The Annals of Probability
1974, Vol. 2, No. 3, 530-534

ON THE APPROXIMATION OF STATIONARY MEASURES BY PERIODIC AND ERGODIC MEASURES

BY J. C. KIEFFER

University of Missouri-Rolla

Let $(\Omega, \mathscr{F})$ be the measurable space consisting of Ω, the set of sequences $(x_1, x_2, \cdots)$ from a finite set A, and $\mathscr{F}$, the usual product sigma-field. Let $X_1, X_2, \cdots$ be the usual coordinate random variables defined on Ω. For $n = 1, 2, \cdots$, let $\mathscr{F}_n$ be the sub sigma-field of $\mathscr{F}$ generated by $X_1, X_2, \cdots, X_n$. We prove the following: if P is a probability measure on $\mathscr{F}$ stationary with respect to the one-sided shift transformation on Ω and if N is a positive integer, then there is a periodic measure Q on $\mathscr{F}$ such that $Q = P$ over $\mathscr{F}_N$. This is a stronger result than the known fact that the periodic measures are dense in the set of stationary measures under the weak topology. We also show that if P assigns positive measure to every non-empty set in $\mathscr{F}_N$, it is possible to find an ergodic measure Q such that $P = Q$ over $\mathscr{F}_N$. We investigate the entropies of all such ergodic measures Q which approximate P in this sense, and show that there is a unique ergodic measure Q of maximal entropy such that $P = Q$ over $\mathscr{F}_N$.

Let $(\Omega, \mathscr{F})$ be the measurable space consisting of Ω, the set of sequences $(x_1, x_2, \cdots)$ from a finite set A, and $\mathscr{F}$, the usual product sigma-field. We take A to have $a > 1$ elements, which without loss of generality we take to be $1, 2, \cdots, a$. Let $X_1, X_2, \cdots$ be the usual coordinate random variables mapping Ω onto A; i.e., $X_i(x_1, x_2, \cdots) = x_i$, $i = 1, 2, \cdots$. For $n = 1, 2, \cdots$, let $\mathscr{F}_n$ be the sub sigma-field of $\mathscr{F}$ generated by $X_1, X_2, \cdots, X_n$, and let $Y_n = (X_1, X_2, \cdots, X_n)$. If T is the one-sided shift transformation on Ω, let $\mathscr{P}$ be the collection of all probability measures on $\mathscr{F}$ which are stationary with respect to T. A measure $P \in \mathscr{P}$ is periodic if it is discrete and concentrated on a finite number of points of Ω.

It is known [3] that the periodic measures are dense in $\mathscr{P}$ under the weak topology; i.e., for any positive integer n, $P \in \mathscr{P}$, and $\varepsilon > 0$, there exists a periodic measure Q such that $|P(E) - Q(E)| < \varepsilon$, $E \in \mathscr{F}_n$. (In fact, in [3] the statement of the density of periodic measures is proven in the more general situation of a countable Cartesian product of a fixed complete separable metric space. The complete separable metric space we are considering here is just the finite set A, with the discrete topology.) In this paper we prove the stronger result:

THEOREM 1. *Given $P \in \mathscr{P}$ and n, there is a periodic measure Q such that $Q = P$ over $\mathscr{F}_n$.*

Received February 8, 1973; revised August 24, 1973.

AMS 1970 *subject classifications.* Primary 60B05, 28-A65, 28-A35; Secondary 94A15.

Key words and phrases. Stationary measures, ergodic measures, periodic measures, entropy, shift transformation.

Institute of Mathematical Statistics is collaborating with JSTOR to digitize, preserve, and extend access to *The Annals of Probability.*
www.jstor.org

For the proof of this theorem, we need the following two lemmas.

LEMMA 1. *Let r, s be given positive integers; let $M_1, M_2, \cdots, M_r$ and $N_1, N_2, \cdots, N_s$ be given nonnegative integers. Then there exists an $r \times s$ matrix $[a_{ij}]$ such that*

(1a) *The elements a_{ij} are nonnegative integers.*
(1b) $\sum_{i=1}^{r} a_{ij} \leqq N_j$, $j = 1, 2, \cdots, s$.
(1c) $\sum_{j=1}^{s} a_{ij} \leqq M_i$, $i = 1, 2, \cdots, r$.
(1d) $\sum_{i,j} a_{ij} = \min [\sum_{k=1}^{r} M_k, \sum_{k=1}^{s} N_k]$.

PROOF. A simple induction argument, which we omit, sufficies to prove this.

LEMMA 2. *Let k be a positive integer. Let p be any probability measure on A^k satisfying the following properties:*

(2a) *There is a positive integer N such that Np is a measure with integral values.*
(2b) *If* $k > 1$, $\sum_{i=1}^{a} p(B, i) = \sum_{i=1}^{a} p(i, B)$, $B \in A^{k-1}$.

Then there exists a probability measure q on A^{k+1} such that

(2c) *Nq is an integral-valued measure.*
(2d) $\sum_{i=1}^{a} q(B, i) = \sum_{i=1}^{a} q(i, B) = p(B)$, $B \in A^k$.

PROOF. We consider the equations (2d) as a system of $2a^k$ linear equations in the a^{k+1} unknowns $q(B)$, $B \in A^{k+1}$. We seek nonnegative solutions for the $q(B)$'s so that (2c) is satisfied. We may partition the system (2d) into subsystems, one subsystem for each $B \in A^{k-1}$, as follows:

(2e) $\sum_{j=1}^{a} q(i, B, j) = p(i, B)$; $\sum_{j=1}^{a} q(j, B, i) = p(B, i)$, $i = 1, 2, \cdots, a$.

(If $k = 1$, we adopt the convention that $B \in A^{k-1}$ means B is just the null sequence, the sequence of length 0.) For each $B \in A^{k-1}$ we solve the above subsystem of $2a$ equations in the a^2 unknowns $p(j, B, i)$, $i, j = 1, 2, \cdots, a$. If we solve (2e) we get an expression of $2a - 1$ of the unknowns in terms of the remaining $(a - 1)^2$ unknowns:

$$q(1, B, 1) = \sum_{i,j=2}^{a} q(i, B, j) + p(B, 1) + p(1, B) - \sum_{i=1}^{a} p(i, B)\,;$$
$$q(1, B, m) = -\sum_{i=2}^{a} q(i, B, m) + p(B, m)\,, \qquad m = 2, 3, \cdots, a\,;$$
$$q(m, B, 1) = -\sum_{i=2}^{a} q(m, B, i) + p(m, B)\,, \qquad m = 2, 3, \cdots, a\,.$$

From Lemma 1, it follows that we may choose the unknowns $q(i, B, j)$, $i, j = 2, 3, \cdots, a$, so that

(2f) the numbers $Nq(i, B, j)$ are nonnegative integers.
(2g) $\sum_{i=2}^{a} q(i, B, m) \leqq p(B, m)$, $m = 2, \cdots, a$.
(2h) $\sum_{i=2}^{a} q(m, B, i) \leqq p(m, B)$, $m = 2, \cdots, a$.
(2i) $\sum_{i,j=2}^{a} q(i, B, j) = \min [\sum_{m=2}^{a} p(B, m), \sum_{m=2}^{a} p(m, B)] = \sum_{m=1}^{a} p(m, B) - \max [p(B, 1), p(1, B)]$.

With this choice of the unknowns $q(i, B, j)$, $i, j = 2, \cdots, a$, the other unknowns are also nonnegative numbers which when multiplied by N give integers.

PROOF OF THEOREM 1. Let $P \in \mathscr{P}$ and the positive integer n be given. Let S be the set of all probability measures q on A^n such that if $n > 1$, $\sum_{i=1}^{a} q(B, i) = \sum_{i=1}^{a} q(i, B)$, $B \in A^{n-1}$. S considered as a subset of the Euclidean space of all real-valued functions on A^n is a compact convex set. There are a finite number of extreme points of S, which we shall designate by $p_1, p_2, \cdots, p_m$. These extreme points are each rational-valued measures. Let p be the measure in S induced by P; i.e., $p(B) = P(Y_n = B)$, $B \in A^n$. Then there are nonnegative numbers $\alpha_1, \alpha_2, \cdots, \alpha_m$, which sum to one, such that $p = \alpha_1 p_1 + \alpha_2 p_2 + \cdots + \alpha_m p_m$. If for each p_i, $i = 1, 2, \cdots, m$, we can find a periodic $Q_i \in \mathscr{P}$ such that $Q_i(Y_n = B) = p_i(B)$, $B \in A^n$, then $Q = \sum_{i=1}^{m} Q_i$ is periodic, $Q(Y_n = B) = P(Y_n = B)$, $B \in A^n$, and the theorem is proved.

Consequently, let $q_n \in S$ and suppose there is a positive integer N such that Nq_n is integral-valued. Using Lemma 2, we may construct a sequence $\{q_i\}_{i=n+1}^{\infty}$ such that

(a) q_i is a probability measure on A^i, $i = n + 1, n + 2, \cdots$;
(b) $\sum_{j=1}^{a} q_{i+1}(B, j) = \sum_{j=1}^{a} q_{i+1}(j, B) = q_i(B)$, $B \in A^i$, $i = n, n + 1, \cdots$;
(c) Nq_i is integral-valued, $i = n + 1, n + 2, \cdots$.

There exists $Q \in \mathscr{P}$ such that $Q(Y_i = B) = q_i(B)$, $B \in A^i$, $i = n, n + 1, \cdots$ (See [1] page 5). NQ is integral-valued over $\mathscr{F}$. Therefore Q is discrete because if Q had a nonzero continuous component, Q would have to assign an irrational value to some set in $\mathscr{F}$. Q is concentrated on no more than N points of Ω since any point of Ω with nonzero Q measure has measure at least N^{-1}. Therefore Q is periodic.

Approximation by ergodic measures. In the result from [3] mentioned earlier in this paper, which states that stationary measures on a countable Cartesian product of a fixed complete separable metric space can be approximated arbitrarily closely by periodic measures, the approximating periodic measures can be chosen to be ergodic. It is natural then to inquire whether in Theorem 1 the measure Q can be chosen to be ergodic as well as periodic. This, however, is not possible in general since periodic ergodic measures on $(\Omega, \mathscr{F})$ take rational values. However, if we eliminate the requirement that Q be periodic, the following theorem is true.

THEOREM 2. *Let $n > 0$, and $P \in \mathscr{P}$ be given. If the numbers $P(Y_n = B)$, $B \in A^n$, are all positive then there exists an ergodic $Q \in \mathscr{P}$ such that $Q = P$ over $\mathscr{F}_n$.*

PROOF. This theorem is implied by Theorem 4.

The entropies of approximating ergodic measures Q. We investigate now the possible entropies of measures $Q \in \mathscr{P}$ approximating P in the sense of Theorem 2. First, we present some notation. If Y is a discrete measurable function defined on Ω and P is a measure on $\mathscr{F}$, let $P(Y)$ denote the discrete measurable function defined on Ω as follows: $P(Y)(\omega) = P[Y = Y(\omega)]$, $\omega \in \Omega$. If Z is also a discrete measurable function on Ω let $P(Y \mid Z)$ denote the discrete measurable function defined on Ω as follows: For each $\omega \in \Omega$, $P(Y \mid Z)(\omega) = P[Y = Y(\omega) \mid Z = Z(\omega)] = P[Y = Y(\omega), Z = Z(\omega)]/P[Z = Z(\omega)]$, if $P[Z = Z(\omega)] > 0$; otherwise $P(Y \mid Z)(\omega) = 0$. For $n > 0$, $P \in \mathscr{P}$, let $H_n(P) = -\int_\Omega \log P(Y_n)\, dP$. Let $H_0(P) = 0$. If $P \in \mathscr{P}$, the sequence $H_n(P) - H_{n-1}(P)$, $n = 1, 2, \cdots$, is non-negative and non-increasing. Let $H(P) = \lim_{n\to\infty} [H_n(P) - H_{n-1}(P)]$. $H(P)$ is called the entropy of P. The mapping $P \to H(P)$ is a linear functional on $\mathscr{P}$ [4].

THEOREM 3. *Let $N > 0$, $P \in \mathscr{P}$. Let $\mathscr{S} = \{Q \in \mathscr{P}: Q = P$ over $\mathscr{F}_n\}$.*

(3a) *The image of $\mathscr{S}$ under H is the closed interval* $[0, H_N(P) - H_{N-1}(P)]$;

(3b) *H attains its maximum value over $\mathscr{S}$, namely $H_N(P) - H_{N-1}(P)$, at a unique $Q \in \mathscr{S}$.*

PROOF. Let Q be the measure on $\mathscr{F}$ such that

(3c) $Q(X_1, X_2, \cdots, X_N) = P(X_1, X_2, \cdots, X_N)$;

(3d) $Q(X_1, X_2, \cdots, X_n) = P(X_1, X_2, \cdots, X_{N-1}) \prod_{i=N}^{n} P(X_i \mid X_{i-N+1}, X_{i-N+2}, \cdots, X_{i-1})$, $n = N + 1, N + 2, \cdots$.

It can be verified that $Q \in \mathscr{S}$, and that for $n \geqq N$, $H_n(Q) - H_{n-1}(Q) = H_N(P) - H_{N-1}(P)$. Therefore $H(Q) = H_N(P) - H_{N-1}(P)$, and if $Q' \in \mathscr{S}$, then $H(Q') \leqq H(Q)$. Let $\mathscr{T} = \{Q' \in \mathscr{S}: H(Q') = H(Q)\}$. $\mathscr{T}$ is a convex set and if $Q' \in \mathscr{T}$ then for $n \geqq N$, $H_n(Q') - H_{n-1}(Q') = H(Q)$. Suppose that there is a $Q' \in \mathscr{T}$ such that $Q' \neq Q$. Let M be the least integer n such that $n > N$ and $Q' \neq Q$ over $\mathscr{F}_n$. Since the function H_M is strictly concave, if $Q'' = \frac{1}{2}Q + \frac{1}{2}Q'$ then $H_M(Q'') > \frac{1}{2}H_M(Q) + \frac{1}{2}H_M(Q')$. But $H_{M-1}(Q'') = H_{M-1}(Q') = H_{M-1}(Q)$, so

$$H_M(Q'') - H_{M-1}(Q'') > \tfrac{1}{2}[H_M(Q) - H_{M-1}(Q)] + \tfrac{1}{2}[H_M(Q') - H_{M-1}(Q')] .$$

However, since $Q, Q', Q'' \in \mathscr{T}$, we have then that $H(Q) > \frac{1}{2}H(Q) + \frac{1}{2}H(Q)$, a contradiction. This proves (3b). (3a) is true because there is a periodic measure in $\mathscr{S}$. Periodic measures have zero entropy. The linearity of H on $\mathscr{S}$ is then used.

REMARK. The measure Q constructed in the proof of Theorem 3 was also used by Krieger ([2] page 458) in the proof of his Theorem 3.4.

THEOREM 4. *Let $N > 0$. Let $P \in \mathscr{P}$ be such that P assigns positive measure to every nonempty set in $\mathscr{F}_N$. Then if H is any number such that $0 < H \leqq H_N(P) - H_{N-1}(P)$, there is an ergodic $Q \in \mathscr{P}$ such that $Q = P$ over $\mathscr{F}_N$ and $H(Q) = H$. If $H = H_N(P) - H_{N-1}(P)$, Q is unique.*

PROOF. First of all, $H_N(P) - H_{N-1}(P) > 0$, so this theorem is not a vacuous statement. There are measures $Q, Q' \in \mathscr{S}$ such that $H(Q) = 0$, $H_n(Q') - H_{n-1}(Q') = H_N(P) - H_{N-1}(P)$, $n \geqq N$. Choose $M > N$ so large that $H_M(Q) - H_{M-1}(Q) < H$. An argument using the intermediate value theorem for continuous functions will show that there is a number λ, $0 \leqq \lambda < 1$, such that if $Q'' = \lambda Q + (1 - \lambda)Q'$, $H_M(Q'') - H_{M-1}(Q'') = H$. Now Q' is the unique measure constructed in the proof of Theorem 3; an examination of that proof will show that Q' assigns positive measure to every nonempty set in $\mathscr{F}_M$. Therefore, so does Q''. There is a unique $P' \in \mathscr{P}$ such that $P' = Q''$ over $\mathscr{F}_M$ and $H(P') = H_M(Q'') - H_{M-1}(Q'') = H$. Again appealing to the proof of Theorem 3 we see that P' is Markovian with respect to the shift T^{M-1}, and since Q'' is positive on nonempty subsets of $\mathscr{F}_M$, P' must therefore be ergodic. Finally, $P' = P$ over $\mathscr{F}_N$.

Final remarks. Given $N > 0$, and P as in Theorem 4, we have determined the entropies of those ergodic measures Q which approximate P in the sense that $P = Q$ over $\mathscr{F}_N$, with one possible exception: there may be such an ergodic measure Q with entropy zero. The construction of such a Q appears to be difficult. It may also be possible to relax the condition on P in Theorem 4.

REFERENCES

[1] Billingsley, P. (1965). *Ergodic Theory and Information.* Wiley, New York.

[2] Krieger, W. (1970). On entropy and generators of measure-preserving transformations. *Trans. Amer. Math. Soc.* **149** 453–464.

[3] Parthasarathy, K. R. (1961a). On the category of ergodic measures. *Illinois J. Math.* **5** 648–656.

[4] Parthasarathy, K. R. (1961b). On the integral representation of the rate of transmission of a stationary channel. *Illinois J. Math.* **5** 299–305.

Department of Mathematics
326 Mathematics-Computer Science Building
University of Missouri
Rolla, Missouri 65401