EFFECTIVE ENTROPY RATE AND TRANSMISSION OF INFORMATION THROUGH CHANNELS WITH ADDITIVE RANDOM NOISE

By K. R. PARTHASARATHY

Indian Statistical Institute

SUMMARY. Transmission of information through channels with additive noise is considered. Coding theorem and its converse are established for these channels with a certain notion of capacity. This capacity is explicitly computed for this class of channels.

1. Introduction

The famous McMillan's theorem regarding ergodic sources can be reformulated as follows. Consider the minimum number of n-length sequences which have a total probability exceeding $1-\epsilon$. If μ is the measure describing the source denote this minimum by $N_n(\epsilon, \mu)$. McMillan's theorem states that for every ergodic source μ the limit $\lim_{n\to\infty} \frac{\log N_n(\epsilon, \mu)}{n}$ exists and is always equal to the entropy rate of the source as defined by Shannon. The question arises as to what happens to the sequence $\frac{\log N_n(\epsilon, \mu)}{n}$ as $n\to\infty$ when the source is not necessarily ergodic but stationary. We show that except for a countable number of ϵ's the limit always exists and in general depends on ϵ. We construct two functions $A(\epsilon)$ and $B(\epsilon)$, $(0 < \epsilon < 1)$ which coincide except on a countable set and both $\underline{\lim}$ and $\overline{\lim}$ of $\frac{\log N_n(\epsilon, \mu)}{n}$ lie between $A(\epsilon)$ and $B(\epsilon)$. Further, as $\epsilon\to 0$, both the functions $A(\epsilon)$ and $B(\epsilon)$ converge to a unique limit $\overline{H}(\mu)$. The precise description of the functional $\overline{H}(\mu)$ is also given.

In the last section we introduce the notion of a channel with additive noise. Here the input and output alphabets coincide with a finite abelian group A and the noise is distributed according to an arbitrary stationary measure on the product space A'. When a message sequence is sent through the channel the noise gets added to the message independently of the message. The disturbed message is received at the output. The binary symmetric channel is a typical example. For the channel with additive noise distributed according to a stationary measure μ, we consider $M_n(\epsilon, \mu)$, the supremum of the length of all possible codes with probability of error less than or equal to ϵ (for transmission of messages during the time period $1, 2, \ldots, n$). A code of length N and probability of error less than or equal to ϵ is defined in the sense of Wolfowitz (1961). Then we analyse the asymptotic behaviour of the sequence $\frac{\log M_n(\epsilon, \mu)}{n}$. We show that the limit of this sequence exists for all ϵ except on a countable set. We also show that the $\underline{\lim}$ and $\overline{\lim}$ of this sequence lie between $\log a - A(\epsilon)$ and $\log a - B(\epsilon)$, where $A(\epsilon)$ and $B(\epsilon)$ are the functions mentioned in the previous paragraph and a is the number of elements in the alphabet A. As $\epsilon\to 0$

$\log a - A(\epsilon)$ and $\log a - B(\epsilon)$ converge to the same limit $\log a - \bar{H}(\mu)$. Thus in this case the capacity of the channel is not described by a single number but by the two functions $A(\epsilon)$ and $B(\epsilon)$.

The idea of studying the asymptotic properties of the sequence $\frac{\log N_n(\epsilon, \mu)}{n}$ is due to Winkelbauer. He gave the description of the function

$$\bar{H}(\mu) = \lim_{\epsilon \to 0} \overline{\lim_{n \to \infty}} \frac{\log N_n(\epsilon, \mu)}{n}$$

in terms of the entropies of the ergodic components of μ. It was stated by him without proof in his lecture at the Indian Statistical Institute.

2. Preliminaries

Throughout this paper A will denote a finite alphabet, A' the space of sequences of elements from A, T the shift transformation in A' and μ a measure which is defined on the usual σ-field of A' and invariant under T. We denote by $[x_1, x_2, \ldots, x_n]$ the cylinder set in A' of all sequences whose i-th coordinate is x_i for $i = 1, 2, \ldots, n$. Any n-length sequence $x_1, x_2, \ldots, x_n$ is referred to as a u-sequence. We denote by $N_n(\epsilon, \mu)$ the smallest number of u-sequences whose total probability is greater than or equal to $1-\epsilon$. This smallest set may not be unique. We choose one of them arbitrarily and denote it by $A_n(\epsilon, \mu)$.

If we assign the discrete topology to A and the product topology to A' then A' becomes a compact metric space. We shall now follow the notation of Oxtoby (1952). If $f(p)$ is a real valued function on A', let

$$M(f, p, k) = f_k(p) = \frac{1}{k} \sum_{i=1}^{k} f(T^i p) \qquad (k = 1, 2, \ldots)$$

and

$$M(f, p) = f^*(p) = \lim_{k \to \infty} M(f, p, k)$$

in case this limit exists. A Borel subset E of A' is said to have invariant measure one if $\mu(E) = 1$ for every invariant probability measure μ. Let Q be the set of points p for which $M(f, p)$ exists for every $f \in C(A')$ where $C(A')$ is the space of continuous functions on A'. It follows easily from Riesz's representation theorem that corresponding to any point $p \in Q$ there exists a unique invariant probability measure μ_p such that

$$M(f, p) = \int f \, d\mu_p.$$

Let $R \subset Q$ be the set of those points for which μ_p is ergodic. R is called the set of regular points. Then we have the following representation theorem of Kryloff and Bogolíouboff which can be found in Oxtoby (1952).

Theorem 2.1 : *The set R of regular points is a Borel measurable set of invariant measure one. For any Borel set $E \subset A'$, $\mu_p(E)$ is Borel measurable on R and*

$$\mu(E) = \int_R \mu_p(E) \, d\mu(p)$$

for any invariant probability measure μ.

Let $H(\mu)$ denote the entropy of any invariant probability measure μ.

Let

$$\overline{H}(\mu) = \text{ess sup} \quad H(\mu_p) \qquad \dots \quad (2.1)$$

$$\underline{H}(\mu) = \text{ess inf} \quad H(\mu_p) \qquad \dots \quad (2.2)$$

where the essential supremum and the essential infimum are taken relative to μ and μ_p denotes the ergodic measure corresponding to the regular point p.

3. Asymptotic properties of the function $N_n(\epsilon, \mu)$

In this section we shall prove the following theorem and two corollaries.

Theorem 3.1 : *Let* $[A', \mu]$ *be an arbitrary stationary source,*

then

$$A(\epsilon) \leqslant \underline{\lim} \frac{\log N_n(\epsilon, \mu)}{n} \leqslant \overline{\lim} \frac{\log N_n(\epsilon, \mu)}{n} \leqslant B(\epsilon) \qquad \dots \quad (3.1)$$

where

$$A(\epsilon) = \lim_{\downarrow} \eta(\delta), \qquad \dots \quad (3.2)$$

$$B(\epsilon) = \lim_{\delta \uparrow \epsilon} \eta'(\delta), \qquad \dots \quad (3.3)$$

$\eta(\delta)$ *is the greatest number with the property*

$$\mu[p : H(\mu_p) \geqslant \eta] \geqslant \delta$$

and $\eta'(\delta)$ *is the smallest number with the property*

$$\mu[p : H(\mu_p) \leqslant \eta'] \geqslant 1-\delta.$$

Corollary 3.1 : *For any stationary source* $[A', \mu]$. $\lim_{n\to\infty} \frac{\log N_n(\epsilon, \mu)}{n}$ *exists for all* $0 < \epsilon < 1$ *except for a countable set.*

Corollary 3.2 : *For any stationary source* $[A', \mu]$,

$$\lim_{\epsilon\to 0} \underline{\lim}_{n\to\infty} \frac{\log N_n(\epsilon, \mu)}{n} = \lim_{\epsilon\to 0} \overline{\lim}_{n\to\infty} \frac{\log N_n(\epsilon, \mu)}{n} = \overline{H}(\mu).$$

Before proceeding to the proof of Theorem 3.1 we need to establish two lemmas.

Lemma 3.1 : *For any stationary source* $[A', \mu]$, *the limit*

$$\lim_{n\to\infty} -\frac{1}{n} \log \mu[x_1 \dots x_n] = g_\mu(x)$$

exists in measure and $\qquad g_\mu(p) = H(\mu_p)$ a.e. $p\,(\mu)$.

Proof: The existence of the limit is the famous McMillan's theorem. In the course of Khinchin's proof of McMillan's theorem (1957) it can be seen that $g_\mu(x)$ can be obtained as follows. Define $h_\mu(x)$ as the conditional probability of x_0 given $x_{-1}, x_{-2}, \dots$ under μ. (Here it is assumed that $x = (\dots x_{-1}, x_0, x_1, \dots)$).

Then

$$g_\mu(x) = \lim_{n\to\infty} \frac{h_\mu(x)+h_\mu(Tx)+\ldots+h_\mu(T^{n-1}x)}{n} \quad \text{a.e. } (\mu).$$

But by Theorem 2.6 of the author (1961),

we have $$h_\mu(x) = h_{\mu_p}(x) \quad \text{a.e.} \quad x(\mu_p)$$

for almost all $p(\mu)$.

Further $$\lim_{n\to\infty} \frac{h_{\mu_p}(x)+h_{\mu_p}(Tx)+\ldots+h_{\mu_p}(T^{n-1}x)}{n} = H(\mu_p) \text{ a.e. } (\mu_p)$$

for any regular point p.

Thus $$g_\mu(x) = H(\mu_p) \quad \text{a.e.} \quad x(\mu_p)$$

for almost all $p(\mu)$. From the Kryloff-Bogoliouboff theory of regular points in a dynamical system we have

$$\mu_x = \mu_p \qquad \text{a.e.} \quad x(\mu_p).$$

Thus $$g_\mu(x) = H(\mu_x) \quad \text{a.e.} \quad (\mu_p)$$

for almost all $p(\mu)$. Thus the set $E = [x : g_\mu(x) \neq H(\mu_x)]$ has measure zero under μ_p for almost all $p(\mu)$. An application of Theorem 2.1 shows that

$$\mu(E) = \int \mu_p(E) d\mu(p) = 0.$$

This completes the proof of Lemma 3.1.

Lemma 3.2 : *For any stationary source* $[A', \mu]$ *and any* $\epsilon > 0$

$$\underline{H}(\mu) \leqslant \varliminf_{n\to\infty} \frac{\log N_n(\epsilon, \mu)}{n} \leqslant \varlimsup_{n\to\infty} \frac{\log N_n(\epsilon, \mu)}{n} \leqslant \bar{H}(\mu).$$

Proof : From Lemma 3.1 it is clear that

$$\underline{H}(\mu) \leqslant g_\mu(x) \leqslant \bar{H}(\mu)$$

with probability one. Thus if we write, for any fixed $\eta > 0$,

$$\mu[x : -\frac{1}{n} \log \mu(x_1 \ldots x_n) \geqslant \underline{H}(\mu)-\eta] = 1-\delta_n \qquad \ldots \quad (3.4)$$

then $\delta_n \to 0$ as $n\to\infty$. The complement of the set written within braces in (3.4) has probability δ_n. Any set with probability $> 1-\epsilon$ must have a subset with probability greater than $1-\epsilon-\delta_n$ whose elements satisfy the inequality within braces in (3.4). Suppose this subset has N' u-sequences. For these sequences

$$\mu[x_1, x_2, \ldots, x_n] \leqslant 2^{-n(\underline{H}(\mu)-\eta)}.$$

Summing up over this subset, we get

$$1-\epsilon-\delta_n \leqslant N' 2^{-n(\underline{H}(\mu)-\eta)}.$$

Thus
$$1-\epsilon-\delta_n \leqslant N_n(\epsilon, \mu) 2^{-n(\underline{H}(\mu)-\eta)}.$$

Since $\delta_n \to 0$ as $n \to \infty$ after some stage

$$1-\epsilon-\delta_n \geqslant \frac{1-\epsilon}{2}.$$

Thus
$$\frac{\log N_n(\epsilon, \mu)}{n} \geqslant \frac{\log (1-\epsilon)/2}{n} + \underline{H}(\mu) - \eta$$

which implies
$$\underline{\lim} \frac{\log N_n(\epsilon, \mu)}{n} \geqslant \underline{H}(\mu) - \eta.$$

Since η is arbitrary we have

$$\underline{\lim} \frac{\log N_n(\epsilon, \mu)}{n} \geqslant \underline{H}(\mu).$$

In order to prove the other inequality, consider, for any fixed $\eta > 0$, the sequence of numbers

$$\mu[x : -\frac{1}{n} \log \mu(x_1 \ldots x_n) \leqslant \underline{H}(\mu) + \eta] = 1 - \delta_n'. \qquad \ldots \quad (3.5)$$

By Lemma 3.1 and (2.1) we have $\lim_{n \to \infty} \delta_n' = 0.$

Thus there exists a subset A_n of u-sequences satisfying the inequality within braces in (3.5) whose probability exceeds $1-\epsilon$ for all sufficiently large n. If the inequality within braces in (3.5) is satisfied, then

$$\mu[x_1 \ldots x_n] > 2^{-n(\bar{H}(\mu)+\eta} \qquad \ldots \quad (3.6)$$

If N' u-sequences satisfying (3.6) are required to make up a probability greater than $1-\epsilon$ then

$$N_n(\epsilon, \mu) \leqslant N' \leqslant 2^{n(\bar{H}(\mu)+\eta)}$$

Thus
$$\overline{\lim} \frac{\log N_n(\epsilon, \mu)}{n} \leqslant \bar{H}(\mu) + \eta.$$

The arbitrariness of η implies the validity of the lemma.

We shall now turn to the proof of Theorem 3.1. Choose any $\delta > \epsilon$. Choose the largest η such that

$$\mu[p : H(\mu_p) \geqslant \eta] \geqslant \delta.$$

Let it be $\eta(\delta)$.

If $$E = [p : H(\mu_p) \geqslant \eta(\delta)]$$

then $$\mu(E) \geqslant \delta.$$

Define $$\mu_1(B) = \frac{\mu(B \cap E)}{\mu(E)}, \qquad \mu_2(B) = \frac{\mu(B \cap E')}{\mu(E')}.$$

We assume that $\mu(E') > 0$.

Then $$\mu = a\mu_1 + (1-a)\mu_2$$

where $$a = \mu(E) \geqslant \delta > \epsilon.$$

If we consider the set $A_n(\epsilon, \mu)$ (the smallest set of u-sequences with probability $> 1-\epsilon$)

then $$\mu_1(A_n(\epsilon, \mu)) \geqslant \frac{1-\epsilon-(1-a)}{a} \geqslant \frac{a-\epsilon}{a} \geqslant \frac{\delta-\epsilon}{a}.$$

Thus $$N_n(\epsilon, \mu) \geqslant N_n\left(1-\frac{\delta-\epsilon}{a}, \mu_1\right)$$

and $$0 < 1-\frac{\delta-\epsilon}{a} < 1.$$

An application of Lemma 3.1 shows that

$$\underline{\lim} \frac{\log N_n(\epsilon, \mu)}{n} \geqslant \underline{H}(\mu_1) \geqslant \eta(\delta).$$

If $\mu(E') = 0$ this inequality is trivially valid. Since δ is any number $> \epsilon$ and $\eta(\delta)$ increases to $A(\epsilon)$ as δ descends to ϵ

we have $$\underline{\lim} \frac{\log N_n(\epsilon, \mu)}{n} \geqslant A(\epsilon). \qquad \ldots \quad (3.7)$$

For proving the other inequality choose $\delta < \epsilon$ and then choose the smallest η' with the property

$$\mu[p : H(\mu_p) \leqslant \eta'] \geqslant 1-\delta.$$

Let it be $\eta'(\delta)$.

Let $$F = [p : H(\mu_p) \leqslant \eta'(\delta)],$$

and $$\mu_1'(B) = \frac{\mu(B \cap F)}{\mu(F)}, \qquad \mu_2'(B) = \frac{\mu(B \cap F')}{\mu(F')}.$$

Then $$\mu = b\mu_1' + (1-b)\mu_2'$$

where $b = \mu(F) \geqslant 1-\delta$. If $\mu_1'(B) > \frac{1-\epsilon}{b}$ then $\mu(B) > 1-\epsilon$.

Thus
$$N_n(\epsilon, \mu) \leqslant N_n\left(1-\frac{1-\epsilon}{b}, \mu_1'\right),$$

$$0 < 1-\frac{1-\epsilon}{b} < 1.$$

Thus by Lemma 3.2
$$\overline{\lim}\ \frac{\log N_n(\epsilon, \mu)}{n} \leqslant \bar{H}(\mu_1') \leqslant \eta'(\delta).$$

As δ increases to ϵ, $\eta'(\delta)$ decreases to a limit $B(\epsilon)$.

Thus
$$\overline{\lim}\ \frac{\log N_n(\epsilon, \mu)}{n} \leqslant B(\epsilon). \qquad \ldots \quad (3.8)$$

(3.7) and (3.8) complete the proof of Theorem 3.1. Corollary 3.1 is an immediate consequence of the fact that in the real line there cannot be more than a countable collection of disjoint open intervals and hence $A(\epsilon) = B(\epsilon)$ except for a countable set. Corollary 3.2 follows immediately from the fact that $A(\epsilon)$ and $B(\epsilon)$ converge to $\bar{H}(\mu)$ as $\epsilon \to 0$.

Remark: From Theorem 3.1 and Corollary 3.2 it is clear that the number $\bar{H}(\mu)$ defined by (2.1) can be rightly called the effective entropy rate of the stationary source $[A', \mu]$. The result that

$$\lim_{\epsilon \to 0}\ \overline{\lim_{n\to\infty}}\ \frac{\log N_n(\epsilon, \mu)}{n} = \bar{H}(\mu)$$

is due to K. Winkelbauer. It was stated by him without proof in one of his lectures at the Indian Statistical Institute. It was his conjecture that

$$\lim_{n\to\infty} \frac{\log N_n(\epsilon, \mu)}{n} \text{ exists for every } \epsilon.$$

4. Channels with additive noise

In this section we introduce the notion of a stationary channel with additive noise, define its capacity and prove the coding theorem as well as its converse.

Let the input and output alphabets of a channel coincide with a finite abelian group A. In a natural way the space A' becomes an abelian group. We denote by $+$ and $-$ the addition and inverse operation in the group A'. For any set E we write

$$E-x = [z : z \in A', z+x \in E].$$

Let μ_0 be an invariant measure defined on A'. Then the probability distributions

$$\nu_x(F) = \mu(F-x) \qquad \ldots \quad (4.1)$$

where F is any set in the usual σ-field of A' and x is any point in A' define a stationary channel. The distribution at the output of this channel corresponding to any input distribution λ is obtained by convoluting λ with μ. Even if μ is ergodic this channel need not be of finite memory in the sense of Feinstein. If the group A consists of two elements 0 and 1 only, the addition is done modulo 2 and μ is the product measure obtained by assuming probability p for one, then we get the so-called binary symmetric channel. We shall call a channel whose distributions are specified by (4.1) as a channel with additive noise and noise distribution μ.

A code with error ϵ and length N is a collection of u-sequences $u_1, u_2, \ldots, u_N$ and sets $V_1, V_2, \ldots, V_N$ of u-sequences with the properties

$$\text{(i)} \quad \mu(V_i - u_i) > 1-\epsilon$$

$$\text{(ii)} \quad V_i \cap V_j = \phi \text{ for } i \neq j. \qquad \ldots \quad (4.2)$$

Let $M_n(\epsilon, \mu)$ be the maximal length possible for a code with error ϵ for a channel with additive noise and noise distribution μ. Then we have the following theorem.

Theorem 4.1 : *For a stationary channel with additive noise and noise distribution* μ

$$\log_2 a - B(\epsilon) \leqslant \underline{\lim} \frac{\log M_n(\epsilon, \mu)}{n} \leqslant \overline{\lim} \frac{\log M_n(\epsilon, \mu)}{n} \leqslant \log_2 a - A(\epsilon)$$

where $A(\epsilon)$ *and* $B(\epsilon)$ *are the functions occurring in the statement of Theorem* 3.1.

Corollary 4.1 : *Except for a countable set of* ϵ*'s the limit*

$$\lim_{n\to\infty} \frac{\log M_n(\epsilon, \mu)}{n} = \log_2 a - A(\epsilon) = \log_2 a - B(\epsilon)$$

exists. Further

$$\lim_{\epsilon\to 0} \log a - A(\epsilon) = \lim_{\epsilon\to 0} \log a - B(\epsilon) = \log a - \overline{H}(\mu).$$

Remark : Corollary 4.1 justifies our calling $\log a - \overline{H}(\mu)$ as the capacity of the additive channel with noise distribution μ.

Proof of Theorem 4.1 : Let $u_1, u_2, \ldots, u_N, V_1, \ldots, V_N$ be a code with error ϵ. For any set E of u-sequences let $m(E)$ be the number of u-sequences in E. By property (i) of (4.2) we have

$$m(V_i) \geqslant N_n(\epsilon, \mu). \qquad \ldots \quad (4.3)$$

For any $\delta > 0$ and all sufficiently large n, we have by Theorem 3.1 and (4.3)

$$m(V_i) \geqslant 2^{n[A(\epsilon)-\delta]}.$$

Since V_i's are disjoint and the total number of u-sequences is a^n, we obtain

$$a^n \geqslant m(\bigcup_1^N V_i) \geqslant N \cdot 2^{n[A(\epsilon)-\delta]}.$$

Thus $$\frac{\log N}{n} \leqslant \log a - A(\epsilon) + \delta. \quad \ldots \quad (4.4)$$

Since (4.4) is true for any code of error ϵ, we get

$$\frac{\log M_n(\epsilon, \mu)}{n} \leqslant \log a - A(\epsilon) + \delta.$$

Allowing n to tend to infinity and then noting the arbitrariness of δ, we get

$$\overline{\lim} \frac{\log M_n(\epsilon, \mu)}{n} \leqslant \log a - A(\epsilon).$$

For proving the other inequality in the theorem we follow Takano (1957). Let $\epsilon' > 0$ be an arbitrary number less than ϵ. Choose the set V_1 to be the set with the smallest nun)f u-sequences whose probability exceeds $1-\epsilon'$. In the notation given in Section 1, $V_1 = A_n(\epsilon', \mu)$. Let u_1 be the u-sequence all of whose elements coincide with the identity of the group A. Now choose u_2 such that

$$\mu[(V_1+u_2)V_1' - u_2] > 1-\epsilon$$

where the prime is used to denote the complement. If no such u_2 exists stop. Write

$$V_2 = (V_1+u_2)V_1'.$$

Choose u_3 such that

$$\mu[(V_1+u_3)V_2'V_1' - u_3] > 1-\epsilon.$$

If no such u_3 exists stop. At the r-th stage u_r is chosen such that

$$\mu[(V_1+u_r)V_{r-1}'V_{r-2}' \ldots V_1' - u_r] > 1-\epsilon.$$

Then we write $$V_r = (V_1+u_r)V_{r-1}' V_{r-2}' \ldots V_1'.$$

Let the process terminate after N stages. Let $V = \bigcup_i V_i$. Then we have u-sequences $u_1, u_2, \ldots u_N$ and sets $V_1, V_2, \ldots, V_N$ of u-sequences with the properties

(1) $\mu(V_i - u_i) > 1-\epsilon$,

(2) $V_i \subseteq V_1 + u_i$,

(3) $V_i \cap V_j = \phi$,

(4) $V = \bigcup_i V_i = \bigcup_i (V_1+u_i)$,

(5) For any u-sequence $\mu((V_1+u)V' - u) \leqslant 1-\epsilon$.

Since V_1 has probability greater than $1-\epsilon'$, we have from property (5) above,

$$\begin{aligned} 1-\epsilon' \leqslant \mu(V_1) &= \mu(V_1+u-u) \\ &\leqslant \mu((V_1+u)V-u) + \mu((V_1+u)V'-u) \\ &\leqslant \mu(V-u) + (1-\epsilon). \end{aligned}$$

This inequality can be rewritten as

$$\mu(V-u) \geqslant \epsilon-\epsilon'. \qquad \ldots \quad (4.5)$$

Since (4.5) is true for every u, we obtain by multiplying both sides of (4.5) by a^{-n} and adding over all the u-sequences

$$m(V)\, a^{-n} \geqslant \epsilon-\epsilon'$$

or

$$m(V) \geqslant a^{n}.\,(\epsilon-\epsilon'). \qquad \ldots \quad (4.6)$$

Since

$$V = \bigcup_{1}^{N} V_i \text{ and } V_i \subset V_1+u_i,$$

we have

$$m(V) \leqslant N.m\,(V_1). \qquad \ldots \quad (4.7)$$

We recall that $V_1 = A_n(\epsilon', \mu)$. By an application of Theorem 3.1 we have, for any $\delta > 0$ and all sufficiently large n

$$m(V_1) \leqslant 2^{n[B(\epsilon')+\delta]}. \qquad \ldots \quad (4.8)$$

Combining (4.6), (4.7) and (4.8)

$$N \geqslant (\epsilon-\epsilon')\cdot a^{n}\cdot 2^{-n[B(\epsilon')+\delta]}.$$

Since $M_n(\epsilon, \mu) \geqslant N$ we have

$$\frac{\log M_n(\epsilon, \mu)}{n} \geqslant \frac{\log(\epsilon-\epsilon')}{n} + [\log a - B(\epsilon')]-\delta.$$

Allowing $n \to \infty$ and then $\delta \to 0$, we get

$$\varliminf_{n\to\infty} \frac{\log M_n(\epsilon, \mu)}{n} \geqslant \log a - B(\epsilon').$$

From the definition of the function $B(\epsilon)$ we see that it is left continuous. Since ϵ' is any number less than ϵ, we get by letting ϵ' increase to ϵ

$$\varliminf_{n\to\infty} \frac{\log M_n(\epsilon, \mu)}{n} \geqslant \log a - B(\epsilon).$$

Corollary 4.1 is an immediate consequence of Corollaries 3.1 and 3.2.

References

Khinchin, A. I. (1957): *Mathematical Foundations of Information Theory*, Dover Publications, New York.

Oxtoby, J. C. (1952): Ergodic sets. *Bull. Amer. Math. Soc.*, 58, 116-136.

Parthasarathy, K. R. (1961): On the integral representation of the rate of transmission of a stationary channel. *Illinois J. Math.*, 5, 299-305.

Takano, K. (1957): On the basic theorems of information theory. *Ann. Inst. Stat. Math.*, 9, 53-77.

Wolfowitz, J. (1961): *Coding Theorems of Information Theory*, Springer-Verlag, Berlin.

Paper received : September, 1962.