CONNECTIONS BETWEEN SOME CRITERIA TO COMPARE FUZZY INFORMATION SYSTEMS

by

P. Gil, M. A. Gil, M. L. Menendez, and L. Pardo

Memorandum No. UCB/ERL M89/105

29 August 1989

CONNECTIONS BETWEEN SOME CRITERIA TO COMPARE FUZZY INFORMATION SYSTEMS

by

P. Gil, M. A. Gil, M. L. Menendez, and L. Pardo

Memorandum No. UCB/ERL M89/105

29 August 1989

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

CONNECTIONS BETWEEN SOME CRITERIA TO COMPARE FUZZY INFORMATION SYSTEMS

P. GIL

Departamento de Matemáticas.Universidad de Oviedo.33071 Oviedo,SPAIN.

M.A. GIL (*)

Departamento de Matemáticas.Universidad de Oviedo.33071 Oviedo,SPAIN.

M.L. MENENDEZ

Departamento de Matemática Aplicada.E.T.S.A..Universidad. Politécnica de Madrid. 28040 Madrid ,SPAIN.

L. PARDO

Departamento de Estadística e Investigación Operativa . Facultad de Matemáticas . Universidad Complutense de Madrid. 28040 Madrid, SPAIN.

SUMMARY

In previous papers we have examined the extension of some methods to compare random experiments to the case in which the available experimental information is fuzzy.

In this paper we are now going to analyze the relationships between the criteria based on Blackwell's sufficiency and some criteria based on well-known information measures

Keywords :Comparison of Experiments; Fuzzy Sets; Information Theory ; Statistics.

1. INTRODUCTION

The theory of comparing two "classic experiments" is usually concerned with the problem of comparing two random experiments whose probability distributions depend on the same state or parameter value, so that the "most informative" (with respect to such a state)is preferred to the other one. The usual purpose of that comparison is that inferences regarding the state will be later based on the sampling from the "best" experiment.

(*) María Angeles Gil is currently a Research Associate at the University of California,Berkeley, for a period of two years.Her research is supported in part by NASA Grant NCC-275,AFOSR Grant 89-0084 and a Grant from the Spanish MEC.

In the traditional criteria suggested to compare experiments one first consider two random experiments, that may be characterized by means of two probability spaces $E \equiv (X, \beta_X, P_\theta)$ and $F \equiv (Y, \beta_Y, Q_\theta)$, $\theta \in \Theta$, where the sample spaces X and Y are subsets of a euclidean space (usually $\mathbb{R}$), β_X and β_Y are the smallest Borel σ-fields on the sample spaces X and Y, respectively, $\{P_\theta, \theta \in \Theta\}$ and $\{Q_\theta, \theta \in \Theta\}$ are families of probability measures on (X, β_X) and (Y, β_Y), respectively. More precisely , it is usually assumed that the "observable events" from E and F may be described by means of ordinary subsets of the samples space X and Y, respectively . In addition, the "elementary observable events" are all the singletons from X and Y, respectively.

In previous papers ,[9-12],[19-23], we have considered the problem of comparing two random experiments, E and F, when the available experimental information on which these conclusions will be based is not exact, but rather it may be described by means of fuzzy events of the spaces X and Y, respectively, [30]. In other words, we have extended well-known criteria to compare experiments when the "previous information" concerning the experimental outcomes involves probabilistic uncertainty due to randomness (formalized through the probability measures P_θ and Q_θ), and the "current available information" after the experimental performance contains fuzzy imprecision.

Thus, for instance, assume that a drug manufacturer has developed a drug that supplies an unknown fraction θ of cured patients . To make posterior inferences about θ, the director of a clinic consider the experiment E consisting in observing the drug effectivenness in a patient drawn at random from the populations of patients in the clinic. This Bernoulli experiment may be characterized in terms of the probability space (X, β_X, P_θ), where $X = \{0,1\}$ (0=non-cured patient , 1=cured patient) , β_X = smallest Borel σ-field on $\{0,1\}$, $P_\theta(0) = 1-\theta$, $P_\theta(1) = \theta$. If the director has not time enough to obtain an exact conclusion about the effectiveness of the drug, but he can only indicate that $\mathscr{C}$= « the patient is more or less cured » , or $\overline{\mathscr{C}}$ = « the patient is more or less not » , then the available experimental information could be easily assimilated with fuzzy events on X.

In the present paper, we are first going to recall previous extended criteria and to introduce new ones. Then, we will discuss the possible connections between the preference relations defining these criteria.

2.- PRELIMINARY CONCEPTS

Consider an experiment $E \equiv (X, \beta_X, P_\theta)$, $\theta \in \Theta$, and suppose that the statistician can base his decision making or drawing of conclusions regarding θ, on the observation of the performance of E . Assume now that the ability to observe the experimental outcome only allows the statistician to assimilate each elementary observable event with fuzzy information , [29], [32], where

Definition 2.1

A fuzzy event $\mathcal{X}$ on X characterized by a Borel-measurable membership function $\mu_{\mathcal{X}}$ from X to [0,1], where $\mu_{\mathcal{X}}(x)$ represents the "grade of compatibility" of x and $\mathcal{X}$, is called *fuzzy information associated with the experiment* E .

As the attention is focussed on the state or parameter value θ governing the distribution of the exact information from E, but the present available information is fuzzy, it would be interesting to obtain the probability of the fuzzy information in terms of θ. For this purpose, we will use the well-Known probabilistic definition stated by Zadeh [31] as follows:

Definition 2.2

The *probability of the fuzzy information* $\mathcal{X}$ induced by the probability measure P_θ on (X, β_X) is given by the value

$$\mathcal{P}_\theta(\mathcal{X}) = \int_X \mu_{\mathcal{X}}(x)\, dP_\theta(x)$$

(the integral being the Lebesgue-Stieltjes integral)

In the definition of measures of information and risk associated with an experiment when the available experimental information is given by grouping of experimental observations, [14], [15], the set of all elementary events associated with the experiment is a classical partition of the sample space. In a similar way, and for the sake of operativeness, we will hereafter assume that the available "elementary observable events" determine a partition of fuzzy sets on the sample space or "fuzzy partition", [1], [26-27], which is called fuzzy information system according to the notion introduced by Tanaka et al. [29] :

Definition 2.3

A *fuzzy information system* (FIS), $\mathscr{X}^*$, *associated with* E is a partition of fuzzy events on X that is, a finite set of fuzzy events on X, satisfying the orthogonality condition

$$\sum_{\mathscr{X} \in \mathscr{X}^*} \mu_{\mathscr{X}}(x) = 1, \text{ for all } x \in X .$$

On the basis of the preceding notions, we are next going to establish same preference relations between two FIS's associated with two experiments whose probability distributions are governed by the same state or parameter value . All of these relations only require the knowledge of the induced probability distributions on the FIS's . Then, we will analyze the connections between such preference relations, and finally apply them to an illustrative example.

3 .- SOME PREFERENCE RELATIONS BETWEEN TWO FUZZY INFORMATION SYSTEMS

Let $\mathscr{X}^*$ and $\mathscr{Y}^*$ two FIS's associated with the experiments $E \equiv (X, \beta_X, P_\theta)$ and $F \equiv (Y, \beta_Y, Q_\theta)$, $\theta \in \Theta$, respectively . Then,

Definition 3.1

We say that $\mathscr{X}^*$ is *sufficient for* $\mathscr{Y}^*$, written $\mathscr{X}^* \geq \mathscr{Y}^*$, if there exists a nonnegative function $h(\mathscr{Y}/\mathscr{X})$ on $\mathscr{X}^* \times \mathscr{Y}^*$, satisfying the relations

$$\text{i) } \sum_{\mathscr{Y} \in \mathscr{Y}^*} h(\mathscr{Y}/\mathscr{X}) = 1 \text{ , for all } \mathscr{X} \in \mathscr{X}^*$$

$$\text{ii) } Q_\theta(\mathscr{Y}) = \sum_{\mathscr{X} \in \mathscr{X}^*} h(\mathscr{Y}/\mathscr{X})\, \mathscr{P}_\theta(\mathscr{X}), \text{ for all } \mathscr{Y} \in \mathscr{Y}^* \text{ , } \theta \in \Theta^*$$

Let now assume that Θ is a subset of the real line, and let β_Θ be the smallest Borel σ-field on Θ, and suppose that for all $\mathscr{X} \in \mathscr{X}^*$ and $\mathscr{Y} \in \mathscr{Y}^*$, $\mathscr{P}_\theta(\mathscr{X})$ and $Q_\theta(\mathscr{Y})$ are Borel-measurable with respect to θ . Let Θ^* denote the class of all prior probability distributions on (Θ, β_Θ). Given two prior distributions $\xi_1, \xi_2 \in \Theta^*$, let

$$\mathscr{P}_1(\mathscr{X}) = \int_\Theta \mathscr{P}_\theta(\mathscr{X})\, d\xi_1(\theta)$$

$$Q_1(\mathscr{Y}) = \int Q_\theta(\mathscr{Y})\, d\xi_1(\theta)$$

$\mathcal{X} \in \mathcal{X}^*$, $\mathcal{Y} \in \mathcal{Y}^*$, $i = 1,2$ (the integrals being the Lebesgue-Stieltjes integrals).

<u>Definition 3.2</u>

We say that $\mathcal{X}^*$ is *not less informative than* (or preferred or indifferent to) $\mathcal{Y}^*$ *in the sense of the non parametric f-divergence*, written $\mathcal{X}^* \geq \mathcal{Y}^*$ (NP f-D), if and only if

$$J^f_{\mathcal{X}^*}(\xi_1,\xi_2) \geq J^f_{\mathcal{Y}^*}(\xi_1,\xi_2) \quad \text{for all } \xi_1,\ \xi_2 \in \Theta^* \qquad (2)$$

where
$$J^f_{\mathcal{X}^*}(\xi_1,\xi_2) = \sum_{\mathcal{X} \in \mathcal{X}^*} \mathcal{P}_2(\mathcal{X})\, f\left(\frac{\mathcal{P}_1(\mathcal{X})}{\mathcal{P}_2(\mathcal{X})}\right) ,$$

f being an arbitrary convex function defined on the interval $(0,+\infty)$ and such that

$f(0) = \lim_{u\to 0^+} f(u)$, $0.f(\frac{0}{0}) = 0$, $0.f(\frac{a}{0}) = \lim_{\varepsilon\to 0^+} \varepsilon\, f(\frac{a}{\varepsilon}) = a \lim_{u\to\infty} \frac{f(u)}{u}$, for all $a \in (0,+\infty)$ (The measure $J^f_{\mathcal{Y}^*}(\xi_1,\xi_2)$ is defined analogously).

If $\xi \in \Theta^*$, let

$$\mathcal{P}(\mathcal{X}) = \int_\Theta \mathcal{P}_\theta(\mathcal{X})\, d\xi(\theta)$$

(The integral being the Lebesgue-Stieltjes integral). Then,

<u>Definition 3.3</u>

We say that $\mathcal{X}^*$ *is not less informative than* $\mathcal{Y}^*$ *in the sense of the non-parametric f-information measure* ,written $\mathcal{X}^* \geq \mathcal{Y}^*$ (NP f-I) , if and only if

$$\int_\Theta J^f_{\mathcal{X}^*}(\theta,\xi)\, d\xi(\theta) \geq \int_\Theta J^f_{\mathcal{Y}^*}(\theta,\xi)\, d\xi(\theta), \text{ for all } \xi \in \Theta^* \quad (3)$$

where

$$J^f_{\mathcal{X}^*}(\theta,\xi) = \sum_{\mathcal{X} \in \mathcal{X}^*} \mathcal{P}(\mathcal{X})\, f\left(\frac{\mathcal{P}_\theta(\mathcal{X})}{\mathcal{P}(\mathcal{X})}\right)$$

(f a function satisfying the conditions indicated in definition 3.2) and $J^f_{\mathcal{X}^*}(\xi_1,\xi_2)$ defined analogously.

If we assume that Θ is either the real line or an open interval on the real line, we can state :

Definition 3.4

We say that $\mathcal{X}^*$ *is not less informative than* $\mathcal{Y}^*$ in the sense of the *parametric f information measure*, written $\mathcal{X}^* \geq \mathcal{Y}^*$ (P f-I) , if and only if

$$I^f_{\mathcal{X}^*}(\theta) \geq I^f_{\mathcal{Y}^*}(\theta) \text{ , for all } \theta \in \Theta \quad (4)$$

where $I^f_{\mathcal{X}^*}(\theta) = \liminf_{(\Delta\theta)^2 \to 0^+} \frac{1}{(\Delta\theta)^2} \sum_{\mathcal{X} \in \mathcal{X}^*} \mathcal{P}_{\theta+\Delta\theta}(\mathcal{X}).f\left[\frac{\mathcal{P}_\theta(\mathcal{X})}{\mathcal{P}_{\theta+\Delta\theta}(\mathcal{X})}\right]$ (if it exists),

and $I^f_{\mathcal{Y}^*}(\theta)$ is defined in a similar way.

Finally, we establish the following criteria:

Definition 3.5

We say that $\mathcal{X}^*$ *is not less informative than* $\mathcal{Y}^*$ *in the sense of Fisher's amount of information* written $\mathcal{X}^* \geq \mathcal{Y}^*$ (F) , if and only if

$$I^F_{\mathcal{X}^*}(\theta) \geq I^F_{\mathcal{Y}^*}(\theta) \text{ , for all } \theta \in \Theta \quad (5)$$

where $I^F_{\mathcal{X}^*}(\theta) = \sum_{\mathcal{X} \in \mathcal{X}^*} \mathcal{P}_\theta(\mathcal{X}) \left[\frac{\partial}{\partial\theta} \log \mathcal{P}_\theta(\mathcal{X})\right]^2$, (if it exists), and

$I^F_{\mathcal{X}^*}(\theta)$ is defined in a similar way.

REMARK:

i)The criterion in Definition 3.1 is an immediate extension of that introduced by Blackwell,[2,3] and exhaustively studied in the literature of comparing experiments (cf.[7],[13],[16-18]Its extension was previously suggested in [20].

ii)The criterion in definition 3.2 was based in the well-known family of f-divergence measures introduced by Csiszár, [4] and extends to the fuzzy case the preference relation in [7] that generalizes the relation based on the Kullback-Leibler directed divergence and suggested in [13].

iii) The criterion in Definition 3.3 extends to the fuzzy case that proposed by Ferentinos and Papaioannou [7], which is a generalization of that in [18] . This extension has been previously examined in [21] for the general case and in [12] for Shannon's information measure.

iv)The criterion in Definition 3.4 extends to the fuzzy case that suggested by Ferentinos and Papaioannou,[7], and is based on a method of constructing parametric measures of information from non-parametric ones they have also proposed [6]. A particularization of this criterion to the non-additive directed divergence of order α has been indicated by Gil in [9].

v)The criterion in Definition 3.5 extends to the fuzzy case that suggested by Stone,[18],and studied by Garcia-Carrasco ,[8] . The extension has been previously developed in [10],where additivity of Fisher's measure makes it very operative when we are concerned with random samples of large size from experiments supplying fuzzy information.

4.- CONNECTIONS BETWEEN THE PREFERENCE RELATIONS

We are now going to investigate the connections between the preference relations we have just established . The following results indicate that, under some regularity conditions, Fisher's criterion is the most widely applicable . Thus,

Theorem 4.1

Let Θ be a subset of the real line and let $\mathcal{X}^*$ and $\mathcal{Y}^*$ be two FIS's associated with experiments whose distributions depend on $\theta \in \Theta$. If $\mathcal{X}^* \geq \mathcal{Y}^*$, there $\mathcal{X}^* \geq \mathcal{Y}^*$ (NP f-D).

Proof.

Indeed, according to Definition 3.1, there exists a nonnegative function $h(\mathcal{Y}/\mathcal{X})$ satisfying conditions i) and ii) of the definition 3.1. Consequently, if f is a function satisfying the conditions in Definition 3.2, a result stated by Csiszár ,[4], guarantee that

$$J^f_{\mathcal{Y}^*}(\xi_1,\xi_2) = \sum_{\mathcal{Y} \in \mathcal{Y}^*} Q_2(\mathcal{Y})\, f\left(\frac{Q_1(\mathcal{Y})}{Q_2(\mathcal{Y})}\right) =$$

$$= \sum_{\mathcal{Y} \in \mathcal{Y}^*} \sum_{\mathcal{X} \in \mathcal{X}^*} h(\mathcal{Y}/\mathcal{X})\, \mathcal{P}_2(\mathcal{X})\, f\left(\frac{\sum_{\mathcal{X}\in\mathcal{X}^*} h(\mathcal{Y}/\mathcal{X})\, \mathcal{P}_1(\mathcal{X})}{\sum_{\mathcal{X}\in\mathcal{X}^*} h(\mathcal{Y}/\mathcal{X})\, \mathcal{P}_2(\mathcal{X})}\right) \leq$$

$$\leq \sum_{\mathcal{X} \in \mathcal{X}^*} \sum_{\mathcal{Y} \in \mathcal{Y}^*} h(\mathcal{Y}/\mathcal{X}) \mathcal{P}_2(\mathcal{X}) f\left[\frac{h(\mathcal{Y}/\mathcal{X}) \mathcal{P}_1(\mathcal{X})}{h(\mathcal{Y}/\mathcal{X}) \mathcal{P}_2(\mathcal{X})} \right] =$$

$$= \sum_{\mathcal{X} \in \mathcal{X}^*} \mathcal{P}_2(\mathcal{X}) f\left[\frac{\mathcal{P}_1(\mathcal{X})}{\mathcal{P}_2(\mathcal{X})} \right] = J^f_{\mathcal{X}^*}(\xi_1, \xi_2)$$

whatever $\xi_1, \xi_2 \in \Theta^*$ may be.

Theorem 4.2

Let Θ be a subset of the real line and let $\mathcal{X}^*$ and $\mathcal{Y}^*$ two FIS's associated with experiments whose distributions depend on $\theta \in \Theta$. If $\mathcal{X}^* \geq \mathcal{Y}^*$ (NP f-D) then , $\mathcal{X}^* \geq \mathcal{Y}^*$ (NP f-I).

Proof.

Indeed, if ξ_1 assigns probability 1 to $\theta = \theta_0$ and $\xi_2 = \xi$, then $\mathcal{X}^* \geq \mathcal{Y}^*$ (NP f-D) implies that

$$J^f_{\mathcal{X}^*}(\theta_0, \xi) \geq J^f_{\mathcal{Y}^*}(\theta_0, \xi) \text{ , for all } \theta_0 \in \Theta \text{ , } \xi \in \Theta^*$$

so that Definition 3.3 holds.

Theorem 4.3

Let Θ be either the real line or an interval of the real line . Assume that $\mathcal{P}_\theta(\mathcal{X})$ and $Q_\theta(\mathcal{X})$ are both twice differentiable with respect to θ and satisfying regularity conditions so that

$$\int_\Theta J^f_{\mathcal{X}^*}(\theta, \xi^{\Delta\theta}) \, d\xi^{\Delta\theta}(\theta) \quad \text{and} \quad \int_\Theta J^f_{\mathcal{Y}^*}(\theta, \xi^{\Delta\theta}) \, d\xi^{\Delta\theta}(\theta)$$

admit double differentiation with respect to $\theta+\Delta\theta$ at $\Delta\theta=0$ for all $\theta\in\Theta$ and every point of Θ is a limit point ($E\xi^{\Delta\theta}$ being prior distribution assigning probability 1/2 to each of θ and $\theta+\Delta\theta$) . If $\mathcal{X}^* \geq \mathcal{Y}^*$ (NP f-I), then $\mathcal{X}^* \geq \mathcal{Y}^*$ (P f-I), for all f such that $f(1)=0$.

Proof.

Indeed,

$$\int_\Theta J^f_{\mathcal{X}^*}(\theta,\xi^{\Delta\theta})\, d\xi^{\Delta\theta}(\theta) =$$

$$= \frac{1}{2} \sum_{\mathcal{X}\in\mathcal{X}^*} \left\{ \left[\frac{1}{2} \mathcal{P}_\theta(\mathcal{X}) + \frac{1}{2} \mathcal{P}_{\theta+\Delta\theta}(\mathcal{X}) \right] f\left[\frac{\mathcal{P}_\theta(\mathcal{X})}{\frac{1}{2} \mathcal{P}_\theta(\mathcal{X}) + \frac{1}{2} \mathcal{P}_{\theta+\Delta\theta}(\mathcal{X})} \right] + \right.$$

$$\left. + \left[\frac{1}{2} \mathcal{P}_\theta(\mathcal{X}) + \frac{1}{2} \mathcal{P}_{\theta+\Delta\theta}(\mathcal{X}) \right] f\left[\frac{\mathcal{P}_{\theta+\Delta\theta}(\mathcal{X})}{\frac{1}{2} \mathcal{P}_\theta(\mathcal{X}) + \frac{1}{2} \mathcal{P}_{\theta+\Delta\theta}(\mathcal{X})} \right] \right\}$$

and the Taylor expansion of the last expression with respect to the argument $\theta+\Delta\theta$ in a neighborhood of θ is equal to

$$f(1) + \frac{f''(1)}{8} I^F_{\mathcal{X}^*}(\theta)\, [\Delta\theta]^2 + O([\Delta\theta]^3)$$

whereas the Taylor expansion of

$$\sum_{\mathcal{X}\in\mathcal{X}^*} \mathcal{P}_{\theta+\Delta\theta}(\mathcal{X}) \cdot f\left[\frac{\mathcal{P}_\theta(\mathcal{X})}{\mathcal{P}_{\theta+\Delta\theta}(\mathcal{X})} \right]$$

with respect to $\theta+\Delta\theta$ in a neighborhood of θ is equal to

$$f(1) + \frac{f''(1)}{2} I^F_{\mathcal{X}^*}(\theta)\, [\Delta\theta]^2 + O([\Delta\theta]^3)$$

Consequently, if $f(1) = 0$, we have

$$J^f_{\mathcal{X}^*}(\theta) = \liminf_{(\Delta\theta)^2\to 0^+} \frac{1}{(\Delta\theta)^2} \left\{ \frac{f''(1)}{2} I^F_{\mathcal{X}^*}(\theta)\, [\Delta\theta]^2 \right\} =$$

$$= \liminf_{(\Delta\theta)^2\to 0^+} \frac{4}{(\Delta\theta)^2} \left\{ \frac{f''(1)}{8} I^F_{\mathcal{X}^*}(\theta)\, [\Delta\theta]^2 \right\} =$$

$$= \liminf_{(\Delta\theta)^2\to 0^+} \frac{4}{(\Delta\theta)^2} \int_\Theta J^f_{\mathcal{X}^*}(\theta,\xi^{\Delta\theta})\, d\xi^{\Delta\theta}(\theta)$$

whence $\mathcal{X}^* \geq \mathcal{Y}^*$ (NP f-I) implies that

$$\frac{4}{(\Delta\theta)^2} \int_\Theta J^f_{\mathcal{X}^*}(\theta,\xi^{\Delta\theta})\, d\xi^{\Delta\theta} \geq \frac{4}{(\Delta\theta)^2} \int_\Theta J^f_{\mathcal{Y}^*}(\theta,\xi^{\Delta\theta})\, d\xi^{\Delta\theta}(\theta), \text{ for all } \theta\in\Theta$$

($\Delta\theta$ in a neighborhood of 0) and hence the relation is preserved by taking limit as $(\Delta\theta)^2 \to 0^+$, so that $J^f_{\mathcal{X}^*}(\theta) \geq J^f_{\mathcal{Y}^*}(\theta)$ for all $\theta \in \Theta$ and f satisfying conditions in Definition 3.2 and $f(1) = 0$.

Theorem 4.4

Under the conditions in theorem 3.3, if $\mathcal{X}^* \geq \mathcal{Y}^*$ (P f-I), then $\mathcal{X}^* \geq \mathcal{Y}^*$ (F).

Proof.

Indeed, the result may be inmediately derived from the relation

$$J^f_{\mathcal{X}^*}(\theta) = \frac{f''(1)}{2} \; I^F_{\mathcal{X}^*}(\theta)$$

obtained in the Taylor expansion of $J^f_{*}(\theta)$ considered above.

5 .- ILLUSTRATIVE EXAMPLE

The following example illustrates the application of the criterion in Definition 3.1 and the existence of a strict preference relation according to such definition entails those for the remainder criteria.

Example

In an immunology process a quarter of a large population of mice received a standard dose of a bacteria determining a character C, whereas a half of the same population received a standard dose of another bacteria determining character D . Consequently, the proportions of mice with characters C and D are respectively 0.25 and 0.5 . Suppose that the proportion θ of mice having both characters is unknown .

On the other hand, assume that the mechanisms of analysis for presence of characters C and D in the population are not quite exact . More precisely, assume that the analysis of each mouse for presence of character C only permits us to distinguish between the fuzzy observations $\mathcal{C}$ ="the mouse seems more or less to have C" and $\bar{\mathcal{C}}$= "the mouse seems more or less not to have C (or have $\bar{C}$)", that the investigator assimilates with the membership functions $\mu_{\mathcal{C}}(C)= 0.75$, $\mu_{\mathcal{C}}(\bar{C})=0.25$, $\mu_{\bar{\mathcal{C}}}(C)= 0.25$, $\mu_{\bar{\mathcal{C}}}(\bar{C})= 0.75$ and the analysis of each mouse for presence of character D only permits us to distinguish between the fuzzy observations $\mathcal{D}$ ="the mouse has D quite sharply" and $\bar{\mathcal{D}}$="the mouse has not D (or has $\bar{D}$) quite sharply", that the investigator assimilates with the membership functions $\mu_{\mathcal{D}}(D)=0.9$, $\mu_{\mathcal{D}}(\bar{D})= 0.1$, $\mu_{\bar{\mathcal{D}}}(D)=0.1$, $\mu_{\bar{\mathcal{D}}}(\bar{D})= 0.9$.

Let X denote the probabilistic information system in which a random individual leading to the fuzzy information $\mathscr{E}$, in the analysis for presence of character C, is observed for presence of character D . Let Y denote the probabilistic information system in which a random individual leading to the fuzzy information $\mathscr{D}$, in the analysis for presence of character D, is observed for presence of character C. Then, the (conditional given $\mathscr{E}$) probabilities associated with X are given by

$$P_\theta(1) = (4\theta + 1)/3 \,, P_\theta(0) = 1- P_\theta(1) = (.2- 4\ \theta) /3$$

(where (X=1) is D, and (X=0) is $\bar{D}$), and the (conditional given $\mathscr{D}$) probabilities associated with Y are given by

$$Q_\theta(1) =(3.2\ \theta + 0.1)/2,\ Q_\theta(0) = 1-Q_\theta(1) = (1.9-3.2\ \theta)/2$$

(where (Y=1) is C, and (Y=0) is $\bar{C}$).

The fuzziness in the available information for the probabilistic information systems X and Y leads, respectively, to the fuzzy information systems $\mathscr{D}^*=\{\mathscr{D},\bar{\mathscr{D}}\}$ and $\mathscr{E}^* = \{\mathscr{E},\bar{\mathscr{E}}\}$, whose probability distribution are given by

$$\mathscr{P}_\theta(\mathscr{D})= (3.2\ \theta + 1.1)/3 \qquad \mathscr{P}_\theta(\bar{\mathscr{D}}) = (1.9 -3.2\ \theta)/3$$

$$\mathscr{P}_\theta(\mathscr{E})= (3.2\ \theta + 1.1) \qquad Q_\theta(\bar{\mathscr{E}}) = (2.9-3.2\ \theta)/4$$

Consequently, the function

$$h(\mathscr{E}/\mathscr{D}) = 3/4 \qquad h(\bar{\mathscr{E}}/\mathscr{D}) = 1/4$$

$$h(\mathscr{E}/\bar{\mathscr{D}}) = 0 \qquad h(\bar{\mathscr{E}}/\bar{\mathscr{D}}) = 1$$

satisfies conditions i) and ii) in definition 3.1, whence $\mathscr{X}^* \geq \mathscr{Y}^*$, and hence the best FIS is that in which individuals with the rarest "character" are observed.

6 .- CONCLUDING REMARKS

The study in the present paper has been carried out on the basis of the model involving the concept of FIS and Zadeh's probabilistic definition . It could be also useful in practice to consider the approach based on the concept of fuzzy random variable [24-25]would be applied to the case in which the previous and the actual information is fuzzy, so that we would be in factinterested in making decisions or drawing statistical conclusions regarding a (fuzzy) state or parameter value.

REFERENCES

[1] J.C. Bezdek , *Pattern Recognition with Fuzzy Objective Function Algorithms* (Plenum Press, New York, 1981).

[2] D.A Blackwell, "Comparison of Experiments", *Proceedings of the Second Berkeley Symposium on Mathemtical Statistics and Probability*

, University of California Press, (1981),93-102.

[3] D.A Blackwell , "Equivalent Comparison of Experiments" , *Annals of Mathematical Statistics* 24, (1953) , 265-272.

[4] I.Csiszár, "Information-type measures of difference of probability distributions and indirect observations", *Studia Scientiarum Mathematicarum Hungarica* 2, (1967) ,299-318.

[5] M.H. DeGroot, *Optimal Statistical Decisions* (McGraw-Hill, New York,1979).

[6] K. Ferentinos and T. Papaioannou ,"New Parametric Measures of Information",*Information and Control* 51,(1981), 193-208.

[7] K. Ferentinos and T. Papaioannou , "Information in Experiments and Sufficiency", *Journal of Statistical Planning and Inference* 6, (1982), 309-317.

[8] M.P. García-Carrasco , "Cantidad de Información de Fisher", *Trabajos de Estadística e Investigación Operativa* 32, (1981),13-42.

[9] M.A.Gil, "On the Loss of information due to fuzziness in experimental observation",*Annals of the Institute of Statistical Mathematics* ,40,4 (1988), 627-639.

[10] M.A.Gil and M.T.López, "The Fisher Information associated with a Fuzzy Information System", *Information Sciences* (accepted for publication 1.989).

[11] M.A.Gil, M.T.López and P.Gil, "Comparison between Fuzzy information Systems", *Kybernetes* 13, (1.984),242-251.

[12] M.A.Gil, M.T. López and P. Gil, "Quantity of Information;Comparison Between Information Systems :1.Nonfuzzy states. 2.Fuzzy States , *Fuzzy Sets and Systems* 15, (1985),65-78, 129-145.

[13] P.K. Goel and M.H. DeGroot,"Comparison of Experiments and Information Measures", *Annals of Statistics* 7,(1979), 1066-1077.

[14] K. Kale, " A note on the Loss of Information due to grouping of observations ", *Biometrika* 51,(1.964), 495-497.

[15] S.Kullback, *Information Theory and Statistics* (John Wiley, New York, 1968).

[16] L. Le Cam , "Sufficiency and Approximate Sufficiency", *Annals of Mathematical Statistics* 35, (1964),1419-1455.

[17] E.L.Lehmann, "Comparing location experiments", *Annals of Statistics* 16, (1988), 521-533.

[18] D.V.Lindley, "On a Measure of the Information Provided by an Experiment ", *Annals of Mathematical Statistics* 27,(1956), 986-1005.

[19] M.L.Menéndez, "Tratamiento de la Informción en Sistemas de Información Difusos", Ph. D. (1986) Thesis , Universidad Politécnica.

[20] M.L.Menéndez, J.A. Pardo and L. Pardo , "Sufficient Fuzzy Information Systems", *Fuzzy Sets and Systems* (accepted for publication),(1989).

[21] L.Pardo, M.L.Menéndez and J.A. Pardo, " The f^*-Divergence as a criterion of comparison between fuzzy information systems", *Kybernetes* 15, (1986),189-194.

[22] L. Pardo, M.L. Menéndez and J.A. Pardo "The Information Energy Gain as a criterion of comparison between Fuzzy Information Systems", *Proceedings Sigth European Meeting on Cybernetics and Systems Research* , Dordreckt, (1986),599-606.

[23] L. Pardo, M.L. Menéndez, and J.A. Pardo, "A Sequential Selection Method of a fixed number of Fuzzy Information Systems based on the Information Energy Gain", *Fuzzy Sets and Systems* 25, 1988, 97-105.

[24] M.L.Puri and D.A. Ralescu, "Fuzzy Random Variables", *Journal of Mathematical Analysis and Applications* 114,(1985) 409-422.

[25] M.L.Puri and D.A. Ralescu, "The concept of Normality for Fuzzy Random Variables", *Annals of Probability* 13, (1986) 1373-1379.

[26] E. Ruspini "A New Approach to clustering", *Information and Control* 15, (1969), 22-32.

[27] E. Ruspini "Numerical Methods for Fuzzy Clustering", *Information Sciences* 2, (1970),319-350.

[28] M.Stone,"Non-equivalent Comparisons of Experiments and their use for Experiments involving Location Parameters",*Annals Mathematical Statistics* 32,(1961), 326-332.

[29] H. Tanaka, T. Okuda and K. Asai, "Fuzzy Information and Decision in Statistical Model", in : *Advances in Fuzzy Sets Theory and Applications* North-Holland, Amsterdam, (1979),303-320.

[30] L.A. Zadeh, "Fuzzy Sets", *Information and Control* 8,(1965), 338-353.

[31] L.A. Zadeh, "Probability Mesures of Fuzzy Events ", *Journal of Mathematical Analysis and Applications* 23, (1968),421-427.

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