Pacific Journal of Mathematics

A NOTE ON GENERALIZATIONS OF SHANNON-MCMILLAN THEOREM

SHU-TEH CHEN MOY

Vol. 11, No. 4 1961

A NOTE ON GENERALIZATIONS OF SHANNON-MCMILLAN THEOREM

SHU-TEH C. MOY

1. Introduction. This paper is a sequel to an earlier paper [6]. All notations in [6] remain in force. As in [6] we shall consider tw probability measures μ, ν an the infinite product σ-algebra of subsets of the infinite product space $\Omega = \pi X$. ν is assumed to be *stationary* and μ to be *Markovian with stationary transition probabilities.* Extensions to K-Markovian μ are immediate. $\nu_{m,n}$, the contraction of ν to $\mathscr{F}_{m,n}$, is assumed to be *absolutely continuous* with respect to $\mu_{m,n}$, the contraction of μ to $\mathscr{F}_{m,n}$, and $f_{m,n}$ is the Radon-Nikodym derivative. In [6] the following theorem is proved. If $\int \log f_{0,0}\, d\nu < \infty$ and if there is a number M such that

$$(1) \qquad \int (\log f_{0,n} - \log f_{0,n-1}) d\nu \leqq M \text{ for } n = 1, 2, \cdots$$

then $\{n^{-1} \log f_{0,n}\}$ converges in $L_1(\nu)$. (1) is also a necessary condition for the $L_1(\nu)$ convergence of $\{n^{-1} \log f_{0,n}\}$. We consider this theorem as a generalization of the Shannon-McMillan theorem of information theory. In the setting of [6] the Shannon-McMillan theorem may be stated as follows. Let X be a finite set of K points. Let ν be any stationary probability measure of $\mathscr{F}$, and μ the equally distributed independent measure on $\mathscr{F}$. Then $\{n^{-1} \log f_{0,n}\}$ converges in $L_1(\nu)$. In fact, the $P(x_0, x_1, \cdots, x_n)$ of Shannon-McMillan is equal to $K^{(n+1)} f_{0,n}$. The convergence with probability one of $\{n^{-1} \log P(x_0, \cdots, x_n)\}$ for a finite set X was proved by L. Breiman [1] [2]. K.L. Chung then extended Breiman's result to a countable set X. [3]. In this paper we shall prove that the convergence with ν-probability one of $\{n^{-1} \log f_{0,n}\}$ follows from the following condition.

$$(2) \qquad \int \frac{f_{0,n}}{f_{0,n-1}} d\nu \leqq L, \; n = 1, 2, \cdots .$$

(2) is a stronger condition than (1) since by Jensen's inequality

$$\log \int \frac{f_{0,n}}{f_{0,n-1}} d\nu \geqq \int \log \frac{f_{0,n}}{f_{0,n-1}} d\nu .$$

An application to the case of countable X is also discussed.

Received November 28, 1960.

2. The convergence theorem. As was proved in [6], condition (1) implies the $L_1(\nu)$ convergence of $\{\log f_{-k,0} - \log f_{-k,-1}\}$ ([6] Theorem 1, 4). The convergence with ν-probability one is automatically true ([6] Theorem 3). Applying a theorem (with obvious modification for T not necessarily ergodic) of Breiman ([1], Theorem 1) the convergence with ν-probability one of $\{n^{-1}\log f_{0,n}\}$ follows from the condition

$$\int \sup_{k\geq 1} |\log f_{-k,0} - \log f_{-k,-1}|\, d\nu < \infty \,. \tag{3}$$

We shall now investigate conditions under which (3) is valid.

Lemma 1. *The following inequality is always true.*

$$\int \sup_{k\geq 1} \log \frac{f_{-k,-1}}{f_{-k,0}} d\nu < \infty \,. \tag{4}$$

Proof. Let $\nu'_{-k,0}$ be as in Lemma 1 [6]. Then

$$\nu_{-k,0} \ll \nu'_{-k,0} \ll \mu_{-k,0}$$

and

$$\frac{d\nu_{-k,0}}{d\nu'_{-k,0}} = \frac{f_{-k,0}}{f_{-k,-1}},\ \frac{d\nu'_{-k,0}}{d\mu_{-k,0}} = f_{-k,-1} \,.$$

Since μ is Markovian, $\nu'_{-k,0}$ are consistent for $k = 1, 2, \cdots$. We shall prove (4) under the assumption that there is a probability measure ν' on $\mathscr{F}_{-\infty,0}$ which is an extension of $\nu'_{-k,0}$ for $k = 1, 2, \cdots$. We shall also prove Lemma 2 under this assumption. If no such ν' exists, the usual procedure of representing Ω into the space of real sequences may be used and the same conclusion follows (cf. the proof of Theorem 4[6]).

Let m be a nonnegative integer and

$$E(m) = [\sup_{k\geq 1} \log \frac{f_{-k,-1}}{f_{-k,0}} > m] \,,$$

$$E_k(m) = [\sup_{1\leq j<k} \log \frac{f_{-j,-1}}{f_{-j\,0}} \leq m,\ \log \frac{f_{-k,-1}}{f_{-k,0}} > m] \,.$$

On $E_k(m)$ we have

$$f_{-k,0} \leqq 2^{-m} f_{-k,-1} \,.$$

Hence

$$\int_{E_k(m)} f_{-k,0}\, d\mu \leqq 2^{-m} \int_{E_k(m)} f_{-k,-1}\, d\mu$$

so that

$$\nu[E_k(m)] \leqq 2^{-m}\nu'[E_k(m)] .$$

Therefore

$$\nu[E(m)] \leqq 2^{-m}\nu'[E(m)] \leqq 2^{-m}$$

and

$$\int \sup_{k>1} \log \frac{f_{-k,-1}}{f_{-k,0}} d\nu \leqq \sum_{m\geqq 0} \nu[E(m)] \leqq \sum_{m\geqq 0} 2^{-m} < \infty .$$

Note that (4) is proved without assuming the integrability of either $\log f_{-k,0}$ or $\log f_{-k,-1}$ or $\log \dfrac{f_{-k,0}}{f_{-k,-1}}$.

LEMMA 2. *If there is a number L such that*

$$\int \frac{f_{-k,0}}{f_{-k,-1}} d\nu \leqq L \text{ for } k = 1, 2, \cdots \tag{5}$$

then

$$\int \sup_{k\geqq 1} \log \frac{f_{-k,0}}{f_{-k,-1}} d\nu < \infty . \tag{6}$$

Proof. It is clear that

$$\int \frac{f_{-k,0}}{f_{-k,-1}} d\nu = \int \left(\frac{f_{-k,0}}{f_{-k,-1}}\right)^2 d\nu'$$

where ν' is defined in the proof of Lemma 1.

Since $\{f_{-k,0}/f_{-k,-1},\ k = 1, 2, \cdots\}$ is a ν'-martingale, $\{(f_{-k,0}/f_{-k,-1})^2,\ k = 1, 2, \cdots\}$ is a ν'-semi-martingale. Hence (5) implies that

$$\nu_{-\infty,0} \ll \nu',\ \int \left(\frac{d\nu_{-\infty,0}}{d\nu'}\right)^2 d\nu' < \infty,\ \left(\frac{f_{-k,0}}{f_{-k,-1}}\right)^2$$

are uniformly ν'-integrable and $\{(f_{-1,0}/f_{-1,-1})^2, (f_{-2,0}/f_{-2,-1})^2 \cdots, (d\nu_{-\infty,0}/d\nu')^2\}$ is a ν'-semi-martingale (Theorem 4.1s, pp. 324[5]).

Hence for any set F defined by $x_0, x_{-1}, \cdots, x_{-k}$

$$\int_F \left(\frac{f_{-k,0}}{f_{-k,-1}}\right)^2 d\nu' \leqq \int_F \left(\frac{f_{-(k+1),0}}{f_{-(k+1),-1}}\right)^2 d\nu' \leqq \int_F \left(\frac{d\nu_{-\infty,0}}{d\nu^1}\right)^2 d\nu'$$

so that

$$\int_F \frac{f_{-k,0}}{f_{-k,-1}} d\nu \leqq \int_F \frac{f_{-(k+1),0}}{f_{-(k+1),-1}} d\nu \leqq \int_F \frac{d\nu_{-\infty,0}}{d\nu^1} d\nu . \tag{7}$$

In fact, we have just proved that

$$\left\{\frac{f_{-1,0}}{f_{-1,-1}}, \frac{f_{-2,0}}{f_{-2,-1}}, \cdots, \frac{d\nu_{-\infty,0}}{d\nu'}\right\}$$

is a ν-semi-martingale. Now let

$$F(m) = [\sup_{k\geqq 1} \log \frac{f_{-k,0}}{f_{-k,-1}} > m]$$

and

$$F_K(m) = [\sup_{1\leqq j<k} \log \frac{f_{-j,0}}{f_{-j,-1}} \leqq m, \log \frac{f_{-k,0}}{f_{-k,-1}} > m] .$$

On $F_k(m)$ we have

$$f_{-k,-1} \leqq 2^{-m} f_{-k,0} .$$

Hence

$$\int_{F_k(m)} f_{-k,-1} \frac{f_{-k,0}}{f_{-k,-1}} d\mu \leqq 2^{-m} \int_{F_k(m)} \left(\frac{f_{-k,0}}{f_{-k,-1}}\right)^2 d\mu$$
$$= 2^{-m} \int_{F_k(m)} \frac{f_{-k,0}}{f_{-k,-1}} d\nu .$$

Applying (7), we obtain

$$\nu[F_k(m)] \leqq 2^{-m} \int_{F_k(m)} \frac{d\nu}{d\nu'} d\nu,$$

therefore,

$$\nu[F(m)] \leqq 2^{-m} \int_{F(m)} \frac{d\nu}{d\nu'} d\nu \leqq 2^{-m} L .$$

Hence

$$\int \sup_{k\geqq 1} \log \frac{f_{-k,0}}{f_{-k,-1}} d\nu \leqq \sum_{m\geqq 0} \nu[F(m)] \leqq \sum_{m\geqq 0} 2^{-m} L < \infty .$$

Combining Lemmas 1, 2 and noting that

$$\int \frac{f_{0,n}}{f_{0,n-1}} d\nu = \int \frac{f_{-n,0}}{f_{-n,-1}} d\nu$$

(cf. Theorem 1, [6]), we obtain the following theorem.

THEOREM 1. *If there is a number L such that*

$$\int \frac{f_{0\,n}}{f_{0,n-1}} d\nu \leqq L \textit{ for } n = 1, 2, \cdots \textit{ then}$$

$$\int \sup_{k \geq 1} |\log f_{-k,0} - \log f_{-k,-1}| d\nu < \infty$$

and $\{n^{-1} \log f_{0,n}\}$ *converges with* ν*-probability one.*

Extensions of Lemma 1, Lemma 2 and Theorem 1 to K-Markovian μ are immediate.

3. **The countable case.** Let X be countable with elements denoted by a. Let ν be an arbitrary stationary probability measure on $\mathscr{F}$. Let

$$P(a_0, a_1, \cdots, a_n) = \nu[x_0 = a_0, x_1 = a_1, \cdots, x_n = a_n] .$$

Let

$$H_1 = -\sum_a P(a) \log P(a) = -\int \log P(x_n) d\nu .$$

Carleson showed that

$$H_1 < \infty \tag{8}$$

implies the $L_1(\nu)$ convergence of $\{n^{-1} \log P(x_0, x_1, \cdots, x_n)\}$ [3]. Chung showed that (8) also implies the convergence with ν-probability one of $\{n^{-1} \log P(x_0, x_1, \cdots, x_n)\}$ [4]. Let μ be defined by

$$\mu[x_m = a_0, x_{m+1} = a_1, \cdots, x_n = a_{n-m}] = P(a_0)P(a_1) \cdots P(a_{n-m}) .$$

μ may be called the independent measure obtained from ν. Then $\nu_{m,n} \ll \mu_{m,n}$ with derivative

$$f_{m,n} = \frac{P(x_m, \cdots, x_n)}{P(x_m) \cdots \cdot P(x_n)}$$

and

$$\log \frac{f_{m,n}}{f_{m,n-1}} = \log \frac{P(x_m, \cdots, x_n)}{P(x_m, \cdots, x_{n-1})} - \log P(x_n) . \tag{9}$$

It follows from (9) that

$$\int (\log f_{0,n} - \log f_{0,n-1}) d\nu \leqq \int - \log P(x_n) d\nu = H_1 .$$

Hence (8) implies that (1) is satisfied, therefore $\{n^{-1} \log f_{0,n}\}$ converges in $L_1(\nu)$ by Theorem 5 [6]. Since

$$\log f_{0,n} = \log P(x_0, \cdots, x_n) + \sum_{k=0}^{n} \log P(x_k) ,$$

Carleson's theorem follows immediately. Furthermore, it follows from (9) and Lemma 1 that

$$\int \sup_{k \geqq 1} [\log \frac{P(x_{-k}, \cdots, x_{-1})}{P(x_{-k}, \cdots, x_0)} + \log P(x_0)] d\nu < \infty .$$

Hence (8) implies

$$\int \sup_{k \geqq 1} \log \frac{P(x_{-k}, \cdots, x_{-1})}{P(x_{-k}, \cdots, x_0)} d\nu < \infty$$

and Chung's theorem [4] follows.

By using a similar approach we shall give a sharpend version of Carleson's and Chung's theorems.

Let

$$P(a_0 \mid a_{-l}, \cdots, a_{-1} = \frac{P(a_{-l}, \cdots, a_{-1}, a_0)}{P(a_{-l}, \cdots, a_{-1})}$$

and let

$$\begin{aligned} H_l &= - \sum_{a_{-l}, \cdots, a_{-1}} P(a_{-l}, \cdots, a_0) \log P(a_0 \mid a_{-l}, \cdots, a_{-1}) \\ &= - \int \log P(x_n \mid x_{n-l}, \cdots, x_{n-1}) d\nu . \end{aligned}$$

H_l is nonnegative but may be $+\infty$. It is known that

$$H_1 \geqq H_2 \geqq H_3 \geqq \cdots \cdots$$

Let

$$H = \lim_{l \to \infty} H_l .$$

The limit is taken to be $+\infty$ if all H_l are $+\infty$.

THEOREM 2. *If $H < \infty$ then $\{n^{-1} \log P(x_0, \cdots, x_n)\}$ converges both in $L_1(\nu)$ and with ν-probability one.*

Proof. There is an l such that $H_l < \infty$. We define an l-Markovian measure μ on $\mathscr{F}$ as follows.

$$\mu[x_m = a_0, x_{m+1} = a_1, \cdots, x_n = a_{n-m}] = P(a_0, \cdots, a_{n-m})$$

if $n - m \leqq l$,

$$\begin{aligned} &\mu[x_m = a_0, x_{m+1} = a_1, \cdots, x_n = a_{n-m}] \\ &= P(a_0, \cdots, a_l) P(a_{l+1} \mid a_1, \cdots, a_l) \cdots \cdots P(a_{n-m} \mid a_{n-m-l}, \cdots \cdots, a_{n-m-1}) \end{aligned}$$

if $n - m > l$. It is easy to check that μ is well defined and $\nu_{m,n} \ll \mu_{m,n}$. It is clear that, if $n - m > l$,

$$\log \frac{f_{m,n}}{f_{m,n-1}} = \log \frac{P(x_m, \cdots, x_n)}{P(x_m, \cdots, x_{n-1}} - \log P(x_n \mid x_{n-l}, \cdots, x_{n-1}) .$$

The rest of the proof goes in the same manner as for the case $H_1 < \infty$ since Theorem 5 [6] and Lemma 1 of this paper remain true for l-Markovian μ.

REFERENCES

1. L. Breiman, *The individual ergodic theorem of information theory*, Ann. Math. Stat. **28** (1957), 809-811.

2. ———, *A correction to the individual ergodic theorem of information theory*, Ann. Math. Stat., **31** (1960), 809-810.

3. L. Carleson, *Two remarks on the basic theorems of information theory*, Math. Scand., **6** (1958), 175-180.

4. K. L. Chung, *A note on the ergodic theorem of information theory*, Ann. Math. Stat., **32** (1961), 612-614.

5. J. L. Doob, *Stochastic Processes*, John Wiley and Sons Inc., New York, 1953.

6. Shu-Teh C. Moy, *Generalizations of Shannon-McMillan theorem*, Pacific J. Math., 11 1961), 705-714.

SYRACUSE UNIVERSITY

PACIFIC JOURNAL OF MATHEMATICS

Mathematical papers intended for publication in the *Pacific Journal of Mathematics* should be typewritten (double spaced), and the author should keep a complete copy. Manuscripts may be sent to any one of the four editors. All other communications to the editors should be addressed to the managing editor, L. J. Paige at the University of California, Los Angeles 24, California.

50 reprints per author of each article are furnished free of charge; additional copies may be obtained at cost in multiples of 50.

The *Pacific Journal of Mathematics* is published quarterly, in March, June, September, and December. The price per volume (4 numbers) is \$12.00; single issues, \$3.50. Back numbers are available. Special price to individual faculty members of supporting institutions and to individual members of the American Mathematical Society: \$4.00 per volume; single issues, \$1.25.

Subscriptions, orders for back numbers, and changes of address should be sent to Pacific Journal of Mathematics, 103 Highland Boulevard, Berkeley 8, California.

Printed at Kokusai Bunken Insatsusha (International Academic Printing Co., Ltd.), No. 6, 2-chome, Fujimi-cho, Chiyoda-ku, Tokyo, Japan.

PUBLISHED BY PACIFIC JOURNAL OF MATHEMATICS, A NON-PROFIT CORPORATION

Reprinted 1966 in the United States of America

Pacific Journal of Mathematics

Vol. 11, No. 4 , 1961