ON THE INTERPRETATION OF GRADES OF MEMBERSHIP IN FUZZY DATA FROM RANDOM EXPERIMENTS

by

Maria Angeles Gil

Memorandum No. UCB/ERL M90/75

22 August 1990

COVER PAGE

ON THE INTERPRETATION OF GRADES OF MEMBERSHIP IN FUZZY DATA FROM RANDOM EXPERIMENTS

by

Maria Angeles Gil

Memorandum No. UCB/ERL M90/75

22 August 1990

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

TITLE PAGE

ON THE INTERPRETATION OF GRADES OF MEMBERSHIP IN FUZZY DATA FROM RANDOM EXPERIMENTS

by

Maria Angeles Gil

Memorandum No. UCB/ERL M90/75

22 August 1990

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

ARTICLE TITLE : ON THE INTERPRETATION OF GRADES OF MEMBERSHIP IN FUZZY DATA FROM RANDOM EXPERIMENTS *.

AUTHOR'S NAME : María Angeles GIL **.

AFFILIATION : Departamento de Matemáticas
Edif. Facultades de Ciencias, Planta 2
C/ Calvo Sotelo, s/n
Universidad de Oviedo
33005 Oviedo
SPAIN

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

** María Angeles Gil is currently a Research Associate of the EECS Department (Computer Science Division) in the University of California at Berkeley.

ON THE INTERPRETATION OF GRADES OF MEMBERSHIP IN FUZZY DATA FROM RANDOM EXPERIMENTS

MARIA ANGELES GIL*

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

This paper first presents cases involving different sources of fuzzy imprecision from random experiments. Then, the interpretation of grades of membership in terms of probabilities in those cases is discussed.

1. INTRODUCTION

Statistical problems often concern the drawing of conclusions or making of decisions about a random experiment, on the basis of the information supplied by the performance of that experiment (and, eventually, on the basis of other available information). A *random experiment* is a process by which something is made, resulting in one outcome that cannot be previously predicted. In Statistics it is assumed that the random experiment can be repeated under (more or less) identical conditions, and there is a predictable long-run pattern (what is referred to as statistical regularity). To characterize a random experiment we need: i) to identify all experimental outcomes, ii) to identify all events of interest, and iii) to assign probabilities to these events.

In "traditional" Statistics it is supposed that the observer is able to obtain the exact outcome after each experimental performance, and an *event of interest* is intended as a statement or question regarding the experimental outcome and so that after the experiment has been conducted one can determine if it is true or false. Then, the model describing a random experiment is given by a probability space $(\Omega,\mathcal{F},P)$, where Ω is the sample space (or set of all possible experimental outcomes), $\mathcal{F}$ is the σ-field of all events of interest, and P is a probability measure on $\mathcal{F}$. Furthermore, we often assume that in the experiment we incorporate a numerical quantity $\mathbf{X}$ whose value is determined by the preceding experiment, that is, we consider a random variable (or vector) is to be observed, so that we can describe the random experiment by means of a new probability space $(X,\mathcal{B}_X,P)$, where X is the set of all variable (or vector) values and contained in a Euclidean space, $\mathcal{B}_X$ is the smallest Borel σ-field on X, and P is the induced probability.

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

Consequently, due to the assumption of the structure of σ-field for the set of all events of interest, Probability Theory guarantees that every element in $\mathcal{F}$ can be identified with a "classical" subset of Ω, and the induced probability of a Borel set $B \in \mathcal{B}_X$ is given by the Lebesgue-Stieltjes integral

$$P(B) = \int_B dP(x) = \int_X \chi_B(x)\, dP(x) \qquad (1)$$

(where χ_B = indicator or characteristic function of B).

The scheme in Figure 1 explains the mechanism in traditional random experiments involving a numerical quantification process.

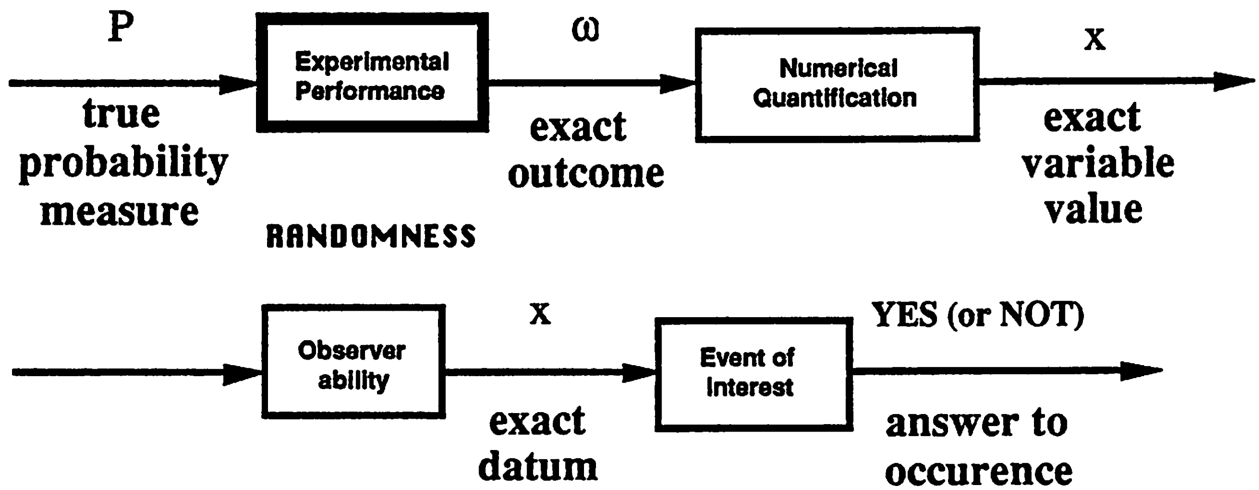

Fig. 1. Process in traditional random experiments

Nevertheless, one frequently encounters situations in which the quantification process or the ability of the observer do not allow to express the available experimental information in terms of an exact variable value or an exact datum (or, alternatively, an event of interest could be intended in some situations as a statement or question regarding the experimental outcome, so that after the experiment has been conducted one can determine the "degree to which it is true").

We next consider the usual ways according to which fuzziness arises in random experiments.

2. SOURCES OF FUZZY IMPRECISION IN RANDOM EXPERIMENTS

In this section we are going to discuss the most common arguments to incorporate fuzziness in random experiments.

We first recall the notions of fuzzy subset and measurable fuzzy subset ([14], [15]) we will employ later:

DEFINITION 2.1. A *fuzzy subset* χ of X (universe of discourse) is characterized by a membership function μ_χ from X to [0,1], where $\mu_\chi(x)$ is the degree to which x belongs to χ (or degree to which x agrees or is compatible with χ). When μ_χ is a Borel-measurable function, then χ is called a *measurable fuzzy subset* or *fuzzy event.*

As we have emphasized in Figure 1, a random experiment involves the following steps: experimental performance, numerical quantification, observation, and definition of the events of interest. When the random experiment is understood in the traditional sense, the first step is developed under randomness and the other ones are carried out under precise conditions leading to exact elements. However, in practice fuzziness could arise in each of the last three steps. We are now going to describe the situations corresponding to the presence of fuzziness in each of those steps, and some suitable mathematical models for such situations we can found in the literature of Fuzzy Sets.

2.1. Fuzzy random variable

In this first case, we are going to consider that fuzziness is incorporated through the second step (numerical quantification). More precisely, it is assumed that the outcome from the original experiment $(\Omega,\mathcal{F},P)$ is exact, but the random variable (or vector) may be described in terms of a quantification process assigning an imprecise (fuzzy) value to each outcome. The scheme in Figure 2 explains the mechanism in this case.

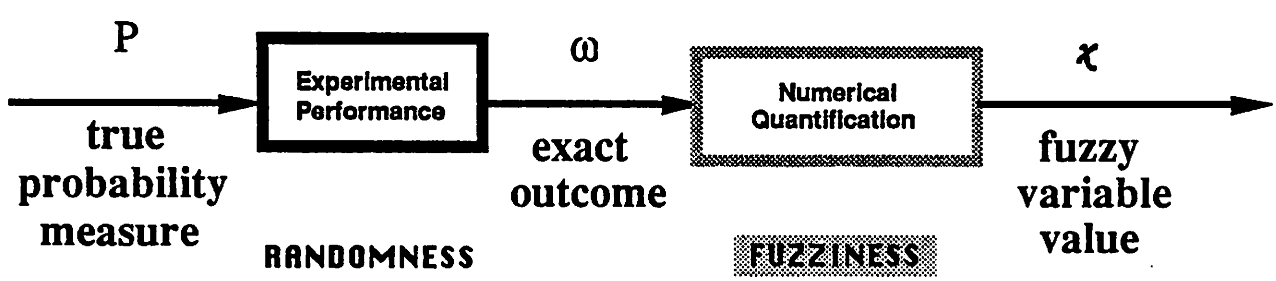

Fig. 2. Process leading to obtain fuzzy variable values

A mathematical model that could suitably describe this situation is that resumed in the concept of *fuzzy random variable* introduced by Puri and Ralescu (cf., [9], [10]). It should be pointed out that in this case the induced probability for the fuzzy random variable can be immediately defined.

The source of fuzzy imprecision we have just examined may be illustrated by means of the following example (cf., [4]).

Example 2.1. In Decision-Making literature we can often find examples involving decision problems in which utilities or losses are quantified in terms of exact values, but it would be more natural to quantify them by means of fuzzy values. The example we are now going to analyze has been taken from an introductory text of Statistics ([13]): suppose a neurologist has to classify his most serious patients as requiring exploratory brain surgery (action a_1) or not (action a_2). From past autopsies, it has been found that 60% of the examined people needed the operation, while 40% did not. The losses of right classifications are null. The losses of wrong classifications are obvious: an unnecessary operation means resources are wasted and the patient may be hurt. The other loss may be higher: if a patient requiring surgery does not get it, the time lost until clear symptoms appear may be crucial. In Wonnacott and Wonnacott ([13]), the preceding problem is regarded as a decision problem with state space $\Omega = \{\omega_1,\omega_2\}$ (ω_1 = the patient requires surgery, ω_2 = the patient does not require surgery), action space $A = \{a_1,a_2\}$, and loss function L on $\Omega \times A$ given by the random variables $\mathsf{L}_1\colon \Omega \to \mathbb{R}$, $\mathsf{L}_2\colon \Omega \to \mathbb{R}$, such that $\mathsf{L}_1(\omega_1) = \mathsf{L}(\omega_1,a_1) = \mathsf{L}_2(\omega_2) = \mathsf{L}(\omega_2,a_2) = 0$, $\mathsf{L}_1(\omega_2) = \mathsf{L}(\omega_2,a_1) = 5\ \mathsf{L}_2(\omega_1) = 5\ \mathsf{L}(\omega_1,a_2)$ (with $\mathsf{L}_2(\omega_1) > 0$).

However, the preceding assessment of losses seems to be extremely precise, due to the nature of the states and actions in the problem. Thus, the following assessment could express better the decision-maker (neurologist) "preferences": $\mathcal{L}(\omega_1,a_1) = \mathcal{L}(\omega_2,a_2) = 0$, $\mathcal{L}(\omega_2,a_1)$ = « inconvenient », $\mathcal{L}(\omega_1,a_2)$ = « dangerous ». This "loss function" $\mathcal{L}$ could be defined in terms of the fuzzy random variables $\mathcal{L}_1\colon \Omega \to F(\mathbb{R})$, $\mathcal{L}_2\colon \Omega \to F(\mathbb{R})$ (with $F(\mathbb{R})$ = collection of fuzzy sets of $\mathbb{R}$), such that $\mathcal{L}_1(\omega) = \mathcal{L}(\omega,a_1)$, $\mathcal{L}_2(\omega) = \mathcal{L}(\omega,a_2)$, for all $\omega \in \Omega$, and $\mathcal{L}_1(\omega_2)$ and $\mathcal{L}_2(\omega_1)$ are described by means of the (measurable) fuzzy subsets characterized by the membership functions in Figure 3.

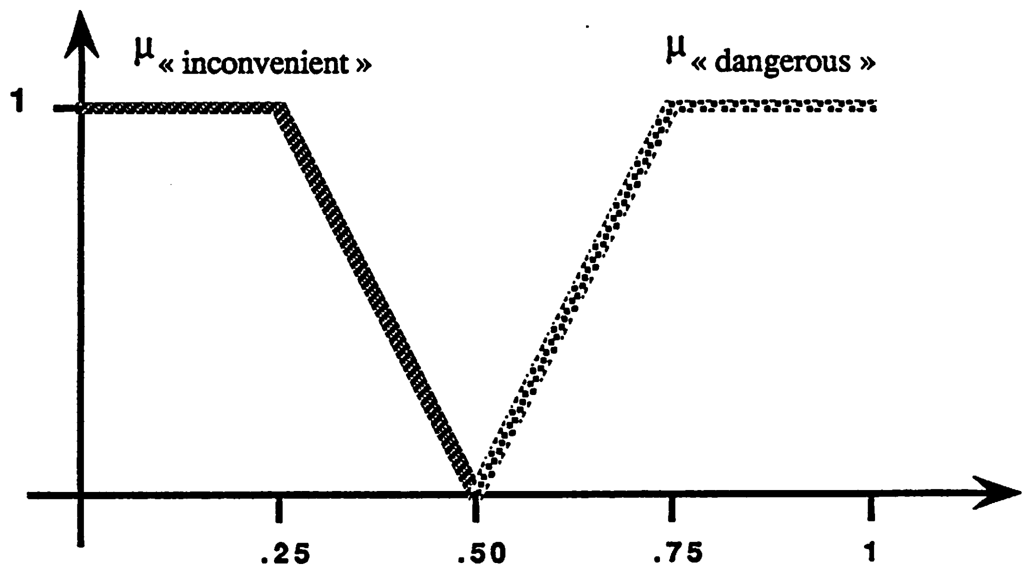

Fig. 3. Membership functions of the fuzzy losses « dangerous » and « inconvenient »

2.2. Fuzzy experimental information

In this second case, we are going to consider that fuzziness is incorporated through the third step (observation) in random experiments, that is, fuzziness arises because of the lack of precision in observing the variable (or vector) values. More precisely, it is assumed that the random variable (or vector) may be described in terms of a quantification process $\mathbf{X}$ assigning a numerical value to

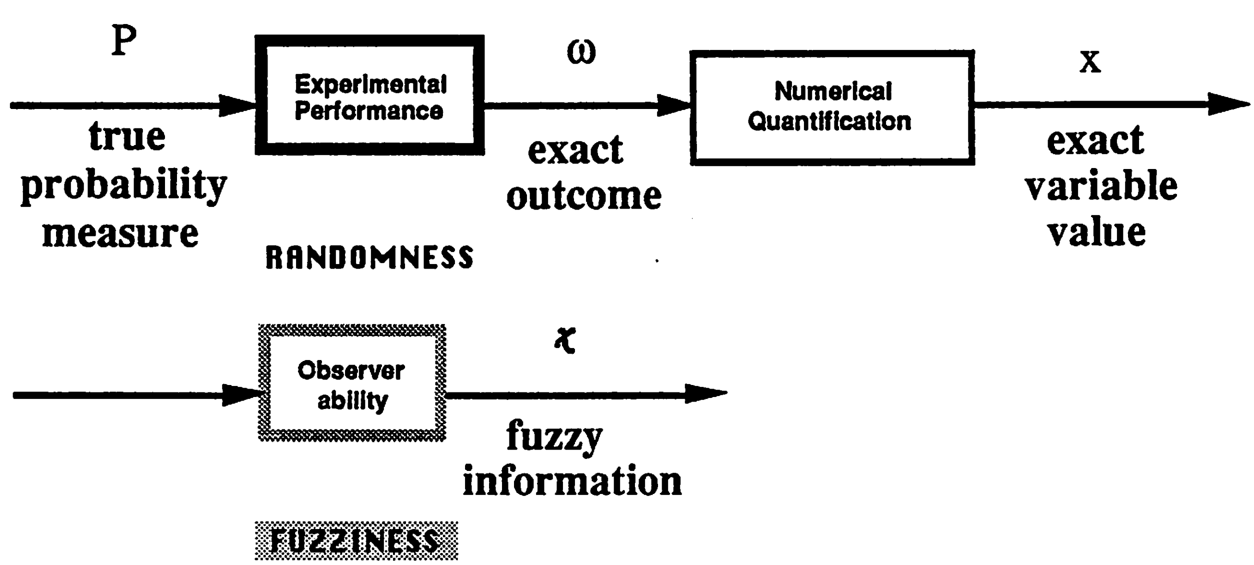

Fig. 4. Process leading to obtain fuzzy information

each experimental outcome, leading to $(X,\mathcal{B}_X,P)$, but this value cannot be exactly reported by the observer, so that he cannot answer YES or NOT to the occurrence of events assimilable with elements in $\mathcal{B}_X$, and a new type of information (fuzzy experimental information) regarding **X** has to be considered. The scheme in Figure 4 explains the mechanism in this case.

A mathematical model that could serve to suitably manage this problem is that involving the definition of *fuzzy information* (as intended by Okuda *et al.*, [8], Tanaka *et al.*, [12], Zadeh, [16], and analyzed in some of our previous studies, [2], [3]). Unlike the situation in 2.1, the induced probability for the fuzzy information is not immediate to define, since there is not a rule associating with each exact variable value a specified observation.

An illustrative example for this situation is the following one:

Example 2.2. Water in a lake is examined to determine if it is drinkable. It is known that the water may contain a type of microorganisms, so that if the mean number of microorganisms per milliliter, θ, is greater than 7 water cannot be regarded as drinkable. Consider the random experiment consisting in drawing a milliliter of lake water, ω, and observing the number of microorganisms in it, $\mathbf{X}(\omega)$, and assume random distribution of microorganisms in lake water. Then, if exact information were available, we could describe this situation by means of a Poisson experiment with mean θ, $(\mathbb{N},\mathcal{B}_{\mathbb{N}},P_\theta)$, $P_\theta(x) = \theta^x e^{-\theta}/x!$ for $x \in \mathbb{N} \equiv X$.

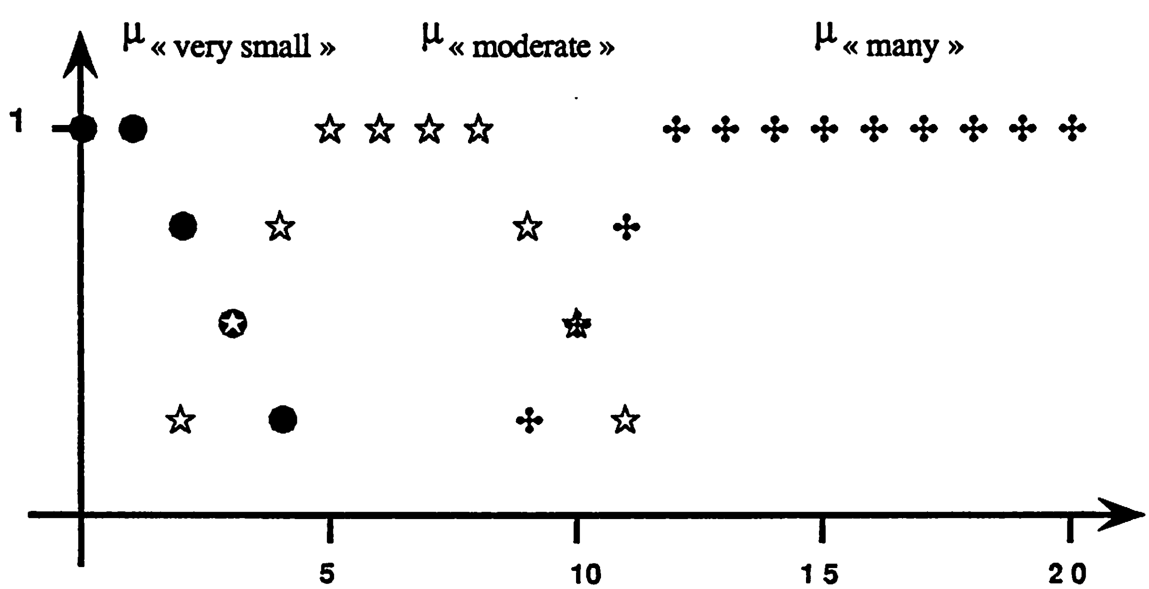

Fig. 5. Membership functions of the fuzzy observations « a very small number of microorganisms » (●), « a moderate number of microorganisms » (☆), and « many microorganisms » (✣)

Suppose that a biologist is interested in concluding if lake water is drinkable or not, but the microorganisms are actually very difficult to be identified, and based on some features he can only report after each experimental performance that there is a « very small number of microorganisms », or a « moderate number of microorganisms », or « many microorganisms », and so on. This type of observations can be easily described in terms of fuzzy subsets of X, rather than in terms of classical ones. Thus, the preceding observations could be described, for instance, by means of the (measurable) fuzzy subsets characterized by the membership functions in Figure 5.

2.3. Fuzzy events of interest

Finally, we can consider the situation in which fuzziness is incorporated through the fourth step (definition of events of interest).Thus, it is assumed that both, the outcome from $(\Omega,\mathcal{F},P)$ and the associated variable (or vector) value from **X**, are exact and the value is exactly perceived by the observer, but the events of interest are statements or questions regarding **X** so that after knowing the value from **X** the observer cannot answer YES or NOT to their occurrence but he is able to specify the degree with which each of them is true (or false). The scheme in Figure 6 explains the mechanism in this case.

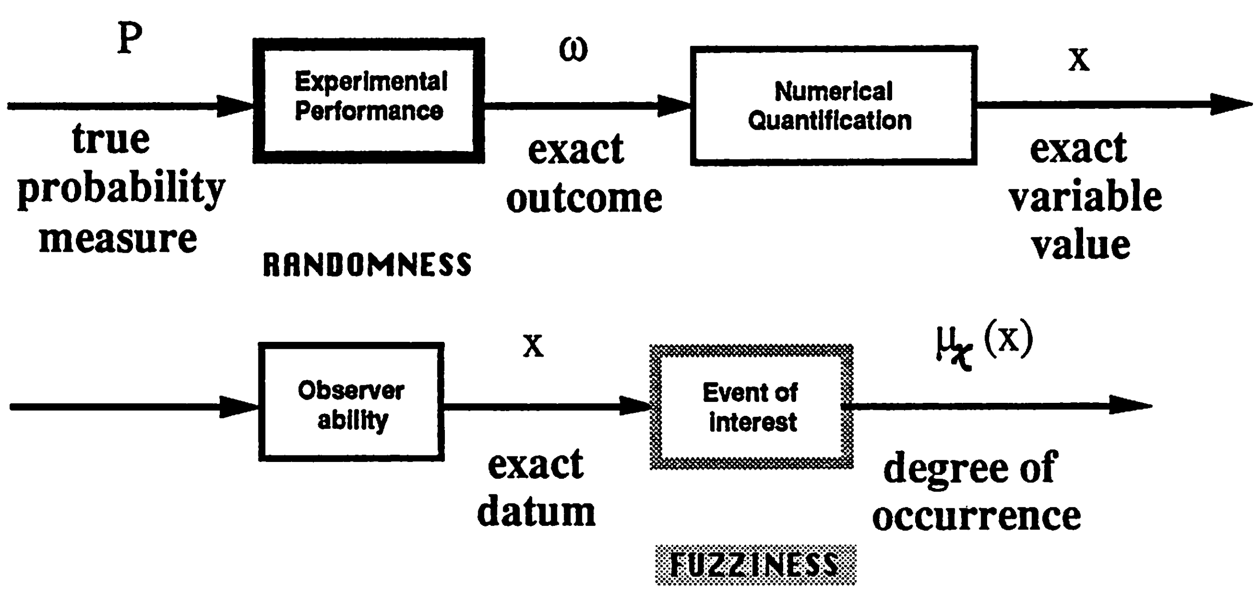

Fig. 6. Process leading to obtain fuzzy events of interest

A mathematical model that could suitably describe this situation is that based on identifying the events of interest with *(measurable) fuzzy subsets* of X, so that after knowing the value from **X** the degree with which each of the events is true is the grade of membership of that value to the corresponding fuzzy subset. As in 2.2, the induced probability for the fuzzy events is not immediate to define, since a fuzzy event cannot be identified with an element of $\mathcal{B}_X$.

An example for this source of fuzziness is the following one:

Example 2.3. Suppose that an observer is measuring the height of a given person, ω, having 70 inches of height, **X**(ω), and assume that to the occurrence of the event "the person is tall" the observer is not able to answer YES or NOT, but he prefers to answer « more or less » (or, in other words, he prefers to answer that the event is true with degree .5). Consequently, the event « being tall » could be described by means of a fuzzy event characterized, for instance, by the membership function in Figure 7.

Fig. 7. Membership functions of the fuzzy event « being tall »

3. DISCUSSING AN INTERPRETATION OF GRADES OF MEMBERSHIP AS PROBABILITIES

In previous papers ([5], [11], [12]) it was emphasized that $\mu_{\chi}(x)$ could be interpreted or estimated as the probability of χ given x.

In accordance with the analysis of the sources of fuzziness we have just developed, we can assert that in Section 2.1 the interpretation above mentioned does not make sense, since the exact value x is not given. In fact, unlike cases 2.2 and 2.3, there is not an exact numerical quantification process underlying the situation in 2.1, so it is not possible to talk about the probability with which something happens when the observer obtains x.

On the other hand, the interpretation of $\mu_\chi(x)$ as a probability in 2.3 could make sense, but it would mean that the criterion to classify the exact value x as belonging or not to χ varies from observer to observer (so that a $100\mu_\chi(x)$ % of them would agree that x belongs to χ, or leads to the occurrence of χ), but any given observer may answer YES or NOT to the occurrence of χ when x is obtained. However, in accordance with the "fuzzy approach" in Section 2.3, $\mu_\chi(x) \neq 0$ or 1 would indicate that the given observer cannot answer YES or NOT to the occurrence of χ, but he prefers to answer that it is true according to a degree equal to $\mu_\chi(x)$. Consequently, this interpretation of $\mu_\chi(x)$ in terms of probabilities would lead to a problem definitively different in nature.

In fact, the only case in which such an interpretation could make some sense is the one considered in Section 2.2. Thus, 2.2 is the only case in which for a given input in the step under fuzziness the unicity of outputs cannot be guaranteed, that is, in which a "kind" of randomness justifying the interpretation above could implicitly be involved, since as we have already remarked there is not a rule associating a specified fuzzy observation with each exact variable value. In other words, in this case we cannot guarantee that whenever a given value x of $\mathbf{X}$ arise, the observer will always report the same fuzzy information χ (since should this be the case χ could be identified with a classical subset of X, or grouped datum). That interpretation may be illustrated in Example 2.2 by saying that $\mu_{\text{« moderate »}}(9) = .75$ indicates that in 75 % of milliliters from the lake water in which there really are 9 microorganisms, the observer would perceive that there would be a moderate number of them.

We are now going to discuss the interpretation of the grades of membership as probabilities for random experiments involving fuzziness because of the reasons explained in Section 2.2. To formalize the discussion, we have to remark again that the induced probability for the fuzzy information associated with the experiment $(X,\mathcal{B}_X,P)$ is not immediate to define. Zadeh, ([15]), suggested to quantify this "induced probability" of the fuzzy information $\tilde{x}$ associated with the experiment $(X,\mathcal{B}_X,P)$ as follows:

DEFINITION 3.1. The probability of $\tilde{x}$ induced by P is given by the Lebesgue-Stieltjes integral

$$\mathcal{P}(\tilde{x}) = \int_X \mu_{\tilde{x}}(x)\, dP(x) \qquad (2)$$

According to Zadeh, ([15]), the value $\mathcal{P}(\tilde{x})$ could be viewed as the "degree of consistency" of the probability distribution P with the possibility distribution associated with the membership function $\mu_{\tilde{x}}$.

Although (2) is introduced as a definition (not a result), it may be justified through the two following arguments:

* it is the most *immediate extension* of the non-fuzzy case in (1) (in which we replace the indicator or characteristic function by the membership function),

and

* it is coherent with *Le Cam's definition* of the probability of bounded numerical functions in a *single stage experiment* (cf., [6], [7]). Thus, Le Cam suggested to replace the structure of a classical experiment, by a weaker structure $(X,\mathcal{V}_X,P)$, called single stage experiment, where $\mathcal{V}_X$ is a vector lattice for the usual operations (sum, product by real numbers, pointwise supremum and infimum) that contains the indicator or characteristic function of X and complete for the norm sup $|\cdot|$, P is a normalized linear functional on $\mathcal{V}_X$, and for $f \in \mathcal{V}_X$ the value of P at f may be considered as the Lebesgue-Stieltjes integral given by $P(f) = \int_X f(x)\, dP(x)$. Consequently, if the fuzzy event $\tilde{x}$ is such that $\mu_{\tilde{x}}$ belongs to the vector lattice $\mathcal{V}_X$ in a single stage experiment, then $P(\mu_{\tilde{x}}) = \mathcal{P}(\tilde{x})$.

On the basis of this probabilistic definition we can now discuss the interpretation of grades of membership we have previously commented on.

Let $(X,\mathcal{B}_X,P)$ be a random experiment and let $\tilde{x}$ denote a fuzzy event associated with it. Given $x^\circ \in X$, we can define a new experiment characterized by the probability space

$(\{x^\circ\},\mathcal{B}_{\{x^\circ\}},P_{x^\circ})$ (with $P_{x^\circ}(\{x^\circ\}) = 1$) in which a new random variable (or vector) degenerated at x° is to be observed. This experiment would be characterized by the probability space . If we now consider the restriction of the fuzzy event χ to $\{x^\circ\}$ (or, more precisely, the restriction of the mapping μ_χ from X to $\{x^\circ\}$), then

PROPOSITION 3.1. The probability of the restriction of χ induced by P_{x° is given by

$$\mathcal{P}_{x^\circ}(\chi) = \mu_\chi(x^\circ) \qquad (3) \qquad \square$$

This last result indicates that, when we use Zadeh's probabilistic definition, $\mu_\chi(x^\circ)$ could be intuitively interpreted as "a kind" of (induced) probability with which the observer gets χ when he really has obtained x°.

Nevertheless, it should be emphasized that such an interpretation is just an intuitive but it does not mean a rigorous approximation to quantify the grade of membership, because of the following reasons:

i) It depends on the particular probabilistic definition for fuzzy events we have considered.

ii) The process of restricting the fuzzy event χ from $(X,\mathcal{B}_X,P)$ to $(\{x^\circ\},\mathcal{B}_{\{x^\circ\}},P_{x^\circ})$ does not mean an intuitive step in computing $\mu_\chi(x^\circ)$, since χ makes only sense when it is defined over all X.

4. CONCLUDING REMARKS

The preceding arguments indicate that when we try to use the probability of χ given x as an alternative way to approximate $\mu_\chi(x)$ in the case of fuzzy experimental information, we have first to formalize χ and its probability. This formalization can be carried out in a framework involving Probability Theory and Fuzzy Sets Theory, but it is not possible to develop it only in a classical probabilistic framework. Consequently, we can conclude that such an approximation can be sometimes viewed as an intuitive but not a formal way to assign grades of membership to fuzzy information associated with a random experiment.

On the other hand, to accept such an interpretation implies to treat χ as a *random set*. This identification would allow us to formally compute the probability of χ given x, but it presents an inconvenient in practice: operations between random sets are usually much more complex than those between fuzzy sets.

It should be pointed out that sources of fuzzy imprecision we have considered in this paper can be easily combined, leading to a general model involving the three cases above considered. In this sense, it is worth remarking that the discussion regarding the sources of fuzziness in random experiments has been developed in this paper attending to the step in which fuzziness arises.An alternative interesting discussion could also be developed attending to the nature of the fuzziness. In other words, all kinds of fuzziness could possibly be modeled by combining to aspects: i) *uncertainty in observation* (error, imprecision, lack of information), and ii) *subjective judgement* (vagueness, cognition, personal bias). As special cases, uncertainty in observation could be zero or nonzero, and subjective judgement could be consistent or vary. According to this alternative discussion, situations in 2.1 and 2.3 would correspond to the particular case in which the uncertainty in observation is zero and subjective judgement varies, whereas the situation in 2.2 would correspond to the one in which uncertainty in observation is nonzero and subjective judgement is consistent.When the uncertainty in observation is nonzero and probabilitistic in nature, then the interpretation of grades of membership as probabilities can make sense, but otherwise that interpretation does not make sense. Consequently, this alternative discussion leads also to obtain the conclusions in Section 3.

Finally, probabilities in the original probability space could also involve fuziness in practice, leading to a new kind of statistical problems.

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 the ideas in the alternative discussion suggested in Section 4.

REFERENCES

1. J.C. Bezdek, Partition Structures: A tutorial, *Analysis of Fuzzy Information,* III, 81-107, CRC Press, Inc., Boca Raton, 1987.

2. M.A. Gil, Fuzziness and Loss of Information in Statistical problems, *IEEE Trans. Systems Man Cybernet.*, 17 (1987) 1016-1025.

3. M.A. Gil, On the Loss of Information due to Fuzziness in Experimental Observations, *Ann. Inst. Statis. Math.*, 40, 4 (1988) 627-639.

4. M.A. Gil and P. Jain, The effects of Perfect and Sample Information on Fuzzy Utilities in Decision-Making, *Proc. 5th Workshop on Uncertainty in Artificial Intelligence,* (1989) 126-133.

5. H. Hisdal, Are grades of membership probabilities, *Fuzzy Sets and Systems*, 25, 3 (1988) 325-348.

6. L. Le Cam, Sufficiency and Approximate Sufficiency, *Ann. Math. Statist.*, 35 (1964) 1419-1455.

7. L. Le Cam, *Asymptotic Methods in Statistical Decision Theory*. Springer-Verlag, Berlin, 1986.

8. T. Okuda, H. Tanaka and K. Asai, A Formulation of Fuzzy Decision problems with Fuzzy Information, using Probability Measures of Fuzzy Events, *Inform. Contr.*, 38 (1978) 135-147.

9. M.L. Puri and D.A. Ralescu, The concept of Normality for Fuzzy Random Variables, *Ann. Probab.*, 13 (1985), 1373-1379.

10. M.L. Puri and D.A. Ralescu, Fuzzy Random Variables, *J. Math. Anal. Appl.*, 114 (1986), 409-422.

11. E. Ruspini, A new approach to clustering, *Inform. Contr.*, 15 (1969) 22-32.

12. H. Tanaka, T. Okuda and K. Asai, Fuzzy Information and Decision in Statistical Model. In *Advances in Fuzzy Sets Theory and Applications*, 1979, 303-320, North-Holland. Amsterdam.

13. R.J. Wonnacott and T.H. Wonnacott, *Introductory Statistics*, 1985, J. Wiley & Sons. New York.

14. L.A. Zadeh, Fuzzy Sets, *Inform. Contr.*, 8 (1965) 338-353.

15. L.A. Zadeh, Probability Measures of Fuzzy Events, *J. Math. Anal. Appl.*, 23 (1968) 421-427.

16. L.A. Zadeh, Fuzzy Sets as a basis for a Theory of Possibility, *Fuzzy Sets and Systems*, 1 (1978) 3-28.