INFORMATION AND CONTROL 16, 36–51 (1970)

Generalized Information Functions

ZOLTÁN DARÓCZY

Department of Mathematics, University of L. Kossuth, Debrecen, Hungary

The concept of information functions of type β ($\beta > 0$) is introduced and discussed. By means of these information functions the entropies of type β are defined. These entropies have a number of interesting algebraic and analytic properties similar to Shannon's entropy. The capacity of type β ($\beta > 1$) of a discrete constant channel is defined by means of the entropy of type β. Examples are given for the computation of the capacity of type β, from which the Shannon's capacity can be derived as the limiting case $\beta = 1$.

1. INTRODUCTION

The concept of information functions has been introduced by the author [2].

DEFINITION 1. We call the real function f defined in [0, 1] an *information function* if it satisfies the boundary conditions

$$f(0) = f(1); \qquad f(\tfrac{1}{2}) = 1, \tag{1.1}$$

and the functional equation

$$f(x) + (1 - x) f\left(\frac{y}{1 - x}\right) = f(y) + (1 - y) f\left(\frac{x}{1 - y}\right) \tag{1.2}$$

for all $(x, y) \in D$, where

$$D = \{(x, y): 0 \leqslant x < 1, 0 \leqslant y < 1, x + y \leqslant 1\}. \tag{1.3}$$

If f is an information function and $(p_1, p_2, \dots, p_n)$ $(p_i \geqslant 0, \sum_{i=1}^{n} p_i = 1)$ is a finite discrete probability distribution, then we define the *entropy* of the distribution $(p_1, p_2, \dots, p_n)$ with respect to f by the quantity

$$H_n^f(p_1, p_2, \dots, p_n) = \sum_{i=2}^{n} s_i f\left(\frac{p_i}{s_i}\right), \quad (s_i = p_1 + \cdots + p_i;\ i = 2, \dots, n). \tag{1.4}$$

In (1.4) the convention $0f(0/0) = 0$ is adopted. We summarize the known results on the information functions in the following

THEOREM 1. *Let f be an information function. If f is*

(a) *measurable in the open interval* $(0, 1)$, *or*

(b) *continuous at the point* $x = 0$, *or*

(c) *nonnegative bounded in* $[0, 1]$, *then we have*

$$f(x) = S(x) \qquad \text{for all} \quad x \in [0, 1],$$

where $S(x)$ is the Shannon's information function defined by

$$S(x) = \begin{cases} -x \log_2 x - (1 - x) \log_2(1 - x) & \text{if} \quad x \in (0, 1) \\ 0 & \text{if} \quad x = 0 \quad \text{or} \quad x = 1. \end{cases} \tag{1.5}$$

The entropy of a probability distribution $(p_1, p_2, ..., p_n)$ with respect to S is the Shannon's entropy

$$H_n^S(p_1, p_2, ..., p_n) = H_n(p_1, p_2, ..., p_n) = -\sum_{i=1}^{n} p_i \log_2 p_i . \tag{1.6}$$

This theorem was proved by Lee [9] under assumption (a). The proof of the theorem can be found in Daróczy [2] under assumption (b). Finally, the theorem was proved by Daróczy and Kátai [4] in case (c). This theorem containes the results of Fadeev [6], Tverberg [11], Kendall [8] and Borges [1]. It is important to remark that there are information functions different from S (see Lee [9] or Daróczy [2]).

In this paper we generalize the concept of information functions as following.

DEFINITION 2. Let β be a positive number. We call the real function f defined in $[0, 1]$ an *information function of type β* if it satisfies the boundary conditions (1.1) and the functional equation

$$f(x) + (1 - x)^\beta f\left(\frac{y}{1 - x}\right) = f(y) + (1 - y)^\beta f\left(\frac{x}{1 - y}\right) \tag{1.7}$$

for all $(x, y) \in D$.

On the analogy of the entropy H_n^f we define the *entropy of type* β of a probability distribution $(p_1, p_2, \ldots, p_n)$ by the quantity

$$H_n^\beta(p_1, p_2, \ldots, p_n) = \sum_{i=2}^{n} s_i^\beta f\left(\frac{p_i}{s_i}\right), \quad (s_i = p_1 + \cdots + p_i;\ i = 2, \ldots, n) \quad (1.8)$$

where f is an information function of type β.

It is clear that an information function is an information function of type β with $\beta = 1$.

In Section 2 of this paper we determine all information functions of type β with $\beta \neq 1$. In the discussion of Section 3, we summarize the algebraic properties of the entropy of type β and we give an another characterization for these entropies. In Section 4 we consider the analytic properties of the entropy of type β. In Section 5 we define the joint and conditional entropy of type β of two discrete finite random variables and we discuss the properties of these quantities. In Section 6 we define and consider the capacity of type β of a discrete constant channel by means of the conditional entropy of type β, where $\beta > 1$.

2. Information Functions of Type β

It may be seen, that we have a number of interesting results on the information functions. Therefore, the following theorem is very unexpected.

THEOREM 2. *Let f be an information function of type β with $\beta \neq 1$. Then we have*

$$f(x) = S_\beta(x) \qquad \textit{for all} \quad x \in [0, 1],$$

where

$$S_\beta(x) = (2^{1-\beta} - 1)^{-1} [x^\beta + (1 - x)^\beta - 1] \quad (2.1)$$

for all $x \in [0, 1]$.

Proof. Let f be an information function of type β $(\beta > 0)$. If we take $x = 0$ in (1.7), then we have $f(0) = 0$. Taking $y = 1 - x$ into (1.7), we obtain by $f(1) = f(0) = 0$

$$f(x) = f(1 - x) \quad (2.2)$$

for all $x \in [0, 1]$.

Let p, q be two arbitrary numbers from the open interval (0, 1). We take $p = 1 - x$ and $q = y/1 - x$ in (1.7), then we obtain by (2.2)

$$f(p) + p^\beta f(q) = f(pq) + (1 - pq)^\beta f\left(\frac{1 - p}{1 - pq}\right). \tag{2.3}$$

We shall prove the following assertion. If $p, q \in (0, 1)$ are arbitrary, then the function $F(p, q)$ defined by

$$F(p, q) = f(p) + [p^\beta + (1 - p)^\beta] f(q) \tag{2.4}$$

is *symmetric*, i.e.

$$F(p, q) = F(q, p). \tag{2.5}$$

This assertion is trivial in the case $\beta = 1$. Let us put $\beta \neq 1$ ($\beta > 0$), then we have by (2.3)

$$\begin{aligned} F(p, q) &= f(p) + p^\beta f(q) + (1 - p)^\beta f(q) \\ &= f(pq) + (1 - pq)^\beta f\left(\frac{1 - p}{1 - pq}\right) + (1 - p)^\beta f(q) \\ &= f(pq) + (1 - pq)^\beta \left[f\left(\frac{1 - p}{1 - pq}\right) + \left(\frac{1 - p}{1 - pq}\right)^\beta f(q)\right] \end{aligned} \tag{2.6}$$

for all $p, q \in (0, 1)$. On the other hand, we obtain from (2.3) and (2.2) with the notation $p^* = 1 - p/1 - pq$

$$\begin{aligned} A(p, q) &= f\left(\frac{1 - p}{1 - pq}\right) + \left(\frac{1 - p}{1 - pq}\right)^\beta f(q) \\ &= f(p^*) + p^{*\beta} f(q) = f(p^* q) + (1 - p^* q)^\beta f\left(\frac{1 - p^*}{1 - p^* q}\right) \\ &= f(1 - p^* q) + (1 - p^* q)^\beta f\left(\frac{1 - p^*}{1 - p^* q}\right) \\ &= f\left(\frac{1 - q}{1 - pq}\right) + \left(\frac{1 - q}{1 - pq}\right)^\beta f(p) = A(q, p). \end{aligned}$$

Therefore, it follows from (2.6)

$$F(p, q) - F(q, p) = (1 - pq)^\beta [A(p, q) - A(q, p)] = 0.$$

Thus, the assertion (2.5) is proved. We now take $q = \frac{1}{2}$ in (2.5), then we have by the definition of $F(p, q)$ and by the boundary condition $f(\frac{1}{2}) = 1$

$$\begin{aligned} 0 &= F(p, \tfrac{1}{2}) - F(\tfrac{1}{2}, p) \\ &= f(p) + [p^\beta + (1-p)^\beta] - 1 - \frac{1}{2^{\beta-1}} f(p), \end{aligned}$$

from which it follows

$$f(p) = (2^{1-\beta} - 1)^{-1} [p^\beta + (1-p)^\beta - 1]$$

for all $p \in (0, 1)$. This formula is true in the case $p = 0$ or $p = 1$ by the boundary conditions (1.1), too. Thus the theorem 2 is proved.

It is very simple to see that

$$\lim_{\beta \to 1} S_\beta(x) = S(x) \qquad \text{for all} \quad x \in [0, 1]. \tag{2.7}$$

This remark shows that $S_\beta(x)$ is a natural generalization of the Shannon's information function (1.5).

3. Entropies of Type β

From the theorem 2 we obtain the following

THEOREM 3. *Let β be a positive number with $\beta \neq 1$. Then we have for the entropy of type β of a probability distribution* $(p_1, p_2, \ldots, p_n)$

$$H_n^\beta(p_1, p_2, \ldots, p_n) = (2^{1-\beta} - 1)^{-1} \left(\sum_{i=1}^{n} p_i^\beta - 1 \right). \tag{3.1}$$

From (2.7) it follows that the Shannon's entropy $H_n(p_1, p_2, \ldots, p_n)$ is a limiting function of $H_n^\beta(p_1, p_2, \ldots, p_n)$, i.e.

$$H_n(p_1, p_2, \ldots, p_n) = \lim_{\beta \to 1} H_n^\beta(p_1, p_2, \ldots, p_n).$$

Previously Rényi [10] has extended the concept of Shannon's entropy by defining the *entropy of order* β ($\beta > 0, \beta \neq 1$) of a probability distribution $(p_1, p_2, \ldots, p_n)$ as

$${}_\beta H_n(p_1, p_2, \ldots, p_n) = (1 - \beta)^{-1} \log_2 \sum_{i=1}^{n} p_i^\beta. \tag{3.2}$$

From (3.1) and (3.2) we have the following relations between the entropy of *order* β and the entropy of *type* β:

$$_{\beta}H_n = (1-\beta)^{-1} \log_2[(2^{1-\beta}-1)\, H_n{}^{\beta} + 1] \tag{3.3}$$

and

$$H_n{}^{\beta} = (2^{1-\beta}-1)^{-1}[2^{(1-\beta)_{\beta}H_n} - 1]. \tag{3.4}$$

We denote by

$$\Delta_n = \left\{(p_1, p_2, ..., p_n) : p_i \geqslant 0, \sum_{i=1}^{n} p_i = 1\right\} \tag{3.5}$$

the set of all n-ary probability distributions. The entropy of type β is a real function defined on $\Delta_n (n = 2, 3,...)$, i.e.,

$$H_n{}^{\beta} : \Delta_n \to R \qquad (n = 2, 3,...),$$

where R is the set of real numbers. In the following theorem we summarize the algebraic properties of the entropy of type β.

THEOREM 4. *The entropies* $H_n{}^{\beta} : \Delta_n \to R$ $(n = 2, 3,...; \beta > 0)$ *have the following properties*:

1^0 *Symmetric*: $H_n{}^{\beta}(p_1, p_2, ..., p_n)$ *is a symmetric function of its variables*;

2^0 *Normalized*: $H_2{}^{\beta}(\frac{1}{2}, \frac{1}{2}) = 1$;

3^0 *Expansible*: $H_n{}^{\beta}(p_1, p_2, ..., p_n) = H_{n+1}^{\beta}(p_1, p_2, ..., p_n, 0)$;

4^0 *Strongly additive type* β:

$$H_{nm}^{\beta}(p_1 q_{11}, ..., p_1 q_{m1}; p_2 q_{12}, ..., p_2 q_{m2}; \cdots; p_n q_{1n}, ..., p_n q_{mn})$$
$$= H_n{}^{\beta}(p_1, p_2, ..., p_n) + \sum_{i=1}^{n} p_i{}^{\beta} H_m{}^{\beta}(q_{1i}, q_{2i}, ..., q_{mi})$$

for all $(p_1, p_2, ..., p_n) \in \Delta_n$ *and* $(q_{1i}, q_{2i}, ..., q_{mi}) \in \Delta_m$ $(i = 1, 2,..., n)$;

5^0 *Recursive type* β:

$$H_n{}^{\beta}(p_1, p_2, p_3, ..., p_n) - H_{n-1}^{\beta}(p_1 + p_2, p_3, ..., p_n)$$
$$= (p_1 + p_2)^{\beta} H_2{}^{\beta}([p_1/p_1 + p_2], [p_2/p_1 + p_2])$$

for all $(p_1, p_2, ..., p_n) \in \Delta_n$ *with* $p_1 + p_2 > 0$ $(n \geqslant 3)$.

Proof. The properties 1^0, 2^0 and 3^0 are obvious consequences of Theorem 3. We prove 4^0 by a direct computation:

$$H_n^\beta(p_1, p_2, \ldots, p_n) + \sum_{i=1}^{n} p_i^\beta H_m^\beta(q_{1i}, q_{2i}, \ldots, q_{mi})$$

$$= (2^{1-\beta} - 1)^{-1} \left(\sum_{i=1}^{n} p_i^\beta - 1\right) + (2^{1-\beta} - 1)^{-1} \sum_{i=1}^{n} p_i^\beta \left(\sum_{k=1}^{m} q_{ki}^\beta - 1\right)$$

$$= (2^{1-\beta} - 1)^{-1} \left(\sum_{i=1}^{n} p_i^\beta - 1 + \sum_{i=1}^{n} \sum_{k=1}^{m} p_i^\beta q_{ki}^\beta - \sum_{i=1}^{n} p_i^\beta\right)$$

$$= (2^{1-\beta} - 1)^{-1} \left[\sum_{i=1}^{n} \sum_{k=1}^{m} (p_i q_{ki})^\beta - 1\right]$$

$$= H_{nm}^\beta(p_1 q_{11}, \ldots, p_1 q_{m1}; \cdots; p_n q_{1n}, \ldots, p_n q_{mn}).$$

The proof of 5^0 is very easy

$$H_n^\beta(p_1, p_2, p_3, \ldots, p_n) - H_{n-1}^\beta(p_1 + p_2, p_3, \ldots, p_n)$$

$$= (2^{1-\beta} - 1)^{-1} \left[p_1^\beta + p_2^\beta + \sum_{i=3}^{n} p_i^\beta - 1 - (p_1 + p_2)^\beta - \sum_{i=3}^{n} p_i^\beta + 1\right]$$

$$= (2^{1-\beta} - 1)^{-1}[p_1^\beta + p_2^\beta - (p_1 + p_2)^\beta]$$

$$= (p_1 + p_2)^\beta (2^{1-\beta} - 1)^{-1} \left[\left(\frac{p_1}{p_1 + p_2}\right)^\beta + \left(\frac{p_2}{p_1 + p_2}\right)^\beta - 1\right]$$

$$= (p_1 + p_2)^\beta H_2^\beta \left(\frac{p_1}{p_1 + p_2}, \frac{p_2}{p_1 + p_2}\right).$$

Now, we shall give an another characterization of the entropy of type β. The problem is the following. What properties have to be imposed upon a sequence

$$I_n : \Delta_n \to R \qquad (n = 2, 3, \ldots)$$

of functions in order that the following identical equality should hold $I_n(p_1, \ldots, p_n) = H_n^\beta(p_1, \ldots, p_n)$ for all $(p_1, \ldots, p_n) \in \Delta_n$, where $\beta > 0$ and $\beta \neq 1$. The following theorem is a generalization of a result given by the author (see Daróczy [3]).

THEOREM 5. *Let* $I_n : \Delta_n \to R$ $(n = 2, 3,...)$ *be a sequence of mappings and let* β *be a positive number different from one. If* I_n *satisfies the following conditions*:

(i) $I_3(p_1, p_2, p_3)$ *is a symmetric function of its variables*;

(ii) $I_2(\frac{1}{2}, \frac{1}{2}) = 1$;

(iii) $I_n(p_1, p_2, p_3, ..., p_n) - I_{n-1}(p_1 + p_2, p_3, ..., p_n)$

$$= (p_1 + p_2)^\beta I_2\left(\frac{p_1}{p_1+p_2}, \frac{p_2}{p_1+p_2}\right)$$

for all $(p_1, p_2, ..., p_n) \in \Delta_n$ $\quad (n = 3, 4, ...; p_1 + p_2 > 0)$,

then we have

$$I_n(p_1, p_2, ..., p_n) = H_n^\beta(p_1, p_2, ..., p_n)$$

for all $(p_1, p_2, ..., p_n) \in \Delta_n$ $(n = 2, 3, ...)$.

Proof. First we prove that the function f defined by $f(x) = I_2(x, 1 - x)$ $(x \in [0, 1])$ is an information function of type β. Let $(x, y) \in D$ be arbitrary, then we have by (i)

$$I_3(y, 1 - x - y, x) = I_3(x, 1 - y - x, y),$$

from which it follows by (iii)

$$f(1 - x) + (1 - x)^\beta f\left(\frac{y}{1 - x}\right) = f(1 - y) + (1 - y)^\beta f\left(\frac{x}{1 - y}\right). \quad (3.6)$$

We take $x = 0, y = \frac{1}{2}$ in (3.6), then we have

$$f(1) = 2^{-\beta} f(0). \quad (3.7)$$

By (iii) we have $I_3(1, 0, 0) = 2f(1)$ and $I_3(0, 1, 0) = f(1) + f(0)$, from which we obtain by (i)

$$f(1) = f(0). \quad (3.8)$$

From (3.7) and (3.8) we get $f(1) = f(0) = 0$. Therefore, we have

$$f(1 - x) = f(x) \qquad \text{for all} \quad x \in [0, 1] \quad (3.9)$$

from (3.6) with the substitution $y = 1 - x$. By (3.9) the functional Eq. (3.6) yields Eq. (1.7), that is, the function f is an information function of type β.

By Theorem 2 we have $f(x) = S_\beta(x)$, hence, the theorem is proved by (iii) with induction.

4. Analytic Properties of the Entropy of Type β

We begin with the following

THEOREM 6. *For all* $(p_1, p_2, ..., p_n) \in \Delta_n$ $(n = 2, 3, ...)$

$$0 \leqslant H_n^\beta(p_1, p_2, ..., p_n) \leqslant H_n^\beta\left(\frac{1}{n}, \frac{1}{n}, ..., \frac{1}{n}\right) = (2^{1-\beta} - 1)^{-1}(n^{1-\beta} - 1). \quad (4.1)$$

Proof. We define the function $l_\beta(x)$ by

$$l_\beta(x) = (2^{1-\beta} - 1)^{-1}(x^\beta - x) \qquad (\beta > 0, \beta \neq 1) \quad (4.2)$$

for all $x \in [0, 1]$. The function $l_\beta(x)$ is a *nonnegative* and *concave* function in $[0, 1]$. By the concavity of l_β we have

$$H_n^\beta(p_1, p_2, ..., p_n) = \sum_{i=1}^n l_\beta(p_i) \leqslant n l_\beta\left(\frac{1}{n}\sum_{i=1}^n p_i\right)$$

$$= n l_\beta\left(\frac{1}{n}\right) = H_n^\beta\left(\frac{1}{n}, \frac{1}{n}, ..., \frac{1}{n}\right) = (2^{1-\beta} - 1)^{-1}(n^{1-\beta} - 1),$$

which proves the theorem.

It is interesting to remark that the function

$$\varphi_\beta(n) = H_n^\beta\left(\frac{1}{n}, \frac{1}{n}, ..., \frac{1}{n}\right)$$

is monotonic, i.e., $\varphi_\beta(n) \leqslant \varphi_\beta(n + 1)$. This is a simple consequence of Theorems 6 and 5:

$$\varphi_\beta(n) = H_n^\beta\left(\frac{1}{n}, ..., \frac{1}{n}\right) = H_{n+1}^\beta\left(\frac{1}{n}, ..., \frac{1}{n}, 0\right)$$

$$\leqslant H_{n+1}^\beta\left(\frac{1}{n+1}, ..., \frac{1}{n+1}\right) = \varphi_\beta(n + 1).$$

If $\beta > 1$, then we have by the monotonicity of φ_β : *For all* $(p_1, p_2, ..., p_n) \in \Delta_n$ $(n = 2, 3, ...)$

$$H_n^\beta(p_1, ..., p_n) \leqslant \varphi_\beta(n) \leqslant \lim_{n\to\infty} \varphi_\beta(n) = (1 - 2^{1-\beta})^{-1}. \quad (4.3)$$

In the case $0 < \beta < 1$ this assertion is not true, while

$$\lim_{n\to\infty} \varphi_\beta(n) = +\infty.$$

THEOREM 7. *Suppose* $\beta > 1$. *For* $(p_1, p_2, \ldots, p_n) \in \Delta_n$,

$$(q_{1i}, q_{2i}, \ldots, q_{mi}) \in \Delta_m \qquad (i = 1, 2, \ldots, n)$$

$$\sum_{i=1}^{n} p_i^\beta H_m^\beta(q_{1i}, q_{2i}, \ldots, q_{mi}) \tag{4.4}$$

$$\leqslant H_m^\beta\left(\sum_{i=1}^{n} p_i q_{1i}, \sum_{i=1}^{n} p_i q_{2i}, \ldots, \sum_{i=1}^{n} p_i q_{mi}\right).$$

Proof. Using the concavity of $l_\beta(x)$ defined by (4.2), we have

$$\sum_{i=1}^{n} p_i l_\beta(q_{ki}) \leqslant l_\beta\left(\sum_{i=1}^{n} p_i q_{ki}\right) \qquad (k = 1, 2, \ldots, m).$$

Let us add these inequalities with respect to k, we have

$$\sum_{i=1}^{n} p_i \sum_{k=1}^{m} l_\beta(q_{ki}) \leqslant \sum_{k=1}^{m} l_\beta\left(\sum_{i=1}^{n} p_i q_{ki}\right). \tag{4.5}$$

Using the assumption $\beta > 1$,

$$p_i^\beta \leqslant p_i \qquad (i = 1, 2, \ldots, n),$$

therefore, by the nonnegativity of l_β

$$p_i^\beta \sum_{k=1}^{m} l_\beta(q_{ki}) \leqslant p_i \sum_{k=1}^{m} l_\beta(q_{ki}).$$

If we add these inequalities with respect to i, then we obtain

$$\sum_{i=1}^{n} p_i^\beta \sum_{k=1}^{m} l_\beta(q_{ki}) \leqslant \sum_{i=1}^{n} p_i \sum_{k=1}^{m} l_\beta(q_{ki}). \tag{4.6}$$

From (4.6) and (4.5) we get

$$\sum_{i=1}^{n} p_i^\beta \sum_{k=1}^{m} l_\beta(q_{ki}) \leqslant \sum_{k=1}^{m} l_\beta\left(\sum_{i=1}^{n} p_i q_{ki}\right).$$

This exactly is (4.4).

5. Entropies of Type β of Discrete Random Variables

Let ξ be a discrete finite random variable with the set $x_1, x_2, \dots, x_n$ of possible values of ξ. We define the *entropy of type β of the random variable ξ* by

$$H^\beta(\xi) = H_n^\beta(p_1, p_2, \dots, p_n), \tag{5.1}$$

where

$$p_i = P(\xi = x_i) \qquad (i = 1, 2, \dots, n). \tag{5.2}$$

Correspondingly, for a two-dimensional discrete finite random variable (ξ, η) with the joint discrete probability distribution

$$\pi_{ik} = P(\xi = x_i, \eta = y_k) \qquad (i = 1, \dots, n;\ k = 1, \dots, m),$$

we have the following notions:

$$H^\beta(\xi, \eta) = H_{nm}^\beta(\pi_{11}, \dots, \pi_{1m}; \dots; \pi_{n1}, \dots, \pi_{nm}) \tag{5.3}$$

and

$$H^\beta(\eta \mid \xi) = \sum_{i=1}^{n} p_i^\beta H_m^\beta(q_{1i}, q_{2i}, \dots, q_{mi}), \tag{5.4}$$

where

$$p_i = P(\xi = x_i) = \sum_{k=1}^{m} \pi_{ik} \qquad (i = 1, 2, \dots, n)$$

and

$$q_{ki} = P(\eta = y_k \mid \xi = x_i) = \frac{\pi_{ik}}{p_i} \qquad (i = 1, 2, \dots, n;\ k = 1, 2, \dots, m).$$

We call the quantity (5.3) *joint entropy of type β of (ξ, η)* and we call the quantity (5.4) *conditional entropy of type β of the random variable η with respect to ξ*. It is clear that in the limiting case $\beta = 1$ we have the usual quantities of the information theory (see Fano [7]).

By the algebraic and analytic properties of the entropy of type β, we have the following

Theorem 8. *If ξ and η are discrete finite random variables, then*

$$H^\beta(\xi, \eta) = H^\beta(\xi) + H^\beta(\eta \mid \xi), \tag{5.5}$$

and

$$H^\beta(\eta \mid \xi) \leqslant H^\beta(\eta) \qquad \text{if} \quad \beta > 1. \tag{5.6}$$

Proof. Equation (5.5) comes from Theorem 4. The inequality (5.6) is equivalent with the inequality (4.1) presented in Theorem 6.

It is simple to see that the statements of Theorem 8 are generalizations of the known equality and inequality for the Shannon's entropy. The following problem is very natural: what is the situation in the case of independent random variables?

THEOREM 9. *If ξ and η are discrete independent random variables, then*

$$H^\beta(\xi, \eta) = H^\beta(\xi) + H^\beta(\eta) + (2^{1-\beta} - 1)\, H^\beta(\xi)\, H^\beta(\eta). \tag{5.7}$$

Proof. In this case we have

$$P(\eta = y_k \mid \xi = x_i) = q_{ki} = q_k \qquad (i = 1,..., n;\, k = 1,..., m),$$

where

$$q_k = P(\eta = y_k).$$

By (5.5) we get

$$\begin{aligned} H^\beta(\xi, \eta) &= H^\beta(\xi) + H^\beta(\eta \mid \xi) \\ &= H^\beta(\xi) + \sum_{i=1}^{n} p_i^\beta (2^{1-\beta} - 1)^{-1} \left(\sum_{k=1}^{m} q_k^\beta - 1\right) \\ &= H^\beta(\xi) + H^\beta(\eta) \left(\sum_{i=1}^{n} p_i^\beta\right) \\ &= H^\beta(\xi) + H^\beta(\eta) + H^\beta(\eta) \left(\sum_{i=1}^{n} p_i^\beta - 1\right) \\ &= H^\beta(\xi) + H^\beta(\eta) + (2^{1-\beta} - 1)\, H^\beta(\xi)\, H^\beta(\eta), \end{aligned}$$

which proves the theorem. In the limiting case $\beta \to 1$, we have the known additivity property of Shannon's entropy.

6. Capacity of Type β

A discrete constant channel with the space $X = \{x_1, \ldots, x_n\}$ of input symbols and with the space $Y = \{y_1, \ldots, y_m\}$ of output symbols is characterized by the $(m \times n)$ transition matrix

$$\left.\begin{aligned} Q &= (q_{ki}) \qquad (k = 1, \ldots, m; i = 1, \ldots, n) \\ \text{with} \qquad q_{ki} &\geqslant 0, \qquad \sum_{k=1}^{m} q_{ki} = 1 \quad (i = 1, \ldots, n). \end{aligned}\right\} \tag{6.1}$$

q_{ki} represents the conditional probability for receiving the k-th output symbol if the j-th input symbol has been transmitted.

Consider an arbitrary input probability distribution $(p_1, p_2, \ldots, p_n) \in \Delta_n$ on the space of input symbols, which induces the distribution

$$(q_1, q_2, \ldots, q_m) \in \Delta_m$$

on the space of output symbols given by

$$q_k = \sum_{i=1}^{n} p_i q_{ki} \qquad (k = 1, \ldots, m). \tag{6.2}$$

The spaces of input and output symbols can now be viewed as the space of values for discrete random variables, say, ξ and η, respectively. Now, we define the *capacity of type* β of the discrete constant channel characterized by Q as the quantity

$$C_\beta = \max_{(p_1, \ldots, p_n) \in \Delta_n} [H^\beta(\eta) - H^\beta(\eta \mid \xi)], \tag{6.3}$$

where $\beta > 1$ and the distributions of η and ξ are given by

$$\begin{aligned} P(\xi = x_i) &= p_i \qquad (i = 1, \ldots, n), \\ P(\eta = y_k) &= q_k = \sum_{i=1}^{n} p_i q_{ki}, \qquad (k = 1, \ldots, m), \\ P(\eta = y_k \mid \xi = x_i) &= q_{ki}, \qquad (k = 1, \ldots, m; i = 1, \ldots, n). \end{aligned} \tag{6.4}$$

If $\beta \to 1+$, then we have the known concept of the Shannon's capacity of a discrete constant channel (see Fano [7]):

$$C = \lim_{\beta \to 1+} C_\beta. \tag{6.5}$$

THEOREM 10. *If the transition matrix of a discrete constant channel has the form*

$$q_{ki} = \begin{cases} 1 - p & \text{if } i = k \\ \dfrac{p}{n-1} & \text{if } i \neq k \end{cases} \qquad (n = m), \tag{6.6}$$

then we have

$$C_\beta = (2^{1-\beta} - 1)^{-1}(n^{1-\beta} - 1) - n^{1-\beta}(2^{1-\beta} - 1)^{-1}[(1-p)^\beta + p^\beta(n-1)^{1-\beta} - 1]. \tag{6.7}$$

Proof. By (6.6) and by Theorem 6, we have

$$H^\beta(\eta) - H^\beta(\eta \mid \xi) \leqslant (2^{1-\beta} - 1)^{-1}(n^{1-\beta} - 1) - \sum_{i=1}^{n} p_i^\beta (2^{1-\beta} - 1)^{-1} \left[(1-p)^\beta + (n-1)\left(\frac{p}{n-1}\right)^\beta - 1\right].$$

By the convexity of t^β ($\beta > 1$), we obtain

$$\sum_{i=1}^{n} p_i^\beta \geqslant n \left(\frac{1}{n} \sum_{i=1}^{n} p_i\right)^\beta = n^{1-\beta}.$$

From these inequalities it follows that

$$H^\beta(\eta) - H^\beta(\eta \mid \xi) \leqslant (2^{1-\beta} - 1)^{-1}(n^{1-\beta} - 1) - n^{1-\beta}(2^{1-\beta} - 1)^{-1}[(1-p)^\beta + p^\beta(n-1)^{1-\beta} - 1] = K_\beta .$$

But K_β is the value of the function $H^\beta(\eta) - H^\beta(\eta \mid \xi)$ at the point $(p_1, p_2, ..., p_n) = (1/n, 1/n, ..., 1/n) \in \Delta_n$, from which we obtain

$$K_\beta = \max_{(p_1,...,p_n) \in \Delta_n} [H^\beta(\eta) - H^\beta(\eta \mid \xi)] = C_\beta .$$

Thus, the theorem is proved.

This theorem is a generalization of a known result for the Shannon's capacity (see Fano [7])

$$\lim_{\beta \to 1+} C_\beta = C = \log_2 n - p \log_2(n-1) + p \log_2 p + (1-p) \log_2(1-p).$$

For a binary symmetric channel we have with $n = 2$

$$C_\beta = 1 + (2^{\beta-1} - 1)^{-1} [(1-p)^\beta + p^\beta - 1], \tag{6.8}$$

from which it follows

$$C = \lim_{\beta\to 1+} C_\beta = 1 + p \log_2 p + (1-p) \log_2(1-p). \tag{6.9}$$

A further result is given by

THEOREM 11. *For an arbitrary* $(m \times 2)$ *transition matrix of a discrete constant channel we have*

$$C_2 = 1 - \sum_{k=1}^{m} q_{k1} q_{k2} . \tag{6.10}$$

Proof. We take $(p_1, p_2) = (x, 1-x) \in \Delta_2$, then

$$H^2(\eta) - H^2(\eta \mid \xi)$$
$$= 2 \left\{ \sum_{k=1}^{m} \{(xq_{k1})^2 + [(1-x)\, q_{k2}]^2 - [xq_{k1} + (1-x)\, q_{k2}]^2\} \right.$$
$$\left. + x^2 + (1-x)^2 - 1 \right\} = T(x).$$

It is easily shown that the function $T(x)$ has an unique maximum value at the point $x = 1/2$. From this assertion we obtain

$$C_2 = \max_{(p_1,p_2)\in\Delta_2} [H^2(\eta) - H^2(\eta \mid \xi)]$$
$$= \max_{x\in(0,1)} T(x) = T(\tfrac{1}{2}) = 1 - \sum_{k=1}^{m} q_{k1} q_{k2} .$$

Thus, the theorem is proved.

For a binary symmetric channel ($q_{11} = q_{22} = 1 - p$ and $q_{12} = q_{21} = p$), we get from (6.10)

$$C_2 = 1 - 2p(1-p) = (1-p)^2 + p^2,$$

which is (6.8) in the case $\beta = 2$. It is clear that we can determine C_2, generally, with the method of Lagrange multipliers. The computation C_β is in general an open question. This problem can be solved with an iterative method which is similar to the method for calculation of the Shannon's capacity (see Eisenberg [5]).

RECEIVED: June 3, 1969

REFERENCES

1. T. BORGES, Zur Herleitung der Shannonschen Information, *Math. Z.* **96** (1967), 282–287.
2. Z. DARÓCZY, Az információ Shannon-féle mértékéről (On the Shannon's measure of information) MTA III. Oszt. Közleményei, Budapest, (1968) (in press).
3. Z. DARÓCZY, On the energy of finite probability distributions, *Bull. Math.* (Bucarest) (1969) (in press).
4. Z. DARÓCZY AND I. KÁTAI, Additive zahlentheoretische Funktionen und das Mass der Information, *Ann. Univ. Sci. Budapest, Sect. Math.* (1969) (in press).
5. E. EISENBERG, "On Channel Capacity," Internal Technical Memorandum M-35, Electronics Research Laboratory, University of California, Berkeley, California, 1963.
6. D. K. FADEEV, On the concept of entropy of a finite probabilistic scheme (Russian), *Uspehi Mat. Nauk* 11 (1956), 227–231.
7. R. M. FANO, "Transmission of Information," M.I.T. Press, Cambridge, Mass., 1961.
8. D. G. KENDALL, Functional equations in information theory, *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* 2 (1963), 225–229.
9. P. M. LEE, On the axioms of information theory, *Ann. Math. Statist.* **35** (1964), 414–418.
10. A. RÉNYI, "On Measures of Entropy and Information," Vol. I, p. 547–561, Proc. 4-th Berkeley Symp. Math. Statist. and Probability 1960, Univ. of Calif. Press, Berkeley, Calif., 1961.
11. H. TVERBERG, A new derivation of the information function, *Math. Scand.* **6** (1958), 297–298.