GENERALIZATIONS OF SHANNON-MCMILLAN THEOREM

SHU-TEH C. MOY

1. Introduction. Let X be a non-empty set and $\mathscr{S}$ be a σ-algebra of subsets of X. Consider the infinite product space $\Omega = \prod_{n=-\infty}^{\infty} X_n$ where $X_n = X$ for $n = 0, \pm 1, \pm 2, \cdots$ and the infinite product σ-algebra $\mathscr{F} = \prod_{n=-\infty}^{\infty} \mathscr{S}_n$ where $\mathscr{S}_n = \mathscr{S}$ for $n = 0, \pm 1, \pm 2, \cdots$. Elements of Ω are bilateral infinite sequences $\{\cdots, x_{-1}, x_0, x_1, \cdots\}$ with $x_n \in X$. Let us denote the elements of Ω by ω. If $\omega = \{\cdots, x_{-1}, x_0, x_1, \cdots\}$ x_n is called the nth coordinate of ω and shall be considered as a function on Ω to X. Let T be the shift transformation on Ω to Ω: the nth coordinate of $T\omega$ is equal to the $n+1$th coordinate of ω. For any function g on Ω, Tg is the function defined by $Tg(\omega) = g(T\omega)$ so that $Tx_n = x_{n+1}$. We shall consider two probability measures μ, ν defined on $\mathscr{F}$. Let $\Omega_n = \prod_{i=1}^{n} X_i$ where $X_i = X$, $i = 1, 2, \cdots, n$ and $\mathscr{F}_n = \prod_{i=1}^{n} \mathscr{S}_i$ where $\mathscr{S}_i = \mathscr{S}$, $i = 1, 2, \cdots, n$. Then $\Omega_1 = X$ and $\mathscr{F}_1 = \mathscr{S}$. Let $\mathscr{F}_{m,n}$, $m \leqq n$, $n = 0, \pm 1, \pm 2, \cdots$, be the σ-algebra of subsets of Ω consisting of sets of the form

$$[\omega = \{\cdots, x_{-1}, x_0, x_1, \cdots\} : (x_m, x_{m+1}, \cdots, x_n) \in E]$$

where $E \in \mathscr{F}_{n-m+1}$. Let $\mathscr{F}_{-\infty,n}$ be the σ-algebra generated by $\bigcup_{m=-1}^{-\infty} \mathscr{F}_{m,n}$. Let $\mu_{m,n}, \nu_{m,n}$ be the contractions of μ, ν, respectively, to $\mathscr{F}_{m,n}$ and $\mu_{-\infty,n}$, $\nu_{-\infty,n}$ be the contractions of μ, ν, respectively, to $\mathscr{F}_{-\infty,n}$. Throughout this paper $\nu_{m,n}$ *is assumed to be absolutely continuous with respect to* $\mu_{m,n}$, $\nu_{m,n} \ll \mu_{m,n}$, *for* $m < n$, $n = 0, \pm 1, \pm 2, \cdots$. Let $f_{m,n}$ be the derivative of $\nu_{m,n}$ with respect to $\mu_{m,n}$, $f_{m,n} = d\nu_{m,n}/d\mu_{m,n}$. $f_{m,n}$ is $\mathscr{F}_{m,n}$ measurable and nonnegative. $f_{m,n}$ is also positive with ν probability one. Hence $1/f_{m,n}$ is well defined with ν probability one. A fundamental theorem of *Information Theory* by Shannon and McMillan may be considered as a theorem concerning the asymptotic properties of $f_{m,n}$ as $n \to \infty$. The theorem may be stated as follows: Let X be a finite set of K points and $\mathscr{S}$ be the σ-algebra of all subsets of X. Let ν be any stationary (T invariant) probability measure on $\mathscr{F}$ and μ be the equally distributed independent (product) measure. Then $n^{-1} \log f_{1,n}$ converges in $L_1(\nu)$. In particular, if ν is ergodic, the limit function is equal to $\log K - H$ with ν probability one where H is the entropy of ν measure [3] [8]. Generalizations to arbitrary $X, \mathscr{S}$ were first studied by A. Pérez. He introduced an A_μ condition on ν as follows. ν is said to satisfy A_μ condition if $\nu_{-\infty,n}$ is absolutely continuous with respect to $\nu_{-\infty,0}, \mu_{1,n}$ for $n = 1, 2, \cdots$. He proved the following theorem. If ν, μ are stationary and μ is the product (independent) measure on $\mathscr{F}$ and if

Received May 31, 1960. This work was supported in part by a summer research grant from Wayne State University.

(a) $\lim_{n\to\infty} n^{-1}\int \log f_{1,n} d\nu$ exists and is finite,

(b) ν satisfies condition A_μ,

then $\{n^{-1}\log f_{1,n}\}$ converges in $L_1(\nu)$ [6]. Later Pérez announced that the theorem remains to be true for any stationary measures μ, ν [8]. The present writer proved that for Markovian μ, ν with ν being stationary and μ having stationary transition probabilities the ν-integrability of $\log f_{1,2}$ implies the $L_1(\nu)$ convergence of $\{n^{-1}\log f_{1,n}\}$. The proof is based on an iteration formula for $f_{1,n}$ [4]. In this paper we shall study the case that ν is stationary and μ is Markovian with stationary transition probabilities. It shall be proved that the condition

(c) $\int(\log f_{1,n} - \log f_{1,n-1})d\nu \leqq M < \infty$ for $n = 1, 2, 3, \cdots$ implies the $L_1(\nu)$ convergence of $\{n^{-1}\log f_{1,n}\}$. In fact the conditions (c) and (a) are equivalent for this case, so that the theorem is a generalization of the theorem of Pérez given in [6]. The proof is conducted along similar lines used by McMillan. The crucial step is proving the $L_1(\nu)$ convergence of $\{\log f_{-n,0} - \log f_{-n,-1}\}$. The condition (c) is shown to be necessary and sufficient for this convegence.

2. Generalizations of Shannon-McMillan theorem. Let $x, \mathscr{S}, \Omega, \mathscr{F}, \Omega_n, \mathscr{F}_n, \mathscr{F}_{m,n}, \mu_{m,n}, \nu_{m,n}, f_{m,n}$ be as in *I*. Notations for conditional probabilities and conditional expectations relative to one or several random variables will be as in [2], Chapter 1, § 7. A probability measure on $\mathscr{F}$ is *Markovian* if, for any $A \in \mathscr{S}$, $m < n$ $n = 0, \pm 1, \pm 2, \cdots$

$$P[x_n \in A \mid x_m, \cdots, x_{n-1}] = P[x_n \in A \mid x_{n-1}]$$

with probability one. A Markovian measure is said to have *stationary transition* probabilities if for any $A \in \mathscr{S}$ and any integer n

$$P[x_n \in A \mid x_{n-1}] = T^n P[x_0 \in A \mid x_{-1}]$$

with probability one. In this paper, since we have two probability measures μ, ν, we need to use subscripts μ, ν to indicate conditional probabilities and conditional expectations taken under μ, ν respectively. For any $E \subset \Omega$, I_E, the indicator of E, is the real valued function on Ω defined by

$$\begin{aligned} I_E(\omega) &= 1 \quad \text{if} \quad \omega \in E \\ &= 0 \quad \text{if} \quad \omega \notin E\,. \end{aligned}$$

The log in this paper is the logarithm with base 2.

LEMMA 1. *Define* $\nu'_{m,n}$ *on* $\mathscr{F}_{m,n}$ *by*

$$\nu'_{m,n}(E) = \int P_\mu[E \mid x_m, \cdots, x_{n-1}]d\nu\,, \tag{1}$$

then $\nu'_{m,n}$ *is a probability measure on* $\mathscr{F}_{m,n}$ *with* $\nu'_{m,n}(E) = \nu_{m,n}(E)$ *for* $E \in \mathscr{F}_{m,n-1}$. *Furthermore* $\nu_{m,n} \ll \nu'_{m,n}$ *with*

$$d\nu_{m,n}/d\nu'_{m,n} = f_{m,n}/f_{m,n-1} \ .$$

Proof.

$$\begin{aligned}\nu'_{m,n}(E) &= \int P_\mu[E \mid x_m, \cdots, x_{n-1}]d\nu \\ &= \int P_\mu[E \mid x_m, \cdots, x_{n-1}]f_{m,n-1}d\mu \\ &= \int E_\mu[I_E f_{m,n-1} \mid x_m, \cdots, x_{n-1}]d\mu \\ &= \int_E f_{m,n-1}d\mu \ .\end{aligned}$$

Hence $\nu'_{m,n}$ is a probability measure on $\mathscr{F}_{m,n}$. Furthermore, for $E \in \mathscr{F}_{m,n}$

$$\begin{aligned}\nu_{m,n}(E) &= \int_E f_{m,n}d\mu = \int_E (f_{m,n}/f_{m,n-1})f_{m,n-1}d\mu \\ &= \int_E (f_{m,n}/f_{m,n-1})d\nu'_{m,n} \ .\end{aligned}$$

Hence $\nu_{m,n}$ is absolutely continuous with respect to $\nu'_{m,n}$ and $d\nu_{m,n}/d\nu'_{m,n} = f_{m,n}/f_{m,n-1}$.

THEOREM 1. *If* ν *is stationary and* μ *is Markovian with stationary transition probabilities then*

$$f_{m,n}/f_{m,n-1} = T^n(f_{m-n,0}/f_{m-n,-1}) \tag{2}$$

with ν *probability one for all* $m < n$, $n = 0, \pm 1, \pm 2, \cdots$.

Proof. If μ is Markovian and has stationary transition probabilities then for any $A \in \mathscr{S}$,

$$\begin{aligned}P_\mu[x_n \in A \mid x_m, \cdots, x_{n-1}] &= P_\mu[x_n \in A \mid x_{n-1}] \\ &= T^n P_\mu[x_0 \in A \mid x_{-1}]\end{aligned}$$

with μ probability one and, therefore, also with ν probability one. Hence for any $A \in \mathscr{S}, B \in \mathscr{F}_{n-m}$

$$\begin{aligned}&\nu'_{m,n}[x_n \in A, (x_m, \cdots, x_{n-1}) \in B] \\ &\quad = \int_{[(x_m, \cdots, x_{n-1}) \in B]} P_\mu[x_n \in A \mid x_m, \cdots, x_{n-1}]d\nu \\ &\quad = \int_{[(x_m, \cdots, x_{n-1}) \in B]} P_\mu[x_n \in A \mid x_{n-1}]d\nu\end{aligned}$$

$$= \int_{[(x_m,\cdots,x_{n-1})\in B]} T^n P_\mu[x_0 \in A \mid x_{-1}] d\nu$$
$$= \int_{[(x_{m-1},\cdots,x_{-1})\in B]} P_\mu[x_0 \in A \mid x_{-1}] d\nu$$
$$= \int_{[(x_{m-n},\cdots,x_{-1})\in B]} P_\mu[x_0 \in A \mid x_{m-n}, \cdots, x_{-1}] d\nu$$
$$= \nu'_{m-n,0}[x_0 \in A, (x_{m-n}, \cdots, x_{-1}) \in B].$$

It follows that

$$\nu'_{m,n}[(x_m, \cdots, x_n) \in C] = \nu'_{m-n,0}[(x_{m-n}, \cdots, x_0) \in C]$$

for every $C \in \mathscr{F}_{n-m+1}$. Since by Lemma 1

$$d\nu_{m,n}/d\nu'_{m,n} = f_{m,n}/f_{m,n-1}, \quad d\nu_{m-n,0}/d\nu'_{m-n,0} = f_{m-n,0}/f_{m-n,-1}$$

(2) follows easily.

LEMMA 2. *If μ is Markovain and $m_1 < m_2 < 0$ then $\nu'_{m_1,0}$ is an extension of $\nu'_{m_2,0}$ to $\mathscr{F}_{m_1,0}$.*

Proof. For any $A \in \mathscr{S}, \beta \in \mathscr{F}_{-m_2}$

$$\nu'_{m_1,0}[x_0 \in A, (x_{m_2}, \cdots, x_{-1}) \in B]$$
$$= \int_{[(x_{m_2},\cdots,x_{-1})\in B]} P_\mu[x_0 \in A \mid x_{m_1}, \cdots, x_{-1}] d\nu$$
$$= \int_{[(x_{m_2},\cdots,x_{-1})\in B]} P_\mu[x_0 \in A \mid x_{-1}] d\nu$$
$$= \int_{[(x_{m_2},\cdots,x_{-1})\in B]} P_\mu[x_0 \in A \mid x_{m_2}, \cdots, x_{-1}] d\nu$$
$$= \nu'_{m_2,0}[x_0 \in A, (x_{m_2}, \cdots, x_{-1}) \in B] .$$

It follows that

$$\nu_{m_1,0}(E) = \nu_{m_2,0}(E)$$

for every $E \in \mathscr{F}_{m_2,0}$.

THEOREM 2. *If μ is Markovian and $m_1 < m_2 < 0$ then*

$$(3) \qquad \int(\log f_{m_1,0} - \log f_{m_1,-1}) d\nu \geqq \int(\log f_{m_2,0} - \log f_{m_2,-1}) d\nu \geqq 0 .$$

Proof. By Lemma 2 $\nu'_{m_1,0}$ is an extension of $\nu'_{m_2,0}$ to $\mathscr{F}_{m_1,0}$. Since $\nu_{m_1,0} \ll \nu'_{m_1,0}, \nu_{m_2,0} \ll \nu'_{m_2,0}$ by Lemma 1, $d\nu_{m_2,0}/d\nu'_{m_2,0}$ is the conditional expectation of $d\nu_{m_1,0}/d\nu'_{m_1,0}$ relative to $\mathscr{F}_{m_2,0}$ under the measure $\nu'_{m_1,0}$. Jensen's inequality for conditional expectation implies that

$$0 \leqq \int (d\nu_{m_2,0}/d\nu'_{m_2,0}) \log (d\nu_{m_2,0}/d\nu'_{m_2,0}) d\nu'_{m_1,0}$$

$$\leqq \int (d\nu_{m_1,0}/d\nu'_{m_1,0}) \log (d\nu_{m_1,0}/d\nu'_{m_1,0}) d\nu'_{m_1,0} .$$

Hence

$$0 \leqq \int \log (d\nu_{m_2,0}/d\nu'_{m_2,0}) d\nu \leqq \int \log (d\nu_{m_1,0}/d\nu_{m_1,0}) d\nu \tag{4}$$

and (3) follows from (4) and Lemma 1.

THEOREM 3. *If μ is Markovian then $\{\log f_{m,0} - \log f_{m,-1}\}$ converges with ν probability one as $m \to -\infty$. The limit function may take $\pm\infty$ as its values.*

Proof. It is sufficient to prove that $\{f_{m,-1}/f_{m,0}\}$ converges with ν probability one as $m \to -\infty$. Since $\nu_{m,0}$ is absolutely continuous with respect to $\nu'_{m,0}$ and $d\nu_{m,0}/d\nu'_{m,0} = f_{m,0}/f_{m,-1}$ by Lemma 1, $f_{m,-1}/f_{m,0}$ is the derivative of $\nu_{m,0}$ continuous part of $\nu'_{m,0}$ with respect to $\nu_{m,0}$. Since, by Lemma 2, $\nu'_{m_1,0}$ is an extension of $\nu'_{m_1,0}$ if $m_1 < m_2$, $\{-f_{-k,-1}/f_{-k,0}, \mathscr{F}_{-k,0}, k \geqq 1\}$ is a ν semimartingale ([2] pp. 632). Since

$$\int | -f_{-k,-1}/f_{-k,0} | d\nu = \int f_{-k,-1}/f_{-k,0} d\nu \leqq 1$$

the semimartingale convergence theorem implies that $\{f_{-k-1}/f_{-k,0}\}$ converges with ν probability one as $k \to \infty$.

The following lemma may be considered as an improvement of a theorem by A. Pérez ([6] Theorem 7; pp. 194).

LEMMA 3. *Let $\beta_1 \subset \beta_2 \subset \cdots$ be a sequence of σ-algebras of subsets of Ω and β be the σ-algebra generated by $\bigcup_k \beta_k$. Let ϕ, λ be two probability measures defined on β and ϕ_k, λ_k be the contractions of ϕ, λ, respectively, to β_k. If ϕ_k is absolutely continuous with respect to λ_k for $k = 1, 2, \cdots$ and if there is a finite number M such that*

$$\int \log (d\phi_k/d\lambda_k) d\phi \leqq M$$

for $k = 1, 2, \cdots$ then

(i) *ϕ is absolutely continuous with respect to λ,*

(ii) *$\log (d\phi/d\lambda)$ is ϕ integrable and there exists*

$$\lim_{k\to\infty} \int \log (d\phi_k/d\lambda_k) d\phi = \int \log (d\phi/d\lambda) d\phi ,$$

(iii) $\{\log (d\phi_k/d\lambda_k)\}$ *converges in* $L_1(\phi)$ *to* $\log (d\phi/d\lambda)$.

Proof.

(i) Let $h_k = d\phi_k/d\lambda_k$. Then $\{h_k, \beta_k, k \geqq 1\}$ is a martingale under λ measure. Now

$$M \geqq \int \log (d\phi_k/d\lambda_k) d\phi = \int (\log h_k) h_k d\lambda \ .$$

and

$$(5) \qquad M + \tfrac{1}{2} \geqq \int (h_k \log h_k + \tfrac{1}{2}) d\lambda \geqq (\log n) \int_{(h_n \geqq n)} h_k d\lambda \ .^{1}$$

Hence

$$\int_{(h_k \geqq n)} h_k d\lambda \leqq (\log n)^{-1}(M + \tfrac{1}{2})$$

so that $\int_{(h_k \geqq n)} h_k d\lambda \to 0$ as $n \to \infty$, uniformly in k. Hence $\{h_k\}$ converges with λ probability one and also in $L_1(\lambda)$ ([2] Theorem 4.1, pp. 319). Let the limit function be h. Then $\int_A h d\lambda = \phi(A)$ for all $A \in \bigcup_k \beta_k$ and so for all $A \in \beta$. This proves that ϕ is absolutely continuous and that $h = (d\phi/d\lambda)$.

(ii) The sequence $\{h_k \log h_k\}$ converges with λ probability one to $h \log h$. Since the functions $h_k \log h_k$ are bounded below uniformly by the number $\frac{1}{2}$,

$$\int h \log h d\lambda \leqq \underline{\lim} \int h_k \log h_k d\lambda = \underline{\lim} \int \log h_k d\phi \leqq M \ .$$

Hence $h \log h$ is λ integrable. Since the real valued function $\xi \log \xi$ is continuous and convex, $h_1 \log h_1, h_2 \log h_2, \cdots, h \log h$ constitute a semi-martingale under the measure λ([2], Theorem 1.1, pp. 295). Hence

$$\int h_1 \log h_1 d\lambda \leqq \int h_2 \log h_2 d\lambda \leqq \cdots \leqq h \log h d\lambda \ ,$$

so that $\lim_{k \to \infty} \int h_k \log h_k d\lambda$ exists and is equal to $\int h \log h d\lambda$. Now

$$\int |\log h| \, d\phi = \int h \, |\log h| \, d\lambda = \int |h \log h| \, d\lambda \ ,$$

hence $\log h$ is ϕ integrable and

$$(6) \qquad \int \log h d\phi = \int h \log h d\lambda = \lim_{k \to \infty} \int h_k \log h_k d\lambda = \lim_{k \to \infty} \int \log h_k d\phi \ .$$

[1] Inequality (5) was pointed out by the referee. The proof of Lemma 3 was much shortened by following his suggestions.

(iii) Since $h_1 \log h_1, h_2 \log h_2, \cdots, h \log h$ constitute a semimartingale under the measure λ, we have, for $E \in \beta_k$,

$$\int_E h_k \log h_k \, d\lambda \leqq \int_E h_{k+1} \log h_{k+1} \, d\lambda \leqq \int_E h \log h \, d\lambda .$$

Hence

$$\int_E \log h_k \, d\phi \leqq \int_E \log h_{k+1} \, d\phi \leqq \int_E \log h \, d\phi ,$$

so that $\log h_1, \log h_2, \cdots, \log h$ constitute a semimartingale under the measure ϕ. Hence (ii) implies that $\log h_k$ are uniformly ϕ integrable and $\{\log h_k\}$ converges to $\log h$ in $L_1(\phi)$ ([2], Theorem 4.1s, pp. 324).

THEOREM 4. *If μ is Markvian and there is a finite number M such that*

$$\int [\log f_{m,0} - \log f_{m,-1}] d\nu \leqq M$$

for $m = -1, -2, \cdots$ then $\{\log f_{m,0} - \log f_{m,-1}\}$ converges in $L_1(\nu)$ as $m \to -\infty$.

Proof. By Lemma 2 $\nu'_{m_1,0}$ is an extension of $\nu'_{m_2,0}$ if $m_1 < m_2 < 0$ and

$$d\nu_{m,0}/d\nu'_{m,0} = f_{m,0}/f_{m,-1} .$$

If there is a probability measure ν' defined on the σ-algebra generated by $\bigcup_{m=-1}^{-\infty} \mathscr{F}_{m,0}$ which is an extension of $\nu'_{m,0}$ for $m = -1, -2, \cdots$, then the conclusion of the theorem follows easily from Lemma 3. If X is the real line and if $\mathscr{S}$ is the σ-algebra of Borel sets then the existence of ν' follows from the Consistency Theorem of Kolmogorov. For the general case we shall proceed by using the usual representation by space Ω' of sequences of real numbers as follows:
Let

$$g_k = f_{-k,0}/f_{-k,-1} .$$

Let G be the map of Ω into the space Ω' of real sequences $\{\xi_1, \xi_2, \cdots\}$ defined by

$$G(\omega) = \{g_1(\omega), g_2(\omega), \cdots\} .$$

Considering ξ_k as functions on Ω' we have

$$\xi_k(G(\omega)) = g_k(\omega) .$$

Let β_k be the collection of Borel subsets of Ω' which are determined by conditions on $\xi_1, \xi_2, \cdots, \xi_k$ and β be the collection of all Borel subsets of Ω'. Let ϕ be the probability measure on β and ϕ_k, λ_k be the probability measures on β_k defined by

$$\phi(E) = \nu(G^{-1}E) ,$$
$$\phi_k(E) = \nu_{-k,0}(G^{-1}E) ,$$
$$\lambda_k(E) = \nu'_{-k,0}(G^{-1}E) .$$

$\{g_k\}$ converges in $L_1(\nu)$ if and only if $\{\xi_k\}$ converges in $L_1(\phi)$. Now λ_k are consistent; Kolmogorov's Consistency Theorem implies the existence of a probability measure λ on β which is an extension of every λ_k and $d\phi_k/d\lambda_k = \xi_k$. Hence Lemma 3 is applicable and the $L_1(\phi)$ convergence of $\{\xi_k\}$ is obtained.

THEOREM 5. *If ν is stationary and μ is Markovian with stationary transition probabilities and if*

$$\int \log f_{0,0} d\nu < \infty$$

and if there is a finite number M such that

$$\int(\log f_{0,n} - \log f_{0,n-1})d\nu \leqq M$$

for $n = 1, 2, \cdots$ then $n^{-1}\log f_{0,n}$ converges in $L_1(\nu)$ as $n \to \infty$. In particular, if ν is ergodic, the limit is equal to a nonnegative constant with ν probability one.

Proof. By Theorem 4 $\{\log f_{m,0} - \log f_{m,-1}\}$ converges in $L_1(\nu)$ as $m \to -\infty$. Let h be the $L_1(\nu)$ limit of the sequence. Let $\bar{h}$ be the $L_1(\nu)$ limit of the sequence $\{n^{-1}\sum_{i=1}^{n} T^i h\}$. By Theorem 1 $f_{0,n}/f_{0,n-1} = T^n(f_{-n,0}/f_{-n,-1})$, hence

$$n^{-1}\log f_{0,n} = n^{-1}\log f_{0,0} + n^{-1}\sum_{i=1}^{n} T^i \log(f_{-i,0}/f_{-i,-1})$$

$$\begin{aligned}\int\left|n^{-1}\sum_{i=1}^{n} T^i \log(f_{-i,0}/f_{-i,-1}) - \bar{h}\right|d\nu &\leqq n^{-1}\sum_{i=1}^{n}\int | T^i \log(f_{-i,0}/f_{-i,-1}) - T^i h | d\nu \\ &\quad + \int | n^{-1}\sum T^i h - \bar{h} | d\nu \\ &= n^{-1}\sum_{i=1}^{n}\int | \log(f_{-i,0}/f_{-i,-1}) - h | d\nu \\ &\quad + \int | n^{-1}\sum T^i h - \bar{h} | d\nu \to 0 \quad \text{as} \quad n \to \infty .\end{aligned}$$

Thus the $L_1(\nu)$ convergence of $\{n^{-1}\log f_{0,n}\}$ is proved. The limit is $\bar{h}$

which is the $L_1(\nu)$ limit of $\{n^{-1}\sum_{i=1}^{n} T^i h\}$. If ν is ergodic

$$\bar{h} = \int h\, d\nu$$

with ν probability one and

$$\int h d\nu = \lim_{m\to-\infty}\int[\log f_{m,0} - \log f_{m,-1}]d\nu \geqq 0 .$$

COROLLARY 1. *Under the hypothesis of Theorem 5 if ν is stationary and ergodic but not Markovian then ν is singular to μ.*

Proof. If μ is Markovian but ν is not Markovian then there is a positive integer n_0 such that

$$\mu[f_{0,n_0-1} \neq f_{0,n_0}] > 0 .$$

For, if for every positive integer n

$$\mu[f_{0,n-1} \neq f_{0,n}] = 0$$

then

$$P_\nu[x_n \in A \mid x_0, \cdots, x_{n-1}] = P_\mu[x_n \in A \mid x_{n-1}]$$

with ν probability one for every $A \in \mathscr{S}$ and ν is Markovian instead. Now since

$$f_{0,n_0-1} = E_\mu[f_{0,n_0} \mid x_0, \cdots, x_{n_0-1}]$$

and the function $\xi \log \xi$ is strictly convex, hence

$$\int f_{0,n_0} \log f_{0,n_0} d\mu - \int f_{0,n_0-1} \log f_{0,n_0-1} d\mu > 0$$

so that

$$\int[\log f_{0,n_0} - \log f_{0,n_0-1}]d\nu > 0 .$$

Since $\int[\log f_{0,n} - \log f_{0,n-1}]d\nu$ is non-decreasing in n,

$$\lim_{n\to\infty}\int[\log f_{0,n} - \log f_{0,n-1}]d\nu = a > 0 .$$

Now ν is ergodic; the $L_1(\nu)$ limit $\bar{h}$ of $\{n^{-1}\log f_{0,n}\}$ is equal to a with ν probability one. Let $n_1, n_2, \cdots$ be a sequence of positive integers for which $\{n_k^{-1}\log f_{0,n_k}\}$ converges with ν probability one to a so that $\{1/f_{0,n_k}\}$ converges to 0 as $n_k \to \infty$. Let $\mathscr{F}'$ be the σ-algebra generated by $\bigcup_n \mathscr{F}_{0,n}$ and let $\mu_{\mathscr{F}'}, \mu_{\mathscr{F}'}$ be the contractions of μ, ν, respectively, to $\mathscr{F}'$. Since $1/f_{0,n}$ is the derivative of ν-continuous part of $\mu_{0,n}$ with respect to $\nu_{0,n}$, $\{1/f_{0,n}\}$ converges with ν probability one to the derivative of ν-continuous part of μ' with respect to ν' by a theorem of Anderson and Jessen [1]. Now we have

$$\lim_{n\to\infty} 1/f_{1,n} = 0$$

with ν probability one and μ' is singular to ν'. Hence μ, ν are singular to each other.

Extensions of Theorem 5 and Corollary 1 to K-Markovian μ are immediate.

3. **Discussion.** As was mentioned in the introduction the crucial step in establishing Theorem 5 is to prove the $L_1(\nu)$ convergence of $\{\log f_{-n,0} - \log f_{-n,-1}\}$. If μ is the product (independent) measure on $\mathscr{F}$ the measure ν' in the proof of Theorem 4 is actually $\nu_{-\infty-1} \times \mu_{0,0}$. Thus condition (c) or, equivalently, condition (a) implies condition (b) in the introduction. In [7] it is stated that the condition (b) is necessary for the $L_1(\nu)$ convergence of $\{\log f_{-n,0} - \log f_{-n,-1}\}$ ([7] Theorem 2(b)). A simple is as follows. Let X be the real line and $\mathscr{S}$ be the collection of all Borel sets. Let $\nu = \mu$ and distribution of x_0 be Gaussian. Let $\nu(x_0 = x_1) = \mu(x_0 = x_1) = 1$. Then $\nu_{-1,0}$ is singular to $\nu_{-1,-1} \times \nu_{0,0}$, however the $L_1(\nu)$ convergence of $\{\log f_{-n,0} - \log f_{-n-1}\}$ is trivially true since $f_{m,n} \equiv 1$.

References

1. Erik Sparre Anderson and Borge Jessen, *Some limit theorems in an abstract set*, Danske Vic. Selsk. Nat.-ys. Medd. 22, No. 14 (1946).
2. J. L. Doob, *Stochastic Processes*, John Wiley and Sons, Inc., New York.
3. B. McMillan, *The basic theorems of information theory*, Annals of Math. Statistics, **24** (1953), 196-219.
4. Shu-Teh C. Moy, *Asymptotic properties of derivatives of stationary measures*, Pacific J. Math., **10** (1960), 1371-1383.
5. A. Pérez, *Sur la théorie de l'information dans le cas d'un alphabet abstrait*, Transactions of the First Prague Conference on information theory, statistical decision functions, random processes, Publishing House of the Czechoslavak Academy of Sciences, Prague, (1957) 183-208.
6. ———, *Sur la convergence des incertitudes, entropies et informations échantillon (sample) vers leurs vraies*, Transactions of the First Prague Conference on information theory, statistical decisions functions, random processes, Publishing House of Czechoslavak Academy of Sciences, Prague (1957), 209-243.
7 ———, *Information theory with an abstract alphabet, generalized forms of McMillan's limit theorem for the case of discrete and continuous times*, Theory of Probability and its Applications, (An English translation of the Soviet Journal Teoriya Veryoatnostei i ee Primeneniya) Vol. **4**, No. 1 (1959), 99-102.
8. C. E. Shannon, *The mathematical theory of communication*, Bell System Technical Journal, **27** (1948), 379-423, 623-656.

Syracuse University