SEQUENTIAL OBSERVATION SCHEME BASED ON THE f*-DIVERGENCE

Mª LUISA MENENDEZ
Departamento de Matemáticas
E.T.S.Arquitectura
U. Politécnica de Madrid
28040 Madrid
ESPAÑA

J.A. PARDO
Escuela Universitaria de Estadística
U. Complutense
28015 Madrid
ESPAÑA

ABSTRACT

In this paper we suggest a sequential observation scheme based on the measure of f*-Divergence when the available information about the state is vague. We further study the behaviour of this sequential observation scheme when the state space have only two elements.

Keywords: probabilistic information system, fuzzy information, f*-Divergence. sequential observation scheme based on the f*-Divergence

1. INTRODUCTION

Let S be the state space and let p(s) be the density of a prior probability distribution on S with respect to a σ-finite measure λ on (S, β_S). Assume the existence of a set of probabilistic information systems (or experiments) E. From a probabilistic information system $A \in E$, the exact information x may be obtained by the conditional density p(x/s) with respect a σ-finite measure ν on (X, β_X) when s is the true state of nature.

In this framework, following Csiszar (1), a measure of the amount of information about the state $s \in S$, provided by the exact information x is defined by

$$D_{f^*}(p(s/x),p(s)) = \int_S p(s)f^*(p(s/x)/p(s))\, d\lambda(s)$$

where p(s/x) is the posterior probability density on S, given x , and f* denotes an arbitrary convex function defined on the interval $(0,\infty)$ satisfying the conditions $f^*(0)=\lim_{u\to 0^+} f^*(u)$, $0f^*(0/0)=0$, $f^*(a/0)=\lim_{\varepsilon\to 0^+} f^*(a/\varepsilon) = \lim_{u\to\infty} (f^*(u)/u)$.

This measure is called "f*-Divergence concerning S , provided by the exact information x with prior knowledge p(s)"

Assume that we can only perceive vague information about the state space S, so that the available information may be characterized by an element in the set of fuzzy information systems, E*. A fuzzy information system $\mathcal{X}^*$ on A is a fuzzy partition (orthogonal system) of X by means of fuzzy events. Each fuzzy event $\mathcal{X} \in \mathcal{X}^*$ is called fuzzy information from X. Menéndez (2) establishes the following definition:

Definition 1

Let $\mathcal{X}$ be a fuzzy information on $\mathcal{X}^*$, fuzzy information system on A. The f*-Divergence associated with $\mathcal{X}$, when there is a prior knowledge p(s), is defined by

$$D_{f^*}(\mathcal{P}(s/\mathcal{X}) = \int_S p(s)\ f^*(\mathcal{P}(s/\mathcal{X})/p(s))\ d\lambda(s) \qquad (1)$$

where $\mathcal{P}(s/\mathcal{X}) = \int_X \mu_{\mathcal{X}}(x) f(x/s).p(s)/\mathcal{P}(\mathcal{X})\ d\nu(x)$,

$\mu_{\mathcal{X}}$ being the membership function of the fuzzy event $\mathcal{X}$ and

$$\mathcal{P}(\mathcal{X}) = \int_X \mu_{\mathcal{X}}(x) f(x)\ d\nu(x) .$$

The expression (1) can be viewed as a measure of the amount of information about the state of S provided by the fuzzy information $\mathcal{X} \in \mathcal{X}^*$

In the following paragraph we stablish a sequential observation scheme based on the expression (1).

2. A SEQUENTIAL OBSERVATION SCHEME BASED ON THE f*-DIVERGENCE

Let $A_1, A_2, \ldots, A_n, \ldots$ be probabilistic information systems associated with the same probability space $(X, \beta_X; P_s)_{s \in S}$. Assume that the probabilistic information systems are independent. We shall suppose that at each stage the information about s, is given by a fuzzy event $\mathcal{X}^j$ on X, satisfying $\sum_{\mathcal{X}^j \in \mathcal{X}^*} \mu_{\mathcal{X}^j}(x) = 1$ for each $x \in X$.

Under the hypothesis of independence we may state that $\forall s \in S, \forall n \in N$,

$$\mathcal{P}(\mathcal{X}^1, \ldots, \mathcal{X}^n/s) = \prod_{j=1}^{n} \mathcal{P}(\mathcal{X}^j/s)$$

for all $(\mathcal{X}^1, \ldots, \mathcal{X}^n) \in \mathcal{X}^*_{(n)}$, where $\mathcal{X}^*_{(n)}$ is the fuzzy information system formed by all algebraic products of the fuzzy events $\mathcal{X}^1, \ldots, \mathcal{X}^n \in \mathcal{X}^*$ and is called random simple of size n from the probabilistic information system $A = (X, \beta_X, P_s)_{s \in S}$. If the statistician can sequentially take observations and, at each stage, he must decide, at the sight of the amount of information obtained about S whether to stop or to continue and take the next observation, then the following rule based on the f*-Divergence is defined:

Definition 2

The sequential observation rule which states to stop observing after the fuzzy events $\mathcal{X}^1,\dots,\mathcal{X}^j$ have been observed if

$$D_{f^*}(\mathcal{P}(s/\mathcal{X}^1,\dots,\mathcal{X}^j),p(s)) = \int_S p(s)f^*(\mathcal{P}(s/\mathcal{X}^1,\dots,\mathcal{X}^j)/p(s))\, d\lambda(s) \geq c \quad (2)$$

(c is a constant which depends on the amount of information required, in each particular problem, by the statiscian, according to subjective criteria) and to continue observing if

$$D_{f^*}(\mathcal{P}(s/\mathcal{X}^1,\dots,\mathcal{X}^n),p(s)) \geq c$$

is called sequential observation scheme based on the f*-Divergence (S.O.S-f*-D)

Now we shall study the behaviour of the sequential observation scheme based on the f*-Divergence when $S=\{s_0,s_1\}$. After the fuzzy events $\mathcal{X}^1,\dots,\mathcal{X}^j$ have been observed the expression (2) described as follows:

$$D_{f^*}(\mathcal{P}(s/\mathcal{X}^1,\dots,\mathcal{X}^j),p(s))=f^*(\mathcal{P}(s_0/\mathcal{X}^1,\dots,\mathcal{X}^j)/p(s_0)).p(s_0) +$$

$$+f^*(1-\mathcal{P}(s_0/\mathcal{X}^1,\dots,\mathcal{X}^j)/(1-p(s_0))).(1-p(s_0))$$

Theorem 1

The sequential observation scheme based on the f*-Divergence when $S=\{s_0,s_1\}$ indicates after the fuzzy events $\mathcal{X}^1,\dots,\mathcal{X}^j$ have been observed:

a)To stop observing probabilistic information systems if $\mathcal{P}(s_0/\mathcal{X}^1,\dots,\mathcal{X}^j) \in [0,t_1]$ or $\mathcal{P}(s_0/\mathcal{X}^1,\dots,\mathcal{X}^j) \in [t_2,1]$ when the constant c belongs to the interval $[f^*(1), D_{f^*}(\mathcal{P}(s_0/\mathcal{X}^1,\dots,\mathcal{X}^j)=1;p(s))]$ where t_1 and t_2 satisfy the equation on x

$$c= f^*(x/p(s_0)).p(s_0) + f^*((1-x)/1-p(s_0))).(1-p(s_0))$$

b)To stop observing probabilistic information systems if $\mathcal{P}(s_0/\mathcal{X}^1,\dots,\mathcal{X}^j) \in [0,t_3]$ when the constant c belongs to the interval $[D_{f^*}(\mathcal{P}(s_0/\mathcal{X}^1,\dots,\mathcal{X}^j)=1;p(s)), D_{f^*}(\mathcal{P}(s_0/\mathcal{X}^1,\dots,\mathcal{X}^j)=0;p(s))]$ where t_3 satisfies the equation

$$c= f^*(x/p(s_0)).p(s_0) + f^*((1-x)/(1-p(s_0))).$$

Proof:

It is immediate to prove that $D_{f*}(\mathcal{P}(s_o/X^1,\dots,X^j);p(s))$ is a convex function of the distribution $\mathcal{P}(s/X^1,\dots,X^j)$ for a fixed prior distribution $p(s)=(p(s_o),p(s_1))$. The general form of this curve is shown in fig. 1

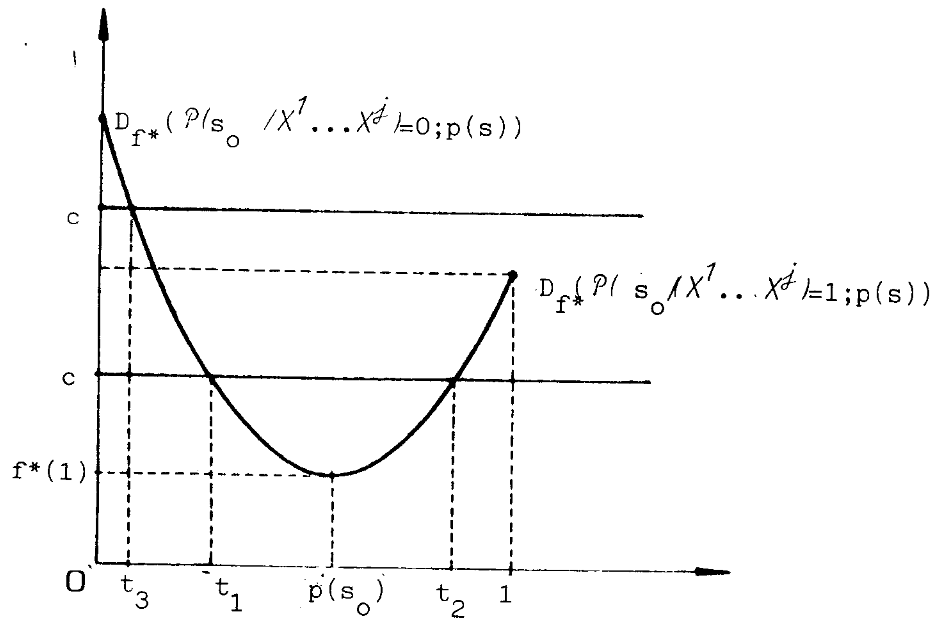

Fig. 1

Assume that $p(s_o)>p(s_1)$. A similar proof can be given when $p(s_o)<p(s_1)$
Let us suppose firstly that the constant c belongs to the interval $[\, f^*(1)\ ,\ D_{f*}(\mathcal{P}(s_o/X^1,\dots,X^j)=1;p(s))\,]$. Since $D_{f*}(\mathcal{P}(s/X^1,\dots X^j;p(s))$ is a convex function in relation to $\mathcal{P}(s_o/X^1,\dots,X^j)$ we will stop observing when $D_{f*}(\mathcal{P}(s_o/X^1,\dots,X^j;p(s))\geq c$, that is, if $\mathcal{P}(s_o/X^1,\dots,X^j)\in[0,t_1]$ or $\mathcal{P}(s_o/X^1,\dots,X^j)\in[t_2,1]$ where t_1 and t_2 verify that

$$f^*(t_1/p(s_o)).p(s_o) + f^*((1-t_1)/(1-p(s_o))).(1-p(s_o)) = c$$

$$f^*(t_2/p(s_o)).p(s_o) + f^*((1-t_2)/(1-p(s_o))).(1-p(s_o)) = c$$

Now we suppose that the constant c belongs to the interval $[D_{f*}(\mathcal{P}(s_o/X^1,\dots,X^j)=1;p(s))\ ,\ D_{f*}(\mathcal{P}(s_o/X^1,\dots,X^j)=0;p(s))]$ .Then we will stop observing when $D_{f*}(\mathcal{P}(s_o/X^1,\dots,X^j);p(s))\geq c$, that is, if $\mathcal{P}(s_o/X^1,\dots,X^j)\in[0,t_3]$ where t_3 verifies that

$$f^*(t_3/p(s_o)).p(s_o) + f^*((1-t_3)/(1-p(s_o))).(1-p(s_o)) = c$$

We now illustrate the sequential observation scheme based on the f*-Divergence by means of an exemple.

3.EXEMPLE

A factory dispose of two types of machines which yield bars of steel. The bars of stell made by machines of the type A have mean breaking strength 50 with standard deviation 10, while the bars of steel made by machines of type B have mean breaking strength 60 with standard deviation 10. The factory has 15 machines of the type A and 10 of the type B. Suppose that, in order to determine the type from which certain lot of bars of steel deriver, we measure the breaking strength of bars in the lot, but the accesible measurement process only provides us with the following fuzzy information:

$\mathcal{X}^1$ = "approximately lower than 45"
$\mathcal{X}^2$ = "approximately 55"
$\mathcal{X}^3$ = "approximately 65"
$\mathcal{X}^4$ = "approximately upper than 75"

to which we associate the following membership functions:

$$\mu_{\mathcal{X}^1}(x) = \begin{cases} 1 & \text{if } x \in (0,45] \\ \frac{55-x}{10} & \text{if } x \in (45,55] \\ 0 & \text{otherwise} \end{cases}$$

$$\mu_{\mathcal{X}^2}(x) = \begin{cases} \frac{x-45}{10} & \text{if } x \in (45,55) \\ \frac{65-x}{10} & \text{if } x \in (55,65) \\ 0 & \text{otherwise} \end{cases}$$

$$\mu_{\mathcal{X}^3}(x) = \begin{cases} \frac{x-55}{10} & \text{if } x \in (55,65] \\ \frac{75-x}{10} & \text{if } x \in (65,75) \\ 0 & \text{otherwise} \end{cases}$$

$$\mu_{\mathcal{X}^4}(x) = \begin{cases} 1 & \text{if } x \geq 75 \\ \frac{x-65}{10} & \text{if } x \in (65,75) \\ 0 & \text{otherwise} \end{cases}$$

Obviously, $\mathcal{X}^1, \mathcal{X}^2, \mathcal{X}^3$ and $\mathcal{X}^4$ make up a fuzzy information system. Assume that $S=\{s_o,s_1\}$ where,

s_o= "The bar of steel has been made by a machine of type A "

s_1= "The bar of steel has been made by a machine of type B "

Obviously $p(s_o)=3/5$ and $p(s_1)=2/5$.

Suppose that the value of the constant c is 0.34 . If at the first stage after the bar of steel has been observed the measurement process provide us with the fuzzy information $\mathcal{X}^1$, using the function $f^*(x)=(1-x)^2$ we have that $D_{f^*}(\mathcal{P}(s_o/\mathcal{X}^1)=0;p(s)) = 3/2$ and $D_{f^*}(\mathcal{P}(s_o/\mathcal{X}^1)=1;p(s)) = 2/3$. Hence c belongs to the interval [0,2/3] where 0=f*(1). Since t_1=0.3143429 ,t_2= = 0.8856571 and $\mathcal{P}(s_o/\mathcal{X}^1)$=0.816577, it follows that $\mathcal{P}(s_o/\mathcal{X}^1) \notin [0,t_1]$ and $\mathcal{P}(s_o/\mathcal{X}^1) \notin [t_2,1]$ and then we continue observing bars of steel.

If at the second stage after the bar of steel has been observed the measurement process provide us with the fuzzy information $\mathcal{X}^3$, it follows that $\mathcal{P}(s_o/\mathcal{X}^1,\mathcal{X}^3)$ = 0.3085162 and hence $\mathcal{P}(s_o/\mathcal{X}^1,\mathcal{X}^3) \in [0,t_1]$, that is, the observation of bars of steel is concluded from theorem 1.

4.REFERENCES

1. I.CSISZAR, "Information type measures difference of probability distributions and indirect observations", Studia Scientarum Mathematicarum Hungarica, 2, p.299-318 (1967)

2. M.R.CASALS, M.A.GIL and P.GIL, "Fuzzy decision and testing hypotheses", Proc. European Journal of Operational Research (1983)

3.M.L.MENENDEZ, "Tratamiento de la información en sistemas de información difusos", Tesis Doctoral, Universidad Politécnica de Madrid (1986)

4. D.V.LINDLEY, "On a measure of information provided by an experiment",Ann. Math. Statist. 27, p.986-1000 (1956)

5.D.V.LINDLEY, "Binomial sampling schemes and the concept of information" , Biometrica 44, p.179-186. (1957)

6. L.PARDO and M.L.MENENDEZ, "Binomial sampling schemes to obtain a prescribed amount of Onicescu's Informational Energy", Proc. 16 th European Meeting of Statisticians, (1984)

7. L.PARDO, "The measure of f*-Divergence as a stoping rule in the sequential random sampling in a bayesian context", Statistica,2, p.243-251 (1986)