Kybernetika

Inder Jeet Taneja; Fernando Guerra

An algorithm for calculating the channel capacity of degree β

Kybernetika, Vol. 21 (1985), No. 2, 123--133

Persistent URL: `http://dml.cz/dmlcz/124803`

KYBERNETIKA — VOLUME 21 (1985), NUMBER 2

AN ALGORITHM FOR CALCULATING THE CHANNEL CAPACITY OF DEGREE β

INDER JEET TANEJA, FERNANDO GUERRA

Arimoto [2] and Blahut [5] proposed a systematic iteration method to compute the channel capacity of a discrete memoryless channel. Arimoto [3, 4] also presented an iteration method for computing the random coding exponent function and channel capacity of order α by defining the mutual information in terms of Rényi [12] entropy of order α. In this paper, we present an algorithm for computing the channel capacity of degree β by defining mutual information in terms of Havrda and Charvát [9] entropy of degree β. Some upper bounds to the channel capacity of degree β have also been derived.

1. INTRODUCTION

The calculation of the capacity of a discrete memoryless channel is well known problem in information theory since this quantity can not be represented in closed form. In order to calculate the channel capacity for a given channel matrix, we must select the necessary and sufficient number of rows needed for the calculation. This remains to be troublesome problem especially in nonregular (nonsquare) channel matrices. A general method for determining the capacity of a discrete memoryless channel has been suggested by Muroga [11], Cheng [6], and Takano [14]. While Meister and Oettli [10] proposed an iterative procedure based upon the method of concave programming and showed that it converges to capacity. Arimoto [2] and Blahut [5] also proposed another iteration method to compute the capacity which is very simple and systematic. Arimoto [3, 4] also presented an iterative algorithm for computing the random coding exponent function and channel capacity of order α by defining the mutual information in terms of the Rényi [12] entropy of order α.

In this paper, we apply Arimoto's technique (cf. [3]) to obtain an algorithm for computing the channel capacity of degree β in which the mutual information has been defined in terms of Havrda and Charvát [9] entropy of degree β. Some upper bounds to the channel capacity of degree β have also been derived. While, the algorithm for computing the channel capacity using generalized γ-entropy of Arimoto [1] has been presented by Taneja and Wanderlinde [16] and using weighted entropy has been presented by Taneja and Flemming [15].

2. CAPACITY OF DEGREE β

Denote a discrete memoryless channel with n input and m output symbols by the stochastic $m \times n$ matrix $\boldsymbol{Q}$:

$$\boldsymbol{Q} = \{Q_{k/j}\}, \quad k = 1, 2, \ldots, m; \quad j = 1, 2, \ldots, n$$

where $Q_{k/j} \geqq 0$ for all i, j and $\sum_{k=1}^{m} Q_{k/j} = 1$.

Let us denote

$$\Delta_n = \{\boldsymbol{P} = (p_1, p_2, \ldots, p_n) : p_j \geqq 0, \quad \sum_{j=1}^{n} p_j = 1\}$$

and

$$\Delta_n^0 = \{\boldsymbol{P} = (p_1, p_2, \ldots, p_n) : p_j > 0, \quad \sum_{j=1}^{n} p_j = 1\}.$$

The mutual information of degree β of the channel matrix $\boldsymbol{Q}$ is defined by

$$I^\beta(\boldsymbol{Q}; \boldsymbol{P}) = H^\beta(\boldsymbol{P}) - H^\beta(\boldsymbol{Q}; \boldsymbol{P}), \tag{2.1}$$

where

$$H^\beta(\boldsymbol{P}) = (2^{1-\beta} - 1)^{-1} \{\sum_{j=1}^{n} p_j^\beta - 1\}, \quad \beta \neq 1, \quad \beta > 0,$$

and

$$H^\beta(\boldsymbol{Q}; \boldsymbol{P}) = (2^{1-\beta} - 1)^{-1} \{\sum_{k=1}^{m} \sum_{j=1}^{n} (p_j Q_{k/j})^\beta - \sum_{k=1}^{m} (\sum_{j=1}^{n} p_j Q_{k/j})^\beta\}, \tag{2.2}$$

where $H^\beta(\boldsymbol{Q}; \boldsymbol{P})$ is the conditional entropy of degree β as defined in [7].

We define the capacity of degree β of a discrete memoryless channel $\boldsymbol{Q}$ as

$$C^\beta(\boldsymbol{P}) = \max_{\boldsymbol{P} \in \Delta_n} I^\beta(\boldsymbol{Q}; \boldsymbol{P}). \tag{2.3}$$

Let us generalize the concept of conditional entropy of degree β given in (2.2).

Introduce a stochastic matrix $\boldsymbol{\Phi}$ such that

$$\boldsymbol{\Phi} = \{\Phi_{j/k}\}, \quad k = 1, 2, \ldots, m; \quad j = 1, 2, \ldots, n, \tag{2.4}$$

where $\Phi_{j/k} \geqq 0$ for all j, k and $\sum_{j=1}^{n} \Phi_{j/k} = 1$ and generalize the conditional entropy of degree β as

$$J^\beta(\boldsymbol{Q}; \boldsymbol{P}; \boldsymbol{\Phi}) = (2^{1-\beta} - 1)^{-1} \{\sum_{k=1}^{m} \sum_{j=1}^{n} p_j^\beta Q_{k/j}^\beta (1 - \Phi_{j/k}^{1-\beta})\}, \tag{2.5}$$

$$\beta \neq 1, \beta > 0.$$

Then, if $\boldsymbol{\Phi}$ is defined by the Bayes formula:

$$\Phi_{j/k} = \frac{p_j Q_{k/j}}{\sum_{i=1}^{n} p_i Q_{k/i}} = Q_{j/k}^*, \tag{2.6}$$

then (2.5) becomes equal to (2.2).

Furthermore, we can easily prove the inequality

$$J^\beta(Q; P; \Phi) \geqq J^\beta(Q; P; Q^*), \tag{2.7}$$

where Q^* is the stochastic matrix whose (j, k)th entry is $Q_{j/k}$ as defined in (2.6). In view of this fact, one obtains another characterization of channel capacity of degree β as

$$C^\beta(Q) = \max_{P \in \Delta_n} \max_{\Phi \in \mathbf{\Phi}} \{H^\beta(P) - J^\beta(Q; P; \Phi)\}, \tag{2.8}$$

where $\mathbf{\Phi}$ denotes the set of all stochastic matrices satisfying (2.4).

The following proposition can be verified easily.

Proposition 2.1. The function $I^\beta(Q; P)$ is a convex $\cap$ function of the input probabilities for all $0 < \beta \leqq 1$.

Proposition 2.2. The probability vector $P^0 = (p_1^0, p_2^0, \ldots, p_n^0) \in \Delta_n$ maximizes $I^\beta(Q; P)$ for all $\beta \leqq 1$ if and only if

$$(2^{1-\beta} - 1)^{-1} \{p_j^{0\beta-1} - 1 - \sum_{i=1}^{m} Q_{k/j}^\beta p_j^{0\beta-1} + \sum_{k=1}^{m} (\sum_{j=1}^{n} p_j^0 Q_{k/j})^{\beta-1} Q_{k/j}\} \tag{2.9}$$

$$\begin{cases} = C^\beta(Q) & \text{if} \quad p_j^0 > 0 \\ \leqq C^\beta(Q) & \text{if} \quad p_j^0 = 0 \end{cases}$$

Proof. We want to maximize the following function:

$$I^\beta(Q; P) = H^\beta(P) - H^\beta(Q; P) = \tag{2.10}$$

$$= (2^{1-\beta} - 1)^{-1} \{\sum_{j=1}^{n} p_j^\beta - 1 - \sum_{k=1}^{m} \sum_{j=1}^{n} p_j^\beta Q_{k/j}^\beta + \sum_{k=1}^{m} (\sum_{j=1}^{n} p_j Q_{k/j})^\beta\},$$

$$\beta \neq 1, \quad \beta > 0.$$

Let us maximize (2.10) with respect to the condition $\sum_{j=1}^{n} p_j = 1$. Using the Langrange method of multipliers and let

$$f(P) = (2^{1-\beta} - 1)^{-1} \{\sum_{j=1}^{n} p_j^\beta - 1 - \sum_{k=1}^{m} \sum_{j=1}^{n} p_j^\beta Q_{k/j}^\beta + \sum_{k=1}^{m} (\sum_{j=1}^{n} P_j Q_{k/j})^\beta +$$

$$+ \lambda(\sum_{j=1}^{n} p_j - 1)\}$$

then, we have

$$\frac{\partial f(P)}{\partial p_j} = (2^{1-\beta} - 1)^{-1} \{\beta p_j^{\beta-1} - \sum_{k=1}^{m} Q_{k/j}^\beta \beta p_j^{\beta-1} + \sum_{k=1}^{m} (\sum_{j=1}^{n} p_j Q_{k/j})^{\beta-1} Q_{k/j}\} + \lambda = 0,$$

$$\lambda = -\beta(2^{1-\beta} - 1)^{-1} \{p_j^{\beta-1} - \sum_{k=1}^{m} p_j^{\beta-1} Q_{k/j}^\beta + \sum_{k=1}^{m} Q_{k/j} (\sum_{j=1}^{n} p_j Q_{k/j})^{\beta-1}\}. \tag{2.11}$$

By maximization lemma (ref. Gallager [8]), as $I^\beta(Q; P)$ is a convex $\cap$ function of $\boldsymbol{P} = (p_1, p_2, \ldots, p_n) \in \Delta_n$ for $0 < \beta \leqq 1$ and the partial derivatives of $I^\beta(\boldsymbol{Q}; \boldsymbol{P})$ are continuous, then the necessary and sufficient conditions at $\boldsymbol{P}^0 = (p_1^0, p_2^0, \ldots, p_n^0) \in \in \Delta_n$ to maximize $I^\beta(\boldsymbol{Q}; \boldsymbol{P})$ are

$$\frac{\partial I^\beta(\boldsymbol{Q}; \boldsymbol{P})}{\partial p_j^0} \begin{cases} = \lambda & \text{if} \quad p_j^0 > 0 \\ \leqq \lambda & \text{if} \quad p_j^0 = 0 \end{cases} \tag{2.12}$$

Expression (2.11) and (2.12) together give

$$(2^{1-\beta} - 1)^{-1} \{p_j^{0\beta-1} - 1 - \sum_{k=1}^m Q_{k/j} p_j^{0\beta-1} + \sum_{k=1}^m Q_{k/j} (\sum_{j=1}^n p_j^0 Q_{k/j})^{\beta-1}\}$$

$$\begin{cases} = C^\beta(\boldsymbol{Q}) & \text{if} \quad p_j^0 > 0 \\ \leqq C^\beta(\boldsymbol{Q}) & \text{if} \quad p_j^0 = 0 \end{cases} \tag{2.13}$$

where $C^\beta(\boldsymbol{Q}) = (\lambda/\beta) - (2^{1-\beta} - 1)^{-1}$.

Let us prove now that $C^\beta(\boldsymbol{Q})$ is the channel capacity. In order to prove this, multiply (2.13) by p_j^0 and taking sum over all j, $j = 1, 2, \ldots, n$ at which $p_j^0 > 0$, we have

$$I^\beta(\boldsymbol{Q}; \boldsymbol{P}^0) = C^\beta(\boldsymbol{Q}),$$

i.e.,

$$\max_{\boldsymbol{P} \in \Delta_n} I^\beta(\boldsymbol{Q}; \boldsymbol{P}) = C^\beta(\boldsymbol{Q}).$$

3. COMPUTATION OF THE CAPACITY OF DEGREE β

Based upon the double-maximum form in (2.8), an iterative algorithm for computing $C^\beta(\boldsymbol{Q})$ is composed of the following steps:

i) Initially, choose an arbitrary probability vector $\boldsymbol{P}^1 \in \Delta_n^0$ (in practice the uniform distribution $p_j^1 = 1/n$ for all $j = 1, 2, \ldots, n$ generally suitable);

ii) Then, iterate the following steps for $t = 1, 2, \ldots$.

a) Maximize $H^\beta(\boldsymbol{P}^t) - J^\beta(\boldsymbol{Q}; \boldsymbol{P}^t; \boldsymbol{\Phi})$ with respect to $\boldsymbol{\Phi} \in \boldsymbol{\Phi}$ with $\boldsymbol{P}^t$ fixed.

According to (2.7) the maximizing $\boldsymbol{\Phi}$ is

$$\Phi_{j/k}^t = \frac{Q_{k/j} p_j^t}{\sum_{i=1}^n Q_{k/i} p_i^t}, \tag{3.1}$$

i.e.

$$C^\beta(t, t) = \max_{\boldsymbol{\Phi} \in \boldsymbol{\Phi}} \{H^\beta(\boldsymbol{P}^t) - J^\beta(\boldsymbol{Q}; \boldsymbol{P}^t; \boldsymbol{\Phi})\} = H^\beta(\boldsymbol{P}^t) - J^\beta(\boldsymbol{Q}; \boldsymbol{P}^t; \boldsymbol{\Phi}^t); \tag{3.2}$$

b) Maximize $H^\beta(\boldsymbol{P}) - J^\beta(\boldsymbol{Q}; \boldsymbol{P}; \boldsymbol{\Phi}^t)$ with respect to $\boldsymbol{P} \in \Delta_n$ while fixing $\boldsymbol{\Phi}^t$. This maximizing probability vector denoted by $\boldsymbol{P}^{t+1}$ is given by

$$p_j^{t+1} = \frac{(s_j^t)^{\frac{1}{1-\beta}}}{\sum_{i=1}^n (s_i^t)^{\frac{1}{1-\beta}}}, \quad \beta \neq 1, \quad \beta > 0, \tag{3.3}$$

where

$$(3.4)\qquad s_j^t = 1 - \sum_{k=1}^{m} Q_{k/j}^{\beta}\{1 - (\Phi_{j/k}^t)^{1-\beta}\}\,, \quad \beta \neq 1\,, \quad \beta > 0\,.$$

In fact, the following lemma is true:

Lemma 3.1. For any fixed $\boldsymbol{\Phi} \in \mathbf{\Phi}$,

$$(3.5)\qquad \max_{\boldsymbol{P}\in\Delta_n} \{H^{\beta}(\boldsymbol{P}) - J^{\beta}(\boldsymbol{Q}; \boldsymbol{P}; \boldsymbol{\Phi})\} = H^{\beta}(\boldsymbol{P}^*) - J^{\beta}(\boldsymbol{Q}; \boldsymbol{P}^*; \boldsymbol{\Phi}) =$$

$$= (2^{1-\beta} - 1)^{-1}\{(\sum_{j=1}^{n} s_j^{\frac{1}{1-\beta}})^{1-\beta} - 1\} \leqq C^{\beta}(\boldsymbol{Q})\,, \quad 0 < \beta \leqq 1\,.$$

where $\boldsymbol{P}^* \in \Delta_n$ is given by

$$(3.6)\qquad p_j^* = \frac{s_j^{\frac{1}{1-\beta}}}{\sum_{i=1}^{n} s_i^{\frac{1}{1-\beta}}}\,,$$

and

$$(3.8)\qquad s_j = 1 - \sum_{k=1}^{m} Q_{k/j}^{\beta}(1 - \Phi_{j/k}^{1-\beta})\,.$$

Proof. The function which we want to maximize is of the following form:

$$H^{\beta}(\boldsymbol{P}) - J^{\beta}(\boldsymbol{Q}; \boldsymbol{P}; \boldsymbol{\Phi}) \quad \text{with} \quad \boldsymbol{\Phi} \in \mathbf{\Phi} \quad \text{fixed}\,.$$

Using the Lagrange method of multipliers, we have

$$f(\boldsymbol{P}) = H^{\beta}(\boldsymbol{P}) - J^{\beta}(\boldsymbol{Q}; \boldsymbol{P}; \boldsymbol{\Phi}) + (1 - \sum_{j=1}^{n} p_j) =$$

$$= (2^{1-\beta} - 1)^{-1}\{\sum_{j=1}^{n}(p_j^{\beta} - p_j) - \sum_{k=1}^{m}\sum_{j=1}^{n} Q_{k/j}^{\beta} p_j^{\beta} +$$

$$+ \sum_{k=1}^{m}\sum_{j=1}^{n} Q_{k/j}^{\beta} p_j^{\beta} \Phi_{j/k}^{1-\beta}\} + \lambda(1 - \sum_{j=1}^{n} p_j).$$

Now

$$\frac{\partial f(\boldsymbol{P})}{\partial p_j} = (2^{1-\beta} - 1)^{-1}\{\beta p_j^{\beta-1} - 1 - \beta\sum_{k=1}^{m} Q_{k/j}^{\beta} p_j^{\beta-1} + \sum_{\kappa=1}^{m} Q_{k/j}^{\beta} p_j^{\beta-1} \Phi_{j/k}^{1-\beta}\} - \lambda = 0\,.$$

This gives

$$\lambda = (2^{1-\beta} - 1)^{-1}\{\beta p_j^{\beta-1} - 1 - \beta\sum_{k=1}^{m} Q_{\beta}^{k/j} p_j^{\beta-1} + \sum_{k=1}^{m} Q_{k/j}^{\beta} p_j^{\beta-1} \Phi_{j/k}^{1-\beta}\}\,.$$

After simplifying, we get

$$p_j^{\beta-1} s_j = \frac{\lambda(2^{1-\beta} - 1) + 1}{\beta}$$

where s_j is as given in (3.7).

Thus,

$$p_j = \left\{\frac{\lambda(2^{1-\beta} - 1) + 1}{\beta s_j}\right\}^{\frac{1}{\beta-1}}.$$

Using the fact that $\sum_{j=1}^{n} p_j = 1$, we get (3.6).

This completes the proof of the lemma. □

At step iib), let

$$C^\beta(t + 1, t) = \max_{P \in \Delta_n} \{H^\beta(P) - J^\beta(Q; P; \Phi^t)\} = \tag{3.8}$$
$$= H^\beta(P^{t+1}) - J^\beta(Q; P^{t+1}; \Phi^t),$$

and from Lemma 3.1, we have

$$C^\beta(t + 1, t) = (2^{1-\beta} - 1)^{-1} \{[\sum_{j=1}^{n} (s_j^t)^{\frac{1}{1-\beta}}]^{1-\beta} - 1\}, \tag{3.9}$$

where s_j^t is as given in (3.4). Thus, from the definitions of $C^\beta(t, t)$ and $C^\beta(t + 1, t)$, we have following lemma and theorem.

Lemma 3.2.

(3.10)

$$C^\beta(1, 1) \leqq C^\beta(2, 1) \leqq C^\beta(2, 2) \leqq \ldots \leqq C^\beta(t, t) \leqq C^\beta(t + 1, t) \leqq \ldots \leqq C^\beta(Q).$$

Theorem 3.2. Let $P^0 \in \Delta_n$ be any probability vector that achieves the maximum of $I^\beta(Q; P)$. Then for all $0 < \beta \leqq 1$, we have

$$C^\beta(Q) - C^\beta(t + 1, t) \leqq (2^{1-\beta} - 1) \sum_{j=1}^{n} p_j^0 \{(p_j^t)^{\beta-1} - (p_j^{t+1})^{\beta-1}\}. \tag{3.11}$$

Theorem 3.3. The sequences $C^\beta(t, t)$ or $C^\beta(t + 1, t)$ defined in (3.2) and (3.9) respectively converges monotonically from below to $C^\beta(Q)$ as $t \to \infty$ for all $0 < < \beta \leqq 1$.

Proof. From Theorem 3.1, we have

$$C^\beta(Q) - C^\beta(t + 1, t) \leqq (2^{1-\beta} - 1)^{-1} \sum_{j=1}^{n} p_j^0 \{(p_j^0)^{\beta-1} - (p_j^{t+1})^{\beta-1}\}. \tag{3.12}$$

Summing (3.12) from $t = 1$ to $t = N$, we have

$$\sum_{t=1}^{N} \{C^\beta(Q) - C^\beta(t + 1, t)\} \leqq (2^{1-\beta} - 1)^{-1} \sum_{t=1}^{N} \sum_{j=1}^{n} p_j^0 \{(p_j^0)^{\beta-1} - (p_j^{t+1})^{\beta-1}\} = \tag{3.13}$$
$$= (2^{1-\beta} - 1) \sum_{j=1}^{n} p_j^0 \{(p_j^1)^{\beta-1} - (p_j^{N+1})^{\beta-1}\} \leqq (2^{1-\beta} - 1)^{-1} \sum_{j=1}^{n} p_j (p_j^1)^{\beta-1},$$

for all $N \geqq 1$. Note that the right hand side of (3.13) is finite and constant since $P^1 \in \Delta_n^0$. Thus the value $C^\beta(Q) - C^\beta(t + 1, t)$ is nonnegative and nonincreasing with increasing t, this clearly gives

$$\lim_{t \to \infty} C^\beta(t + 1, t) = \lim_{t \to \infty} C^\beta(t, t) = C^\beta(Q).$$ □

Corollary 3.1. The approximation error $e^{\beta}(t) = C^{\beta}(Q) - C^{\beta}(t+1, t)$ is inversely proportional to the number of iterations. In particular, if P^1 is chosen as the uniform distribution, then

$$C^{\beta}(Q) - C^{\beta}(t+1, t) \leqq \frac{(2^{1-\beta}-1)^{-1}\, n^{1-\beta}}{t}.$$

4. UPPER BOUNDS ON THE CAPACITY OF DEGREE β

In this section, we shall derive some properties of $C^{\beta}(Q)$ that give upper bounds on the capacity of degree β.

First, let

$$C^{\beta}(Q; \Phi) = \max_{P \in \Delta_n} \{H^{\beta}(Q) - J^{\beta}(Q; P; \Phi)\} \tag{4.1}$$

and from (3.9), we have

$$C^{\beta}(Q; \Phi) = (2^{1-\beta} - 1)^{-1} \{[\sum_{j=1}^{n} s_j^{\frac{1}{1-\beta}}]^{1-\beta} - 1\}, \quad \beta \neq 1, \beta > 0, \tag{4.2}$$

where s_j is as given in (3.7).

From the Lemma 3.1, we have

$$\max_{\Phi \in \Phi} C^{\beta}(Q; \Phi) = C^{\beta}(Q). \tag{4.3}$$

Moreover, we can prove the following:

Theorem 4.1. Let Q_1 and Q_2 be $m \times n$ channel matrices respectively, α an arbitrary number such that $0 \leqq \alpha \leqq 1$, and Φ and arbitrary $n \times m$ stochastic matrix. Then, we have

$$C^{\beta}(\alpha Q_1 + (1-\alpha) Q_2; \Phi) \leqq \alpha\, C^{\beta}(Q_1; \Phi) + (1-\alpha)\, C^{\beta}(Q_2; \Phi), \tag{4.4}$$

for all $0 < \beta \leqq 1$.

Proof. From (4.2), we have

$$C^{\beta}(\alpha Q_1 + (1-\alpha) Q_2; \Phi) = (2^{1-\beta} - 1)^{-1} \{[\sum_{j=1}^{n} (\alpha s_j^1 + (1-\alpha) s_j^2)^{\frac{1}{1-\beta}}]^{1-\beta} - 1\}, \tag{4.5}$$

where s_j is as given in (3.7).

Now from Minkowski inequality, we have

$$\begin{aligned} &\{\sum_{j=1}^{n} [\alpha s_j^1 + (1-\alpha) s_j^2]^{\frac{1}{1-\beta}}\}^{1-\beta} \lesseqqgtr \\ &\lesseqqgtr \alpha \{\sum_{j=1}^{n} (s_j^1)^{\frac{1}{1-\beta}}\}^{1-\beta} + (1-\alpha) \{\sum_{j=1}^{n} (s_j^2)^{\frac{1}{1-\beta}}\}^{1-\beta}, \end{aligned} \tag{4.6}$$

according as $\beta \gtreqqless 0$. Also

$$(2^{1-\beta} - 1)^{-1} \gtreqqless 0 \tag{4.7}$$

according as $\beta \lesseqqgtr 1$. From (4.6) and (4.7), we have (4.4) for all $0 < \beta \leqq 1$. □

Using (4.3) and (4.4), we can easily prove the following inequality:

Corollary 4.1. For $0 < \beta \leqq 1$, we have

$$C^\beta(\alpha Q_1 + (1-\alpha) Q_2) \leqq \alpha\, C^\beta(Q_1) + (1-\alpha)\, C^\beta(Q_2)\,. \tag{4.8}$$

Theorem 4.2.

(i)
$$C^\beta(Q) \geqq H^\beta\left(\frac{1}{n}, \frac{1}{n}, \dots, \frac{1}{n}\right) - H^\beta\left(\frac{\pi}{n}, 1 - \frac{\pi}{n}\right) + $$
$$+ (2^{1-\beta} - 1)^{-1}\left\{\left[\left(1 - \frac{\pi}{n}\right) + \left(\frac{\pi}{n}\right)(n-1)^{\frac{1-\beta}{\beta}}\right]^\beta - 1\right\}$$

where $\pi = \sum_{k=1}^{m} Q_{k/j_k}$, and j_k denotes one of the integers arbitrarily chosen from 1 to n corresponding to each k.

(ii)
$$C^\beta(Q) \geqq (2^{1-\beta} - 1)^{-1}\left\{\left[\sum_{j=1}^{n}\left(1 - \sum_{k=1}^{m} Q^\beta_{k/j} + \sum_{k=1}^{m} \frac{Q_{k/j}}{(\sum_{i=1}^{n} Q_{k/i})^{1-\beta}}\right)^{\frac{1}{1-\beta}}\right]^{1-\beta} - 1\right\}$$

(iii)
$$C^\beta(Q) \geqq H^\beta\left(\frac{\sum_{j=1}^{n} Q_{(./j)}}{n}\right) - \left(\frac{1}{n}\right)^\beta \sum_{j=1}^{n} H^\beta(Q_{(./j)})\,,$$

where $H^\beta(Q_{(./j)}) = (2^{1-\beta} - 1)^{-1}\{\sum_{k=1}^{m} Q^\beta_{k/j} - 1\}$, is the conditional entropy of degree β of $X = k$ when $Y = j$ is given.

Proof. (i) Let ε be an arbitrary number such that $0 \leqq \varepsilon \leqq 1$ and define

$$p_j = 1/n\,, \quad j = 1, 2, \dots, n \tag{4.9}$$
$$\Phi_{j/k} = \begin{cases} 1 - \varepsilon\,, & j = j_k \\ \dfrac{\varepsilon}{n-1}\,, & j \neq j_k \end{cases}$$

Then from (4.1) and (4.3), we have

$$C^\beta(Q) \geqq H^\beta(P) - J^\beta(Q; P; \Phi) =$$
$$= (2^{1-\beta} - 1)^{-1}\left\{n^{1-\beta} - 1 - \sum_{k=1}^{n}\left(\frac{Q_{k/j_k}}{n}\right)^\beta - \sum_{k=1}^{n}\sum_{j \neq j_k}\left(\frac{Q_{k/j}}{n}\right)^\beta + \right.$$
$$\left. + \sum_{k=1}^{n}\left(\frac{Q_{k/j_k}}{n}\right)^\beta (1-\varepsilon)^{1-\beta} + \sum_{k}\sum_{j = j_k}\left(\frac{Q_{k/j}}{n}\right)^\beta\left(\frac{\varepsilon}{n-1}\right)^{1-\beta}\right\} = \tag{4.10}$$
$$= (2^{1-\beta} - 1)^{-1}\left\{(n^{1-\beta} - 1) - \sum_{k=1}^{n}\left(\frac{Q_{k/j_k}}{n}\right)^\beta [1 - (1-\varepsilon)^{1-\beta}] - \right.$$

$$-\sum_{k=1}^{n}\sum_{j\neq j_k}\left(\frac{Q_{k/j}}{n}\right)^{\beta}\left[1-\left(\frac{\varepsilon}{n-1}\right)^{1-\beta}\right]\Bigg\}\geqq$$

$$\geqq (2^{1-\beta}-1)^{-1}\left\{(n^{1-\beta}-1)-\left(\sum_{k=1}^{n}\frac{Q_{k/j_k}}{n}\right)^{\beta}[1-(1-\varepsilon)^{1-\beta}]-\right.$$

$$\left.-\left(1-\sum_{k=1}^{n}\frac{Q_{k/j_k}}{n}\right)^{\beta}\left[1-\left(\frac{\varepsilon}{n-1}\right)^{1-\beta}\right]\right\}$$

(ref. Gallager [8] for $0<\beta\leqq 1$)

$$=(2^{1-\beta}-1)^{-1}\left\{(n^{1-\beta}-1)-\left(\frac{\pi}{n}\right)^{\beta}[1-(1-\varepsilon)^{1-\beta}]-\left(1-\frac{\pi}{n}\right)^{\beta}\left[1-\left(\frac{\varepsilon}{n-1}\right)^{1-\beta}\right]\right\}.$$

Maximizing right hand side of (4.10) with respect to ε, $0\leqq\varepsilon\leqq 1$, we obtain

$$\varepsilon=\frac{1-\frac{\pi}{n}}{\left(1-\frac{\pi}{n}\right)+\left(\frac{\pi}{n}\right)(n-1)^{\frac{1-\beta}{\beta}}}\,. \tag{4.11}$$

In fact, let

$$F(\varepsilon)=(2^{1-\beta}-1)^{-1}\left\{(n^{1-\beta}-1)-\left(\frac{\pi}{n}\right)^{\beta}[1-(1-\varepsilon)^{1-\beta}]-\right.$$

$$\left.-\left(1-\frac{\pi}{n}\right)^{\beta}\left[1-\left(\frac{\varepsilon}{n-1}\right)^{1-\beta}\right]\right\},$$

then

$$F'(\varepsilon)=\frac{1-\beta}{2^{1-\beta}-1}\left\{-\left(\frac{\pi}{n}\right)^{\beta}(1-\varepsilon)^{-\beta}+\left(1-\frac{\pi}{n}\right)^{\beta}\frac{\varepsilon^{-\beta}}{(n-1)^{1-\beta}}\right\}=0\,,$$

gives (4.11).

Substituting this value of ε from (4.11) in (4.10), we get the required result.

(ii) We have

$$\max_{P\in\Delta_n}\{H^{\beta}(\boldsymbol{P})-J^{\beta}(\boldsymbol{Q};\boldsymbol{P};\boldsymbol{\Phi})\}= \tag{4.12}$$

$$=(2^{1-\beta}-1)^{-1}\{(\sum_{j=1}^{n}s_j^{\frac{1}{1-\beta}})^{1-\beta}-1\}\leqq C^{\beta}(\boldsymbol{Q})\,,$$

where

$$s_j=1-\sum_{k=1}^{m}Q_{k/j}^{\beta}(1-\Phi_{j/k}^{1-\beta})\,. \tag{4.13}$$

Substituting $\Phi_{j/k}=Q_{k/j}/\sum_{i=1}^{m}Q_{k/i}$ in (4.13) and using (4.12), we get the required result.

(iii) From (4.1) and (4.13), we have

$$C^{\beta}(Q) \geqq H^{\beta}(P) - J^{\beta}(Q; P; \Phi) = \tag{4.14}$$

$$= (2^{1-\beta} - 1)^{-1} \{ \sum_{j=1}^{n} p_j^{\beta} - 1 - \sum_{k=1}^{m} \sum_{j=1}^{n} Q_{k/j}^{\beta} p_j^{\beta} + \sum_{k=1}^{m} \sum_{j=1}^{n} Q_{k/j}^{\beta} p_j^{\beta} \Phi_{j/k}^{1-\beta} \} .$$

Substituting in (4.14),

$$p_j = 1/n , \quad j = 1, 2, \ldots, n ,$$

and

$$\Phi_{j/k} = \frac{Q_{k/j}}{\sum_{i=1}^{m} Q_{k/i}} , \quad k = 1, 2, \ldots, m$$

we get

$$C^{\beta}(Q) \geqq (2^{1-\beta} - 1)^{-1} \left\{ n^{1-\beta} - 1 - \sum_{k=1}^{m} \sum_{j=1}^{n} Q_{k/j}^{\beta} \left(\frac{1}{n}\right)^{\beta} + \right.$$

$$\left. + \sum_{k=1}^{m} \sum_{j=1}^{n} Q_{k/j}^{\beta} \left(\frac{1}{n}\right)^{\beta} \left(\frac{Q_{k/j}}{\sum_{i=1}^{m} Q_{k/i}} \right)^{1-\beta} \right\} =$$

$$= (2^{1-\beta} - 1)^{-1} \left\{ -1 - \left(\frac{1}{n}\right)^{\beta} [\sum_{k=1}^{m} \sum_{j=1}^{n} Q_{k/j}^{\beta} - n] + \left(\frac{1}{n}\right)^{\beta} \sum_{k=1}^{m} (\sum_{j=1}^{n} Q_{k/j})^{\beta} \right\} \geqq$$

$$\geqq \frac{\left(\frac{1}{n}\right)^{\beta} \sum_{k=1}^{m} (\sum_{j=1}^{n} Q_{k/j})^{\beta} - 1}{2^{1-\beta} - 1} - \left(\frac{1}{n}\right)^{\beta} \sum_{j=1}^{n} \left\{ \frac{\sum_{k=1}^{m} Q_{k/j}^{\beta} - 1}{2^{1-\beta} - 1} \right\} =$$

$$= H^{\beta} \left(\frac{\sum_{j=1}^{n} \boldsymbol{Q}_{(./j)}}{n} \right) - \left(\frac{1}{n}\right)^{\beta} \sum_{j=1}^{n} H^{\beta}(\boldsymbol{Q}_{(./j)}) ,$$

which completes the proof of part (iii). □

ACKNOWLEDGEMENTS

Authors are thankful to the refree for his valuable suggestions on an earlier version of the paper.

(Received June 6, 1983.)

REFERENCES

[1] S. Arimoto: Information-theoretical considerations on estimation problems. Inform. and Control *19* (1971), 3, 181—194.

[2] S. Arimoto: An algorithm for computing the capacity of arbitrary discrete memoryless channels. IEEE Trans. Inform. Theory *IT-14* (1972), 14—20.

[3] S. Arimoto: Information measures and capacity of order α for discrete memoryless channels. In: Proc. Colloq. on Inform. Theory, Keszthely (Hungary), 1975, 41—52.

[4] S. Arimoto: Computation of random coding exponent functions. IEEE Trans. Inform. Theory *IT-22* (1976), 666—671.

[5] R. E. Blahut: Computation of channel capacity and rate distortion functions. IEEE Trans. Inform. Theory *IT-18* (1972), 460—473.

[6] M. C. Cheng: On the computation of capacity of a discrete memoryless channel. Inform. and Control *24* (1974), 292—298.

[7] Z. Daróczy: Generalized information functions. Inform. and Control *16* (1970), 46—51.

[8] R. G. Gallager: Information Theory and Reliable Communications. J. Wiley and Sons, New York 1968.

[9] J. Havrda and F. Charvát: Quatification method of classification processes: concept of structural a-entropy. Kybernetika *3* (1967), 1, 30—35.

[10] B. Meister and W. Oettli: On the capacity of a discrete, constant channel. Inform. and Control *11* (1967), 341—351.

[11] S. Muroga: On the capacity of a discrete channel, I. J. Phys. Soc. Japan *8* (1953), 484—494.

[12] A. Rényi: On measures of entropy and information. In: Proc. 4th Berkley Symp. Math. Statist. Prob., Univ. of Calif. Press, Berkeley, Calif., vol. I (1951), 547—561.

[13] B. D. Sharma and D. P. Mittal: New non-additive measures of entropy for discrete probability distributions. J. Math. Sci. *10* (1975), 28—40.

[14] S. Takano: On a method of calculating the capacity of a discrete memoryless channel. Inform. and Control *29* (1975), 327—336.

[15] I. J. Taneja and D. M. Flemming: An iterative method for calculating the weighted capacity of a discrete constant channel. 13° Colloq. Brasileira de Met., 20—24 July 1982, Poços de Caldas, MG.

[16] I. J. Taneja and M. J. Wanderlinde: An algorithm for computing the generalized γ-capacity of a discrete memoryless channel. J. Inform. and Optim. Sci. *4* (1983), 3, 269—282.

Professor Inder Jeet Taneja, Professor Fernando Guerra, Departamento de Matématica, Universidade Federal de Santa Catarina, 88.000 — Florianópolis, SC, Brasil.