KYBERNETIKA — VOLUME 14 (1978), NUMBER 3

Functional Equations and Information Measures with Preference

N. L. AGGARWAL, C. F. PICARD

In Information Theory and Questionnaire Theory we come across three measures of information: weighted entropy (entropy of Shannon with utility) of Belis and Guiaşu [3], [7], inaccuracy of Kerridge [10], and preference of Questionnaires [11]. These measures, though different, have analogous properties. In this paper we propose to unify these measures and to characterize them through functional equations while modifying certain axioms.

The measure of information of Shannon verifies the well known branching property; the quantitative and qualitative generalization of this measure, given by Belis and Guiaşu with the help of mapping of the events into the reals does not verify a branching type property without a boundary condition. Because the applications of this second measure could be numerous in various sciences, we had to solve the following problem.

To find a new measure of information with the same kind of applications as the measure of Belis and Guiaşu and verifying a constructive property analogous to the branching without boundary conditions. To resolve this problem, we are resolving functional equations giving a measure of information on a probability space with two measures; a second solution generalizes the measure of degree β.

1. INTRODUCTION

Let $(\Omega, \mathscr{S}, \mathsf{P})$ be a probability space.

Let us consider an experiment Π, i.e. a finite measurable partition $\{A_1, \ldots, A_k, \ldots \ldots, A_n\}$ $(n > 1)$ of Ω such that $\mathrm{P}(A_k) > 0$ for every A_k.

The different events A_k depend more or less relevantly upon the experimenter's goal or upon some qualitative characteristic of the physical system taken into account; that is, they have different utilities (or weights). Belis and Guiasu [3], [7] have ascribed to each event A_k a non-negative number U_{A_k} directly proportional to its importance. U_{A_k} is called the utility of the event A_k.

The weighted entropy (or the entropy of Shannon with utility) of the experiment Π

is defined as:

$$H(\Pi) = -\sum_{k=1}^{n} U_{A_k} \mathsf{P}(A_k) \log \mathsf{P}(A_k) \tag{1}$$

where the base of the logarithm is taken as 2. (The same base of the logarithm will be taken throughout this paper.)

The entropy of degree β with utility of the experiment Π is defined as [6]:

$$H_\beta(\Pi) = \frac{\sum_{k=1}^{n} U_{A_k} \mathsf{P}(A_k) \left[1 - \mathsf{P}^{\beta-1}(A_k)\right]}{1 - 2^{1-\beta}} \tag{2}$$

where β is a positive number different from 1.

It is easy to see that

$$\lim_{\beta \to 1} H_\beta(\Pi) = H(\Pi).$$

The utility being not additive we need a law of composition for U_{A_k}. Two laws of composition have been proposed:

L.C.1: The information $I(\mathsf{P}(A_k), U_{A_k})$ supplied by an event A_k is proportional to the utility U_{A_k} and to the information supplied by this event if we neglect its utility, i.e.

$$I(\mathsf{P}(A_k), U_{A_k}) = U_{A_k} \,.\, I(\mathsf{P}(A_k)).$$

L.C.2: The utility of two incompatible events A_k and A_j is the mean value of the utilities of the respective events, i.e.

$$U_{A_k + A_j} = \frac{\mathsf{P}(A_k)\, U_{A_k} + \mathsf{P}(A_j)\, U_{A_j}}{\mathsf{P}(A_k) + \mathsf{P}(A_j)}. \tag{3}$$

Along with other axioms, the weighted entropy is characterized by L.C.1 in [3] where as by L.C.2 in [7].

Applications of the entropies with utility in Questionnaire Theory [4], [11] have shown that what is important for an event A_k it is not U_{A_k} alone but the product $\mathsf{P}(A_k) \,.\, U_{A_k}$, called the *preference* of the event A_k and defined for every event A_k of Π.

If we put

$$V_{A_k} = \mathsf{P}(A_k) \,.\, U_{A_k}$$

we get:

$$H(\Pi) = -\sum_{k=1}^{n} V_{A_k} \log \mathsf{P}(A_k), \tag{4}$$

$$H_\beta(\Pi) = \frac{\sum_{k=1}^{n} V_{A_k}\left[1 - \mathsf{P}^{\beta-1}(A_k)\right]}{1 - 2^{1-\beta}}. \tag{5}$$

These are the inaccuracies of Kerridge and of degree β with the only difference that V_{A_k} and $\mathsf{P}(A_k)$ are functionally related whereas they are defined in an independent way in the usual definitions of inaccuracies.

To give a systematic treatment of these measures of information we define preference V as a positive finite set function on $(\Omega, \mathscr{S})$. Then the modified definitions of entropies with preference are as:

$$H(\Pi) = -\sum_{k=1}^{n} V(A_k) \log \mathsf{P}(A_k), \tag{6}$$

$$H_\beta(\Pi) = \frac{\sum_{k=1}^{n} V(A_k)\left[1 - \mathsf{P}^{\beta-1}(A_k)\right]}{1 - 2^{1-\beta}}, \tag{7}$$

where β is a real number different from 1.

As before, we have

$$\lim_{\beta \to 1} H_\beta(\Pi) = H(\Pi).$$

Let $\Pi' = \{B_1, \ldots, B_j, \ldots, B_m\}$ $(m > 1)$ be another experiment. If we suppose that Π and Π' are independent with respect to P and, moreover, that $V(A_k \cap B_j) = V(A_k) \,.\, V(B_j)$ for $k = 1, 2, \ldots, n$ and $j = 1, 2, \ldots, m$, then we have the following properties:

Weighted Additivity –

$$H(\Pi \wedge \Pi') = V_0\, H(\Pi) + U_0\, H(\Pi'), \tag{8}$$

Weighted β-Additivity –

$$H_\beta(\Pi \wedge \Pi') = V_0\, H(\Pi) + \sum_{k=1}^{n} V(A_k)\, \mathsf{P}^{\beta-1}(A_k)\, H_\beta(\Pi'), \quad \beta \neq 1, \tag{9}$$

where U_0 and V_0 are the preferences of the experiments Π and Π':

$$U_0 = \sum_{k=1}^{n} V(A_k),$$

$$V_0 = \sum_{j=1}^{m} V(B_j).$$

In this paper we propose to study the inverse problem.

We suppose that the entropy with preference of an experiment Π possesses the following property:

Simple Sum Property – There exists a positive function f, continuous on $]0, 1[\times [0, \infty[$ such that

$$H(\Pi) = \sum_{k=1}^{n} f(\mathsf{P}(A_k), \; V(A_k)). \tag{10}$$

Then we wish to find all the functions f and therefore all the entropies possessing property (8) or (9).

If we put

$$p_k = \mathsf{P}(A_k)\,, \quad u_k = V(A_k)\,, \quad k = 1, \ldots, n\,, \quad \sum_{k=1}^{n} u_k = U_0\,,$$

$$q_j = \mathsf{P}(B_j)\,, \quad v_j = V(B_j)\,, \quad j = 1, \ldots, m\,, \quad \sum_{j=1}^{m} v_j = V_0$$

then the properties (8) and (9) can be written in the form of functional equations:

$$(8') \qquad \sum_{k=1}^{n} \sum_{j=1}^{m} f(p_k q_j, u_k v_j) = V_0 \sum_{k=1}^{n} f(p_k, u_k) + U_0 \sum_{j=1}^{m} f(q_j, v_j)\,,$$

$$(9') \qquad \sum_{k=1}^{n} \sum_{j=1}^{m} f(p_k q_j, u_k v_j) = V_0 \sum_{k=1}^{n} f(p_k, u_k) + \sum_{k=1}^{n} u_k p_k^{\beta-1} \sum_{j=1}^{m} f(q_j, v_j)$$

with

$$u_k \geqq 0 \ \forall k\,, \quad v_j \geqq 0 \ \forall j\,, \quad U_0 = \sum_{k=1}^{n} u_k\,, \quad V_0 = \sum_{j=1}^{m} v_j\,;$$

$$p_k > 0 \ \forall k\,, \quad \sum_{k=1}^{n} p_k = 1\,; \quad q_j > 0 \ \forall j\,, \quad \sum_{j=1}^{m} q_j = 1\,.$$

The functional equation (8′) has been studied by Kannappan [8], [9] under quite different conditions.

2. WEIGHTED ENTROPY

In this section we will find all the continuous solutions of the equation (8′) and therefore all the entropies possessing the property (8), under the boundary conditions:

(i) For an experiment $\Pi_0 = \{A_1, A_2\}$ with $\mathsf{P}_0(A_1) = \mathsf{P}_0(A_2) = \frac{1}{2}$ and $U_0(A_1) = U_0(A_2) = \frac{1}{2}$, $H(\Pi_0) = 1$;

(ii) $f(\frac{1}{2}, 1) = 1$;

(iii) $f(p, 0)$ is finite $\forall p \in \,]0, 1[$.

Lemma 1. A necessary and sufficient condition that a continuous positive function f satisfies the equation (8′) under the conditions (i), (ii), (iii) is

$$f(p, u) = -u \log p\,, \quad u \geqq 0\,, \quad p \in \,]0, 1[\,.$$

Proof. The sufficient part is a mere verification.

To prove the necessary part, we note that in writing $V_0 = \sum_{k=1}^{m} v_j$ and $U_0 = \sum_{k=1}^{n} u_k$, the equation (8′) takes the form:

$$\sum_{k=1}^{n} \sum_{j=1}^{m} f(p_k q_j, u_k v_j) = \sum_{k=1}^{n} \sum_{j=1}^{m} v_j f(p_k, u_k) + \sum_{k=1}^{n} \sum_{j=1}^{m} u_k f(q_j, v_j) . \tag{11}$$

This gives

$$f(pq, uv) = v f(p, u) + u f(q, v) \tag{12}$$

for all $p, q \in]0, 1[$ and $u, v \geqq 0$.

$f(p, 0)$ being finite for all $p \in]0, 1[$ we see easily $f(p, 0) = 0$; so we can put

$$f(p, u) = u \, \Phi(p, u) , \quad u > 0 . \tag{13}$$

Then (12) yields: if $uv = 0$ then $f(pq, 0) = 0$, else

$$\Phi(pq, uv) = \Phi(p, u) + \Phi(q, v) . \tag{14}$$

But we know from Aczél ([1], chap. 5) that the most general continuous solution of the equation (15), called 2-variables Cauchy equation

$$F(x_1 + x_2, y_1 + y_2) = F(x_1, y_1) + F(x_2, y_2) \tag{15}$$

is

$$F(x, y) = c_1 x + c_2 y .$$

Making the necessary changes we prove easily that the most general continuous solution of:

$$G(x_1 x_2, y_1 y_2) = G(x_1, y_1) + G(x_2, y_2)$$

is

$$G(x, y) = c_1 \ln x + c_2 \ln y .$$

Finally the most general solution of (14) is:

$$\Phi(p, u) = c_1 \ln p + c_2 \ln u$$

where c_1 and c_2 are two arbitrary constants.

These constants are determined by the conditions (i) and (ii) which imply:

$$\Phi(\tfrac{1}{2}, \tfrac{1}{2}) = 1 ,$$

$$\Phi(\tfrac{1}{2}, 1) = 1 .$$

This gives

$$c_1 = -\frac{1}{\ln 2} \quad \text{and} \quad c_2 = 0 .$$

So we get

$$f(p, u) = -u \log p \quad (u \neq 0)$$

and we note that

$$\lim_{u \to 0} f(p, u) = 0\,, \quad \forall p \in]0, 1[$$

so that the lemma is proved.

Finally we get the result:

Theorem 1. The only measure of information possessing the properties (10) (simple sum) and (8) (weighted additivity) and satisfying the conditions (i), (ii), (iii) is the weighted entropy (6).

Corollary. If V is a probability measure different from P, $H(\Pi)$ is the inaccuracy of Kerridge. If $V \equiv$ P, $H(\Pi)$ is the entropy of Shannon.

3. ENTROPY OF β-TYPE WITH PREFERENCE

In this section we will find the most general continuous solutions of the equation (9′) for all real β ($\beta \neq 1$) and therefore the measures of information possessing the properties (10) and (9) under the boundary conditions (i) and (iii) of the Section 2.

Lemma 2. The most general continuous positive solution f of the functional equation (9′) under the conditions (i) and (iii) is

$$f(p, u) = \frac{u(1 - p^{\beta - 1})}{1 - 2^{1-\beta}}\,, \quad \beta \neq 1\,, \quad u \geqq 0\,, \quad p \in]0, 1[\,.$$

Proof. If we write $V_0 = \sum_{j=1}^{m} v_j$, then the equation (9′) takes the form:

$$\sum_{k=1}^{n} \sum_{j=1}^{m} f(p_k q_j, u_k v_j) = \sum_{k=1}^{n} \sum_{j=1}^{m} v_j f(p_k, u_k) + \sum_{k=1}^{n} \sum_{j=1}^{m} u_k p_k^{\beta-1} f(q_j, v_j)\,.$$

This gives

$$(16) \qquad f(pq, uv) = v f(p, u) + u p^{\beta-1} f(q, v)$$

for all $p, q \in]0, 1[$; $u, v \geqq 0$ and $\beta \in \mathbb{R}$.

If $\beta = 1$, (16) reduces to (12). So we suppose $\beta \neq 1$. Now we will use the method employed in [12] to solve the equation (16). Without any supplementary hypothesis, we can write

$$f(qp, vu) = f(pq, uv)$$

that is

$$f(qp, vu) = u\, f(q, v) + vq^{\beta-1} f(p, u) . \tag{17}$$

Thus we get

$$[v - vq^{\beta-1}] f(p, u) = [u - up^{\beta-1}] f(q, v) \tag{18}$$

with $p, q \in]0, 1[$; $u, v \geqq 0$ and $\beta \neq 1$.

As before, if we put

$$f(p, u) = u(1 - p^{\beta-1})\, \Phi(p, u)$$

we obtain $(uv \neq 0)$

$$\Phi(p, u) = \Phi(q, v) .$$

So Φ is constant c and

$$f(p, u) = c\, u(1 - p^{\beta-1}) \quad (u \neq 0) ;$$

moreover we note that

$$\lim_{u \to 0} f(p, u) = 0 \quad \forall p \in]0, 1[$$

so that the most general continuous solution of (16) is

$$f(p, u) = c\, u(1 - p^{\beta-1}) .$$

The constant c is determined by the condition (i); so we have

$$c = \frac{1}{1 - 2^{1-\beta}} .$$

Therefore

$$f(p, u) = \frac{u(1 - p^{\beta-1})}{1 - 2^{1-\beta}} , \quad \beta \neq 1 .$$

This completes the proof of the Lemma 2.

Theorem 2. The only measure of information possessing the properties (10) (simple sum) and (9) (weighted β-additivity) and satisfying the conditions (i), (iii) is the entropy of degree β with preference, $H_\beta(\Pi)$.

Corollary. If $V = \mathsf{P}$, then $H_\beta(\Pi)$ is the entropy of degree β of Havrda and Charvát.

(Received November 15, 1977.)

[1] J. Aczél: Lectures on Functional Equations and their Applications. Academic Press, New York 1966.

[2] J. Aczél, Z. Daroczy: On Measures of Information and their characterizations. Academic Press, London 1975.

[3] M. Belis, S. Guiaşu: A quantitative-qualitative measure of information in cybernetic systems. IEEE Trans. Information Theory *14* (1968), 593—594.

[4] B. Bouchon: Useful Information and Questionnaires. Information and Control *13* (1976), 368—378.

[5] T. W. Chaundy, J. B. McLeod: On a functional Equation. Proc. Edinburgh Math. Soc. [2] *12* (1960) (Edinburgh Math. Notes 43, 7—8).

[6] H. Emptoz: Information de type β intégrant un concept d'utilité. C. R. Acad. Sci Paris *282A* (1976), 911—914.

[7] S. Giaşu: Weighted Entropy. Reports on Mathematical Physics *2* (1971), 165—179.

[8] Pl. Kannappan: On Shannon's Entropy, Directed Divergence and Inaccuracy. Z. Wahrscheinlichkeitstheorie und verw. Gebiete *22* (1972), 95—100.

[9] Pl. Kannappan: On Directed Divergence and Inaccuracy. Z. Wahrscheinlichkeitstheorie und verw. Gebiete *25* (1972), 49—55.

[10] D. F. Kerridge: Inaccuracy and Inference. J. Roy. Statist. Soc. ser. B *23* (1961), 184—194.

[11] C. F. Picard: Graphes et Questionnaires. Gauthier-Villars, Paris 1972.

[12] C. F. Picard: Propriétés d'additivité de l'information de type β. C. R. Acad. Sci. Paris 282*A* (1976), 915—918.

Prof. N. L. Aggarwal, Laboratoire de Mécanique Théorique, Faculté des Sciences et des Techniques, 25030 Besançon Cedex. France.

Prof. C. F. Picard, Structures de l'Information, Groupe de Recherche du CNRS, associé à l'Université Pierre et Marie Curie, Tour 45, 4 Place Jussieu, 75230 Paris Cedex 05. France.