KYBERNETIKA — VOLUME 10 (1974), NUMBER 6

Information and Entropy of Countable Measurable Partitions* I

M. BEHARA, P. NATH

In ergodic theory, the notions of information and entropy are separated from each other. In the existing literature, it is usual to assume the additive nature of information. In this paper, we have proposed a general definition of information in § 2 and studied its properties extensively in § 3. In § 4, information and entropy of countable measurable partitions of a Lebesgue probability space have been defined.

1. MEASURABLE PARTITIONS

Let (Ω, S, μ) be a Lebesgue probability space, i.e., a measure space which is isomorphic modulo zero to the unit interval $[0, 1]$ with its usual Lebesgue measure. A countable measurable partition $\mathscr{A} = \{A_i\}_{i\in I}$ is a collection of non-empty measurable sets such that

$$\mu(A_i \cap A_j) = 0\,, \quad i \neq j\,, \quad \mu(\Omega - \bigcup_{i\in I} A_i) = 0\,.$$

Two partitions $\mathscr{A}$ and $\mathscr{B}$ are called equivalent: $\mathscr{A} \sim \mathscr{B}$ (mod 0) if $\forall A \in \mathscr{A}$, there exists a $B \in \mathscr{B}$ such that A and B differ by a set of measure zero. A partition $\mathscr{B}$ is called a *refinement* of $\mathscr{A}$, written as $\mathscr{A} \subset \mathscr{B}$, if every element of $\mathscr{A}$ is a disjoint union of elements of $\mathscr{B}$. Obviously, by definition, every partition is equivalent to itself and also a refinement of itself. The sum of two partitions $\mathscr{A}$ and $\mathscr{B}$, written as $\mathscr{A} \vee \mathscr{B}$, is defined as the *least common refinement* of $\mathscr{A}$ and $\mathscr{B}$ i.e.

$$\mathscr{A} \vee \mathscr{B} = \{A_i \cap B_j\}_{i\in I, j\in J}$$

and the definition may be extended to any finite number of partitions. The operation $\vee$ is both commutative and associative. Two partitions $\mathscr{A}$ and $\mathscr{B}$ are called *indepen-*

* Research supported by a grant from the National Research Council of Canada.

dent, if $\mu(A \cap B) = \mu(A)\,\mu(B)$, $A \in \mathscr{A}$, $B \in \mathscr{B}$. By a *σ-field*, we shall mean a collection of measurable sets closed under complementation and countable unions. For a given partition $\mathscr{A}$, the collection of unions of sets from $\mathscr{A}$, together with the null set $\emptyset$, is clearly a σ-field which we shall denote by $\tilde{\mathscr{A}}$. Obviously $\tilde{\Omega} = \{\Omega, \emptyset\}$ and is called the trivial field corresponding to the trivial partition $\{\Omega\}$. The elements of $\mathscr{A}$ are called atoms of $\tilde{\mathscr{A}}$.

Throughout the discussion, all the partitions of Ω will be such that there corresponds to each partition a sub-σ-field of S.

2. INFORMATION

We investigate the following question*: How much information is obtained when we are told that a point $\omega \in \Omega$ belongs to a non-empty subset E of Ω? It is obvious that the answer should depend upon $\mu(E) \in [0, 1]$. Let $F(\mu(E))$ denote the amount of information obtained when we are told that $\omega \in E$. Clearly the domain of F is $(0,1]$.

Definition 2.1. A function F, defined on $(0, 1]$ is called an *information function* if it satisfies the following properties:

(a_1) F is a continuous function of $p \in (0, 1]$.

(a_2) $F(\frac{1}{2}) = 1$, $F(1) = 0$.

(a_3) $F(pq) = \phi(F(p), F(q))$ where ϕ is a polynomial of its arguments.

Theorem 2.1. *Let F be a function defined on $(0, 1]$ and satisfy (a_1), (a_2), and (a_3). Then $F = z_\alpha^*$ where*

$$(2.1) \qquad z_\alpha^*(t) = \frac{1 - t^{\alpha-1}}{1 - 2^{1-\alpha}}, \quad t \in (0, 1], \quad \alpha \neq 1,$$

$$= -\log t, \quad t \in (0, 1], \quad \alpha = 1.$$

Proof. Let

$$F(pq) = \phi(F(p), F(q)) = F(p) \square F(q).$$

Then, it can be easily seen that the operation '$\square$' is commutative and associative. Hence, by following the arguments as in §2.2.2 and §2.2.4 in [1], it follows that ϕ can only be a symmetric polynomial of degree one in each of its arguments. The possibility of ϕ being a constant polynomial is of no use because it leads to the fact that F is

* This question occurs on page 19 in Halmos [6] where instead of (a_3), the additivity of F is assumed.

a constant function which is a contradiction to (a_2). Hence, the only admissible form of ϕ can be

$$\phi(F(p), F(q)) = a\,F(p)\,F(q) + b\,F(p) + b\,F(q) + c \tag{2.2}$$

where $ac = b^2 - b$. From (a_3) and (2.2),

$$F(pq) = a\,F(p)\,F(q) + b\,F(p) + b\,F(q) + c\,. \tag{2.3}$$

Let $a = 0$. Then $b = 1$, and (2.3) reduces to

$$F(pq) = F(p) + F(q) + c\,. \tag{2.4}$$

Putting

$$g(p) = F(p) + c\,, \tag{2.5}$$

(2.4) reduces to Cauchy's functional equation

$$g(pq) = g(p) + g(q)\,. \tag{2.6}$$

By (a_1) and (2.5), g is a continuous function of $p \in (0, 1]$. Hence, the continuous solutions of (2.6) are of the form $g(p) = K \log_2 p$ where K is an arbitrary constant. Obviously, $F(p) = K \log_2 p + c$. By (a_2), $F(\frac{1}{2}) = 1$ and $F(1) = 0$. Hence, $c = 0$, $K = -1$, so that $F = z_1^*$ where

$$z_1^*(p) = \log_2 \frac{1}{p}\,. \tag{2.7}$$

If $a \neq 0$, then the substitution

$$h(p) = a\,F(p) + b \tag{2.8}$$

reduces (2.3) to

$$h(pq) = h(p)\,h(q)\,. \tag{2.9}$$

Again, by (a_1) and (2.8), h is a continuous function of $p \in (0, 1]$. Hence the *non-identically vanishing continuous* solutions of (2.9) are of the form $h(p) = p^{\alpha-1}$, $\alpha \in R$, $R = (-\infty, +\infty)$ so that $F(p) = (p^{\alpha-1} - b)/a$. But, by (a_2), it can be seen that $b = 1$, $a = 2^{1-\alpha} - 1$. Thus $F = z_\alpha^*$ where

$$z_\alpha^*(p) = \frac{1 - p^{\alpha-1}}{1 - 2^{1-\alpha}}\,, \quad p \in (0, 1],\ \alpha \in R,\ \alpha \neq 1. \tag{2.10}$$

Note that $\alpha = 1$ leads to $F(p) = (1 - b)/a$, a situation which is contrary to (a_2). However, it can be seen that $\lim_{\alpha \to 1} z_\alpha^*(p) = z_1^*(p)$.

Remark 1. Whereas assumptions (a_1) and (a_2) are self-explanatory, (a_3) needs some justification. It is customatory to assume additive nature of information, i.e.

$$F(pq) = F(q) + F(q)\,, \quad p \in (0, 1]\,, \quad q \in (0, 1]\,. \tag{2.11}$$

But the R.H.S. in (2.11) is a particular case of (a_3) when $\phi(x, y) = x + y$. Obviously ϕ is a polynomial (symmetric) of degree one in each of its arguments x and y. Since we have not assumed non-negative nature of information, therefore $x \in R$, $y \in R$. But (a_1), (a_2) and (a_3) make F non-negative.

If, the R.H.S. in (2.11) is assumed to be an arbitrary polynomial in $F(p)$ and $F(q)$, it is natural to expect some more measures of information and this is the justification to assume (a_3).

As mentioned above, the operation '$\square$' is both cummutative and associative. Moreover, if $F(1) = e$, then

$$F(p) = F(p) \square F(1) = F(1) \square F(p).$$

Also, from (a_3), the range of ϕ is the same as that of F. Thus, it is clear that the functional equation

$$F(pq) = \phi(F(p), F(q)) \tag{2.12}$$

admits of *at least one non-constant continuous solution* provided the range of F forms a commutative monoid under the operation '$\square$' and the identity of the monoid is $F(1)$. By (a_2), $F(1) = 0$. By Theorem 2.1, $F = z_\alpha^*$ which is non-negative. Thus the range of F is $\bar{R}^+ = \{x : x \geqq 0\}$ and $(\bar{R}^+, \square)$ is a commutative monoid under the operation '$\square$' and with identity 0. On the other hand, if $U = (0, 1]$, then $(U, .)$ is also a commutative monoid under ordinary multiplication '.' and with identity 1. Thus the functions z_α^*, $\alpha \in R$, constitute the set Δ of all mappings from $(U, .)$ into $(\bar{R}^+, \square)$ satisfying (a_1), (a_2) and (a_3).

From (2.1), it is easily seen that

$$z_\alpha^*(pq) = z_\alpha^*(p) + z_\alpha^*(q) + (2^{1-\alpha} - 1)\, z_\alpha^*(p)\, z_\alpha^*(q), \quad \alpha \in R. \tag{2.13}$$

On the other hand

$$z_\alpha^*(pq) = z_\alpha^*(p) + p^{\alpha-1} z_\alpha^*(q) = z_\alpha^*(q) + q^{\alpha-1} z_\alpha^*(p), \quad \alpha \in R, \tag{2.14}$$

Hence, Definition 2.1 is equivalent to the following:

Definition 2.2. A function F, defined on $(0, 1]$ is called an *information function* if it satisfies (a_1), and the following properties:

(a_4) $F(\frac{1}{2}) = 1$,

(a_5) For $p \in (0, 1]$, $q \in (0, 1]$,

$$F(pq) = F(p) + p^{\alpha-1} F(q), \quad \alpha \in R. \tag{2.15}$$

It should be noted that, without making any regularity assumptions on F, the general solutions of (2.15) are of the form

$$F(p) = \lambda(p^{\alpha-1} - 1), \quad \alpha \neq 1, \quad p \in (0, 1],$$

and (a_4) implies $\lambda = 2^{1-\alpha} - 1$ so that $F(p) = z_\alpha^*(p)$, $\alpha \neq 1$.

However, when $\alpha = 1$, (2.15) reduces to

$$F(pq) = F(p) + F(q) \tag{2.16}$$

and (a_1) is needed to ensure continuous solutions of (2.16).

Remark 2. From information-theoretic point of view, $z_\alpha^*(t)$ can be interpreted as the entropy of order α of a generalized singleton distribution $\{t\}$ where $t \in (0, 1]$. The conditions (a_1), (a_2), (a_3) stated in Definition 2.1 are enough to characterize it. From (2.1), it follows that the entropy of a generalized singleton distribution need not be *necessarily additive always*. A detailed study of z_α^* is desirable in order to study the entropy of a non-sigleton probability distribution because according to Renyi [8] a non-singleton distribution with elements $n \geqq 2$ can be written as the union of $n \geqq 2$ singleton probability distributions.

3. THE FUNCTION z_α^*

It is obvious from (a_1) that continuity of F has not been assumed at $p = 0$. Hence, nothing can be said, in general, regarding the continuity of z_α^* at $t = 0$. But, one can see from (2.1) that when $\alpha > 1$, the definition of z_α^* can be extended to $t = 0$ so that the new function will be right continuous at $t = 0$. Hence, whenever $\alpha > 1$, we shall include $t = 0$. From the physical point of view, $z_\alpha^*(0)$ means the amount of information obtained when we are told that a point $\omega \in \Omega$ belongs to a set of measure zero. *Is this amount of information finite or infinite? It seems difficult to decide this question as we notice that* $z_\alpha^*(0) = \infty, 0 \leqq \alpha \leqq 1$ *and* $z_2^*(0) = 2$; $z_3^*(0) = \frac{4}{3}$, etc.

From ergodic theory point of view, taking into consideration the above difficulty, it is desirable to assume that the elements of the countable measurable partitions under consideration have positive measures. Now we discuss some properties of z_α^*.

(b_1) (i) $z_\alpha^*(t) \geqq 0$.

(ii) $\lim_{t \to 0+} z_\alpha^*$ lies between 0 and $+\infty$.

(b_2) z_α^* *is strictly monotonically decreasing* continuous function of t.

(b_3) (i) z_α^* is *convex* function of t for $0 \leqq \alpha \leqq 2$.

(ii) z_α^* is a *concave* function of t for $2 < \alpha < \infty$.

(iii) z_α^* is a *strictly sub-additive* function of t, i.e.

$$z_\alpha^*(t_1 + t_2) < z_\alpha^*(t_1) + z_\alpha^*(t_2), \quad (t_1, t_2, t_1 + t_2 \in (0, 1]).$$

(b_4) z_α^* satisfies functional equation (2.13).

 In particular,

$$z_\alpha^*(xy) \leqq z_\alpha^*(x) + z_\alpha^*(y)\,, \quad x, y \in [0, 1]\,, \quad 1 < \alpha < \infty\,, \tag{3.1}$$

$$z_\alpha^*(xy) \geqq z_\alpha^*(x) + z_\alpha^*(y)\,, \quad x, y \in (0, 1]\,, \quad 0 \leqq \alpha < 1\,, \tag{3.2}$$

$$z_\alpha^*(xy) = z_\alpha^*(x) + z_\alpha^*(y)\,, \quad x, y \in (0, 1]\,, \quad \alpha = 1\,, \tag{3.3}$$

equality in (3.1) and (3.2) being true for $x = 1$ or $y = 1$.

(b_5) (i) $z_0^*(t) = (1 - t)/t$, $t \in (0, 1]$, represents a branch, lying in the first and fourth quadrants, of the rectangular hyperbola $xy + x = 1$, $x \in R$, $y \in R$. Since $t \in (0, 1]$, we shall be concerned with only that part of the branch which commences from the point (1, 0) and becomes asymptotic to y-axis. The asymptotes of the hyperbola are the lines $t = 0$ and $y = -1$.

The function z_0^* also satisfies the functional equation

$$t\, z_0^*(t) + (1 - t)\, z_0^*(1 - t) = 1\,, \quad t \in (0, 1]\,. \tag{3.4}$$

In addition to $z_0^*(t) = (1 - t)/t$, $t \in (0, 1]$, there are other solutions of (3.4) which are continuous for $t \in (0, 1]$. For example,

$$z_0^*(t) = 2t^2 - 3t + \frac{1}{t}\,.$$

Let us define another function z_α such that

$$\begin{aligned} z_\alpha(t) &= t\, z_\alpha^*(t)\,, & t &\in (0, 1]\,, & \alpha &\in [0, \infty) \\ &= 1\,, & t &= 0\,, & \alpha &= 0\,, \\ &= 0\,, & t &= 0\,, & \alpha &\in (0, \infty)\,. \end{aligned} \tag{3.5}$$

Clearly (3.4) and (3.5) give

$$z_0(t) + z_0(1 - t) = 1\,, \quad t \in [0, 1]\,. \tag{3.6}$$

From the definition of z_α, it is obvious that $z_0(t) = 2t^3 - 3t^2 + 1$ is also a continuous solution of (3.6) in addition to $z_0(t) = 1 - t$. All these solutions satisfy the conditions $z_0(1) = 0$, $z_0(\frac{1}{2}) = \frac{1}{2}$, the latter being an obvious consequence of (3.6).

In fact, infinitely many continuous solutions of (3.6) can be obtained by choosing an arbitrary continuous graph joining the points $(\frac{1}{2}, \frac{1}{2})$ and (1, 0) and then extending this continuous function from $(\frac{1}{2}, \frac{1}{2})$ to (0, 1) by defining

$$z_0(t) = 1 - z_0(1 - t)\,, \quad t \in [0, \tfrac{1}{2}]\,.$$

Once continuous solutions of (3.6) are known, the corresponding continuous solutions of (3.4) can be found easily.

Besides continuous solutions, (3.4) and (3.6) also admit discontinuous solutions. For example,

$$g(t) = \begin{cases} 1, & t \in [0, \frac{1}{2}], \\ 0, & t \in (\frac{1}{2}, 1], \end{cases} \qquad h(t) = \begin{cases} 0, & t = 0, \\ 1, & t \neq 0, \end{cases}$$

clearly satisfy (3.6) and (3.4) respectively.

(ii) $y_1 = z_1^*(t) = -\log_2 t$. Both z_1^* and z_0^* pass through the points (1, 0) and $(\frac{1}{2}, 1)$ and as $t \to 0$, both become asymptotic to their y-axis. Also

$$z_0^*(t) \log_2 e \geqq z_1(t) \geqq t\, z_0^*(t) \log_2 e\,, \quad t \in (0, 1]\,, \tag{3.7}$$

equality being true only when $t = 1$.

(iii) $y_2 = z_2^*(t) = 2(1 - t)$ represents the equation of a straight line passing through (1, 0) and (0, 2). It intersects both z_0^* and z_1^* only at (1, 0) and $(\frac{1}{2}, 1)$. In addition to (2.13), z_2^* also satisfies the functional equation

$$z_2^*(t) + z_2^*(1 - t) = 2\,, \quad t \in [0, 1]\,. \tag{3.8}$$

Let

$$g(t) = \tfrac{1}{2} z_2^*(t)\,, \qquad t \in [0, 1]\,.$$

Clearly (3.8) reduces to

$$g(t) + g(1 - t) = 1\,, \qquad t \in [0, 1]\,. \tag{3.9}$$

Due to the fact that z_α^* is *well-defined* at $t = 0$ *whenever* $\alpha > 1$, (3.9) differs from (3.6) in the sense that it holds at $t = 0$. Since $z_2^*(1) = 0$, $z_2^*(0) = 2$, therefore, $g(1) = 0$, $g(0) = 1$. Hence the continuous solutions of (3.9) are the same as those of (3.6). Consequently, the solutions of (3.8) can be found out easily.

It is obvious that

$$z_2^*(t) = 2t\, z_0^*(t)\,, \quad t \in (0, 1]\,.$$

Thus

$$\begin{aligned} z_2^*(t) &< z_0^*(t)\,, \qquad t \in (0, \tfrac{1}{2})\,, \\ z_2^*(t) &\geqq z_0^*(t)\,, \qquad t \in [\tfrac{1}{2}, 1]\,, \end{aligned} \tag{3.10}$$

equality holding only when $t = \frac{1}{2}, 1$.

(iv) $y_3 = z_3^*(t) = \frac{4}{3}(1 - t^2)$ represents the equation of a parabola with vertex $(0, \frac{4}{3})$, focus (0, 1), latus rectum 1. Also $z_2(t) = 2(t - t^2)$ represents a parabola passing through the points (0, 0), (1, 0) and $(\frac{1}{2}, \frac{1}{2})$. These two parabolae intersect only at one point (1, 0) and

$$z_2(t) \leqq z_2^*(t)\,, \quad t \in [0, 1]\,.$$

(v) For positive integral values of α, $z_\alpha^*(t)$ is a polynomial of degree $(\alpha - 1)$ in t, and hence $z_\alpha(t)$ is a polynomial of degree α in t. But

$$z_\alpha(t) \leqq z^*(t), \quad \alpha \in [0, \infty].$$

(b_6) $\alpha \to z_\alpha^*$ is a continuous mapping such that

$$\alpha_1 < \alpha_2 \Rightarrow z_{\alpha_2}^*(t) \leqq z_{\alpha_1}^*(t), \quad t \in (0, \tfrac{1}{2}],$$

$$\alpha_1 < \alpha_2 \Rightarrow z_{\alpha_1}^*(t) \leqq z_{\alpha_2}^*(t), \quad t \in [\tfrac{1}{2}, 1].$$

(b_7) (i) $\alpha \to z_\alpha^*(t) \in [1, 2], \quad \alpha \in [2, \infty], \quad t \in [0, \tfrac{1}{2}]$.

(ii) $\alpha \to z_\alpha^*(t) \in [0, 1], \quad \alpha \in [0, \infty], \quad t \in [\tfrac{1}{2}, 1]$.

(iii) It is only in the region $t \in (0, \tfrac{1}{2}]$ that $z_\alpha^*(t)$, for all $\alpha \in [0, 1]$, becomes infinite as $t \to 0$.

(iv) $z_\infty^*(t) = 1, \quad t \in [0, 1]$.

(b_8) Since z_α^* is defined even at $t = 0$ for $\alpha \geqq 2$, it follows that

$$z_\alpha^*(t) \geqq 1, \quad \alpha \geqq 2, \quad t \in [0, \tfrac{1}{2}],$$

$$z_\alpha^*(t) \leqq 1, \quad \alpha \geqq 2, \quad t \in [\tfrac{1}{2}, 1].$$

Let us consider the Euclidean metric $d(x, y) = \sqrt{[(x_1 - x_2)^2 + (y_1 - y_2)^2]}$ where $x = (x_1, x_2)$ and $y = (y_1, y_2)$. Then, it can be easily seen that, with respect to this metric, the sets

$$E_\alpha = \{(t, z_\alpha^*(t)), \ t \in [0, 1]\}, \quad \alpha \in G = [2, \infty],$$

are both connected as well as compact. Obviously, for any $S \subset G$, $\bigcap_{\alpha \in S} E_\alpha \neq \emptyset$. Hence $\bigcup_{\alpha \in G} E_\alpha$ *is connected but not compact*. Let

Δ_1 = set of all points lying inside and on the triangle with verticles $(0, 2)$, $(0, 1)$, $(\tfrac{1}{2}, 1)$,

Δ_2 = set of all points lying inside and on the triangle with vertices $(\tfrac{1}{2}, 1)$, $(1, 0)$, $(1, 1)$.

Clearly Δ_1 and Δ_2 are connected and compact sets. Since $\Delta_1 \cap \Delta_2 = \{(\tfrac{1}{2}, 1)\}$, it follows that $\Delta_1 \cup \Delta_2$ is also connected and compact. Moreover $\bigcup_{\alpha \in G} E_\alpha \subset (\Delta_1 \cup \Delta_2)$. The sets Δ_1 and Δ_2 can be transformed into each other because to each $(X, Y) \in \Delta_2$ there corresponds $(x, y) \in \Delta_1$ such that

$$X = 1 - x, \quad Y = 2 - y.$$

Also $(\Delta_1 \cup \Delta_2) \setminus \bigcup_{\alpha \in G} E_\alpha = \{(1, y) : 0 < y < 1\}$, and $\overline{\bigcup_{\alpha \in G} E_\alpha} = \Delta_1 \cup \Delta_2$.

(b_9) (i) $z_2^*(t) = 2t \, z_0^*(t), \quad t \in (0, 1]$.

(ii) For $\alpha > 2$,

$$z_\alpha^*(t) = \left(\frac{1 - 2^{2-\alpha}}{1 - 2^{1-\alpha}}\right) z_{\alpha-1}^*(t) + (1 - 2^{1-\alpha})^{-1} t^{\alpha-1} z_0^*(t) .$$

(b_{10}) $z_\alpha^*(t) \neq z_\alpha^*(1 - t)$, $t \in (0, 1)$ except when $t = \frac{1}{2}$.
However, if we define a function ψ_α such that

$$\psi_\alpha(t) = z_\alpha^*(t) + z_\alpha^*(1 - t), \quad t \in (0, 1), \quad \alpha \in [0, \infty),$$

then

$$\psi_\alpha(t) = \psi_\alpha(1 - t), \quad t \in (0, 1).$$

Clearly ψ_α is a symmetric function of $t \in (0, 1)$. Also, ψ_α is *monotonic decreasing function* of t, $\frac{1}{2} \leqq t < 1$, and *monotonic increasing function* of t, $0 < t < \frac{1}{2}$. For $0 < \alpha \leqq 1$, $\psi_\alpha(0)$ and $\psi_\alpha(1)$ are not defined because of the difficulty that $z_\alpha^*(0) = \infty$. However, for $\alpha > 1$, $\psi_\alpha(0)$ and $\psi_\alpha(1)$ are both finite and further $\psi_\alpha(0) = \psi_\alpha(1)$ and consequently $\psi_\alpha(t)$ is a symmetric function of $t \in [0, 1]$.

It is easily seen that $\psi_\alpha(t)$ is not necessarily a polynomial of degree $\alpha - 1$ for positive integral values of α. Rather, $\psi_\alpha(t)$ is a polynomial of degree $\alpha - 1$ for positive odd integral values of α and of degree $\alpha - 2$ for positive even integral values of α. Thus $\psi_2(t) = 2$, a straight line and $\psi_3(t) = \frac{4}{3}(1 + 2t - 2t^2)$, a parabola.

(b_{11}) Since

$$\log_2 t^{\alpha-1} \leqq (t^{\alpha-1} - 1) \log_2 e, \quad \alpha \geqq 0, \quad t \in (0, 1],$$

equality being true when $\alpha = 1$ or $t = 1$, it follows that

$$z_1^*(t) \overset{\leqq}{\underset{\geqq}{\text{or}}} (\log_2 e) \left(\frac{1 - 2^{1-\alpha}}{\alpha - 1}\right) z_\alpha^*(t) \text{ according as } \alpha < 1 \text{ or } \alpha > 1,$$

$$\psi_1(t) \overset{\leqq}{\underset{\geqq}{\text{or}}} (\log_2 e) \left(\frac{1 - 2^{1-\alpha}}{\alpha - 1}\right) \psi_\alpha(t) \text{ according as } \alpha < 1 \text{ or } \alpha > 1.$$

(b_{12}) Let us define

$$(3.11) \qquad z_{\alpha,\beta}^*(t) = \frac{t^{\beta-1} - t^{\alpha-1}}{2^{1-\beta} - 2^{1-\alpha}}, \quad \alpha \neq \beta, \quad t \in (0, 1], \ \alpha \geqq 0, \ \beta \geqq 0.$$

Obviously $z_{\alpha,1}^*(t) = z_\alpha^*(t)$ and

$$z_{\alpha,\beta}^*(t) = \frac{(1 - 2^{1-\alpha}) z_\alpha^*(t) - (1 - 2^{1-\beta}) z_\beta^*(t)}{(1 - 2^{1-\alpha}) - (1 - 2^{1-\beta})}.$$

Also

(i) $z^*_{\alpha,\beta}(t) \geqq 0$.

(ii) $z^*_{\alpha,\beta}(\frac{1}{2}) = 1$.

(iii) $\lim_{\alpha\to\beta} z^*_{\alpha,\beta}(t) = -2^{\beta-1}t^{\beta-1}\log_2 t = \zeta^*_\beta(t)$ (say), $\beta \geqq 0,\ t \in (0, 1]$.

Clearly

$$\zeta^*_1(t) = z^*_1(t)\,.$$

(iv) ζ^*_α satisfies the functional equation

$$\zeta^*_\alpha(xy) = y^{\alpha-1}\,\zeta^*_\alpha(x) + x^{\alpha-1}\,\zeta^*_\alpha(y) \quad (x, y \in (0, 1])\,. \tag{3.12}$$

However, the continuous solutions of (3.12) are of the form

$$\zeta^*_\alpha(x) = K(\alpha)\,x^{\alpha-1}\log_2 x\,.$$

(v) $z^*_{\alpha,\beta}$ satisfies the functional equation

$$z^*_{\alpha,\beta}(xy) = y^{\alpha-1}\,z^*_{\alpha,\beta}(x) + x^{\beta-1}\,z^*_{\alpha,\beta}(y) \quad (x, y \in (0, 1])\,, \tag{3.13}$$

whose solutions are of the form

$$z^*_{\alpha,\beta}(x) = \lambda(x^{\alpha-1} - x^{\beta-1})\,, \quad \alpha \neq \beta\,.$$

However if $z^*_{\alpha,\beta}(\frac{1}{2}) = 1$, then we get (3.11).

Now we state the following theorem.

Theorem 3.1. *Let there be two sets* $\{a_i\}$, $i = 1, 2, \ldots, m$, *and* $\{b_j\}$, $j = 1, 2, \ldots, n$, *of non-negative real numbers such that* $\sum_{i=1}^m a_i = \sum_{j=1}^n b_j = 1$. *Let there be a real-valued continuous function h which satisfies the functional equation*

$$\sum_{i=1}^m \sum_{j=1}^n h(a_ib_j) = \Big(\sum_{j=1}^n b_j^\alpha\Big)\Big(\sum_{i=1}^m h(a_i)\Big) + \Big(\sum_{i=1}^n a_i^\beta\Big)\Big(\sum_{j=1}^m h(b_j)\Big)\,, \tag{3.14}$$

where $\alpha > 0$, $\beta > 0$ *and* $h(\frac{1}{2}) = \frac{1}{2}$. *Then* $h = z_{\alpha,\beta}$ *where*

$$\begin{aligned} z_{\alpha,\beta}(t) &= 0\,, & t &= 0, 1\,,\\ &= t\,z^*_{\alpha,\beta}(t)\,, & \alpha &\neq \beta\,,\quad t \in (0, 1)\,,\\ &= t\,\zeta^*_\beta(t)\,, & \alpha &= \beta\,,\quad t \in (0, 1)\,. \end{aligned} \tag{3.15}$$

We give a sketch of the proof of the theorem. Let $n = m = 1$. Then, (3.14) gives $h(1) = 0$. Now, let $n = m = 2$. Choosing $a_1 = 1$, $a_2 = 0$; $b_1 = 0$, $b_2 = 1$, (3.14) gives $h(1) = h(0)$. Thus

$$h(1) = h(0) = 0\,.$$

Writing $m = u - r + 1$, $n = v - s + 1$, $1 \leqq r < u$, $1 \leqq s < v$; u, r, v, s being positive integers, and choosing

$$a_i = \frac{1}{n}, \quad i = 1, 2, \ldots, u - r; \quad a_{u-r+1} = \frac{r}{u},$$

and

$$b_j = \frac{1}{v}, \quad j = 1, 2, \ldots, v - s; \quad b_{v-s+1} = \frac{s}{v},$$

it can be shown easily by following the technique as in [3] that (3.14) reduces to the functional equation (using continuity of h)

$$\Psi(xy) = y^{1-\alpha}\,\Psi(x) + x^{1-\beta}\,\Psi(y), \tag{3.16}$$

x, y being real number greater than unity,

where

$$\Psi(x) = xh\left(\frac{1}{x}\right), \quad x > 1. \tag{3.17}$$

Only two cases arise i.e. $\alpha \neq \beta$ and $\alpha = \beta$. In both cases, the functional equation (3.16) can be easily solved. Making use of (3.17) and the condition $h(\frac{1}{2}) = \frac{1}{2}$, (3.15) follows immediately.

In [2], the authors have discussed the functional equation

$$\sum_{i=1}^{m} \sum_{j=1}^{n} \hat{g}(a_i b_j) = \sum_{i=1}^{m} \hat{g}(a_i) + \sum_{j=1}^{n} \hat{g}(b_j) + (2^{1-\alpha} - 1)\left(\sum_{i=1}^{m} \hat{g}(a_i)\right)\left(\sum_{j=1}^{n} g(b_j)\right), \tag{3.18}$$

$$a \geqq 0.$$

Both (3.14) and (3.18) are generalizations of the well-known functional equation given by Chaundy and Mcleod [3]

$$\sum_{i=1}^{m} \sum_{j=1}^{n} h(a_i b_j) = \sum_{i=1}^{m} h(a_i) + \sum_{j=1}^{n} h(b_j), \tag{3.19}$$

whose continuous solution satisfying $h(\frac{1}{2}) = \frac{1}{2}$ is of the form

$$h(t) = -t \log_2 t, \quad t \in (0, 1),$$
$$= 0, \qquad t = 0, 1.$$

We shall show in §4, as to how (3.14) and (3.18) are of considerable importance in the development of entropy theory.

Let $\mathscr{F}$ denote set of all countable measurable partitions of Ω and $\mathscr{A} \in \mathscr{F}$. The information $\mathscr{I}_\alpha(\mathscr{A})$ of order α of partition $\mathscr{A}$ is defined as

$$\mathscr{I}_\alpha(\mathscr{A}) = \sum_{A \in \mathscr{A}} \chi(A)\, z_\alpha^*(\mu(A)), \quad \mu(A) > 0, \quad \forall A \in \mathscr{A}, \tag{4.1}$$

where $\chi(A)$ denotes the characteristic function of set A.

The entropy of order α of the partition $\mathscr{A}$ is defined as

$$I_\alpha(\mathscr{A}) = \int_\Omega \mathscr{I}_\alpha(\mathscr{A})\, d\mu. \tag{4.2}$$

Both the functions $\mathscr{I}_\alpha(\mathscr{A})$ and $I_\alpha(\mathscr{A})$ are measurable with respect to $\bar{\mathscr{A}}$. If $\mathscr{A}$ is a finite partition, say $\mathscr{A} = (A_1, A_2, \ldots, A_n)$, then it is obvious that

$$I_\alpha(\mathscr{A}) = \sum_{A \in \mathscr{A}} \mu(A)\, z_\alpha^*(\mu(A)) = \sum_{A \in \mathscr{A}} z_\alpha(\mu(A)),$$

that is,

$$I_\alpha(\mathscr{A}) = \frac{1 - \sum_{i=1}^{n} \mu^\alpha(A_i)}{1 - 2^{1-\alpha}}, \quad \alpha \neq 1, \quad \mu(A_i) \in (0, 1], \quad i = 1, 2, \ldots, n, \tag{4.3}$$

$$= -\sum_{i=1}^{n} \mu(A_i) \log_2 \mu(A_i), \quad \alpha = 1, \quad \mu(A_i) \in (0, 1], \quad i = 1, 2, \ldots, n.$$

To every finite measurable partition $\mathscr{A} = (A_1, A_2, \ldots, A_n)$ of Ω, there corresponds a *discrete finite complete probability distribution*

$$\mathscr{P}_\mathscr{A} = (p_1, p_2, \ldots, p_n), \quad p_k > 0, \quad k = 1, 2, \ldots, n, \quad \sum_{k=1}^{n} p_k = 1. \tag{4.4}$$

The entropy $I_1(\mathscr{P}_\mathscr{A})$ is the well-known Shannon's entropy. Recently, Havrda and Charvát [5], *while studying the quantificatory theory of classificatory processes,* introduced $I_\alpha(\mathscr{P}_\mathscr{A})$, $\alpha \neq 1$, by calling it structural α-entropy. Quite independently, *by considering a generalized form of fundamental equation of information theory,* Daroczy [4] introduced $I_\alpha(\mathscr{P}_\mathscr{A})$ and called them *generalized information functions. Their characterizations differ in the sense that they assume the prior existence of parameter* α. In this paper, the authors have given simultaneous characterization of Shannon's entropy and $I_\alpha(\mathscr{A})$, $\alpha \neq 1$, without assuming the prior existence of parameter α.

To the authors, it appears that it might be more general to define entropy of a finite partition $\mathscr{A}$ as

$$\mathscr{H}_{\alpha,\beta}(\mathscr{A}) = \sum_{A \in \mathscr{A}} z_{\alpha,\beta}(\mu(A)), \quad \alpha > 0, \quad \beta > 0, \tag{4.5}$$

of which $I_\alpha(\mathscr{A})$ is a special case i.e. $I_\alpha(\mathscr{A}) = \mathscr{H}_{\alpha,1}(\mathscr{A})$. Moreover,

$$(4.6) \qquad \mathscr{H}_\beta(\mathscr{A}) = \mathscr{H}_{\beta,\beta}(\mathscr{A}) = \sum_{A\in\mathscr{A}} \mu(A)\, \zeta^*_\beta(\mu(A))\,, \quad \beta > 0\,,$$

of which $I_1(\mathscr{A})$ is a special case i.e. $I_1(\mathscr{A}) = \mathscr{H}_1(\mathscr{A})$. From Theorem 3.1, it is obvious that when $\mathscr{A}$ and $\mathscr{B}$ are finite independent partitions,

$$\mathscr{H}_{\alpha,\beta}(\mathscr{A} \vee \mathscr{B}) = \big(\sum_{B\in\mathscr{B}} \mu^\alpha(B)\big)\, \mathscr{H}_{\alpha,\beta}(\mathscr{A}) + \big(\sum_{A\in\mathscr{A}} \mu^\beta(A)\big)\, \mathscr{H}_{\alpha,\beta}(\mathscr{B})\,,$$

so that

$$(4.7) \qquad \mathscr{H}_{\alpha,\beta}(\mathscr{A} \vee \mathscr{B}) \leqq \mathscr{H}_{\alpha,\beta}(\mathscr{A}) + \mathscr{H}_{\alpha,\beta}(\mathscr{B})\,, \quad \alpha > 1\,, \quad \beta > 1\,,$$

$$(4.8) \qquad \mathscr{H}_{\alpha,\beta}(\mathscr{A} \vee \mathscr{B}) \geqq \mathscr{H}_{\alpha,\beta}(\mathscr{A}) + \mathscr{H}_{\alpha,\beta}(\mathscr{B})\,, \quad 0 < \alpha < 1\,, \quad 0 < \beta < 1\,,$$

equality in (4.7) and (4.8) being true when at least one of $\mathscr{A}$ and $\mathscr{B}$ is a trivial partition.

Halmos [6] and Parry [7] have described in considerable details the role of $\mathscr{I}_1(\mathscr{A})$ and $I_1(\mathscr{A})$ in ergodic theory. It is expected that $\mathscr{I}_\alpha(\mathscr{A})$ and $I_\alpha(\mathscr{A})$, $\alpha \neq 1$, may also prove useful in some situations. In this direction, the results for countable measurable partitions shall be presented elsewhere. It will be appropriate to mention that many of the properties of z^*_α discussed in §3 can be utilized to study the properties of $\mathscr{I}_\alpha(\mathscr{A})$ and $I_\alpha(\mathscr{A})$ in great detail.

The authors are extremely grateful to the referees for their valuable comments. Thanks are also due to Professor Janos Aczel for his valuable suggestions.

(Received November 17, 1971.)

REFERENCES

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

[2] Behara, N.; Nath, P.: Additive and non-additive Entropies of Finite Measurable Partitions. Probability and Information Theory, Vol. 2. Springer-Verlag, Berlin 1973. LN in Math. Vol. 296, 102–138.

[3] Chaundy, T. W.; Mcleod, J. B.: On a Functional Equation. Proc. Edin. Math. Soc. Notes *43*, (1960), 7–8.

[4] Daroczy, Z.: Generalized Information Functions. Information and Control *16* (1970), 36–51.

[5] Havrda, J.; Charvát, E.: Quantification Method of Classification Processes. The Concept of Structural a-entropy, Kybernetika *3* (1967), 30–35.

[6] Halmos, P. R.: Entropy in Ergodic Theory. University of Chicago, Chicago 1959.

[7] Parry, W.: Entropy and Generators in Ergodic Theory. W. A. Benjamin, Inc., New York 1959.

[8] Renyi, A.: On Measures of Entropy and Information. Proc. 4th Berk. Symp. on Math. Stat. and Prob. Vol. 1 (1967) 547–561.

Prof. Dr. M. Behara; Department of Mathematics, McMaster University, Hamilton, Ontario. Canada.

Dr. Prem Nath; Faculty of Mathematics, Delhi University, Delhi. India.