The Annals of Probability
1980, Vol. 8, No. 2, 372–376

ON KAMAE'S CONJECTURE CONCERNING THE $\bar{d}$-DISTANCE BETWEEN TWO-STATE MARKOV PROCESSES

BY MARTIN H. ELLIS[1]
Northeastern University

For pairs of two-state Markov processes satisfying certain conditions, the $\bar{d}$-distance is shown to be partition distance.

A (stationary random) process (with discrete time and discrete state space) can be represented as a 1-1 bimeasurable measure-preserving transformation on a probability space together with a finite or countable ordered partition of the space. The sets in the partition are the states of the process. Given two processes $\mathfrak{T}_1$ and $\mathfrak{T}_2$ each with n states, a process $\mathfrak{T}$ is a joining of $\mathfrak{T}_1$ and $\mathfrak{T}_2$ if $\mathfrak{T}$ has n^2 states denoted by (i, j), $1 \leqslant i, j \leqslant n$, and $\mathfrak{T}$'s marginals are $\mathfrak{T}_1$ and $\mathfrak{T}_2$. For each joining $\mathfrak{T}$ of $\mathfrak{T}_1$ and $\mathfrak{T}_2$, let $d_{\mathfrak{T}}$ be the measure of $\{x: x \in (i, j) \text{ and } i \neq j\}$; $d_{\mathfrak{T}}$ is the distance (between $\mathfrak{T}_1$ and $\mathfrak{T}_2$) attained by $\mathfrak{T}$. The $\bar{d}$-distance between $\mathfrak{T}_1$ and $\mathfrak{T}_2$, denoted $\bar{d}(\mathfrak{T}_1, \mathfrak{T}_2)$, is the infimum of $\{d_{\mathfrak{T}}: \mathfrak{T} \text{ a joining of } \mathfrak{T}_1 \text{ and } \mathfrak{T}_2\}$; in fact, the infimum is attained by some joining. The $\bar{d}$-distance between two processes is a natural and useful measure of how much the processes differ (see [4] or [5] for further discussion of $\bar{d}$).

If $\mathfrak{T}_1$ and $\mathfrak{T}_2$ each have states $1, 2, \cdots, n$ and assign to these states probabilities $r_1, \cdots, r_n$ and $s_1, \cdots, s_n$ respectively, the partition distance between $\mathfrak{T}_1$ and $\mathfrak{T}_2$ is $1 - \Sigma_{i=1}^n \min\{r_i, s_i\}$. Clearly the partition distance is a lower bound on the $\bar{d}$-distance, and it is natural to ask when are they equal.

The aim of this paper is to give a partial answer to this question for pairs of two-state Markov processes. It will be shown that for pairs of two-state Markov processes satisfying certain conditions, the $\bar{d}$-distance equals the partition distance.

Let (κ, λ) denote the two state Markov process with transition matrix

$$\begin{pmatrix} 1-\kappa & \kappa \\ \lambda & 1-\lambda \end{pmatrix}$$

(i.e., the first state sends κ of its measure to the second state and the second state sends λ of its measure to the first). All two-state Markov processes considered in this paper are assumed to have positive entropy, which is equivalent to assuming $0 < \min\{\kappa, \lambda\} < 1$; thus the measure of the first state of (κ, λ) is $\lambda/(\kappa + \lambda)$, the measure of the second is $\kappa/(\kappa + \lambda)$. The measure of any set E will be denoted by $\mu(E)$.

When considering a pair $((\alpha, \beta), (\gamma, \delta))$, denote the first state of (α, β) by A, the

Received February 1978; revised February 1979.

[1]While this paper was in press, Martin Herbert Ellis died February 16, 1980 at the age of 30.

AMS 1970 *subject classification.* Primary 28A65, 60J10.

Key words and phrases. Stationary random process, Markov process, Markov chain, $\bar{d}$-distance, partition distance, $\bar{d}$-distance equalling partition distance, joint process.

second state of (α, β) by B, and the first state of (γ, δ) by C, the second state of (γ, δ) by D.

Notice that the pairs $((\alpha, \beta), (\gamma, \delta))$, $((\beta, \alpha), (\delta, \gamma))$, $((\gamma, \delta), (\alpha, \beta))$, and $((\delta, \gamma), (\beta, \alpha))$ all have the same $\bar{d}$-distance and partition distance, for they are just relabellings of one another. Call such pairs *equivalents*. Using this symmetry each pair can be assumed to be in some standard form.

DEFINITION. $((\alpha, \beta), (\gamma, \delta))$ is in *proper form* (abbreviations: PFI, PFII) if it is in one of the following two forms:

I) $((\alpha, \beta), (\alpha - u, \beta + w))$ with $\mu(C) \geqslant \mu(A)$ and $u, w \geqslant 0$.

II) $((\alpha, \beta), (\alpha + u, \beta + w))$ with $\mu(C) \geqslant \mu(A)$ and $u, w > 0$.

Note that every $((\alpha, \beta), (\gamma, \delta))$ has an equivalent in proper form, and no pair has an equivalent in PFI and an equivalent in PFII. Henceforth it will be assumed that all pairs are in proper form. Thus the partition distance between (α, β) and (γ, δ) is $\mu(C) - \mu(A)$, and $\bar{d}((\alpha, \beta), (\gamma, \delta)) = \mu(C) - \mu(A)$ if and only if there is a joining of (α, β) and (γ, δ) which has $A \subseteq C$.

THEOREM 1. *If $((\alpha, \beta), (\gamma, \delta))$ is in* PFI *and any of the following three conditions holds, then $\bar{d}((\alpha, \beta), (\gamma, \delta)) = \mu(C) - \mu(A)$.*

i) $\beta + \gamma + \delta \leqslant 2$ *and* $(\alpha - \gamma)(\beta + \gamma - 1) \leqslant (\delta - \beta)\gamma$.

ii) $\alpha + \beta + \gamma \leqslant 2$ *and* $(\delta - \beta)(\beta + \gamma - 1) \leqslant (\alpha - \gamma)\beta$.

iii) $\alpha + \beta = \gamma + \delta$.

PROOF. Note that if $((\alpha, \beta), (\gamma, \delta))$ is in PFI and satisfies (ii), then $((\delta, \gamma), (\beta, \alpha))$ is in PFI and satisfies (i); hence to prove the theorem it suffices to show that pairs in PFI satisfying (i) or (iii) have $\bar{d}$ equal to the partition distance. For such pairs a joint process will be constructed which has $A \subseteq C$. The states of the joint process will be $p(= A \cap C)$, $q(= B \cap C)$, and $r(= B \cap D)$. If T is the transformation being constructed and e and f are subsets of the measure space, "send κ of the measure in e to f" means $\mu(f \cap Te) = \kappa\mu(e)$, and if P is a partition of the measure space, "send κ of the measure of e to f irrespective of P-history" means that for every positive integer m and for every sequence of partition elements $p_1, \cdots, p_m$ of p,

$$\mu(f \cap T(e \cap T^1p_1 \cap \cdots \cap T^m p_m)) = \kappa\mu(e \cap T^1p_1 \cap \cdots \cap T^m p_m).$$

Let $(\gamma, \delta) = (\alpha - u, \beta + w)$. Construct the joint process as follows: Send γ of the measure in p and γ of the measure in q to r irrespective of $\{p, q, r\}$-history and δ of the measure in r to $p \cup q$ irrespective of $\{p, q, r\}$-history. This insures that the $\{p \cup q, r\}$-process is isomorphic to (γ, δ). Furthermore, send u of the measure in p to q irrespective of $\{p, q, r\}$-history. What remains to be determined is how measure in q and r is sent to p. It will be seen that this construction can be successfully completed (so that the $\{p, q \cup r\}$-process is isomorphic to (α, β)) if and only if (i) or (iii) hold.

If $\beta + \gamma \leqslant 1$, send β of the measure in q and β of the measure in r to p

irrespective of $\{p, q, r\}$-history. One then has a joint process which is Markov on the joint atoms and which attains the partition distance.

Henceforth assume $\beta + \gamma > 1$. Thus q cannot send β of its measure to p since q sends γ of its measure to r; the most q can send to p is $1 - \gamma$ of its measure. Have q send $1 - \gamma$ of its measure to p irrespective of $\{p, q, r\}$-history. The $\{p, q \cup r\}$-process will be isomorphic to (α, β) if and only if $q \cup r$ can be made to send β of its measure to p irrespective of $\{p, q \cup r\}$-history. Since the part in q sends only $1 - \gamma$ of its measure to p, the part in r must "make up the slack." This construction can be successfully completed if and only if for every sequence of $\{p, q \cup r\}$-history, "making up the slack" never forces the part in r with that $\{p, q \cup r\}$-history to send more than δ of its measure to p.

Since measure entering $q \cup r$ from p distributes itself between q and r in ratio $u{:}\gamma$ irrespective of its $\{p, q, r\}$-history, the way measure satisfying any sequence of $\{p, q \cup r\}$-history whose final state (present state) is $q \cup r$ is distributed between q and r depends only on how far back in the sequence a p last occurred (assuming a p appears in the sequence).

Let q_n be the part of q which was last in p n steps ago,
r_n be the part of r which was last in p n steps ago,
v_n be the fraction of r_n which must be sent to p.

Since $q_n \cup r_n$ must send β of its measure to p and q_n sends $1 - \gamma$ of its measure to p,

$$v_n = \beta + \frac{\mu(q_n)}{\mu(r_n)}(\beta + \gamma - 1). \tag{1}$$

Since r_n sends $\delta - v_n$ of its measure to q, by (1)

$$\mu(q_{n+1}) = (\delta - v_n)\mu(r_n) = w\mu(r_n) - (\beta + \gamma - 1)\mu(q_n) \tag{2}$$

and

$$\mu(r_{n+1}) = (1 - \delta)\mu(r_n) + \gamma\mu(q_n). \tag{3}$$

Using (2) and (3) and the fact that $\mu(q_1) = \beta u/(\alpha + \beta)$ and $\mu(r_1) = \beta\gamma/(\alpha + \beta)$, it can be shown by induction that for all $n \in \mathbb{N}$

$$\mu(q_{n+1}) = \frac{\alpha\beta(1 - \beta)^n}{(\alpha + \beta)(\gamma + w)}\left[w - \frac{\gamma(w - u)}{\alpha}\left(\frac{1 - \gamma - \delta}{1 - \beta}\right)^n\right] \tag{4}$$

and

$$\mu(r_{n+1}) = \frac{\alpha\beta(1 - \beta)^n}{(\alpha + \beta)(\gamma + w)}\left[\gamma + \frac{\gamma(w - u)}{\alpha}\left(\frac{1 - \gamma - \delta}{1 - \beta}\right)^n\right]. \tag{5}$$

If (iii) holds then $w - u = 0$, so for all $n \in \mathbb{N}$, $\mu(q_{n+1}){:}\mu(r_{n+1}) = w{:}\gamma$, whence for all $n \in \mathbb{N}$, $v_{n+1} = \beta + (w/\gamma)(\beta + \gamma - 1) \leqslant \delta$, hence the construction is successfully completed; in fact, since v_{n+1} does not depend on n, the joint process constructed is Markov on the joint atoms. Henceforth assume (iii) does not hold, so $w - u \neq 0$. Then, since $\mu(q_{n+1})$ and $\mu(r_{n+1})$ must never be negative, $(1 - \gamma - \delta)/(1 - \beta) \geqslant -1$, that is, $\beta + \gamma + \delta \leq 2$. What must be guaranteed is that all the v_n's are $\leq \delta$. If $u > w$, then (1), (4), and (5) imply the largest of the v_n's is v_1, and

$$v_1 = \beta + \frac{\mu(q_1)}{\mu(r_1)}(\beta + \gamma - 1) = \beta + \frac{u}{\gamma}(\beta + \gamma - 1).$$

Hence, if $\beta + \gamma + \delta \leq 2$ and $u > w$, the construction can be completed if and only if $u(\beta + \gamma - 1) \leq w\gamma$. This inequality is insured by (i). If $u < w$, then the largest of the v_n's is v_2, and

$$\begin{aligned} v_2 &= \beta + \frac{\mu(q_2)}{\mu(r_2)}(\beta + \gamma - 1) \\ &= \beta + \frac{w\gamma - (\beta + \gamma - 1)u}{(1 - \delta + u)\gamma}(\beta + \gamma - 1) \\ &\leq \beta + \frac{w}{1 - \delta}(\beta + \gamma - 1) \\ &\leq \delta. \end{aligned}$$

Hence if $\beta + \gamma + \delta \leq 2$ and $u < w$, the construction can be completed. Note that in this case (i) holds. ▯

In [1] it is shown that a Markov joining of $((\alpha, \beta), (\gamma, \delta))$ can attain the partition distance between them if and only if the pair is in PFI and either $\beta + \gamma \leq 1$ or $\alpha + \beta = \gamma + \delta$. Note for these pairs the joint process constructed in Theorem 1 is a Markov joining. (Furthermore, in [2] it is shown that except for the above-mentioned cases, Markov joinings of two state Markov processes with positive entropy never attain $\bar{d}$.)

For what pairs not covered by Theorem 1 is $\bar{d}$ the partition distance? Kamae gave the following necessary condition.

Let (α_n, β_n) (respectively (γ_n, δ_n)) denote the two-state Markov process whose transition matrix is the nth power of the transition matrix of (α, β) (respectively (γ, δ)).

LEMMA 2. *If* $\bar{d}((\alpha, \beta), (\gamma, \delta)) = \mu(C) - \mu(A)$, *then*

$$(*) \qquad \forall n \in \mathbb{N}^+ \quad ((\alpha_n, \beta_n), (\gamma_n, \delta_n)) \text{ is in PFI.}$$

PROOF. If T is the transformation for a joint process attaining the partition distance, then

$$\begin{aligned} \forall n \in \mathbb{N}^+ \forall k \in \mathbb{N} \quad \mu\left(\cap_{i=0}^{k} T^{-in} A\right) &= \mu(A)(1 - \alpha_n)^k \\ &\leq \mu\left(\cap_{i=0}^{k} T^{-in} C\right) \leq \mu(C)(1 - \gamma_n)^k \end{aligned}$$

and

$$\mu\left(\cap_{i=0}^{k} T^{-in} D\right) = \mu(D)(1 - \delta_n)^k \leq \mu\left(\cap_{i=0}^{k} T^{-in} B\right) = \mu(B)(1 - \beta_n)^k,$$

whence, since $\mu(A) > 0$ and $\mu(D) > 0$, $\forall n \in \mathbb{N}^+$ $\alpha_n \geq \gamma_n$ and $\delta_n \geq \beta_n$. ▯

Kamae conjectured that (*) is sufficient to insure $\bar{d}((\alpha, \beta), (\gamma, \delta)) = \mu(C) - \mu(A)$. Not all pairs satisfying (*) satisfy any of the conditions given in Theorem 1.

EXAMPLES.

By Theorem 1 (i), $\bar{d}((.6, .6), (.6, .8)) = \frac{1}{14}$.

By Theorem 1 (ii), $\bar{d}((.8, .6), (.6, .9)) = \frac{6}{35}$.

By Theorem 1 (iii), $\bar{d}((.8, .8), (.7, .9)) = \frac{1}{16}$.

Pairs ((.65, .65), (.6, .8)) and ((.8, .8), (.68, 1)) satisfy (*) but do not satisfy (i) or (ii) or (iii) of Theorem 1.

Pairs ((.8, .8), (.7, 1)) and ((.8, .6), (.8, .7)) are in PFI but do not satisfy (*), since $((.8, .8)^2, (.7, 1)^2) = ((.32, .32), (.21, .3))$ which has an equivalent in PFII, and likewise $((.8, .6)^2, (.8, .7)^2)$ has an equivalent in PFII.

I do not know whether the $\bar{d}$-distance equals the partition distance for pairs satisfying (*) but not covered by Theorem 1.

REFERENCES

[1] ELLIS, M. H. (1978). Distances between two-state Markov processes attainable by Markov joinings. *Trans. Amer. Math. Soc.* **241** 129-153.

[2] ELLIS, M. H. (1980). Conditions for attaining $\bar{d}$ by a Markovian joining. To appear in *Ann. Probability*.

[3] KAMAE, T., Private communication.

[4] ORNSTEIN, D. S. (1973). An application of ergodic theory to probability theory. *Ann Probability* **1** 43-65.

[5] ORNSTEIN, D. S. (1974). *Ergodic theory, randomness and dynamical systems.* Yale University Press, New Haven.

DEPARTMENT OF MATHEMATICS
NORTHEASTERN UNIVERSITY
360 HUNTINGTON AVENUE
BOSTON, MASSACHUSETTS 02115