COMPARISON OF EXPERIMENTS IN DECISION PROBLEMS WITH FUZZY UTILITIES. A CRITERION BASED ON THE EXPECTED VALUE OF SAMPLE INFORMATION

by

Maria Angeles Gil and Pramod Jain

Memorandum No. UCB/ERL M89/103

29 August 1989

COVER PAGE

COMPARISON OF EXPERIMENTS IN DECISION PROBLEMS WITH FUZZY UTILITIES. A CRITERION BASED ON THE EXPECTED VALUE OF SAMPLE INFORMATION

by

Maria Angeles Gil and Pramod Jain

Memorandum No. UCB/ERL M89/103

29 August 1989

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

COMPARISON OF EXPERIMENTS IN DECISION PROBLEMS WITH FUZZY UTILITIES. A CRITERION BASED ON THE EXPECTED VALUE OF SAMPLE INFORMATION

by

Maria Angeles Gil and Pramod Jain

Memorandum No. UCB/ERL M89/103

29 August 1989

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

COMPARISON OF EXPERIMENTS IN DECISION PROBLEMS WITH FUZZY UTILITIES. A CRITERION BASED ON THE EXPECTED VALUE OF SAMPLE INFORMATION

María Angeles Gil*
Department of Electrical Engineering and Computer Sciences, University of California, Berkeley

Pramod Jain
Department of Mechanical Engineering, University of California, Berkeley

Abstract

The choice of an optimal action in a Decision-Making problem involving fuzzy utilities has been approached in previous studies. This choice reduces, in most of them, to model the fuzzy utilities and the expected fuzzy utility (that is usually considered as a fuzzy number), and to select a technique for ranking fuzzy numbers that fits well the situation. In this paper, we incorporate a new element to the Decision problem above: the sample information supplied by a random experiment associated with it. This new element, along with the use of fuzzy random variables (Puri and Ralescu) to model fuzzy utilities, and a fuzzy preference relation stated by Kolodziejczyk, will allow us first to extend appropriately in the Bayesian framework the concept of Expected Value of Sample Information (or gain in expected fuzzy utility due to the knowledge of the sample information). On the basis of this concept, we will then establish a criterion to compare random experiments associated with the problem, and analyze some interesting properties confirming its suitability. An example will illustrate the application of the suggested procedure. Finally, it will be contrasted with the "pattern criterion", based on statistical sufficiency and introduced by Blackwell.

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

1. INTRODUCTION

Consider two statistical experiments in which two random variables or vectors are to be observed on the same population, so that their distributions depend on the same unknown state of nature or parameter value. The idea of comparing such experiments was suggested by Bohnenblust, Shapley and Sherman in a private communication and their basic results were collected by Blackwell [3]. The Bohnenblust *et al.* procedure of comparison was formulated in a Decision-Making context, whereas Blackwell proposed a method that was based on statistical sufficiency, in which the purpose was to get as much probabilistic information about the state or parameter as possible, without having decisions in mind. Blackwell [3] proved that both methods are equivalent. The comparison of experiments was developed into a theory by Blackwell [3], [4] and other authors established criteria to compare experiments and connected them with Blackwell's sufficiency criterion (see, for instance, papers referenced in Lehmann, [19]).

The aim of this paper is to introduce a method to compare experiments in Decision-Making problems in a Bayesian framework, when utilities (or losses) cannot be assessed in terms of numerical values, but can be described by means of "fuzzy values". The traditional utility assessment procedures usually involve acceptance of some conditions or axioms for the preference relations, in order to guarantee the existence of a numerical utility function (axiomatic approach to the Utility Theory). As remarked in previous papers (see, for instance, Jain [17], Watson *et al.* [28], Adamo [1], Freeling [10], Orlovsky [23], Dubois and Prade [8], Whalen [29], and Kolodziejczyk [18]), the necessity for assessing utilities in terms of numerical values may be, in practice, too restrictive, whereas the use of fuzzy sets to describe utilities is often more realistic.

To justify this assertion we now examine an example that has been taken from an introductory text of Statistics [30], with some modifications.

Example: 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 utilities (intended in this example as opposite to losses) of right classifications are null. The utilities of a wrong classification are obvious: an unnecessary operation means resources are wasted and the patient may be hurt. Yet the other utility may be worse: if a patient requiring surgery does not get it, the time lost until clear symptoms appear may be crucial.

Assume that the neurologist can base his decision on two different scores, **X** and **Y**, each one of them is obtained from combination of several tests. Past experiences have shown that **X** is normally distributed with variance equal to 64 and mean equal to 120 for those who require surgery and 100 for those who do not, and **Y** is normally distributed with variance equal to 81 and mean equal to 120 for those who require surgery and 100 for those who do not.

The question we are now interested in is the following: what is the "most convenient" score the neurologist can base his decision on?.

In [30], without considering the comparison of experiments, the preceding situation, is regarded as a Decision problem in a Bayesian context, with state space $\Theta = \{\theta_1, \theta_2\}$ (θ_1 = the patient requires surgery, θ_2 = the patient does not require surgery), action space $A = \{a_1, a_2\}$, prior distribution ξ, such that $\xi(\theta_1) = .6$, $\xi(\theta_2) = .4$, and utility function $u(\theta_1,a_1) = u(\theta_2,a_2) = 0$, $u(\theta_1,a_2) = 5u(\theta_2,a_1)$, with $u(\theta_2,a_1) < 0$.

However, the preceding assessment of utilities seems to be extremely precise for the nature of the actions and states in the problem. Thus, the following assessment could express decision maker's (neurologist) "preferences" better: $\mathcal{u}(\theta_1,a_1) = \mathcal{u}(\theta_2,a_2) = 0$, $\mathcal{u}(\theta_2,a_1)$ = «inconvenient», $\mathcal{u}(\theta_1,a_2)$ = «dangerous», where $\mathcal{u}(\theta_1,a_2)$ and $\mathcal{u}(\theta_2,a_1)$ could be described by means of some adequate fuzzy sets as we will see later.

We are next going to formalize these kinds of problems, and extend the notion of Expected Value of Sample Information, introduced by Raiffa and Schlaifer [26], and exhaust- ively studied by García-Carrasco [11]. On the basis of this value, we will establish a criterion to compare experiments, whose suitability will be then analyzed. The above example will be considered again to select the "most convenient" score for the neurologist, according to the suggested method. Finally, as it is usual in most of these studies, we will examine the implica- tions of the sufficiency comparison in terms of the new criterion.

2. PRELIMINARY CONCEPTS AND MODELS

In traditional Decision-Making problems, probabilities are numerical representations of the beliefs and the current state of information of the decision maker, whereas utilities are numerical representations of his preferences. Thus, in these problems the decision maker must be able to quantify the relative value of any situation that may arise. In a Bayesian context, the utility function is formalized as follows (cf., DeGroot [6]). Let Θ and A denote the state and action spaces of the Decision problem, and let ξ be the prior distribution on a measurable space defined on Θ, say $(\Theta,\mathcal{C})$. Then, a *utility function* is a real-valued function u on $\Theta \times A$ (where $u(\theta,a)$ is the utility assessed to the consequence of taking action a when θ is the true state of nature), such that

i) for each action $a \in A$, $u(\bullet,a)$ is a *random variable* on $(\Theta,\mathcal{C})$, having a finite expectation with respect to ξ denoted by $E[u(a|\xi)]$.

ii) a is *preferred or indifferent* to a' (depending on decision maker's preferences) if and only if $E[u(a|\xi)] \geq E[u(a'|\xi)]$.

Following this idea, we are now going to formalize the notion of fuzzy utility function by using the concepts of fuzzy random variable and the associated expected value, as defined by Puri and Ralescu, [24, 25] (see, also Negoita and Ralescu [21], for a review of the main results).

Let $(\Theta,\mathcal{C},\xi)$ be a probability space and let $\mathcal{F}_0(\mathbb{R})$ denote the collection of all fuzzy subsets $\mathcal{V}$ of $\mathbb{R}$, characterized by a membership function $\mu_{\mathcal{V}}: \mathbb{R} \to [0,1]$, with the following properties:

(1) supp $\mathcal{V}$ = support of $\mathcal{V}$ = closure of $\{ \omega \in \mathbb{R} \mid \mu_{\mathcal{V}}(\omega) > 0 \}$ is compact (i.e., closed and bounded).

(2) $L_\alpha(\mathcal{V})$ = α-level set of $\mathcal{V} = \{ \omega \in \mathbb{R} \mid \mu_{\mathcal{V}}(\omega) \geq \alpha \}$ is closed for each $0 \leq \alpha \leq 1$.

(3) $L_1(\mathcal{V})$ = modal set of $\mathcal{V} = \{ \omega \in \mathbb{R} \mid \mu_{\mathcal{V}}(\omega) = 1 \} \neq \emptyset$.

Definition 2.1. A (one-dimensional) *fuzzy random variable* is a function $\mathcal{v}: \Theta \to \mathcal{F}_0(\mathbb{R})$, such that $\{ (\theta,\omega) \mid \omega \in L_\alpha(\mathcal{v}(\theta)) \} \in \mathcal{C} \times \mathcal{B}_{\mathbb{R}}$, for all $\alpha \in [0,1]$, where $\mathcal{B}_{\mathbb{R}}$ is the Borel σ-field on

$\mathbb{R}$. (In other words, the random set $L_\alpha(\mathcal{V}(\cdot))$, defined on Θ, is measurable for each $0 \le \alpha \le 1$).

Remark 2.1: Fuzzy random variables (FRV) generalize random variables (by replacing the singleton $\{\mathcal{V}(\theta)\}$ by the α-sets $L_\alpha(\mathcal{V}(\theta))$, $0 \le \alpha \le 1$, and $\mathbb{R}$ by $\mathcal{F}_0(\mathbb{R})$), and random sets (by replacing the set $\mathcal{V}(\theta)$ by the α-sets $L_\alpha(\mathcal{V}(\theta))$, $0 \le \alpha \le 1$, and $\mathcal{K}(\mathbb{R})$ = {non-empty compact subsets of $\mathbb{R}$} by $\mathcal{F}_0(\mathbb{R})$). According to the definition of $\mathcal{F}_0(\mathbb{R})$, condition (1) has been imposed to ensure that the variable "values" are "bounded" in some sense, (2) is a measurability condition for the membership functions describing the fuzzy variable values, and (3) guarantees that all levels of each variable "value" are non-empty.

The expected value of a fuzzy random variable is defined by

Definition 2.2. Let $\mathcal{V}: \Theta \to \mathcal{F}_0(\mathbb{R})$ be a simple FRV, that is, a FRV taking on fuzzy values $\mathcal{V}_j \in \mathcal{F}_0(\mathbb{R})$ on $C_j \in \mathcal{C}$, j=1,...,k, respectively. This variable can be written by

$$\mathcal{V} = \sum_{j=1}^{k} \mathcal{V}_j \chi_{C_j}$$

(where χ_C = indicator function of C). Then, the *expected value of* $\mathcal{V}$, with respect to the probability measure ξ on $(\Theta,\mathcal{C})$, is the fuzzy set $\mathcal{E}(\mathcal{V}|\xi) = \int_\Theta \mathcal{V}(\theta)\, d\,\xi(\theta) \in \mathcal{F}_0(\mathbb{R})$ given by

$$\mathcal{E}(\mathcal{V}|\xi) = \sum_{j=1}^{k} \mathcal{V}_j\, \xi(C_j) \quad .$$

with $\mathcal{V}_j\,\xi(C_j)$ = fuzzy product of $\mathcal{V}_j$ by the constant real value $\xi(C_j)$, and $\sum$ = fuzzy addition (these fuzzy operations are based on the extension principle, [32]).

For a more general FRV, $\mathcal{V}$, it is possible to take a sequence of simple FRV, $\mathcal{V}_n$, such that $\lim_{n\to\infty} d_\infty(\mathcal{V}_n(\theta),\mathcal{V}(\theta)) = 0$, for almost all $\theta \in \Theta$ (where $d_\infty(\mathcal{V}_n(\theta),\mathcal{V}(\theta)) = \sup_{0\le\alpha\le1} d_H(L_\alpha(\mathcal{V}_n(\theta)),L_\alpha(\mathcal{V}(\theta))$ is a metric on $\mathcal{F}_0(\mathbb{R})$ and d_H is the Hausdorff distance defined on the set of all non-empty compact subsets of $\mathbb{R}$, that is, $d_H(\Omega,\Omega') = \max\{\sup_{\omega\in\Omega}\inf_{\omega'\in\Omega'}|\omega-\omega'|, \sup_{\omega'\in\Omega'}\inf_{\omega\in\Omega}|\omega-\omega'|\}$). Then, the expected value of $\mathcal{V}$, with respect to the probability measure ξ on $(\Theta,\mathcal{C})$, is the unique fuzzy set $\mathcal{E}(\mathcal{V}|\xi) = \int_\Theta \mathcal{V}(\theta)\, d\,\xi(\theta) \in \mathcal{F}_0(\mathbb{R})$ such that $\lim_{n\to\infty} d_\infty(\mathcal{E}(\mathcal{V}_n|\xi),\mathcal{E}(\mathcal{V}|\xi)) = 0$.

Remark 2.2: A rigourous and detailed justification of the definition of the expected value above, can be found in papers of Puri and Ralescu, ([24, 25]), and the book of Negoita and Ralescu, [21]. In the original paper (Puri and Ralescu, [25]), the expected value of a FRV is introduced in a different way, as the unique fuzzy set such that $L_\alpha(\mathcal{E}(\mathcal{V}|\xi))$ = Aumann integral ([2]) of the random set $L_\alpha(\mathcal{V}(\bullet))$ with respect to ξ, for all $0 \leq \alpha \leq 1$, but the final definition is equivalent to the last one and more complex to describe.

Remark 2.3: It should be emphasized that, whenever the probability space $(\Theta,\mathcal{C},\xi)$, is non-atomic (that is, for each $C \in \mathcal{C}$ such that $\xi(C) > 0$, there exists a $C' \in \mathcal{C}$ such that $C \supset C'$ and $\xi(C) > \xi(C') > 0$) and $\mathcal{V}$ is *integrably bounded* (that is, there exists $h : \Theta \to \mathbb{R}$, integrable with respect to ξ, such that $\sup_{\omega \in L_\alpha(\mathcal{V}(\theta))} \|\omega\| \leq h(\theta)$, for all $0 < \alpha < 1$)), then $\mathcal{E}(\mathcal{V}|\xi)$ exists and is a *fuzzy number* (intended as a normalized convex fuzzy set).

Remark 2.4: It is worth also pointing out that, in practice, the computation of $\mathcal{E}(\mathcal{V}|\xi)$ for non-simple FRV, obtained from a limiting process (in d_∞), becomes usually complicated.

To extend the notion of utility function to fuzzy utility function, we additionally need to consider a third element: the comparison between expected values. As these expected values will be fuzzy numbers, that comparison has to be a ranking of fuzzy numbers. Several procedures have been proposed in the literature of fuzzy numbers. Kolodziejczyk [18] analyzed different fuzzy preference relations (following Orlovsky ideas, [23]) satisfying some properties that confirm their suitability to rank fuzzy numbers. Some of them were suggested so that the calculations in the set of fuzzy numbers, with respect to the fuzzy addition and product by a constant real number, could be performed in a manner analogous to the operations on real numbers (and, consequently, so that the calculations through the expected value for a FRV could be performed in an analogous way as for random variables, what is very convenient and plausible for our purposes). Two of these preference relations are defined as follows:

Let $\mathcal{V} \in \mathcal{F}_0(\mathbb{R})$ be a fuzzy number. Let $\mathcal{V}^L$ and $\mathcal{V}^P$ denote the fuzzy sets of $\mathbb{R}$ with membership functions

$$\mu_{\mathcal{V}^L}(\omega) = \begin{cases} \mu_{\mathcal{V}}(\omega) & \text{for } \omega \leq z \\ 1 & \text{for } \omega > z \end{cases}$$

$$\mu_{\mathcal{V}^P}(\omega) = \begin{cases} \mu_{\mathcal{V}}(\omega) & \text{for } \omega \geq z \\ 1 & \text{for } \omega < z \end{cases}$$

where $z \in \mathbb{R}$ is such that $\mu_{\mathcal{V}}(z) = 1$.

Definition 2.3. Let $\mathcal{U}, \mathcal{V} \in \mathcal{F}_0(\mathbb{R})$, be two fuzzy numbers. The following values represent "degrees of truth for the expression « $\mathcal{U}$ is not higher than $\mathcal{V}$»"

$$R^*(\mathcal{U},\mathcal{V}) = \frac{d(\mathcal{U}^L \vee \mathcal{V}^L, \mathcal{U}^L) + d(\mathcal{U}^P \vee \mathcal{V}^P, \mathcal{U}^P)}{d(\mathcal{U}^L, \mathcal{V}^L) + d(\mathcal{U}^P, \mathcal{V}^P)}$$

$$R^\circ(\mathcal{U},\mathcal{V}) = \frac{d(\mathcal{U}^L \vee \mathcal{V}^L, \mathcal{U}^L) + d(\mathcal{U}^P \vee \mathcal{V}^P, \mathcal{U}^P) + d(\mathcal{U} \cap \mathcal{V}, \mathit{0})}{d(\mathcal{U}^L, \mathcal{V}^L) + d(\mathcal{U}^P, \mathcal{V}^P) + 2\, d(\mathcal{U} \cap \mathcal{V}, \mathit{0})}$$

(where d = Hamming distance between fuzzy sets : $d(\mathcal{U},\mathcal{V}) = \int_{\mathbb{R}} |\mu_{\mathcal{U}}(z) - \mu_{\mathcal{V}}(z)|\, dz$; $\vee$ = extended maximum of fuzzy sets : $\mu_{\mathcal{U} \vee \mathcal{V}}(z) = \sup\limits_{x \vee y = z} [\min\{\mu_{\mathcal{U}}(x), \mu_{\mathcal{V}}(y)\}]$, for all $z \in \mathbb{R}$; $\cap$ = intersection of fuzzy sets : $\mu_{\mathcal{U} \cap \mathcal{V}}(z) = \min\{\mu_{\mathcal{U}}(x), \mu_{\mathcal{V}}(y)\}$; and $\mathit{0}$ = especial fuzzy set assigning membership function equal to 1 to the value 0, and equal to 0 otherwise). In addition,

- $\mathcal{U}$ is said to be *preferred or indifferent to* $\mathcal{V}$, denoted by $\mathcal{U} \geq^* \mathcal{V}$ (or $\mathcal{U} \geq^\circ \mathcal{V}$), whenever $R^*(\mathcal{U},\mathcal{V}) \leq R^*(\mathcal{V},\mathcal{U}) = 1 - R^*(\mathcal{U},\mathcal{V})$ (respectively, $R^\circ(\mathcal{U},\mathcal{V}) \leq R^\circ(\mathcal{V},\mathcal{U}) = 1 - R^\circ(\mathcal{U},\mathcal{V})$), that is, whenever $R^*(\mathcal{U},\mathcal{V}) \leq .5$ (respectively, $R^\circ(\mathcal{U},\mathcal{V}) \leq .5$).
- $\mathcal{U}$ is said to be *indifferent to* $\mathcal{V}$, denoted by $\mathcal{U} \sim^* \mathcal{V}$ (or $\mathcal{U} \sim^\circ \mathcal{V}$), whenever $R^*(\mathcal{U},\mathcal{V}) = R^*(\mathcal{V},\mathcal{U})$ (respectively, $R^\circ(\mathcal{U},\mathcal{V}) = R^\circ(\mathcal{V},\mathcal{U})$), that is, whenever $R^*(\mathcal{U},\mathcal{V}) = .5$ (respectively, $R^\circ(\mathcal{U},\mathcal{V}) = .5$).

Remark 2.5: Kolodziejczyk [18], proved that if $\mathcal{T}, \mathcal{U}, \mathcal{V}, \mathcal{W} \in \mathcal{F}_0(\mathbb{R})$ are four fuzzy numbers and $R^*(\mathcal{T},\mathcal{U}) \leq .5$ (i.e., $\mathcal{T} \geq^* \mathcal{U}$), and $R^*(\mathcal{V},\mathcal{W}) \leq .5$ (i.e., $\mathcal{V} \geq^* \mathcal{W}$), then $R^*(\mathcal{T}+\mathcal{V}, \mathcal{U}+\mathcal{W}) \leq .5$ (i.e., $\mathcal{T}+\mathcal{V} \geq^* \mathcal{U}+\mathcal{W}$), with $+$ = fuzzy adition. On the other hand, if λ is a real constant, $\lambda > 0$, $R^*(\mathcal{U},\mathcal{V}) = R^*(\lambda\mathcal{U}, \lambda\mathcal{V})$ (i.e., $\mathcal{U} \geq^* \mathcal{V}$ iff $\lambda\mathcal{U} \geq^* \lambda\mathcal{V}$), and $R^*(\mathcal{U}+(-\mathcal{U}), \mathit{0}) = .5$ (i.e., $\mathcal{U}+(-\mathcal{U}) \sim^* \mathit{0}$). Consequently, if $R^*(\mathcal{U},\mathcal{V}) \leq .5$ (i.e., $\mathcal{U} \geq^* \mathcal{V}$), then $R^*(\mathcal{U} - \mathcal{V}, \mathit{0}) \leq .5$ (i.e., $\mathcal{U} - \mathcal{V} \geq^* \mathit{0}$). Analogous results are true for R°. In general, any suitable comparison satisfying these properties would be useful for our purposes (that is, the selection of Kolodziejczyk indices is just a possible option).

Remark 2.6: Kolodziejczyk [18], discussed the advantages of using R° instead of R^*. Thus, although both of them lead to the same preference relation, (since obviously, $R^*(\mathcal{U},\mathcal{V}) \leq .5$ if and only if $R^{\circ}(\mathcal{U},\mathcal{V}) \leq .5$) the index R^* differentiates strongly close fuzzy numbers, whereas R° is less susceptible. In coherence with this remark, when only a preference relation is required we hereafter adopt coefficient R^*, that is slightly easier to compute, but when we need to specify the degree of truth of the assertion « $\mathcal{U}$ is not higher than $\mathcal{V}$», we will take coefficient R°.

Remark 2.7: Obviously, the comparison procedure in Definition 2.3 can only be manageable when the set of fuzzy numbers to be compared is finite. This fact motivates us to assume some constraints (usually irrelevant in practice) on the action space of the Decision problems we will consider next.

We can now model the fuzzy utility function as follows:

Let Θ and A be the state and action spaces of the Decision problem, and let ξ be the prior distribution on the measurable space $(\Theta,\mathcal{C})$. Now,

Definition 2.4. A *fuzzy utility function* is a fuzzy set-valued function u on $\Theta \times A$ such that

i) for each action $a \in A$, $u(\bullet,a)$ is a FRV on $(\Theta,\mathcal{C})$, integrable bounded, and whose expected value with respect to ξ is a fuzzy number denoted by $\mathcal{E}[u(a|\xi)]$.

ii) a is *preferred or indifferent* to a' (according to the decision maker preferences) if and only if $\mathcal{E}[u(a|\xi)] \geq^* \mathcal{E}[u(a'|\xi)]$ (that is, $\mathcal{E}[u(a|\xi)] \geq^{\circ} \mathcal{E}[u(a'|\xi)]$).

Remark 2.8: In accordance with Remark 2.1, to assume that the fuzzy utility function is a FRV, for each action, implies that the utility function is supposed to be "bounded" in some sense. This assumption is not an important constraint, since in many axiomatic developments of the traditional utility theory, the utility function must be bounded (cf., DeGroot [6]).

Remark 2.9: A further constraint, we will assume from now on, is that the action space, A, is finite, so that the selection of the optimal action is possible. In most practical applications the action space is finite, thus such an assumption is not in fact restrictive.

Remark 2.10: It is interesting to emphasize that when the fuzzy utility function is such that $u(\bullet,a)$ is a simple FRV for each action a, then the definition of the expected utility coincides with that of Freeling [10]. However, Definition 2.4 does not require that $u(\theta,a)$ be fuzzy numbers.

For the sake of practical fuzzy utilities assessments, it is useful to extend a well-known result in the non-fuzzy case. This result indicates that if a fuzzy utility function exists, then certain linear transformations of this function will also be utility functions. Thus,

Theorem 2.1. Let u be a fuzzy utility function on $\Theta \times A$. Then, the fuzzy set-valued function on $\Theta \times A$ defined by $v(\theta,a) = \alpha\, u(\theta,a) + \beta$ (where α and β are real constants, $\alpha > 0$) is also a fuzzy utility function.

Proof. Indeed, $v(\bullet,a) = \alpha\, u(\bullet,a) + \beta$ is a FRV for each action $a \in A$. If $a, a' \in A$, we can then verify that $R^*(\mathcal{E}[v(a|\xi)],\mathcal{E}[v(a'|\xi)]) = R^*(\mathcal{E}[u(a|\xi)],\mathcal{E}[u(a'|\xi)])$, due to the properties of the Hamming distance and the operations between fuzzy sets. Consequently, conditions i) and ii) in Definition 2.4 are both satisfied. ❑

Remark 2.11: On the basis of the preceding result, we could arbitrarily constraint, without loss of generality, the support of $u(\theta,a)$ to be contained in a particular bounded real interval, say $[0,1]$ or $[-1,0]$.

3. THE EXPECTED VALUE OF SAMPLE INFORMATION

Let Θ and $A = \{a_1,\dots,a_N\}$ be the state and action spaces of the Decision problem, respectively, and u be a fuzzy utility function on $\Theta \times A$. Let us also assume that the decision maker has an estimate of the prior distribution ξ on the measurable space $(\Theta,\mathcal{C})$. For any action $a \in A$, the fuzzy number $\mathcal{E}[u(a|\xi)]$ will be called the *prior expected fuzzy utility of* a. In accordance with condition ii) in Definition 2.4, the existence of a fuzzy utility function entails the acceptance of the Decision-Making principle based on the "maximization" of expected fuzzy utility. In this way, the Bayes principle of choice is now extended as follows:

Definition 3.1. An action $a_0^* \in A$ is called *optimal prior action* if it maximizes the prior expected fuzzy utility, that is, $\mathcal{E}[u(a_0^*|\xi)] \geq^* \mathcal{E}[u(a_i|\xi)]$, $i=1,\dots,N$. (Obviously, the criterion based on $\geq^{\circ}$ would be equivalent to that using $\geq^*$).

Generally, to increase the "highest" expected fuzzy utility in a Decision problem the decision maker takes advantage of the fact that additional information may reduce his uncertainty about the state in Θ. Thus, the decision maker usually try to get information by performing a random experiment whose distribution depends on the state in Θ.

Let **X** be a random experiment, associated with the observation of a random variable (or vector) in a population, and characterized by the probability space $(X,\mathcal{B}_X,P_\theta)$, $\theta \in \Theta$, where the set of variable (or vector) values, X, is a set in $\mathbb{R}$ (or in a Euclidean space), $\mathcal{B}_X$ is the smallest Borel σ-field on X and P_θ is a probability measure on $(X,\mathcal{B}_X)$, so that θ is the state governing the experimental distribution. If the information obtained by performing of **X** is $x \in X$, then the decision maker can use it to revise the distribution on Θ in light of the experimental information using Bayes' Theorem. This revision leads to the posterior distribution ξ_x on $(\Theta,\mathcal{C})$, characterized by the density function (with respect to a dominating measure on $\mathcal{C}$) $h_x(\theta) = h(\theta)f_\theta(x)/f(x)$ (where h is the density function associated with the prior distribution ξ with respect to the same dominating measure on $\mathcal{C}$; f_θ is the density function associated with the distribution P_θ with respect to a dominating measure on $\mathcal{B}_X$; and f is the density function associated with the marginal probability measure P with respect to the same dominating measure on $\mathcal{B}_X$, that is, f(x) is given by the Lebesgue-Stieltjes integral $\int_\Theta f_\theta(x)\, d\,\xi(\theta)$, for all $x \in X$. The fuzzy number $\mathcal{E}[u(a|\xi_x)]$ will be called the *posterior expected fuzzy utility of the action* a. The application of the Decision-Making principle in Definition 3.1 leads us to

Definition 3.2. An action $a_x^* \in A$, is called *optimal posterior action given* x, if it maximizes the posterior expected fuzzy utility given x, that is, $\mathcal{E}[u(a_x^*|\xi_x)] \geq^* \mathcal{E}[u(a_i|\xi_x)]$, i=1,...,N. (Obviously, the criterion based on $\geq^\circ$ is equivalent to that using $\geq^*$).

Remark 3.1: Theoretically, the set X of variable (or vector) values in the experiment **X** is not required to be finite to develop the present study. Thus, when X is not finite it may be possible to determine in a generic way a_x^* for each $x \in X$ (see, for instance, the example in Section 5). If this generic determination is not easy to carry out, the search of optimum posterior actions becomes in the infinite case unmanageable.

The purpose of the use of sample information is to get a "gain" in expected fuzzy utility on the average. To "quantify" this gain, we can take into account that the "highest" expected fuzzy utility for the decision maker, under the prior information, is equal to $\mathcal{E}[u(a_0^*|\xi)]$. On the other hand, if the decision maker obtains the sample information x in the performance of $\mathbf{X} = (X, \mathcal{B}_X, P_\theta)$, $\theta \in \Theta$, his "highest" expected fuzzy utility would be equal to $\mathcal{E}[u(a_x^* | \xi_x)]$. Thus, his average "highest" expected fuzzy utility, under sample information from **X**, could be measured by the fuzzy number $\int_X \mathcal{E}[u(a_x^* | \xi_x)] \, d\, P(x)$.

Then, the average *worth of the sample information* from **X** could be measured, as in the non-fuzzy case, through the "difference" between the "highest" expected fuzzy utility, under sample information from **X** and the "highest" expected fuzzy utility, under prior information. Thus,

Definition 3.3. The fuzzy set given by

$$\mathcal{U}(\mathbf{X}) = \int_X \mathcal{E}[u(a_x^* | \xi_x)] \, d\, P(x) - \mathcal{E}[u(a_0^* | \xi)]$$

is called *Expected Value of Sample Information* (*EVSI*) *associated with* **X** *in the Decision problem.*

Remark 3.2: Fuzzy operations (Dubois and Prade [7, 9]) guarantee that the *EVSI* could be alternatively computed as follows:

$$\mathcal{U}(\mathbf{X}) = \sum_{i=1}^{N} \mathcal{E}[u(a_i | \xi_{X(a_i)})] \, P(X(a_i)) - \mathcal{E}[u(a_0^* | \xi)]$$

where $X(a_i) = \{ x \in X \mid a_x^* = a_i \} \in \mathcal{B}_X$, and $\xi_{X(ai)}$ being the posterior distribution on $(\Theta, \mathcal{C})$, characterized by the density function (with respect to the dominating measure on $\mathcal{C}$) $h_{X(ai)}(\theta) = h(\theta) P_\theta(X(a_i))/P(X(a_i))$ (with $P_\theta(X(a_i))$ is given by the Lebesgue-Stieltjes integral $\int_{X(ai)} d\, P_\theta(x)$, and $P(X(a_i))$ is given by the Lebesgue-Stieltjes integral $\int_{X(ai)} d\, P(x)$). Thus, in this alternative computation, the *EVSI* can be regarded as the expected value of a simple FRV.

We are next going to define the criterion to compare experiments based on *EVSI*. The analysis of some properties of this criterion will guarantee the suitability of the model and ranking of numbers involved in it.

4. Criterion to compare experiments. Main properties

Let Θ and $A = \{a_1, \ldots, a_N\}$ be the state and action spaces of the Decision problem, respectively, and u be a fuzzy utility function on $\Theta \times A$, and assume that the decision maker assess the prior distributiont ξ on the measurable space $(\Theta, \mathcal{C})$.

Let $\mathbf{X} = (X,\mathcal{B}_X,P_\theta)$, $\theta \in \Theta$, and $\mathbf{Y} = (Y,\mathcal{B}_Y,Q_\theta)$, $\theta \in \Theta$, be two statistical experiments in which two random variables or vectors are to be observed on the same population.

Definition 4.1. We say that $\mathbf{X}$ is *at least as informative as* $\mathbf{Y}$ in the Decision problem, written $\mathbf{X} \geq^{\mathcal{U}} \mathbf{Y}$, if and only if $\mathcal{U}(\mathbf{X}) \geq^* \mathcal{U}(\mathbf{Y})$. We say that $\mathbf{X}$ is *as informative as* $\mathbf{Y}$ in the Decision problem, written $\mathbf{X} \sim^{\mathcal{U}} \mathbf{Y}$, if and only if $\mathbf{X} \geq^{\mathcal{U}} \mathbf{Y}$ and $\mathbf{Y} \geq^{\mathcal{U}} \mathbf{X}$, that is, $\mathcal{U}(\mathbf{X}) \sim^* \mathcal{U}(\mathbf{Y})$. $R^\circ(\mathcal{U}(\mathbf{X}),\mathcal{U}(\mathbf{Y}))$ is the *degree of truth for the assertion « $\mathcal{U}(\mathbf{X})$ is not higher than $\mathcal{U}(\mathbf{Y})$».*

The preceding fuzzy preference relation between two statistical experiments satisfies some natural properties guaranteeing its suitability and that of the involved concepts. We are now going to develop a study similar to that in Lindley, [20] (for Shannon's amount of information).

First of all we verify that provided the distribution of a statistical experiment varies with θ, performing that experiment is at least as informative as not performing it, on the average. Thus,

Theorem 4.1. Let $\mathbf{X} = (X,\mathcal{B}_X,P_\theta)$, $\theta \in \Theta$, be a statistical experiment whose distribution varies with θ. Then $\mathcal{U}(\mathbf{X}) \geq^* \mathit{0}$, whatever the prior distribution on Θ may be.

Let $\mathbf{N} = (N,\mathcal{B}_N,Q_\theta)$, $\theta \in \Theta$, be a statistical experiment whose distribution does not depend on θ (*null* experiment). Then $\mathcal{U}(\mathbf{N}) \sim^* \mathit{0}$, whatever the prior distribution on Θ may be.

Proof. Indeed, for each $x \in X$, we have that $\mathcal{E}[u(a_x^*|\xi_x)] \geq^* \mathcal{E}[u(a_0^*|\xi_x)]$. Consequently, and due to the properties of the R^* index, we can conclude that

$$\int_X \mathcal{E}[u(a_x^*|\xi_x)]\,dP(x) \geq^* \int_X \mathcal{E}[u(a_0^*|\xi_x)]\,dP(x) = \mathcal{E}[u(a_0^*|\xi)]$$

whence, according to Remark 2.5, we obtain $R^*(\mathcal{U}(\mathbf{X}),\mathit{0}) \leq .5$.

On the other hand, for each $n \in N$, we have that ξ_n and ξ coincide, and hence a_n^* and a_0^* are indifferent for all $n \in N$. Therefore, $R^*(\mathcal{U}(\mathbf{N}),\mathit{0}) = .5$. ❑

Suppose that each observation s from a statistical experiment $\mathbf{S} = (S,\mathcal{B}_S,P_\theta)$, $\theta \in \Theta$, consists of a pair of observations (x,y), $x \in X$, $y \in Y$. Consider the experiments, $\mathbf{X} = (X,\mathcal{B}_X,P_\theta^1)$, $\theta \in \Theta$, and $\mathbf{Y} = (Y,\mathcal{B}_Y,P_\theta^2)$, $\theta \in \Theta$, where $\mathcal{B}_X$ and $\mathcal{B}_Y$ are the Borel σ-field over X and Y, induced from $\mathcal{B}_S$ by the projections $g_1(x,y) = x$ and $g_2(x,y) = y$, respectively, and P_θ^1 and P_θ^2 are the probability measures on $\mathcal{B}_X$ and $\mathcal{B}_Y$, respectively, from P_θ. Then, the experiment $\mathbf{S}$ is said to be the *sum* of $\mathbf{X}$ and $\mathbf{Y}$, written $\mathbf{S} = \mathbf{X} \times \mathbf{Y}$.

We are next verifying that performing two statistical experiments is at least as informative as performing only one of them.

Theorem 4.2. Let $\mathbf{S} = \mathbf{X} \times \mathbf{Y}$. Then $\mathbf{S} \geq^{\mathcal{U}} \mathbf{X}$, $\mathbf{S} \geq^{\mathcal{U}} \mathbf{Y}$, whatever the prior distribution on Θ may be.

Proof. Indeed, to prove this result we first discuss the definition of the average *EVSI* associated with the experiment $\mathbf{Y}$ after $\mathbf{X}$ has been performed and x observed. This value may be quantified as follows: given $x \in X$, we first consider the experiment $\mathbf{Y}_x = (Y, \mathcal{B}_Y, P_\theta^2(\cdot|x))$, $\theta \in \Theta$, and the "prior" distribution ξ_x on Θ. From this prior distribution, the posterior distribution given $y \in Y$ would be $\xi_{(x,y)}$, whence the *EVSI* associated with $\mathbf{Y}_x$ for the prior distribution ξ_x is

$$\mathcal{U}(\mathbf{Y}_x|x) = \int_Y \mathcal{E}[u(a_{(x,y)}^*|\xi_{(x,y)})]\, d\,P^2(y|x) - \mathcal{E}[u(a_x^*|\xi_x)]$$

(where the density characterizing $P^2(\cdot|x)$ with respect to a dominating measure on $\mathcal{B}_Y$ is given by the Lebesgue-Stieltjes integral $f^2(y|x) = \int_\Theta f_\theta^2(y|x)\, d\,\xi_x(\theta)$, with $f_\theta^2(y|x)$ being the density characterizing $P_\theta^2(\cdot|x)$ with respect to the same measure). Consequently, the average *EVSI* associated with $\mathbf{Y}$ after $\mathbf{X}$ has been performed is given by $\mathcal{U}(\mathbf{Y}|\mathbf{X}) = \int_X \mathcal{U}(\mathbf{Y}_x|x)\, d\,P^1(x)$, and due to properties of fuzzy operations (Dubois and Prade, [7, 9]),

$$\mathcal{U}(\mathbf{Y}|\mathbf{X}) = \mathcal{U}(\mathbf{X} \times \mathbf{Y}) - \mathcal{U}(\mathbf{X}) = \mathcal{U}(\mathbf{S}) - \mathcal{U}(\mathbf{X})$$

Therefore, as in virtue of Theorem 4.1 $\mathcal{U}(\mathbf{Y}_x|x) \geq^* 0$ for all $x \in X$, then $\mathcal{U}(\mathbf{Y}|\mathbf{X}) \geq^* 0$, and according to Remark 2.5, we obtain $R^*(\mathcal{U}(\mathbf{S}), \mathcal{U}(\mathbf{X})) \leq .5$.

Analogously, it can be shown that $R^*(\mathcal{U}(\mathbf{S}), \mathcal{U}(\mathbf{Y})) \leq .5$. ❑

As an immediate consequence from the last result, we deduce that the greater the size of a random sample from a statistical experiment, the more informative.

Corollary 4.1. Let $\mathbf{X}^{(m)} = (X^m, \mathcal{B}_{X^m}, P_\theta^m)$, $\theta \in \Theta$, be a *random sample* of size m from the statistical experiment $\mathbf{X} = (X, \mathcal{B}_X, P_\theta)$, $\theta \in \Theta$ (that is, $\mathbf{X}^{(m)} = \overbrace{\mathbf{X} \times \mathbf{X} \times \ldots \times \mathbf{X}}^{(m\ \text{times})}$). Then, $\mathbf{X}^{(m+1)} \geq^{\mathcal{U}} \mathbf{X}^{(m)}$, for all $m \in \mathbb{N}$, whatever the prior distribution on Θ may be. ❑

The following result formalizes the fact that when the experiment $\mathbf{Y}$ cannot add probabilistic information about θ to what is contained in $\mathbf{X}$ (that is, $P_\theta^2(\cdot|x)$ does not depend on θ), it is indifferent performing both experiments than only the last one.

Theorem 4.3. If $P_\theta^2(\cdot|x)$ does not depend on θ, and $\mathbf{S} = \mathbf{X} \times \mathbf{Y}$, then $\mathbf{S} \sim^{\mathcal{U}} \mathbf{X}$, whatever the prior distribution on Θ may be.

Proof. Indeed, under the assumed condition and following arguments similar to those in the proof of Theorem 4.1, we obtain that $\mathcal{U}(\mathbf{Y}|\mathbf{X}) \sim^* 0$, so that $R^*(\mathcal{U}(\mathbf{S}), \mathcal{U}(\mathbf{X})) = .5$. ❑

Two statistical experiments $\mathbf{X} = (X, \mathcal{B}_X, P_\theta^1)$, $\theta \in \Theta$, and $\mathbf{Y} = (Y, \mathcal{B}_Y, P_\theta^2)$, $\theta \in \Theta$, are said to be *independent experiments* if the probability measure P_θ associated with $\mathbf{X} \times \mathbf{Y}$ is the product of the probability measures P_θ^1 and P_θ^2, associated with $\mathbf{X}$ and $\mathbf{Y}$, respectively. The property below indicates that the fuzzy preference relation between two statistical experiments $\mathbf{Y}$ and $\mathbf{Z}$ is preserved when each of them is sumed to a statistical experiment $\mathbf{X}$ independent of both, $\mathbf{Y}$ and $\mathbf{Z}$.

Theorem 4.4. Let $\mathbf{X}$, $\mathbf{Y}$, and $\mathbf{Z}$ be three statistical experiments, whose distributions depend on $\theta \in \Theta$. If $\mathbf{Y} \geq^{\mathcal{U}} \mathbf{Z}$, for all prior distribution on Θ, and $\mathbf{X}$ is independent of both, $\mathbf{Y}$ and $\mathbf{Z}$, then $\mathbf{X} \times \mathbf{Y} \geq^{\mathcal{U}} \mathbf{X} \times \mathbf{Z}$, whatever the prior distribution on Θ may be.

Proof. Indeed,

$$\mathcal{U}(\mathbf{X} \times \mathbf{Y}) = \mathcal{U}(\mathbf{Y}|\mathbf{X}) + \mathcal{U}(\mathbf{X}) \quad \text{and} \quad \mathcal{U}(\mathbf{X} \times \mathbf{Z}) = \mathcal{U}(\mathbf{Z}|\mathbf{X}) + \mathcal{U}(\mathbf{X})$$

As $\mathbf{X}$ and $\mathbf{Y}$ are independent experiments, the experiments $\mathbf{Y}_x = (Y, \mathcal{B}_Y, P_\theta^2(\cdot|x))$, $\theta \in \Theta$, and $\mathbf{Y} = (Y, \mathcal{B}_Y, P_\theta^2)$, $\theta \in \Theta$, are equivalent, since $P_\theta^2(\cdot|x)$ does not depend on x. Analogously, the experiments $\mathbf{Z}_x = (Z, \mathcal{B}_Z, P_\theta^3(\cdot|x))$, $\theta \in \Theta$, and $\mathbf{Z} = (Z, \mathcal{B}_Z, P_\theta^3)$, $\theta \in \Theta$, are equivalent.

Consequently, $\mathcal{U}(\mathbf{Y}_x|x) = EVSI$ associated with $\mathbf{Y}$ for the prior distribution ξ_x, and $\mathcal{U}(\mathbf{Z}_x|x) = EVSI$ associated with $\mathbf{Z}$ for the prior distribution ξ_x, so that under the assumptions in this theorem, $\mathcal{U}(\mathbf{Y}_x|x) \geq^* \mathcal{U}(\mathbf{Z}_x|x)$, whence $\mathcal{U}(\mathbf{Y}|\mathbf{X}) \geq^* \mathcal{U}(\mathbf{Z}|\mathbf{X})$, and hence $\mathcal{U}(\mathbf{X} \times \mathbf{Y}) \geq^* \mathcal{U}(\mathbf{X} \times \mathbf{Z})$. ❑

In the following properties we discuss the effects of *grouping of experimental observations* on the *EVSI*.

Thus, the grouping of experimental data through a *partition* of the set of variable (or vector) values, entails a loss of "worth of information", as we formalize in the next result.

Theorem 4.5. Let $\mathbf{X} = (X,\mathcal{B}_X,P_\theta)$, $\theta \in \Theta$. Let $\mathbf{X'} = (X',\mathcal{B}_{X'},P_\theta)$, $\theta \in \Theta$, be a new statistical experiment where $X' = \{B_j, j \in J\}$ (with $B_j \in \mathcal{B}_X$), $\mathcal{B}_{X'}$ is the smallest Borel σ-field on X', and $P_\theta(B_j)$ is given by the Lebesgue-Stieltjes integral $\int_{B_j} d\,P_\theta(x)$, for all $\theta \in \Theta, j \in J$. Then $\mathbf{X} \geq^{\mathcal{U}} \mathbf{X'}$, whatever the prior distribution on Θ may be.

Proof. Indeed, for all $x \in B_j$, we have that $\mathcal{E}[u(a_x^*|\xi_x)] \geq^* \mathcal{E}[u(a_{B_j}^*|\xi_x)]$. Consequently, and due to the properties of the R^* index, and those of fuzzy operations, we can conclude that

$$\mathcal{U}(\mathbf{X}) = \int_X \mathcal{E}[u(a_x^*|\xi_x)]\, d\,P(x) - \mathcal{E}[u(a_0^*|\xi)] \geq^* \int_X \mathcal{E}[u(a_{B_j}^*|\xi_x)]\, d\,P(x) - \mathcal{E}[u(a_0^*|\xi)] = \mathcal{U}(\mathbf{X'})$$

□

Finally, we now analyze the effects of grouping of sampling information from a statistical experiment when this grouping is due to the use of a *statistic*. The result below indicates that given a random sample $\mathbf{X}^{(m)} = (X^m,\mathcal{B}_{X^m},P_\theta{}^m)$, $\theta \in \Theta$, any *statistic* from it, $\mathbf{T}(\mathbf{X}^{(m)})$, (that is, a Borel-measurable function $\mathbf{T}$ from X^m to a subset in a Euclidean space) entails a reduction of the original random sample, involving a loss of "worth of information" about the true state of nature. In particular, if the statistic is a *sufficient statistic* (that is, the conditional distribution of $\mathbf{X}^{(m)}$ given $\mathbf{T}(\mathbf{X}^{(m)}) = t$, does not depend on θ, for almost all $t \in \mathbf{T}(X^m)$), it exhausts all the "worth of information" about the true state that is contained in the original sample $\mathbf{X}^{(m)}$.

Theorem 4.6. Let $\mathbf{X}^{(m)} = (X^m,\mathcal{B}_{X^m},P_\theta{}^m)$, $\theta \in \Theta$, be a random sample of size m from the statistical experiment $\mathbf{X}$ and let $\mathbf{T}(\mathbf{X}^{(m)})$ be a statistic based on that sample. Then $\mathbf{X}^{(m)} \geq^{\mathcal{U}} \mathbf{T}(\mathbf{X}^{(m)})$, whatever the prior distribution on Θ may be.

In particular, if $\mathbf{T}(\mathbf{X}^{(m)})$ is a sufficient statistic, then $\mathbf{X}^{(m)} \sim^{\mathcal{U}} \mathbf{T}(\mathbf{X}^{(m)})$, whatever the prior distribution on Θ may be.

Proof. Indeed, $\mathcal{U}(\mathbf{X}^{(m)}) \geq^* \mathcal{U}(\mathbf{T}(\mathbf{X}^{(m)}))$, whatever the prior distribution on Θ may be, in virtue of Theorem 4.5 and due to the fact that $\mathbf{T}(\mathbf{X}^{(m)})$ determines a partition on X^m (each class of this partition enclosing sample data mapping by $\mathbf{T}$ into the same statistic value).

On the other hand, if $\mathbf{T}(\mathbf{X}^{(m)})$ is a sufficient statistic, for each $t \in \mathbf{T}(X^m)$, we have that the posterior distributions ξ_t and ξ_{x^m} coincide, for all $x^m \in X^m$ such that $\mathbf{T}(x^m) = t$, and hence the corresponding a_t^* and $a_{x^m}^*$ are indifferent. Therefore,

$$\mathcal{U}(\mathbf{X}^{(m)}) = \int_{X^m} \mathcal{E}[u(a_{x^m}^*|\xi_{x^m})]\, d\,P(x^m) - \mathcal{E}[u(a_0^*|\xi)] \sim^*$$

$$\sim^* \int_{X^m} \mathcal{E}[u(a_t^*|\xi_{x^m})]\, dP(x^m) - \mathcal{E}[u(a_0^*|\xi)] = \mathcal{U}(T(\mathbf{X}^{(m)})) \qquad \square$$

We are next going to illustrate the application of the criterion studied in this section.

5. ILLUSTRATIVE EXAMPLE

We will now examine the application of the criterion above suggested by considering the example illustrating the motivation for the present study in the Introduction of this paper.

Example: As we commented in the presentation of the example, the assessment of utilities in that Decision problem seemed to be very precise, because of the nature of the actions and states in the problem. Thus, the following assessment would express better the decision maker (neurologist) "preferences": $u(\theta_1,a_1) = u(\theta_2,a_2) = 0$, $u(\theta_2,a_1)$ = «inconvenient», $u(\theta_1,a_2)$ = «dangerous», where $u(\theta_1,a_2)$ and $u(\theta_2,a_1)$ could be described by means of the fuzzy sets characterized, for instance, by the membership functions in Figure 1.

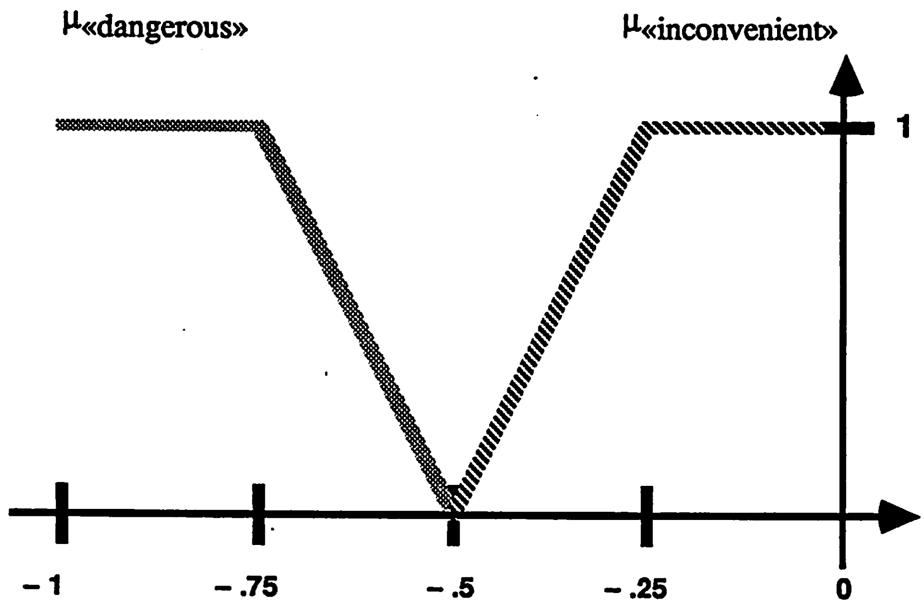

Fig. 1. Membership functions of the fuzzy utilities "inconvenient" ( ) and "dangerous" ( )

The question we were interested in was the following: what is the "most convenient" score the neurologist can base his decision on, **X** or **Y**?.

On the basis of the information from $\mathbf{X}$ we can revise the prior distribution on Θ to obtain the posterior ones. Then, by computing $\mathcal{E}[u(a|\xi_x)]$ for each $a \in A$ and $x \in X = \mathbb{R}$, we conclude that it can be possible to determine $a_x{}^*$ in a generic way for each $x \in \mathbb{R}$, so that $a_x{}^* = a_1$ for $x \geq 110 - 3.2 \log 6.5$, $a_x{}^* = a_2$ otherwise. Consequently, $\mathcal{U}(\mathbf{X}) = .1234\ u(\theta_2,a_1) + .0136\ u(\theta_1,a_2) - .4\ u(\theta_2,a_1)$.

In the same way, on the basis of the information from $\mathbf{Y}$ we can revise the prior distribution on Θ to obtain the posterior ones. Then, by computing $\mathcal{E}[u(a|\xi_y)]$ for each $a \in A$ and $y \in Y = \mathbb{R}$, we conclude that it can be possible to determine $a_y{}^*$ in a generic way for each $y \in \mathbb{R}$, so that $a_y{}^* = a_1$ for $y \geq 110 - 4.1 \log 6.5$, $a_y{}^* = a_2$ otherwise. Consequently, $\mathcal{U}(\mathbf{Y}) = .1575\ u(\theta_2,a_1) + .0596\ u(\theta_1,a_2) - .4\ u(\theta_2,a_1)$.

Therefore, the degree of truth of the proposition «*EVSI* of **X** is not higher than *EVSI* of **Y**» is equal to $R^{\circ}(\mathcal{U}(\mathbf{X}),\mathcal{U}(\mathbf{Y})) = .4811$, so that $\mathbf{X}$ is slightly more informative than $\mathbf{Y}$.

6. CONNECTIONS WITH SUFFICIENCY CRITERION TO COMPARE EXPERIMENTS

The criterion presented in this paper establish a complete preordering among all the statistical experiments associated with the same Decision problem, since under the assumption of a Bayesian framework the *EVSI* is well-defined for all the experiments.

On the contrary, the well-known criterion of comparing experiments based on Blackwell's sufficiency [3, 4] only determines a partial preordering on the set of all statistical experiments associated with the same population. As we have previously remarked, the purpose of this criterion is to get as much probabilistic information about the state or parameter as possible, without having decisions in mind. Thus,

Let $\mathbf{X} = (X,\mathcal{B}_X,P_\theta)$, $\theta \in \Theta$, and $\mathbf{Y} = (Y,\mathcal{B}_Y,Q_\theta)$, $\theta \in \Theta$, be two statistical experiments in which two random variables or vectors are to be observed on the same population.

Definition 6.1. We say that $\mathbf{X}$ is *sufficient for* $\mathbf{Y}$, written $\mathbf{X} \geq^{\mathcal{S}} \mathbf{Y}$, if and only if there exists a nonnegative function h on $X \times Y$, so that the density function associated with Q_θ with respect to a dominating measure ν on $\mathcal{B}_X \times \mathcal{B}_Y$ is given by

$$g_\theta(y) = \int_X h(x,y)\, f_\theta(x)\, d\,\nu(x), \text{ for all } \theta \in \Theta$$

where

$$\int_Y h(x,y)\, d\,\nu(y) = 1, \text{ for all } x \in X$$

and h is integrable with respect to x ($f_\theta(x)$ being the density function associated with P_θ with respect to ν). Since the function h (called stochastic transformation) does not depend on θ, the above sufficiency condition indicates intuitively that an outcome from **Y** could be generated from an observation on **X** and an auxiliary randomization according to h. Consequently, to observe **Y** does not add any information about θ to what is contained in **X**.

By means of the next result we verify that, whenever the comparison through sufficiency is applicable the preference relation in Definition 4.1 and that in Definition 6.1 are coherent, (leading to the same conclusion) although the first one is clearly the most widely applicable (thus, for instance, **X** and **Y** in Section 5 are classical examples of non-comparable experiments *via* sufficiency that are comparable through the preference relation herein proposed). This property means a new guarantee for the suitability of the fuzzy preference relation suggested in this paper.

Theorem 6.1. Let $\mathbf{X} = (X,\mathcal{B}_X,P_\theta)$, $\theta \in \Theta$, and $\mathbf{Y} = (Y,\mathcal{B}_Y,Q_\theta)$, $\theta \in \Theta$, be two statistical experiments in which two random variables or vectors are to be observed on the same population. If $\mathbf{X} \geq^{\mathcal{S}} \mathbf{Y}$, then $\mathbf{X} \geq^{\mathcal{U}} \mathbf{Y}$, whatever the action space, A, the fuzzy utility function on $\Theta \times A$, and the prior distributiont ξ on $(\Theta,\mathcal{C})$, may be.

Proof. Indeed, $\mathcal{E}[u(a_x^*|\xi_x)] \geq^* \mathcal{E}[u(a_y^*|\xi_x)]$. Consequently, and due to the properties of the R^* index, those of fuzzy operations, and those of the stochastic transformation h, if Q is the probability measure associated with h we can conclude that

$$\mathcal{U}(\mathbf{X}) = \int_X \mathcal{E}[u(a_x^*|\xi_x)]\, d\,P(x) - \mathcal{E}[u(a_0^*|\xi)] =$$
$$= \int_Y \int_X \mathcal{E}[u(a_x^*|\xi_x)]\, d\,P(x)\, d\,Q(y/x) - \mathcal{E}[u(a_0^*|\xi)] \geq^*$$
$$\geq^* \int_Y \int_X \mathcal{E}[u(a_y^*|\xi_x)]\, d\,P(x)\, d\,Q(y/x) - \mathcal{E}[u(a_0^*|\xi)] =$$
$$= \int_Y \mathcal{E}[u(a_y^*|\xi_y)]\, d\,Q(y) - \mathcal{E}[u(a_0^*|\xi)] = \mathcal{U}(\mathbf{Y}). \quad \square$$

7. CONCLUDING REMARKS

The study in this paper could be immediately extended to the case in which the prior distribution on the state space was fuzzy. Thus, in order to be able to express the prior available information (non-sample information) in probabilistic terms, most (although not all) Bayesians follow, if necessary, the sujective interpretation of probabilities. The description of these probabilities by means of imprecise propositions (such as, « likely », « improbable », « very likely », and so on), is often more realistic than the numerical one. The Decision-Making problem with fuzzy probabilities and fuzzy utilities, has been examined in previous papers (see, for instance, Freeling [10], Dubois and Prade [8]). We now propose to develop a study similar to the present one by modeling fuzzy utilities through FRV, and using the arithmetic operations on fuzzy probabilities in Jain and Agogino [16].

Fianlly, another immediate extension to be carried out, would be that assuming the presence of fuzziness in sample information, following the ideas in Okuda *et al.* [22], Zadeh [33], Tanaka *et al.* [27], Gil [12, 13, 14], Gil *et al.* [15].

ACKNOWLEDGEMENTS

This research is supported in part by NASA Grant NCC 2-275, AFOSR Grant 89-0084, NSF PYI Grant DMC-84511622 and a Grant from the Spanish MEC. Their financial support is gratefully acknowledged.

REFERENCES

[1] J.M. Adamo, "Fuzzy decision trees," *Fuzzy sets and Systems*, Vol. 4, pp. 1-12, 1980.

[2] R.J. Aumann, "Integrals of set-valued functions," *J. Math. Anal. Appl.*, Vol. 12, pp. 1-12, 1965

[3] D.A. Blackwell, "Comparison of experiments," *Proc. Sec. Berkeley Symp. on Math. Stat. and Probability*,. Berkeley, CA: University California Press, 1951, pp. 93-102.

[4] D.A. Blackwell, "Equivalent Comparison of Experiments," *Ann. Math. Statist.*, Vol. 24, pp. 265-272, 1953.

[5] D.A. Blackwell and M.A. Girshick, *Theory of Games and Statistical Decisions*. New York, NY: Wiley, 1954.

[6] M. H. DeGroot, *Optimal Statistical Decisions*. New York, NY: McGraw-Hill, 1970.

[7] D. Dubois and H. Prade, "Operations on fuzzy numbers," *Int. J. Systems Sci.*, Vol. 9, pp. 613-626, 1978.

[8] D. Dubois and H. Prade, "The use of Fuzzy Numbers in Decision Analysis," *Fuzzy Information and Decision Processes*, M.M. Gupta and E. Sanchez, Eds. Amsterdam: North-Holland, pp. 309-321, 1982.

[9] D. Dubois and H. Prade, "Fuzzy Numbers: An overview," *Analysis of Fuzzy Information*, Vol. I J.C. Bezdek, Ed., Boca Raton: CRC Press, pp. 3-39, 1987.

[10] A.N.S. Freeling, "Fuzzy Sets and Decision Anlysis," *IEEE Trans. on Syst., Man, Cyb.*, Vol. 10, pp. 341-354, 1980.

[11] M.P. García-Carrasco, "Criterios para la Comparación de Experimentos," *Trab. de Est. e Inv. Oper.*, Vol. 29, no. 2, pp. 28-51, 1978.

[12] M.A. Gil, "Fuzziness and Loss of Information in statistical problems," *IEEE Trans. on Syst., Man, Cyb.*, Vol. 17, 1016-1025, 1987.

[13] M.A. Gil, "On the Loss of Information due to Fuzziness in Experimental Observations," *Ann. Inst. Statist. Math.*, Vol. 40, no. 4, pp. 627-639, 1988.

[14] M.A. Gil, "Probabilistic-Possibilistic approach to some Statistical problems with Fuzzy Experimental Observations," *Lecture Notes in Economics, No. 310 – Combining Fuzzy Imprecision with Probabilistic Uncertainty in decision Making*, J. Kacprzyk and M. fedrizzi, Eds., Berlin: Springer-Verlag, pp. 286-306, 1988.

[15] M.A. Gil, N. Corral and P. Gil, "The Minimum Inaccuracy Estimates in χ^2 Tests for Goodness of Fit with Fuzzy Observations," *J. Statist. Plann. Inference*, Vol. 19, pp. 95-115, 1988.

[16] P. Jain and A. Agogino, "Arithmetic operations on Fuzzy Probabilities," Working paper #87-0803-3, Intelligent Systems Research Group-Department of Mechanical Engineering, Berkeley, 1980.

[17] R. Jain, "Decisionmaking in the presence of Fuzzy Variables," *IEEE Trans. on Syst., Man, Cyb.*, Vol. 6, pp. 698-703, 1976.

[18] W. Kolodziejczy, "Orlovsky's concept of Decision-Making with Fuzzy Preference Relation – Further results," *Fuzzy Sets and Systems*, Vol. 19, pp. 11-20, 1986.

[19] E.L. Lehmann, "Comparing Location Experiments," *Ann. Statist.*, Vol. 16, pp. 521-533, 1988.

[20] D.V. Lindley, "On a Measure of the Information provided by an Experiment," *Ann. Math. Statist.*, Vol. 27, pp. 986-1005, 1956.

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

[22] T. Okuda, H. Tanaka and K. Asai, "A formulation of Fuzzy Decision Problems with Fuzzy Information, using Probability Measures of Fuzzy Events," *Inform. Contr.*, Vol. 38, pp. 135-147, 1978.

[23] S.A. Orlovsky, "On formalization of a general fuzzy mathematical problem," *Fuzzy Sets and Systems*, Vol. 3, pp. 311-321, 1980.

[24] M.L. Puri and D.A. Ralescu, "The concept of Normality for Fuzzy Random Variables," *Ann. Probab.*, Vol. 13, no. 4, pp. 1373-1379, 1985.

[25] M.L. Puri and D.A. Ralescu, "Fuzzy Random Variables," *J. Math. Anal. Appl.*, Vol. 114, pp. 409-422, 1986.

[26] H. Raiffa and R. Schlaifer, *Applied Statistical Decision Theory*. Cambridge, MA: MIT Press, 1961.

[27] H. Tanaka, T. Okuda and K. Asai, "Fuzzy Information and Decision in Statistical Model," *Advances in Fuzzy Sets Theory and Applications*, M.M. Gupta, K.K. Ragade and R.R. Yager, Eds., Amsterdam: North-Holland, pp. 303-320, 1979.

[28] S.R. Watson, J.J. Weiss and M.L. Donnell, "Fuzzy Decision Analysis," *IEEE Trans. on Syst., Man, Cyb.*, Vol. 9, pp. 1-9, 1979.

[29] T. Whalen, "Decisionmaking under uncertainty with various assumptions about available information," *IEEE Trans. on Syst., Man, Cyb.*, Vol. 14, pp. 888-900, 1984.

[30] R.J. Wonnacott and T.H. Wonnacott, *Introductory Statistics*. New York, NY: Wiley, 1985.

[31] L.A. Zadeh, "Fuzzy Sets," *Inform. Contr.*, Vol. 8, pp. 338-353, 1965.

[32] L.A. Zadeh, "The concept of a linguistic variable and its application to approximate reasoning," *Inf. Sci.*, Part 1: Vol. 8, pp. 199-249; Part 2: Vol. 8, pp. 301-353; Part 3: Vol. 9, pp. 43-80, 1975

[33] L.A. Zadeh, "Fuzzy Sets as a basis of a Theory of Possibility," *Fuzzy Sets and Systems*, Vol. 1, pp. 3-28, 1978.