{\mathcal{A}} \in \mathbb{R}$ with $a_{\mathcal{A}} \leq b_{\mathcal{A}} \leq c_{\mathcal{A}} \leq d_{\mathcal{A}}$, and

* $\mu_{\mathcal{A}}(x) = 0$ for all $x \in (-\infty, a_{\mathcal{A}}] \cup [d_{\mathcal{A}}, +\infty)$,
* $\mu_{\mathcal{A}}$ is non-decreasing on $[a_{\mathcal{A}}, b_{\mathcal{A}}]$ and non-increasing on $[c_{\mathcal{A}}, d_{\mathcal{A}}]$
* $\mu_{\mathcal{A}}(x) = 1$ for all $x \in [b_{\mathcal{A}}, c_{\mathcal{A}}]$.

As we have previously commented, the value $\mu_{\mathcal{A}}(x)$ is interpreted in this situation as the degree of compatibility of x with $\mathcal{A}$. Thus, if $\mathcal{A}$ is the class of *"tall"* people and $x = 68$ inches, then Figure 1 indicates that $\mu_{\mathcal{A}}(x) = 0.25$, what means that 0.25 is the degree with which a height equal to 68 inches agrees with the property of being *"tall"*.

On the other hand,

DEFINITION 2.2. A *random interval* $\mathfrak{S}$ of $\mathbb{X} = \mathbb{R}$ associated with the probability space $(\Omega, \mathcal{C}, P)$ is a random set of $\mathbb{R}$ associated with that probability space such that it may be described by means of two random variables $\mathbf{Y}$ and $\mathbf{Z}$ (associated with the same space) so that $\mathfrak{S}(\omega) = [\mathbf{Y}(\omega), \mathbf{Z}(\omega)]$, for all $\omega \in \Omega$, and it will be denoted by $\mathfrak{S} = [\mathbf{Y}, \mathbf{Z}]$.

In this situation the value $P(x \in \mathfrak{S})$ admits two possible interpretations: the objective one and the subjective one. According to the first one, $P(x \in \mathfrak{S})$ is the relative frequency of occurrence of the event « $x \in \mathfrak{S}$ » to be expected in a long-run. Thus, if in Example 1.1 $\mathfrak{S}$ = class of *"tall"* people and $x = 68$ inches, the fact that $P(x \in \mathfrak{S}) = 0.25$ means that if N people were requested to indicate if a person having 68 inches of height is *"tall"*, then 0.25N of the inquired people would be expected to answer YES. Nevertheless, according to the subjective interpretation, $P(x \in \mathfrak{S})$ is the degree of belief that « $x \in \mathfrak{S}$ ». Thus, if in Example 1.1 $\mathfrak{S}$ = class of *"tall"* people and $x = 68$ inches, the fact that $P(x \in \mathfrak{S}) = 0.25$

means that 0.25 is the quantified individual's judgement about the fact that a person having 68 inches of height will be classified as *"tall"* (or degree of belief that the next person requested will answer YES to « $x \in \mathfrak{S}$ », that is, will classify as *"tall"* a person having 68 inches of height).

The question we are now interested in is the following one: Given a fuzzy number $\mathcal{A}$, does there exist a random interval $\mathfrak{S}_{\mathcal{A}}$ such that $P(x \in \mathfrak{S}_{\mathcal{A}}) = \mu_{\mathcal{A}}(x)$, for all $x \in \mathbb{R}$?.

To guarantee that the numerical coincidence (2) is true we will introduce a general condition for the fuzzy numbers we will consider in this paper. Later, under such a condition, we will prove that for any fuzzy number there exists a random interval satisfying (2).

3. JUSTIFYING THE RELATIONSHIP BETWEEN FUZZY NUMBERS AND RANDOM INTERVALS

The result we are going to prove entails that the class of all random sets associated with a given fuzzy number, always contains at least a random interval. In addition, the justification for this result is developed in an intuitive way providing us with an easy interpretation.

THEOREM 3.1. Given a fuzzy number $\mathcal{A}$, such that $\mathrm{supp}(\mathcal{A}) = \{x \in \mathbb{R} \mid \mu_{\mathcal{A}}(x) > 0\} = [a_{\mathcal{A}}, d_{\mathcal{A}}]$, and satisfying that

* $\mu_{\mathcal{A}}$ is continuous from the right at every $x \in [a_{\mathcal{A}}, b_{\mathcal{A}}]$,
* $\mu_{\mathcal{A}}$ is continuous from the left at every $x \in [c_{\mathcal{A}}, d_{\mathcal{A}}]$.

there exists a random interval $\mathfrak{S}_{\mathcal{A}}$ such that

$$P(x \in \mathfrak{S}_{\mathcal{A}}) = \mu_{\mathcal{A}}(x), \text{ for all } x \in \mathbb{R} \qquad (2).$$

Proof. To justify intuitively the result above we are going to distinguish two possible situations: when $\mu_{\mathcal{A}}$ is symmetric with respect to the middle point of $\mathrm{supp}(\mathcal{A})$, and when $\mu_{\mathcal{A}}$ is non-symmetric.

i) When $\mu_{\mathcal{A}}$ is symmetric (see Figure 2) with respect to the middle point of $\mathrm{supp}(\mathcal{A})$, $[b_{\mathcal{A}}+c_{\mathcal{A}}]/2$, then the random interval

$$\mathfrak{S}_{\mathcal{A}} = [\, [b_{\mathcal{A}}+c_{\mathcal{A}}]/2 - \mathbf{Z} \,,\, [b_{\mathcal{A}}+c_{\mathcal{A}}]/2 + \mathbf{Z} \,]$$

where $\mathbf{Z}$ is a random variable whose distribution function $\mathbf{F_Z}$ = membership function of the fuzzy set $(\leq \mathcal{A})^c - [b_{\mathcal{A}}+c_{\mathcal{A}}]/2$ (with $(\leq \mathcal{A})^c$ = complement of the fuzzy number "lower than or equal to $\mathcal{A}$") (see Figure 3), is a solution of (2).

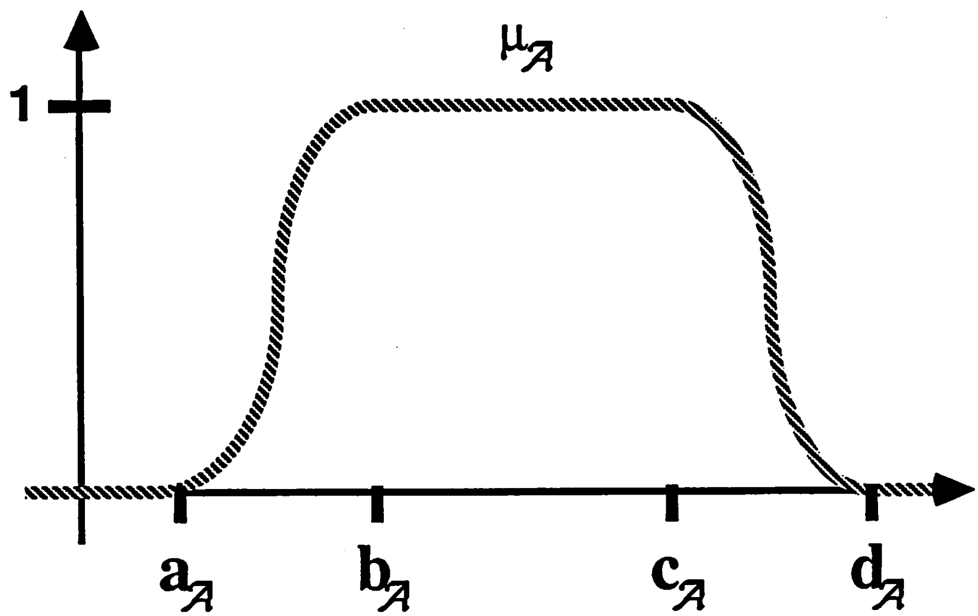

Fig. 2. Membership function of the symmetric fuzzy number $\mathcal{A}$.

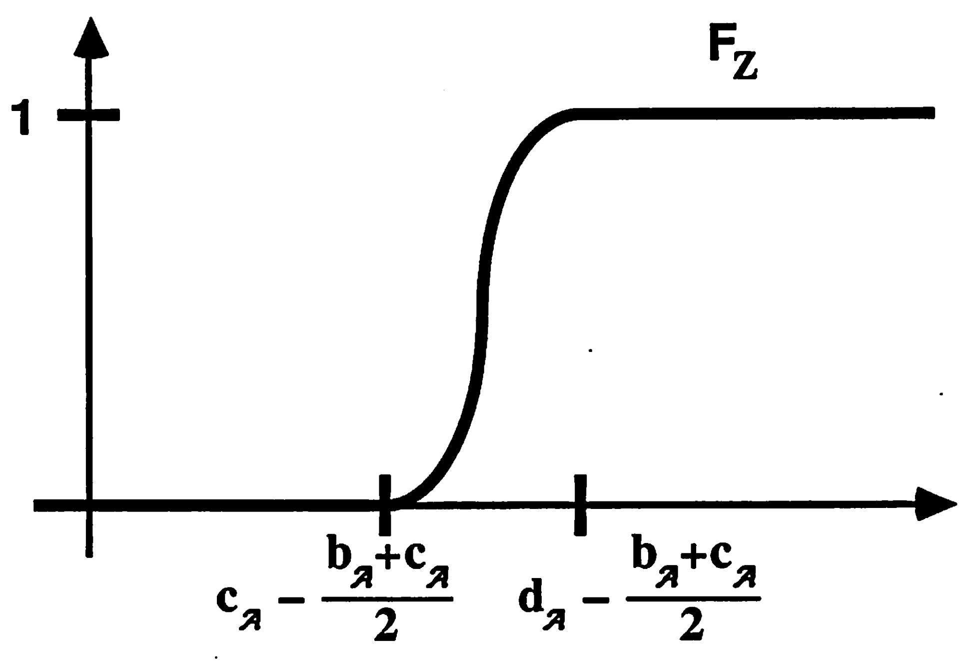

Fig. 3. Distribution function of the random variable Z.

ii) When $\mu_{\mathcal{A}}$ is non-symmetric (see Figure 4) with respect to the middle point of supp($\mathcal{A}$), $[b_{\mathcal{A}}+c_{\mathcal{A}}]/2$, then the random interval

$$\Im_{\mathcal{A}} = [\, Y , Z \,]$$

where $\mathbf{Y}$ is a random variable whose distribution function F_Y = membership function of the fuzzy set $(\geq \mathcal{A})$ (see Figure 5) (with $(\geq \mathcal{A})$ = fuzzy number "higher than or equal to $\mathcal{A}$") , $\mathbf{Z}$ is a random variable whose distribution function F_Z = membership function of the fuzzy set $(\leq \mathcal{A})^c$ (see Figure 6), and $\mathbf{Y}$ and $\mathbf{Z}$ are independent, is a solution of (2).

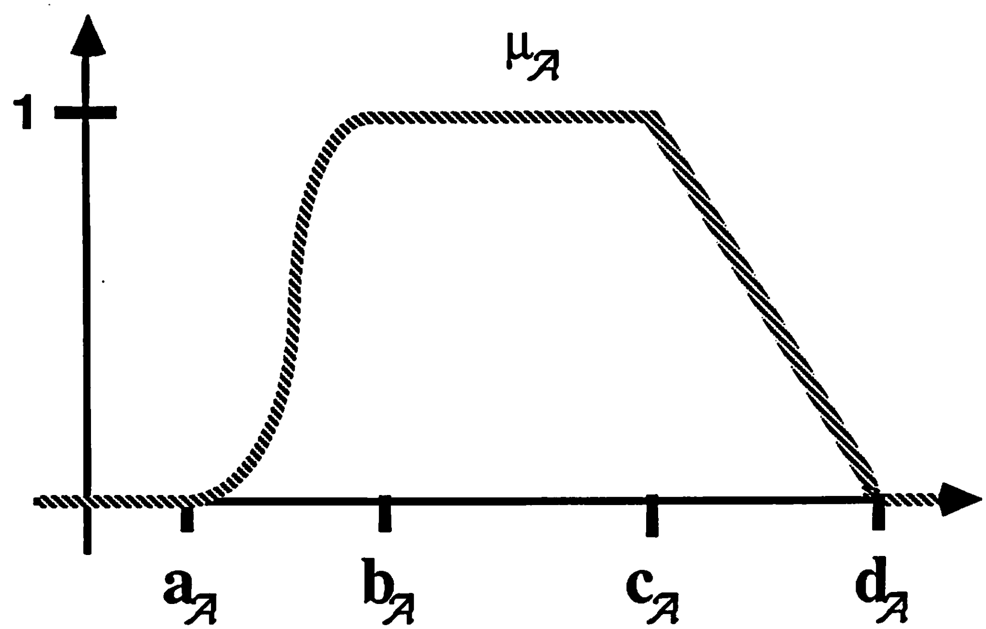

Fig. 4. Membership function of the non-symmetric fuzzy number $\mathcal{A}$.

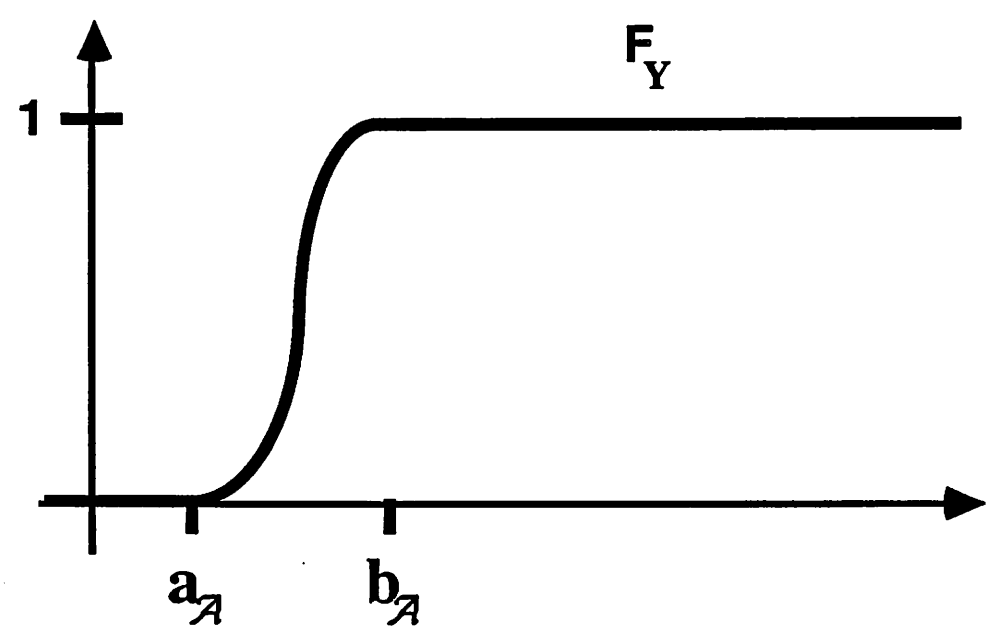

Fig. 5. Distribution function of the random variable $\mathbf{Y}$.

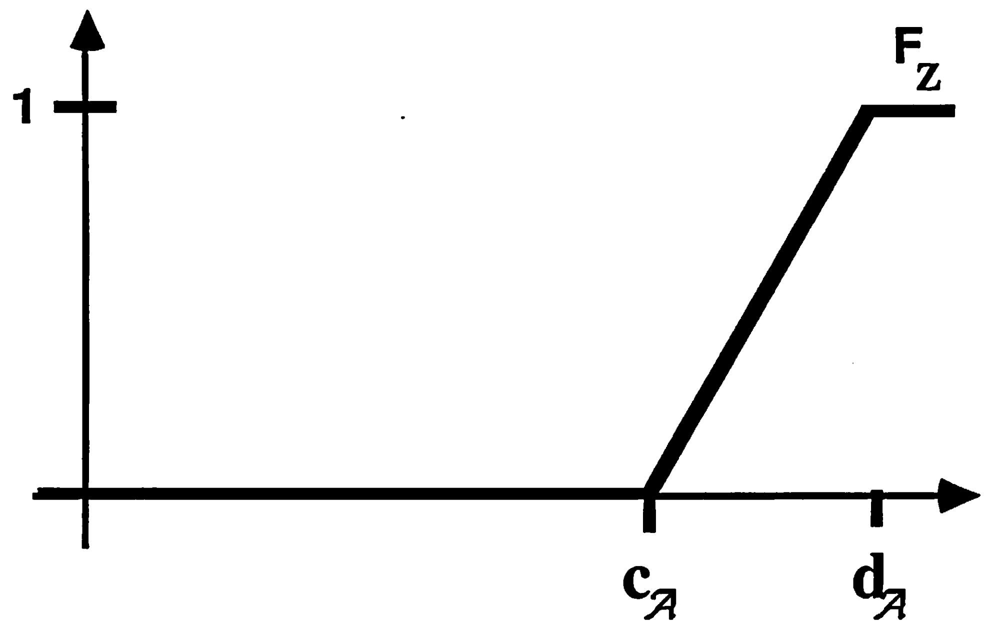

Fig. 6. Distribution function of the random variable $\mathbf{Z}$.

■

REMARK 3.1. The solution given for the non-symmetric case, could also be applied for the symmetric case, but the first one described more intuitively the solution.

4. SOME ILLUSTRATIVE EXAMPLES

The situation in Example 1.1 illustrates the result above established. We now consider more examples illustrating the solutions stated in the proof of Theorem 3.1, and for which the interpretation of the associated random intervals is easy and immediate.

Example 3.1. Let $\mathcal{A}$ be the class of *"very young"* people. If this class is regarded as a fuzzy number characterized by the membership function $\mu_{\mathcal{A}}$ in Figure 7, then the associated random interval according to the proof of Theorem 3.1 is $\mathfrak{I}_{\mathcal{A}} = [\mathbf{Y}, \mathbf{Z}]$ where $\mathbf{Y}$ is a random variable degenerated in 0, and $\mathbf{Z}$ is a random variable having a distribution characterized by the triangular density function given by

$$f(z) = \begin{cases} 0.16(z-10) & \text{if } z \in (10, 12.5] \\ 0.16(15-z) & \text{if } z \in (12.5, 15) \\ 0 & \text{otherwise} \end{cases}$$

(what means that a person is considered as *"being very young"* whenever his age is lower or equal to the random value $\mathbf{Z}$).

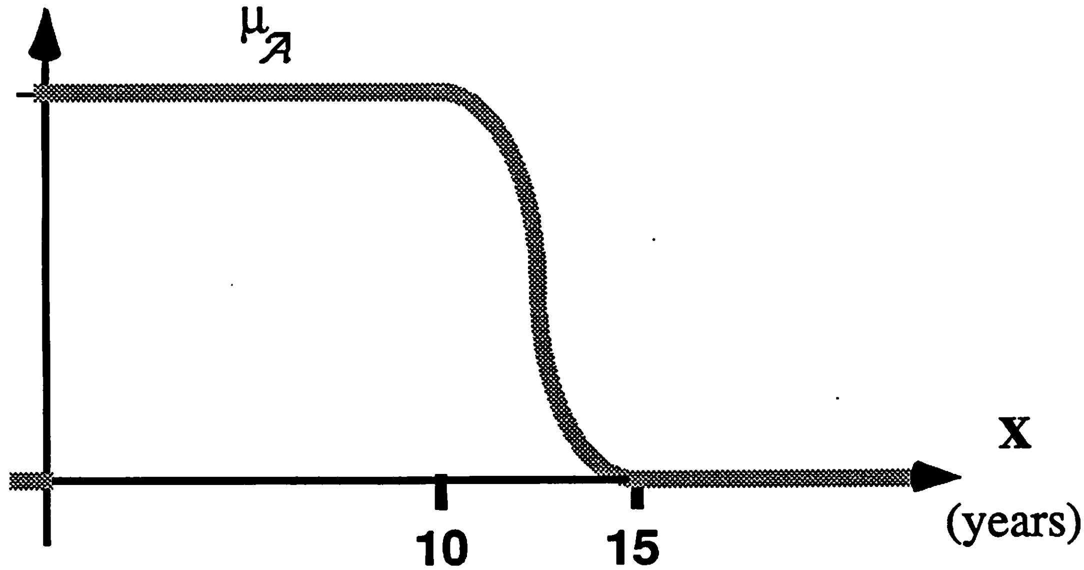

Fig. 7. Membership function of the fuzzy number $\mathcal{A}$ = class of *"very young"* people.

Example 3.2. Let $\mathcal{A}$ be the proposition *"approximately 250 inches"*. If this proposition is regarded as a fuzzy number characterized by the membership function $\mu_{\mathcal{A}}$ in Figure 8, then the associated random interval according to the proof of Theorem 3.1 is $\mathfrak{S}_{\mathcal{A}} = [\ 250 - \mathbf{Z}\ ,\ 250 + \mathbf{Z}]$ where $\mathbf{Z}$ is a random variable having a uniform distribution in $[0,25]$ (what means that a real number is considered as *"being approximately 250 inches"* whenever it is in a neighborhood centered at 250, and whose radium $\mathbf{Z}$ is random).

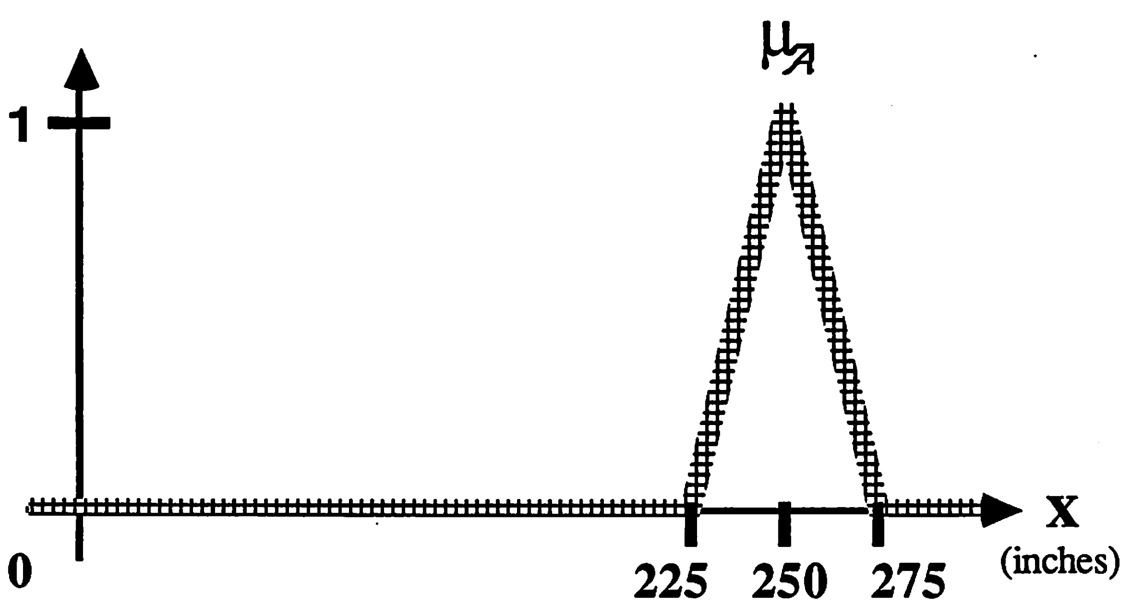

Fig. 8. Membership function of the fuzzy number $\mathcal{A}$ = *"approximately 250 inches"*.

Example 3.3. Let $\mathcal{A}$ be the class of *"moderately expensive"* restaurants (assumed defined on the basis of the average per-person dinner cost). If this class is regarded as a fuzzy number characterized by the membership function $\mu_{\mathcal{A}}$ in Figure 9, then the associated random interval according to the proof of Theorem 3.1 is $\mathfrak{I}_{\mathcal{A}} = [\mathbf{Y},\mathbf{Z}]$ where $\mathbf{Y}$ is a random variable uniformly distributed in $[5,10]$, $\mathbf{Z}$ is a random variable uniformly distributed in $[20,30]$, and $\mathbf{Y}$ and $\mathbf{Z}$ are independent (what means that a restaurant is considered as *"being moderately expensive"* whenever its average per-person dinner cost is $\mathbf{Y}$ to $\mathbf{Z}$, $\mathbf{Y}$ and $\mathbf{Z}$ being random values).

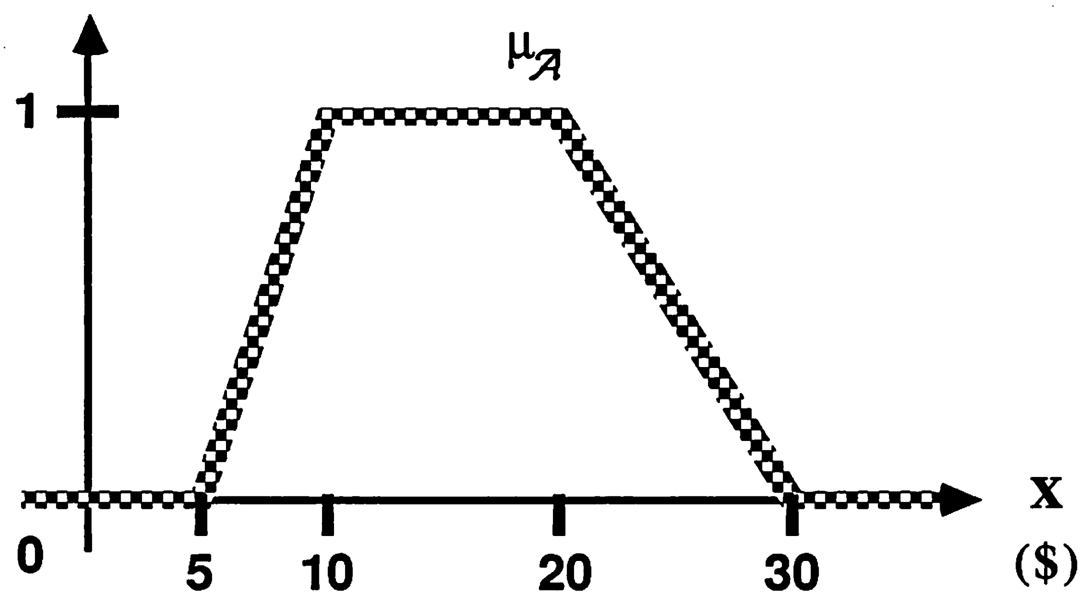

Fig. 9. Membership function of the fuzzy number $\mathcal{A}$ = *"moderately expensive"*.

5. ADVANTAGES AND INCONVENIENCES IN TREATING FUZZY NUMBERS AS RANDOM INTERVALS

Although we have formally established a relationship between fuzzy numbers and random intervals, it is worth briefly analyzing the advantages and inconveniences that could arise in practice if we replace fuzzy numbers by the associated random intervals.

The main advantage we can found in this replacement is the *objectivity* of random intervals in comparison with fuzzy numbers, when using objective probabilities to compute $P(x \in \mathfrak{I}_{\mathcal{A}})$.

On the other hand, it is possible to find several inconveniences in this replacement. Firstly, the employment of objective probabilities to compute $P(x \in \mathfrak{I}_{\mathcal{A}})$ (when available) usually entails a *cost* (for inquiries, observations, etc.). However, the main inconvenience is that related to *operations*. Thus, operations between fuzzy numbers, when based on Zadeh's extension principle, admit a general expression that is often very operative. On the contrary, operations between random intervals are frequently much more complex than those between fuzzy numbers, they additionally require to explicit a dependence relation between the random variables representing the upper and lower extremes of the intervals, and there is not a general formula expressing the distribution of the random variables obtained through the operations in terms of the distributions of the random variables involved in those operations. This last comment is now illustrated through a very simple example:

Example 5.1. Let $\mathcal{A}$ and $\mathcal{B}$ be the classes of *"very cheap"* and *"cheap"* restaurants (assumed defined on the basis of the average per-person dinner cost). If these classes are regarded as the fuzzy numbers characterized by the membership functions $\mu_{\mathcal{A}}$ and $\mu_{\mathcal{B}}$ in Figure 10,

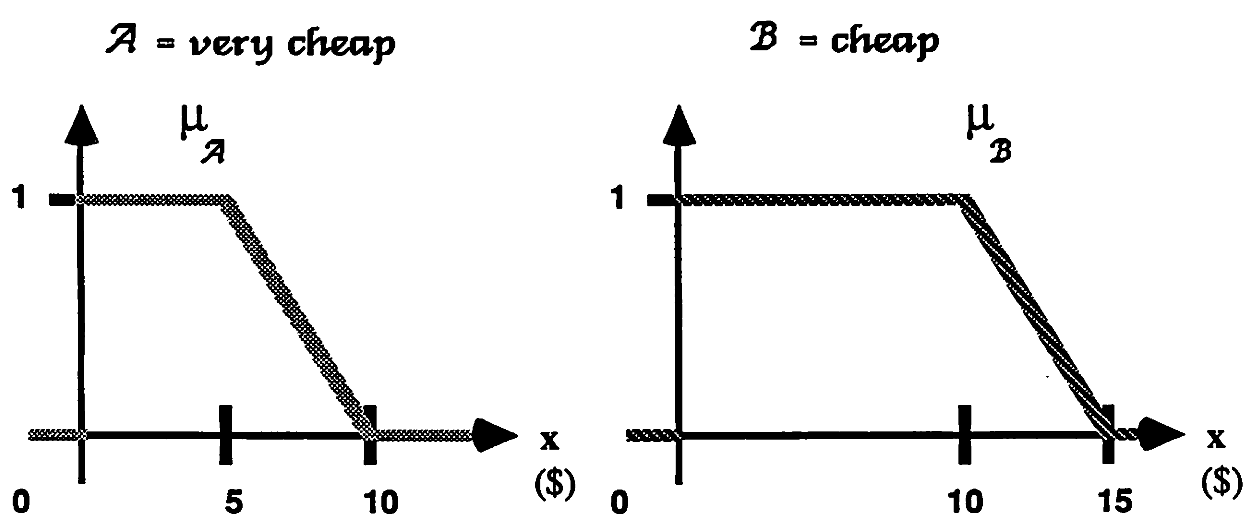

Fig. 10. Membership function of the fuzzy numbers $\mathcal{A}$ = *"very cheap"* and $\mathcal{B}$ = *"cheap"*.

then, the addition of $\mathcal{A}$ and $\mathcal{B}$, based on Zadeh's extension principle, that is,

$$\mu_{\mathcal{A}+\mathcal{B}}(x) = \max_{\substack{(u,v) \\ u+v=x}} \min \{ \mu_{\mathcal{A}}(u), \mu_{\mathcal{B}}(v) \}$$

leads to the fuzzy number characterized by the membership function $\mu_{\mathcal{A}+\mathcal{B}}$ in Figure 11.

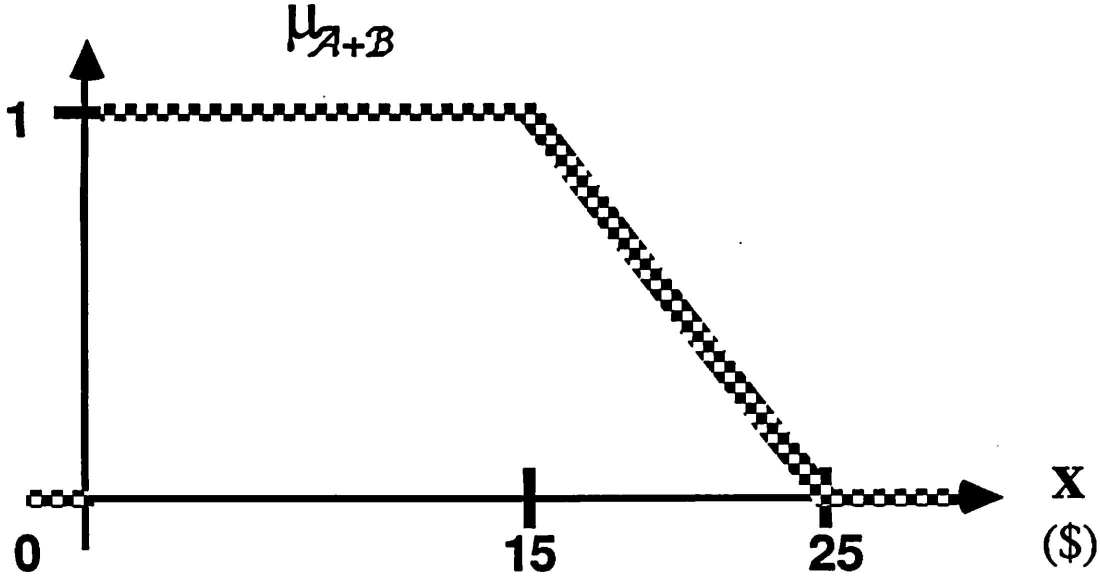

Fig. 11. Membership function of the fuzzy number $\mathcal{A}$ + $\mathcal{B}$.

The associated random intervals according to the proof of Theorem 3.1 are $\mathfrak{I}_{\mathcal{A}} = [Y_1,Z_1]$ and $\mathfrak{I}_{\mathcal{B}} = [Y_2,Z_2]$ where Y_1 and Y_2 are both random variables degenerated in 0, and Z_1 is a random variable uniformly distributed in $[5,10]$, Z_2 is a random variable uniformly distributed in $[10,15]$. If we now assume that Z_1 and Z_2 are independent random variables, then the addition of $\mathcal{A}$ and $\mathcal{B}$, based on Minkowski's addition of ordinary sets, that is,

$$\mathfrak{I}_{\mathcal{A}} + \mathfrak{I}_{\mathcal{B}} = [Y_1+Y_2, Z_1+Z_2]$$

leads to the random interval whose lower extremum is a random variable Y_1+Y_2 degenerated in 0, and whose upper extremum is a random variable Z_1+Z_2 whose distribution is characterized by the triangular density function given by

$$f(z) = \begin{cases} 0.04(z-15) & \text{if } z \in (15,20] \\ 0.04(25-z) & \text{if } z \in (20,25) \\ 0 & \text{otherwise} \end{cases}$$

It should be emphasized that the interval above would be the one associated in the proof of Theorem 3.1 with the fuzzy number C whose membership function is that in Figure 12.

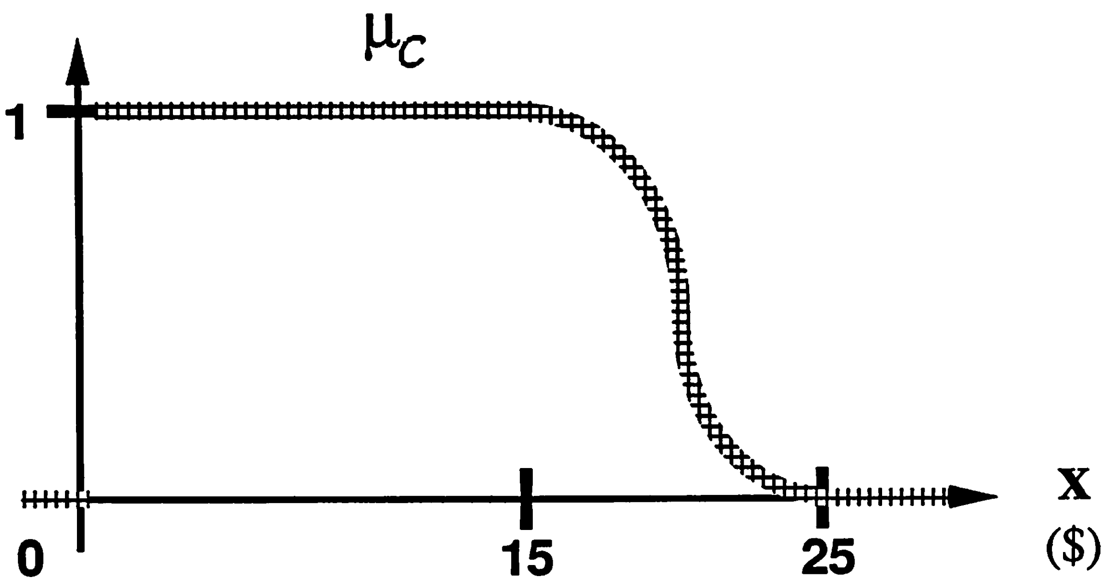

Fig. 12. Membership function of the fuzzy number C.

Consequently, the solutions obtained under the assumption of independence of Z_1 and Z_2 are definitively different, and the time and background required to make the first addition are smaller than those for the second one.

6. CONCLUDING REMARKS

In Section 5 we have discussed the interest of the connection between fuzzy numbers and random intervals, when the purpose is to replace the first ones by the second ones. On the other hand, it is worth remarking that such a replacement is not very coherent if we pay attention to the essential differences in the nature of imprecision in both cases. Thus, when fuzziness arises because of an uncertainty in observation (error, lack of information, etc.) nonprobabilistic in nature, or because of a subjective judgement (vagueness, cognition, personal bias, etc.), then the assimilation of fuzzy numbers with random intervals does not make real sense. To illustrate this last comment, we can consider for instance the situation in Example 1.1, in which according to the fuzzy approach *"tall"* is an imprecise event in the sense that for a given height x each person is assumed not to be able to answer YES or NOT to the occurrence of that event, but that person can only asserts that the event is true with a certain degree. Nevertheless, according to the probabilistic approach *"tall"* is an imprecise event in the sense that the criterion to classify people as being or not *"tall"* varies from person to person. However, it is supposed that for a given height x each person is assumed to be able to answer YES or NOT to the occurrence of that event, although the opinion is not the same for everybody.

ACKNOWLEDGEMENTS

The author likes to acknowledge Professor David A. Blackwell and Professor Enrique H. Ruspini for their invaluable discussions and suggestions. She also likes to thank Professors Lotfi A. Zadeh and Alice M. Agogino, and members of their research groups, for all their helpful comments, specially to Pratap Khedkar for some of the ideas in the last section.

REFERENCES

I. R. Goodman, "Some relations between fuzzy sets and random sets." (Unpublished, 1976).

I. R. Goodman, "Fuzzy Sets as equivalence classes of possibility random sets." In: *Fuzzy Set and Possibility Theory: Recent Developments.* (R.R. Yager, ed.), 1982, Pergamon: Oxford.

A. Kaufmann, *Introduction to the Theory of Fuzzy Subsets*, Vol. 1, 1975, Academic Press: New York.

G. Matheron, *Random Sets and Integral Geometry*, 1975, Wiley: New York.

C.V. Negoita and D.A. Ralescu, *Simulation, Knowledge-based Computing, and Fuzzy Statistics*, 1987, Van Nostrand Reinhold Co.: New York.

H. T. Nguyen, "Some mathematical tools for linguistic probabilities," *Fuzzy Sets and Systems*, 2, pp. 53-65, 1979.

H. T. Nguyen, "Possibility measures and related topics." In: *Advances in Fuzzy Set Theory and Applications*, (MM. Gupta and E. Sanchez, Eds.), 1982, North-Holland: Amsterdam.

Pei-zhuang Wang, "From the fuzzy statistics to the random subsets." In: *Fuzzy Information and decision Processes*, (P. P. Wang, ed.) 1982, Plenum: New York.

Pei-zhuang Wang and E. Sanchez, "Treating a fuzzy subset as a projectable random subset." In: *Advances in Fuzzy Set Theory and Applications*, (MM. Gupta and E. Sanchez, Eds.), 1982, North-Holland: Amsterdam.

W.L. Hays and R.L. Winkler, *Statistics, Vol. I, Probability Inference, and Decision*, 1970, Holt, Rinehart and Winston, Inc.: New York.

L.A. Zadeh, "Fuzzy Sets," *Inform. Contr.*, 8, 338-353, 1965.

L.A. Zadeh, "Fuzzy sets as a basis for a theory of possibility," *Fuzzy Sets and Systems*, 1, 3-28, 1978.