

M. A. GIL

J. M. A. GARRIDO

P. GIL

**On a principle of choice among Bayes actions and
its application for comparing experiments**

*Revue française d'automatique, d'informatique et de recherche
opérationnelle. Recherche opérationnelle*, tome 21, n° 3 (1987),
p. 259-279.

http://www.numdam.org/item?id=RO_1987__21_3_259_0

© AFCET, 1987, tous droits réservés.

L'accès aux archives de la revue « Revue française d'automatique, d'informatique et de recherche opérationnelle. Recherche opérationnelle » implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/legal.php>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
<http://www.numdam.org/>

**ON A PRINCIPLE OF CHOICE
AMONG BAYES ACTIONS
AND ITS APPLICATION
FOR COMPARING EXPERIMENTS (*)**

by M. A. GIL, J. M. A. GARRIDO and P. GIL ⁽¹⁾

Abstract. — *We consider a decision problem having a family of potential experiments which the decision maker wishes to compare.*

This purpose is to be achieved by following an extensive-form analysis, in which we take all possible experimental outcomes, choosing for each one an optimal action, and then comparing all pairs of experiments on the basis of that choice.

Both, the principle of choice among actions and the preference relation for comparing experiments, are founded on the well-known concept of "expected value of sample information" and the new concept of "expected quietness of sample information", determining a lexicographical preordering.

The suitability of the preference relation above stated is proven by the study of its main properties and the contrast with other relations for the same purpose.

Keywords : Decision problem; experiments; expected value of sample information; expected quietness of sample information.

Résumé. — *Nous considérons un problème de décision avec une famille d'expériences parmi lesquelles le décideur veut faire une comparaison.*

Ce propos est atteint avec une analyse de forme extensive, où nous choisissons en premier lieu une action pour chaque résultat expérimental et puis, nous comparons chaque couple d'expériences.

Le choix parmi les actions et la relation de préférence parmi les expériences se sont basés sur la notion de « valeur moyenne de l'information d'échantillonnage » et « quiétude moyenne de l'information d'échantillonnage », en déterminant un préordre lexicographique.

L'aptitude du critère exposé est prouvée par le développement de ses propriétés et le parallèle avec autres critères.

Mots clés : Problème de décision; expériences; valeur moyenne de l'information d'échantillonnage; quiétude moyenne de l'information d'échantillonnage.

(*) Reçu October 1985.

(¹) Departamento de Matemáticas, Universidad de Oviedo, 33071 Oviedo, Spain.

1. INTRODUCTION

The statistical decision problem is concerned with the choice among several courses of action, the consequence of which depends upon the state of nature. The available information about the true state usually permits to specify a prior distribution on the state space. In addition, the person responsible for the choice (decision maker) can evaluate the gain of taking each action, when a concrete state is arisen, by defining a utility function.

On the other hand, to obtain further information about the true state the decision maker may select a single experiment from a set of potential experiments, for which the distributions depend upon the state of nature.

Such a selection is accomplished by Raiffa and Schlaifer [19] following an extensive-form analysis in which the optimal actions are elected by using Bayes principle, and then having recourse to this choice to compare the pairs of experiments through the “expected value of sample information (E.V.S.I.)”.

Nevertheless, the comparison above leads in some situations to indifference between the experiments in a pair. To avoid this indifference, a mixed criterion is formulated, Gil M.A. [10]. In this criterion, an optimal action is chosen after an experiment has already been performed by following the Bayes principle and, on the basis of this choice, a preference relation is stated consisting of considering preferred or indifferent the experiment in each pair which first exhibits more E.V.S.I. and then, if the indifference has happened, provides more “expected quietness”.

However, this mixed criterion can be only applied provided that a unique Bayes action exists (with respect to both, the prior and the posterior distribution) or provided that the actions having the same prior (posterior) expected utility involve the same prior (posterior) unquietness.

The aim of the present paper is to leave out the last constraints by defining a new principle of choice. This principle can be described as follows: we first consider the set of Bayes actions with respect to the prior and posterior distributions on the state space, and then select inside an action providing the smallest unquietness. On the basis of that principle we state a preference relation with a similar structure to that of the mixed criterion.

Remark 1. 1: The concept of “expected quietness of sample information”, introduced below, is based on the *unquietness measure*, [9]. The entropy associated with the observation of a random variable is a measure of the uncertainty which is removed by revealing the observed value of the variable. In [21] A. M. Yaglom and I. M. Yaglom emphasize that “...entropy cannot lay claim to take into account all factors determining the uncertainty of an

experiment in every sense in which it may be encountered in real life. Thus, entropy depends only on the probabilities of the possible values of the variable but in no way depends on what these values are, whether they are in a certain sense "close" to or quite "remote" from each other..." (that is, it does not pay attention to the different nature of the variable values). In order to take into account the nature of each of the variable values, we first quantify this nature by means of certain valuation, called *utility*. Then, unquietness measures are introduced as measures of the uncertainty which is motivated by the variety of utilities (that is, by the variety in the nature or quality of the variable values), before revealing the value of the variable.

The measure in [9] belongs to the family of "inset entropies of the Shannon type", [1]. J. Aczél, [1], points out that "... in order to define measures of uncertainty including probabilities and utilities, it is worthwhile looking at these measures as inset entropies". The suitability of the considered measure in quantifying the unquietness associated with a random variable where the possible values are known as probabilities and present different utilities, has been corroborated by an exhaustive study and characterization, [9]. On the other hand, the unquietness measure above coincides with the "additively decomposable index of order 0" for evaluating the income inequality of a population, [5, 6, 7, 20, 22].

2. PRELIMINARY CONCEPTS

The following are the four basic elements in the considered decision problem:

- the *state space* Θ .
- the *action space* \mathbf{A} .
- the *set of potential experiments*, \mathbf{E} . An *experiment* \mathbf{X} belonging to \mathbf{E} can be represented as a probability space $(X, \mathcal{B}_X, P_\theta)$ where (X, \mathcal{B}_X) is a measurable space and the probability measure P_θ belongs to a specified family $\{P_\theta, \theta \in \Theta\}$ (that is, from \mathbf{X} the outcomes in X can be obtained with a distribution depending on the true state).
- a *positive utility evaluation*, $u(\theta, a)$, defined on $\Theta \times \mathbf{A}$. This evaluation quantifies the gain allocated to the consequence of taking each action a in \mathbf{A} , when each particular θ in Θ is the true state of nature. In order to avoid difficulties it is often supposed that $u(\theta, a)$ and $\log u(\theta, a)$ are bounded \mathcal{F} -measurable and \mathcal{B}_X -measurable functions of θ for each a (whatever $\mathbf{X} \in \mathbf{E}$ may be, and being \mathcal{F} a σ -field on Θ). In addition, and in order to deal with

a less restrained framework, we assume that the utility evaluation in the present problem is not necessarily defined satisfying the classical Von Neumann-Morgenstern axiomatic system.

Moreover, we assume the existence of a fifth element:

– a *prior probability measure* $\pi(\theta)$ defined on a measurable space (Θ, \mathcal{F}) , being \mathcal{F} a σ -field of subsets. Let $p(\theta)$ designate the prior density of $\pi(\theta)$ with respect to a σ -finite measure on Θ .

Usually, we suppose that Θ is a subset in a euclidean space and \mathcal{F} is the smallest Borel σ -field on Θ , as well as, for each experiment X in \mathbf{E} , X is a subset in a euclidean space and \mathcal{B}_X is the smallest Borel σ -field on X .

If $f(x; \theta)$ designates the density function of P_θ with respect to a σ -finite measure on X , the conditional measures P_θ along with the prior probability measure π on Θ determine two other probability measures:

– the marginal measure P on X described by a density function $f(x)$, given by:

$$f(x) = \int_{\Theta} f(x; \theta) d\pi(\theta)$$

a. s. (with respect to the σ -finite measure on X) (the integral being according to the Lebesgue-Stieltjes sense).

– the posterior probability measure on Θ given $x \in X$, $\pi_x(\theta)$, described by a density function $\mathbf{p}(\theta/x)$, given by:

$$\mathbf{p}(\theta/x) = \frac{f(x; \theta) \mathbf{p}(\theta)}{f(x)}$$

a. s. (with respect to the σ -finite measure on X).

We now recall certain well-known definitions:

DEFINITION 2.1: The *expected utility of the action* $a \in \mathbf{A}$ is defined as the value:

$$u(a) = \int_{\Theta} u(\theta, a) d\pi(\theta).$$

DEFINITION 2.2: The *expected utility of the action* $a \in \mathbf{A}$ given the outcome $x \in X$ is defined as the value:

$$u(a/x) = \int_{\Theta} u(\theta, a) d\pi_x(\theta).$$

Following a similar structure, we next introduce two new definitions:

DEFINITION 2.3: The *unquietness associated with the action $a \in \mathbf{A}$* is defined as the value:

$$HU^*(a) = - \int_{\Theta} \log \left[\frac{u(\theta, a)}{u(a)} \right] d\pi(\theta)$$

DEFINITION 2.4: The *unquietness associated with the action $a \in \mathbf{A}$ given the outcome $x \in X$* is defined as the value:

$$HU^*(a/x) = - \int_{\Theta} \log \left[\frac{u(\theta, a)}{u(a/x)} \right] d\pi_x(\theta)$$

(In the preceding definitions the integrals are assumed according to the Lebesgue-Stieltjes sense.)

3. PRINCIPLE OF CHOICE

Because of the extensive form adopted to analyze the available data, the essence of the present paper lies first in the principle introduced for the choice of the action to be taken, before, and after a particular outcome has been observed (terminal analysis).

For an experiment $\mathbf{X} = (X, \mathcal{B}_X, P_\theta)$, $\theta \in \Theta$, that principle is described as follows:

If the prior probability measure π on Θ is such that the set of prior Bayes actions, $\mathbf{A}_0 = \{ a_0 \in \mathbf{A} / u(a_0) = \max_{a \in \mathbf{A}} u(a) \}$, is not empty and for almost all $x \in X$ the posterior probability measure π_x on Θ is such that the set of posterior Bayes actions given x , $\mathbf{A}_x = \{ a_x \in \mathbf{A} / u(a_x/x) = \max_{a \in \mathbf{A}} u(a/x) \}$, is not empty, we can state following definitions:

DEFINITION 3.1: An action $a_0^* \in \mathbf{A}_0$ satisfying:

$$HU^*(a_0^*) = \min_{a_0 \in \mathbf{A}_0} HU^*(a_0)$$

is called *optimal prior action*.

DEFINITION 3.2: An action $a_x^* \in A_x$ satisfying:

$$HU^*(a_x^*/x) = \min_{a_x \in A_x} HU^*(a_x/x)$$

is called *optimal posterior action*.

Remark 3.1: The principle above results from a comparison, within each pair of actions, determining a total preordering in A and satisfying the axiom of absolute preferences. On the other hand, it improves the possibility of discriminating actions in relation to the Bayesian principle.

Remark 3.2: From an intuitive viewpoint, the stated principle will be adopted by a decision maker with a “risk aversion” (or inequality aversion, in economics terms) behavior, since the concavity of the unquietness measure with respect to the utilities. Thus, in the present principle the risk aversion is not necessarily described by the shape of the utility evaluation (since it is defined without restrictions), but rather this aversion is incorporated by means of the introduction of the unquietness.

On the basis of definitions 3.1 and 3.2 we now proceed to recall an essential concept and to introduce a new one:

DEFINITION 3.3: The *expected value of sample information* (E.V.S.I.) of a given experiment $X \in E$ is defined as the value, if exists:

$$V(\Theta; X) = \int_x u(a_x^*/x) dP(x) - u(a_0^*) = \int_x [u(a_x^*/x) - u(a_0^*)] dP(x)$$

[when we must specify the prior distribution on Θ , we shall note $V(\Theta; X; \pi)$ instead of $V(\Theta; X)$].

DEFINITION 3.4: The *expected quietness of sample information* of a given experiment $X \in E$ is defined as the value, if exists:

$$\begin{aligned} Q(\Theta; X) &= HU^*(a_0^*) - \int_x HU^*(a_x^*/x) dP(x) \\ &= \int_{\Theta} \left\{ \int_x \log \left[\frac{u(\theta, a_x^*) u(a_0^*)}{u(\theta, a_0^*) u(a_x^*/x)} \right] dP_{\theta}(x) \right\} d\pi(\theta) \end{aligned}$$

(when we must specify the prior distribution on Θ , we shall note $Q(\Theta; X; \pi)$ instead of $Q(\Theta; X)$).

4. COMPARING EXPERIMENTS

We now proceed to define a preference relation in order to select the experiment which is to be performed (preposterior analysis).

In the preceding decision problem, let X and Y be two potential experiments in E :

We say that X is *preferred or indifferent to* Y , written $X \succeq^* Y$, if and only if:

$$V(\Theta; X) > V(\Theta; Y)$$

or:

$$V(\Theta; X) = V(\Theta; Y) \quad \text{and} \quad Q(\Theta; X) \geq Q(\Theta; Y)$$

for the concrete prior distribution on Θ .

We say that X is *indifferent to* Y , written $X \sim^* Y$, if and only if $X \succeq^* Y$ and $Y \succeq^* X$ for the concrete prior distribution on Θ [that is, if and only if $V(\Theta; X) = V(\Theta; Y)$ and $Q(\Theta; X) = Q(\Theta; Y)$ for the concrete prior distribution on Θ].

Remark 4.1: The preference relation above proposed establishes a comparison between two experiments in E with respect to a prior distribution on Θ , and hence it determines a complete preordering on E . Nevertheless, the experiments belonging to E could be compared, in a similar way, without reference to a prior distribution, even though in this case the preference relation would only determine a partial preordering on E .

Remark 4.2: On the other hand, it is worth noting that the exposed comparison [which establishes a lexicographical preordering on the set of all pairs $(V(\Theta; X), Q(\Theta; X))$, where X belongs to E] will be adopted again by a decision maker with a "risk aversion" behaviour.

5. ILLUSTRATIVE EXAMPLE

To illustrate the interest of the present paper we first expose an example in which the E.V.S.I. criterion is unable to compare certain experiments:

An old machine produces pieces so that it behaves as a Bernoulli process with known fraction defective $q_0 = 1 - p_0$. Some experiments with a similar new machine lead to the assignment of a uniform prior distribution on the unit interval, P , to its unknown fraction defective $q = 1 - p$ (and, consequently,

to the unknown p). The sale price of a good piece is $c_1 (> 0)$ if the piece has been produced by the old machine, and $c_2 > c_1$ in the other case. A lot of size N is to be generated and before it is produced a decision maker can either adopt the old machine (action a_1) or adopt the new machine (action a_2). If the decision maker considers the utilities:

$$\begin{aligned} u(p, a_1) &= c_1 N p_0 & \text{for all } p \in [0, 1] \\ u(p, a_2) &= c_2 N p & \text{for all } p \in [0, 1] \end{aligned}$$

we have:

$$u(a_1) = c_1 N p_0 \quad \text{and} \quad u(a_2) = c_2 N / 2$$

and hence $A_0 = \{a_1, a_2\}$.

Assume that in order to obtain further information on p , allowing for a more meticulous action choice, the decision maker can perform one of the following experiments:

Binomial sampling with the new machine (X):

Two pieces are drawn with reposition from an experimental stock in the new machine. This experiment behaves as a Binomial process, and:

$$u(a_1/x) = c_1 N p_0$$

and

$$u(a_2/x) = c_2 N (x + 1) / 4, \quad x = 0, 1, 2$$

with x = number of good pieces among the two drawings.

Pascal sampling with the new machine (Y):

Pieces are drawn with reposition from an experimental stock in the new machine until a good piece has been found. This experiment behaves as a Pascal process, and:

$$u(a_1/y) = c_1 N p_0$$

and

$$u(a_2/y) = 2 c_2 N / (y + 2), \quad y = 1, 2, \dots$$

with y = number of drawings for obtaining the first good piece.

When the values of c_1 , c_2 and p_0 are given, the discrimination between X and Y can be usually accomplished by means of the E.V.S.I. Nevertheless, certain values of c_1 , c_2 and p_0 can occasionally lead to indifference between

such experiments. Thus, if $c_2 = 2c_1p_0$ (being $p_0 > 1/2$), we have

$$a_x^* = a_1 \quad \text{for } x=0, \quad a_x^* = a_2 \quad \text{for } x=2,$$

and

$$a_x^* \in \{a_1, a_2\} = A_x \quad \text{for } x=1.$$

Then:

$$V(\mathbf{P}; \mathbf{X}) = (2c_1Np_0 + 3c_2N/4)/3 - c_1Np_0 = c_2N/12.$$

On the other hand,

$$a_y^* = a_2 \quad \text{for } y=1, \quad a_y^* = a_1 \quad \text{for } y>2,$$

and

$$a_y^* \in \{a_1, a_2\} = A_y \quad \text{for } y=2.$$

Then:

$$V(\mathbf{P}; \mathbf{Y}) = 2c_2N/6 + c_1Np_0/2 - c_1Np_0 = c_2N/12 = V(\mathbf{P}; \mathbf{X}).$$

Therefore, the E.V.S.I. criterion does not allow for a discrimination between the experiments above.

The application of concepts and the criterion in Section 4 is now to be illustrated by breaking the indifference situation which arises for these particular conditions.

Following notations in the preceding section, we have:

$$HU^*(a_1) = 0, \quad HU^*(a_2) = 0.3069$$

and, hence, $a_0^* = a_1$.

In the binomial sampling, for the case that $x=1$, $A_x = \{a_1, a_2\}$. Then, as in such a case $HU^*(a_1/x) = 0$ and $HU^*(a_2/x) = 0.1402$, $x=1$ entails $a_x^* = a_1$, whence:

$$Q(\mathbf{P}; \mathbf{X}) = HU^*(a_0^*) - \sum_{x=0}^2 HU^*(a_x^*/x)/3 = -0.0152.$$

On the other hand, in the Pascal sampling when $y=2$, $A_y = \{a_1, a_2\}$. But, as in such a case $HU^*(a_1/y) = 0$ and $HU^*(a_2/y) = 0.1402$ then, $y=2$ entails

$a_y^* = a_1$, whence:

$$Q(P; Y) = HU^*(a_0^*) - \sum_{y=1}^{\infty} HU^*(a^*/y)/y(y+1) = -0.0473.$$

In this way, from the criterion in Section 4 the binomial sampling of size 2 is regarded as better than the Pascal sampling.

6. PROPERTIES OF THE COMPARISON

In order to investigate the suitability of the preference relation in the last section, it would be desirable to carry out a research following the study for the relation based on the experimental amount of information, developed by Lindley [18].

For this purpose, we first recall certain concepts:

DEFINITION 6.1: An experiment from which the outcomes can be (almost surely) obtained with a distribution irrespective of the state, is called *null experiment*.

DEFINITION 6.2: If $X_1 = (X_1, \mathcal{B}_{X_1}, P_0^1)$ and $X_2 = (X_2, \mathcal{B}_{X_2}, P_0^2)$ are experiments belonging to E , the *combined experiment*, or *sum experiment*, corresponds to an experiment $X_1 \times X_2 = (X_1 \times X_2, \mathcal{B}_{X_1 \times X_2}, P_0)$ belonging to E , where \mathcal{B}_{X_i} is the σ -field over X_i induced from $\mathcal{B}_{X_1 \times X_2}$ by the projection of index i ($i = 1, 2$), and P_0^i is the probability measure on \mathcal{B}_{X_i} from P_0 , ($i = 1, 2$). Particularly, the experiment $X^{(m)} = X \times X \times \dots \times X$ (m times) $= (X^m, \mathcal{B}_{X^m}, P_0^{(m)})$, $m \in \mathbb{N}$, is called *random sample of size m from X* .

DEFINITION 6.3: Let $X_1 \times X_2$ be the combination, or sum, of two experiments X_1 and X_2 , belonging to E . X_1 and X_2 are said to be *independent experiments* if and only if the probability measure on $X_1 \times X_2$, P_0 , is the product measure of the marginal probability measures P_0^1 and P_0^2 , on X_1 and X_2 , respectively.

DEFINITION 6.4.1: Let $X^{(m)}$ be a random sample of size m from X . A Borel-measurable function T from $X^{(m)}$ to a subset in a euclidean space is called a *statistic from $X^{(m)}$* if it does not depend on the state $\theta \in \Theta$. A statistic T from $X^{(m)}$ is said to be *sufficient* for θ (or for the family $\{P_\theta, \theta \in \Theta\}$ on X) if and only if the conditional distribution of $X^{(m)}$, given $T = t$, does not depend on θ for almost all t .

In the assumed framework, the sufficiency is equivalent to the Bayesian sufficiency, where:

Then, $\mathbf{X} \succeq^* \mathbf{N}$, whatever the prior distribution on Θ may be. ■

Theorem 6.1 allows the following interpretation: Any experiment is preferred or indifferent to an experiment which does not provide statistical information pertinent to the state space.

THEOREM 6.2: *If $\mathbf{X}_1 = (X_1, \mathcal{B}_{X_1}, P_\theta^1)$ and $\mathbf{X}_2 = (X_2, \mathcal{B}_{X_2}, P_\theta^2)$, $\theta \in \Theta$, are two experiments belonging to \mathbf{E} , then:*

$$\mathbf{X}_1 \times \mathbf{X}_2 \succeq^* \mathbf{X}_1$$

whatever the prior distribution on Θ may be.

Moreover, if the conditional density $f(x_2/x_1; \theta)$ does not involve θ , a. s., then:

$$\mathbf{X}_1 \times \mathbf{X}_2 \sim^* \mathbf{X}_1$$

whatever the prior distribution on Θ may be.

Proof: If P_1 and P_2 are the “marginal” probability measures on $\mathbf{X}_1 \times \mathbf{X}_2$, \mathbf{X}_1 and \mathbf{X}_2 , respectively, we have:

$$V(\Theta; \mathbf{X}_1 \times \mathbf{X}_2) \geq \iint_{X_1 \times X_2} u(a_{x_1}^*/x_1, x_2) dP(x_1, x_2) - u(a_0^*) = V(\Theta; \mathbf{X}_1)$$

and, when $V(\Theta; \mathbf{X}_1 \times \mathbf{X}_2) = V(\Theta; \mathbf{X}_1)$, then $u(a_{x_1 x_2}^*/x_1, x_2) = u(a_{x_1}^*/x_1, x_2)$ for almost all $(x_1, x_2) \in X_1 \times X_2$, so that $a_{x_1}^* \in \mathbf{A}_{x_1 x_2}$ (a. s.). Hence $V(\Theta; \mathbf{X}_1 \times \mathbf{X}_2) = V(\Theta; \mathbf{X}_1)$ implies that:

$$Q(\Theta; \mathbf{X}_1 \times \mathbf{X}_2) \geq HU^*(a_0^*) - \iint_{X_1 \times X_2} HU^*(a_{x_1}^*/x_1, x_2) dP(x_1, x_2)$$

As $u(a_{x_1}^*/x_1) = \int_{X_2} f(x_2/x_1) u(a_{x_1}^*/x_1, x_2) dx_2$ [being $f(x_2/x_1)$ the conditional density of \mathbf{X}_2 , given $\mathbf{X}_1 = x_1$, with respect to a σ -finite measure denoted by dx_2] the Jensen inequality implies that

$$\int_{X_1} \log u(a_{x_1}^*/x_1) dP_1(x_1) \geq \iint_{X_1 \times X_2} \log u(a_{x_1}^*/x_1, x_2) dP(x_1, x_2)$$

whence:

$$Q(\Theta; \mathbf{X}_1 \times \mathbf{X}_2) \geq Q(\Theta; \mathbf{X}_1)$$

Then, $\mathbf{X}_1 \times \mathbf{X}_2 \succeq^* \mathbf{X}_1$, whatever the prior distribution on Θ may be.

On the other hand, if $f(x_2/x_1; \theta)$ does not involve θ , a. s., the posterior probability measures $\pi_{x_1x_2}$ and π_{x_1} coincide almost surely and consequently, $u(a/x_1, x_2) = u(a/x_1)$, $HU^*(a/x_1, x_2) = HU^*(a/x_1)$, $\forall a \in \mathbf{A}$ and for almost all $(x_1, x_2) \in X_1 \times X_2$, which entails:

$$u(a_{x_1}^*/x_1) = u(a_{x_1x_2}^*/x_1, x_2) \quad (\text{i. e., } a_{x_1}^* \in \mathbf{A}_{x_1x_2}, a_{x_1x_2}^* \in \mathbf{A}_{x_1})$$

for almost all $(x_1, x_2) \in X_1 \times X_2$. Therefore:

$$HU^*(a_{x_1}^*/x_1) \stackrel{\text{a. s.}}{=} HU^*(a_{x_1x_2}^*/x_1, x_2)$$

and, thus, $\mathbf{V}(\Theta; \mathbf{X}_1 \times \mathbf{X}_2) = \mathbf{V}(\Theta; \mathbf{X}_1)$ and $\mathbf{Q}(\Theta; \mathbf{X}_1 \times \mathbf{X}_2) = \mathbf{Q}(\Theta; \mathbf{X}_1)$. Then:

$$\mathbf{X}_1 \times \mathbf{X}_2 \sim^* \mathbf{X}_1$$

whatever the prior distribution on Θ may be. ■

Theorem 6.2 can be explained as follows: it is preferred or indifferent to perform two experiments than to perform one of them. In addition, when one of them cannot add statistical information about the state space to what is contained in the other experiment, it is indifferent to perform both experiments than to perform only the last one.

COROLLARY 6.1: *If $\mathbf{X}^{(m)}$ represents the experiment corresponding to a random sample of size m from an experiment \mathbf{X} belonging to \mathbf{E} , then:*

$$\mathbf{X}^{(m+1)} \succeq^* \mathbf{X}^{(m)}, \quad \forall m \in \mathbb{N}$$

whatever the prior distribution on Θ may be. ■

This corollary suggests: the greater the size of the random sample, the more preferred it is.

THEOREM 6.3: *Let $\mathbf{X}_1 = (X_1, \mathcal{B}_{X_1}, P_\theta^1)$, $\mathbf{X}_2 = (X_2, \mathcal{B}_{X_2}, P_\theta^2)$ and $\mathbf{X}_3 = (X_3, \mathcal{B}_{X_3}, P_\theta^3)$, $\theta \in \Theta$, be three experiments belonging to \mathbf{E} . If \mathbf{X}_2 is independent of both \mathbf{X}_1 and \mathbf{X}_3 , and $\mathbf{X}_1 \succeq^* \mathbf{X}_3$ for all prior distribution on Θ , then:*

$$\mathbf{X}_1 \times \mathbf{X}_2 \succeq^* \mathbf{X}_3 \times \mathbf{X}_2$$

whatever the prior distribution on Θ may be.

Proof: We first establish a supporting lemma:

LEMMA: If $\mathbf{Y}_1 = (Y_1, \mathcal{B}_{Y_1}, P_0^1)$ and $\mathbf{Y}_2 = (Y_2, \mathcal{B}_{Y_2}, P_0^2)$, $\theta \in \Theta$, are two experiments belonging to \mathbf{E} , where \mathbf{Y}_1 and \mathbf{Y}_2 are independent, then:

$$\begin{aligned} \mathbf{V}(\Theta; \mathbf{Y}_1 \times \mathbf{Y}_2; \pi) &= \mathbf{V}(\Theta; \mathbf{Y}_2; \pi) + \int_{Y_2} \mathbf{V}(\Theta; \mathbf{Y}_1; \pi_{y_2}) dP_2(y_2) \\ \mathbf{Q}(\Theta; \mathbf{Y}_1 \times \mathbf{Y}_2; \pi) &= \mathbf{Q}(\Theta; \mathbf{Y}_2; \pi) + \int_{Y_2} \mathbf{Q}(\Theta; \mathbf{Y}_1; \pi_{y_2}) dP_2(y_2) \end{aligned}$$

(being P_2 the marginal probability measure on \mathbf{Y}_2 , determined by P_0^2 and the prior distribution π on Θ , and being π_{y_2} the posterior probability measure on Θ given $y_2 \in Y_2$).

Proof: If P is the marginal probability measure on $\mathbf{Y}_1 \times \mathbf{Y}_2$ determined by the probability measure P_0 on the sum experiment, when π is the prior distribution on Θ , we have:

$$\begin{aligned} \mathbf{V}(\Theta; \mathbf{Y}_1 \times \mathbf{Y}_2; \pi) &= \int_{Y_2} u(a_{y_2}^*/y_2) dP_2(y_2) - u(a_0^*) \\ &\quad + \int_{Y_2} \left[\int_{Y_1} f(y_1/y_2) u(a_{y_1 y_2}^*/y_1, y_2) dy_1 - u(a_{y_2}^*/y_2) \right] dP_2(y_2) \end{aligned}$$

and:

$$\begin{aligned} \mathbf{Q}(\Theta; \mathbf{Y}_1 \times \mathbf{Y}_2; \pi) &= HU^*(a_0^*) - \int_{Y_2} HU^*(a_{y_2}^*/y_2) dP_2(y_2) \\ &\quad + \int_{Y_2} \left[HU^*(a_{y_2}^*/y_2) - \int_{Y_1} f(y_1/y_2) HU^*(a_{y_1 y_2}^*/y_1, y_2) dy_1 \right] dP_2(y_2) \end{aligned}$$

[being $f(y_1/y_2)$ the conditional density of \mathbf{Y}_1 given $y_2 \in Y_2$, with respect to a σ -finite measure denoted by dy_1].

Given $y_2 \in Y_2$, let $\pi^0 = \pi_{y_2}$. Then:

$$\begin{aligned} \int_{Y_1} f(y_1/y_2) u(a_{y_1 y_2}^*/y_1, y_2) dy_1 - u(a_{y_2}^*/y_2) &= \mathbf{V}(\Theta; \mathbf{Y}_1; \pi^0) \\ HU^*(a_{y_2}^*/y_2) - \int_{Y_1} f(y_1/y_2) HU^*(a_{y_1 y_2}^*/y_1, y_2) dy_1 &= \mathbf{Q}(\Theta; \mathbf{Y}_1; \pi^0) \end{aligned}$$

which proves the lemma.

We turn now to the proof of Theorem 6.3. Applying the lemma above, the assumptions in the theorem yields:

$$V(\Theta; X_1 \times X_2; \pi) = V(\Theta; X_2; \pi) + \int_{X_2} V(\Theta; X_1; \pi_{x_2}) dP_2(x_2)$$

$$V(\Theta; X_3 \times X_2; \pi) = V(\Theta; X_2; \pi) + \int_{X_2} V(\Theta; X_3; \pi_{x_2}) dP_2(x_2)$$

$$Q(\Theta; X_1 \times X_2; \pi) = Q(\Theta; X_2; \pi) + \int_{X_2} Q(\Theta; X_1; \pi_{x_2}) dP_2(x_2)$$

$$Q(\Theta; X_3 \times X_2; \pi) = Q(\Theta; X_2; \pi) + \int_{X_2} Q(\Theta; X_3; \pi_{x_2}) dP_2(x_2)$$

(being P_2 the marginal probability measure on X_2 determined by P_θ^2 , when π is the prior distribution Θ , and being π_{x_2} the posterior distribution on Θ given $x_2 \in X_2$).

From $X_1 \succcurlyeq^* X_3$, for all prior distribution on Θ , two situations can be arisen:

– or there is a non null set in X_2 such that $V(\Theta; X_1; \pi_{x_2}) > V(\Theta; X_3; \pi_{x_2})$ holds when x_2 is any element in that set, and $V(\Theta; X_1; \pi_{x_2}) = V(\Theta; X_3; \pi_{x_2})$ otherwise, which entails:

$$V(\Theta; X_1 \times X_2; \pi) > V(\Theta; X_3 \times X_2; \pi)$$

– or else, $V(\Theta; X_1; \pi_{x_2}) = V(\Theta; X_3; \pi_{x_2})$ holds for almost all $x_2 \in X_2$, and then $Q(\Theta; X_1; \pi_{x_2}) \geq Q(\Theta; X_3; \pi_{x_2})$ for almost all $x_2 \in X_2$, which implies:

$$V(\Theta; X_1 \times X_2; \pi) = V(\Theta; X_3 \times X_2; \pi)$$

$$Q(\Theta; X_1 \times X_2; \pi) \geq Q(\Theta; X_3 \times X_2; \pi)$$

Then, $X_1 \times X_2 \succcurlyeq^* X_3 \times X_2$, whatever the prior distribution on Θ may be. ■

We interpret Theorem 6.3 as saying that the preference relation between two experiments is preserved when each of them is combined with an experiment independent of both of them.

COROLLARY 6.2: *Let X_1, X_2, X_3 and X_4 be four experiments belonging to E . If $X_1 \succcurlyeq^* X_2, X_3 \succcurlyeq^* X_4, X_1$ is independent of X_3 , and X_2 of X_4 , then:*

$$X_1 \times X_3 \succcurlyeq^* X_2 \times X_4$$

whatever the prior distribution on Θ may be.

Proof: Consider four experiments Y_1, Y_2, Y_3 and Y_4 where, for any $\theta \in \Theta$, the experiment Y_i has the same distribution as $X_i (i=1, 2, 3, 4)$ being Y_1 independent of Y_3, Y_2 of Y_4 and Y_2 of Y_3 . Then, it follows from Theorem 6.3 that:

$$X_1 \times X_3 \sim *Y_1 \times Y_3 \gtrsim *Y_2 \times Y_3 \gtrsim *Y_2 \times Y_4 \sim *X_2 \times X_4$$

whatever the prior distribution on Θ may be. ■

THEOREM 6.4: Let $X=(X, \mathcal{B}_X, P_\theta), \theta \in \Theta$, be an experiment belonging to \mathbf{E} , and let $Y=(Y, \mathcal{B}_Y, P'_\theta)$ be an experiment belonging to \mathbf{E} where Y is a partition of X by sets of the σ -field $\mathcal{B}_X, Y=\{B_i, i \in \mathbf{I}\}, \mathcal{B}_Y$ is the smallest Borel σ -field on Y , and $P'_\theta(B_i)=\int_{B_i} dP_\theta(x), \forall \theta \in \Theta, \forall i \in \mathbf{I}$. Then:

$$X \gtrsim *Y$$

whatever the prior distribution on Θ may be.

Proof: If P is the marginal probability measure on X , when π is the prior distribution on Θ :

$$V(\Theta; X) - V(\Theta; Y) = \sum_{i \in \mathbf{I}} \left\{ \int_{B_i} u(a_{x_i}^*/x_i) dP(x_i) - P(B_i) u(a_{B_i}^*/B_i) \right\} \geq 0$$

On the other hand, if $V(\Theta; X) = V(\Theta; Y)$, then $u(a_{x_i}^*/x_i) = u(a_{B_i}^*/B_i)$ for almost all $x_i \in B_i$ and $i \in \mathbf{I}$, whence $a_{B_i}^* \in \mathbf{A}_{x_i}$, and hence:

$$\text{a. s.} \\ HU^*(a_{B_i}^*/x_i) \geq HU^*(a_{x_i}^*/x_i)$$

Consequently:

$$Q(\Theta; X) - Q(\Theta; Y) = \sum_{i \in \mathbf{I}} \int_{B_i} [HU^*(a_{B_i}^*/x_i) - HU^*(a_{x_i}^*/x_i)] dP(x_i) \geq 0$$

Then, $X \gtrsim *Y$, whatever the prior distribution on Θ may be. ■

In Theorem 6.4 we have shown that any experiment obtained from another one by the “condensation” consisting of a partition is less preferred than the last one.

THEOREM 6.5: If $X^{(m)}$ is the experiment corresponding to a random sample of size m from the experiment X , belonging to \mathbf{E} , and T is a statistic from

$\mathbf{X}^{(m)}$, then:

$$\mathbf{X}^{(m)} \succcurlyeq^* \mathbf{T}$$

whatever the prior distribution on Θ may be.

In addition, if \mathbf{T} is sufficient for θ , then:

$$\mathbf{X}^{(m)} \sim^* \mathbf{T}$$

whatever the prior distribution on Θ may be.

Proof: Since a statistic \mathbf{T} from $\mathbf{X}^{(m)} = (X^m, \mathcal{B}_{X^m}, P_\theta^{(m)})$ determines a partition on X , by an argument similar to that followed in Theorem 6.4 (by taking $\int_{\mathbf{T}(X^m)}$ instead of $\sum_{t \in \mathbf{I}}$) the first part in Theorem 6.5 can be proven.

On the other hand, if \mathbf{T} is sufficient for θ , the Bayesian definition implies that, for almost all $t \in \mathbf{T}(X^m)$:

$$\begin{aligned} u(a/z) &= u(a/t) \\ HU^*(a/z) &= HU^*(a/t), \end{aligned}$$

$\forall a \in \mathbf{A}$ and for almost all $z \in X^m$ such that $\mathbf{T}(z) = t$.

Hence:

$$\begin{aligned} u(a_z^*/z) &= u(a_t^*/t) \\ HU^*(a_z^*/z) &= HU^*(a_t^*/t), \end{aligned}$$

for almost all $z \in X^m$ such that $\mathbf{T}(z) = t$.

Then, $\mathbf{V}(\Theta; \mathbf{X}^{(m)}) = \mathbf{V}(\Theta; \mathbf{T})$ and $\mathbf{Q}(\Theta; \mathbf{X}^{(m)}) = \mathbf{Q}(\Theta; \mathbf{T})$ whatever the prior distribution on Θ may be. ■

Theorem 6.5 can be interpreted as follows: any statistic entails a reduction of the original random sample, implying a loss of information about the state space, so that this random sample is preferred or indifferent to the statistic. In addition, any sufficient statistic exhausts all the information about the state space that is contained in the original random sample, so that this random sample is indifferent to the sufficient statistic.

COROLLARY 6.3: *Let \mathbf{R} and \mathbf{T} be two sufficient statistics for θ , from the random samples $\mathbf{X}^{(m)}$ and $\mathbf{Y}^{(m')}$, respectively (\mathbf{X}, \mathbf{Y} experiments belonging to \mathbf{E}).*

Then:

$$\mathbf{X}^{(m)} \succsim^* \mathbf{Y}^{(m')} \quad \text{iff} \quad \mathbf{R} \succsim^* \mathbf{T}$$

whatever the prior distribution on Θ may be. ■

7. CONTRAST WITH PREVIOUS CRITERIA FOR COMPARING EXPERIMENTS

The properties in the preceding section have been also verified for previous criteria in the comparison of experiments. Thus, the criteria given by Blackwell [2], Bohnenblust, Shapley and Shermann [2], Lindley [18], Lehmann [17], and Raiffa and Schlaifer [19], satisfy similar properties. In this sense, a parallel study for the five criteria was developed in [8].

On the other hand, the contrast of the criterion in Section 4 with the previous ones leads to the following conclusions (whose proofs are similar to those in [10]):

THEOREM 7.1: *Let \mathbf{X} and \mathbf{Y} be two potential experiments such that \mathbf{X} is preferred or indifferent to \mathbf{Y} according to the Lehmann's criterion. Then, \mathbf{X} is preferred or indifferent to \mathbf{Y} according to the criterion in Section 4, whatever the prior distribution on the state space may be. ■*

THEOREM 7.2: *Let \mathbf{X} and \mathbf{Y} be two potential experiments. Assume that the parametric family of probability measures associated with \mathbf{X} is a complete class. If \mathbf{X} is preferred or indifferent to \mathbf{Y} according to the Blackwell's criterion, then \mathbf{X} is preferred or indifferent to \mathbf{Y} according to the criterion in Section 4, whatever the prior distribution on the state space and the action space may be. ■*

THEOREM 7.3: *Let \mathbf{X} and \mathbf{Y} be two potential experiments such that \mathbf{X} is preferred or indifferent to \mathbf{Y} according to the criterion in Section 4 for a prior distribution on the state space. Then, \mathbf{X} is preferred or indifferent to \mathbf{Y} according to the criterion of Raiffa and Schlaifer for the same prior distribution on the state space. ■*

THEOREM 7.4: *Let \mathbf{X} and \mathbf{Y} be two potential experiments associated with a finite state space and such that \mathbf{X} is preferred or indifferent to \mathbf{Y} according to the criterion in Section 4, for the uniform prior distribution on the state space, and for all closed bounded convex action space. Then, \mathbf{X} is preferred or indifferent to \mathbf{Y} according to the criterion of Bohnenblust, Shapley and Shermann. ■*

THEOREM 7.5: *Let X and Y be two potential experiments associated with a finite state space of size n and such that X is preferred or indifferent to Y according to the criterion in Section 4, for the uniform prior distribution on the state space, and for all closed bounded convex action space. If $n=2$, or both, X and Y , are finite or bounded experiments, then X is preferred or indifferent to Y according to the Blackwell's criterion. ■*

THEOREM 7.6: *Let X and Y be two potential experiments associated with a finite state space and such that X is preferred or indifferent to Y according to the criterion in Section 4, for the uniform prior distribution on the state space. Then, X is preferred or indifferent to Y according to the Lindley's criterion for the uniform prior distribution on the state space. ■*

8. CONCLUDING REMARKS

Firstly, it is worth emphasizing that the effectiveness of the unquietness measure in Remark 1.1 will be greater if the utility evaluation has been accomplished by paying attention to the ratios between the utilities (because of its invariance by homotheties with respect to them) and this fact permits us to consider them as positive values. In addition, a similar criterion could be stated by using the variance of the utility scheme for measuring the unquietness (Gil M.A. [12]). This criterion satisfies the properties in Section 6 and is sometimes more operative than the criterion herein proposed (thus, the effectiveness of the new criterion will be greater if the utility evaluation has been accomplished by paying attention to the differences between the utilities). Nevertheless, the asymmetry in the concave function $f(x) = -\ln x$ (in comparison with the symmetry in the concave function $f(x) = x^2$) allows for a major discrimination power in favour of the unquietness measure in Remark 1.1 and, consequently, in favour of the criterion in Section 4. On the other hand, and according to the shape of the utility evaluation and the probability distributions in the problem, we can alternatively make use of some of the unquietness measures of order β (which coincide respectively with the additively decomposable indices of order $1-\beta$ for measuring the income inequality of a population, [5, 6, 7, 20, 22]).

Secondly, from an economic viewpoint, an interesting criterion could be stated by reversing the priority in the lexicographical preordering in both, the principle of choice and the preference relation. Properties in Section 6 would be immediately satisfied by this new procedure. Moreover, instead of these lexicographical preorderings, other natural alternative criteria would be

determined by some sort of weighted combination of the expected utility and the unquietness in the principle of choice, and a similar combination of the expected value and quietness in the preference relation. Particularly, a comparison between two potential experiments have been established in [11] on the basis of the unquietness and expected quietness.

Finally, the optimality principle and the preference relation in this paper may be extended to the fuzzy decision problem, in which the decision maker only perceives fuzzy information from the experimental observation. In such a sense, several methods have been extended (for instance, Gil M.A. *et al.* [13, 14, 15]).

ACKNOWLEDGEMENT

The authors are grateful to Prof. M. Pilar García-Carrasco for her suggestions in connection with this paper.

REFERENCES

1. J. ACZÉL, *A New Theory of Generalized Information Measures, Recent Results in the "Old" Theory and Some "Real Life" Interpretations of Old and New Information Measures*, Proc. Sec. World Conf. on Math. at the service of man, Canarias-Spain, 1982, p. 3.
2. D. BLACKWELL, *Comparison of Experiments*, Proc. Sec. Berkeley Symp. on Math. Stat. and Probability, University of California Press, 1951, pp. 93-102.
3. D. BLACKWELL, *Equivalent Comparisons of Experiments*, Ann. Math. Stat., 24, 1953, pp. 265-272.
4. D. BLACKWELL and M. A. GIRSHICK, *Theory of Games and Statistical Decisions*, J. Wiley, New York, 1954.
5. F. BOURGUIGNON, *Decomposable Income Inequality Measures*, Econometrica, 47, 1979, pp. 901-920.
6. F. A. COWELL, *On the Structure of Additive Inequality Measures*, Review of Economic Studies, 47, 1980, pp. 521-531.
7. W. EICHORN and W. GEHRIG, *Measurement of Inequality in Economics*, Modern applied Mathematics-Optimization and Operations Research, B. KORTE Ed., North Holland, Amsterdam, 1982.
8. M. P. GARCÍA-CARRASCO, *Criterios para la comparación de experimentos*, Trab. de Est. e Inv. Oper., Vol. 29, No. 2, 1978, pp. 28-51.
9. M. A. GIL, *Estudio de una medida para la incertidumbre correspondiente a las utilidades*, Trab. de Est. e Inv. Oper., Vol. 32, No. 3, 1981, pp. 45-66.
10. M. A. GIL, *Mixed Criterion of Expected Utility and Quietness (Invariant by Homotheties with Respect to the Utilities)*, Statistica, Anno XLII, 1, 1982, pp. 21-37.

11. M. A. GIL, *Criterion of Maximizing the Expected Quietness (Invariant by Homotheties in Relation to the Utilities)*, R.A.I.R.O., Rech. Opér., Vol. 16, 1982, pp. 319-331.
12. M. A. GIL, *On Two Unquietness Measures for the Finite Experiences*, Proc. Sec. World Conf. on Math. at the service of man, Canarias-Spain, 1982, pp. 299-304.
13. M. A. GIL and T. BREZMES, *Fuzzified Blackwell's Method to Compare Experiments*, R.A.I.R.O., Rech. Opér., Vol. 19, No. 1, 1985, pp. 105-111.
14. M. A. GIL, M. T. LOPEZ and P. GIL, *Comparison Between Fuzzy Information Systems*, Kybernetes, Vol. 13, 1984, pp. 245-251.
15. M. A. GIL, M. T. LOPEZ and P. GIL, *Quantity of Information; Comparison Between Information Systems: 1. Nonfuzzy States; 2. Fuzzy States*, Fuzzy Sets and Systems, Vol. 15, No. 1, 1985, pp. 65-78; Fuzzy Sets and Systems, Vol. 15, No. 2, 1985, pp. 129-145.
16. P. GIL, *Teoría Matemática de la Información*, I.C.E. Ed, Madrid, 1981.
17. E. L. LEHMANN, *Testing Statistical Hypotheses*, J. Wiley, New York, 1959.
18. D. V. LINDLEY, *On a Measure of the Information Provided by an Experiment*, Ann. Math. Stat., Vol. 27, 1956, pp. 986-1005.
19. H. RAIFFA and R. SCHLAIFER, *Applied Statistical Decision Theory*, The M.I.T. Press, 1961.
20. A. F. SHORROCKS, *The Class of Additively Decomposable Inequality Measure*, Econometrica, 48, 1980, pp. 613-625.
21. A. M. YAGLOM and I. M. YAGLOM, *Probability and Information*, Theory and Decision Library, Vol. 35, D. Reidel Publishing Company, Dordrecht, 1983.
22. D. ZAGIER, *On the Decomposability of the Gini Coefficient and Other Indices of Inequality*, Discussion Paper No 108, Projectgruppe Theoretische Modelle, Universität Bonn, 1983.

Imprimerie GAUTHIER-VILLARS-France

Dépôt légal 1987 : Imprimeur : 3018 — Éditeur : 029 — CPPP 63423

Le Directeur de la Publication : J.-M. BOURGOIS

1173-87

septembre 1987

Imprimé en France