J. Appl. Prob. **12**, 661–672 (1975)
Printed in Israel

AN ERGODIC THEOREM FOR EVOLUTION IN A RANDOM ENVIRONMENT

M. FRANK NORMAN *University of Pennsylvania*

Abstract

Let w_1, w_{12} and w_2 be the fitnesses of genotypes A_1A_1, A_1A_2 and A_2A_2 in an infinite diploid population, and let p_n be the A_1 gene frequency in the nth generation. If fitness varies independently from generation to generation, then p_n is a Markov process with a continuum of states. If $E[\ln(w_i/w_{12})]<0$ for $i=1,2$, then there is a unique stationary probability, and the distribution of p_n converges to it as $n\to\infty$.

GENETIC MODELS; RANDOM ENVIRONMENTS; MARKOV PROCESSES; ERGODIC THEORY

1. Introduction

Consider an infinite monoecious diploid population with two alleles, A_1 and A_2, at some locus and A_1 gene frequency p_n in the nth generation. If mating is random, and the genotypes A_1A_1, A_1A_2 and A_2A_2 have fitnesses w_1, w_{12} and w_2, then p_n satisfies the classical difference equation

$$p_{n+1}=\frac{p_n^2w_1+p_nq_nw_{12}}{p_n^2w_1+2p_nq_nw_{12}+q_n^2w_2}$$

($q=1-p$). We shall assume, throughout the paper, that all fitnesses are strictly positive, and $0<p_0<1$. Without loss of generality, we take $w_{12}=1$.

The trajectory of p_n when fitness is constant over time is well understood (see, for example, Ewens (1969)). Recently, several investigators have attempted to model random temporal variation in the environment by introducing random fitnesses w_{1n} and w_{2n}. We mention the work of Gillespie (1973), to whom we owe our interest in this subject, and a paper of Karlin and Lieberman (1974), which contributes a variety of new results and reviews a number of earlier studies. The main thrust of this work is that random variation in fitness greatly enhances the ability of this model to interpret polymorphism.

Having determined that $w_n=(w_{1n},w_{2n})$ is to be regarded as a stochastic process, it remains to specify its properties. A rather natural assumption is stationarity, but we shall impose the stronger condition that the random vectors w_n, $n\geqq 0$, are independent and identically distributed (and independent of p_0).

Received in revised form 26 February 1975.

Their common distribution is denoted Π ($\Pi(A) = P(w_n \in A)$). The dependence between w_{1n} and w_{2n} is unrestricted. It is assumed that, for some $\varepsilon > 0$,

$$E(\exp(\varepsilon \mid \ln w_{in} \mid)) < \infty, \qquad i = 1, 2. \tag{1}$$

This is certainly true if w_{in} is bounded away from 0 and ∞.

The assumption that the random vectors w_n are independent and identically distributed implies that $\{p_n, n \geqq 0\}$ is a Markov process in $(0, 1)$ with stationary transition probabilities. Let $\mathscr{L}(p_0)$ be the distribution of p_0, and let $\Rightarrow$ denote convergence in distribution. The purpose of this paper is to prove the following ergodic theorem.

Theorem. *If $E(\ln w_{1n}) < 0$ and $E(\ln w_{2n}) < 0$, then $\{p_n, n \geqq 0\}$ has a unique stationary distribution, μ. For any $\mathscr{L}(p_0)$, $p_n \Rightarrow \mu$ as $n \to \infty$.*

Gillespie (1973) made the fundamental observation that the behaviour of p_n near 1 (0) hinges on the sign of $E(\ln w_{1n})$ ($E(\ln w_{2n})$). Clearly $E(\ln w_{in}) < 0$ is not inconsistent with $E(w_{in}) > 1$. Thus both conditions $E(\ln w_{in}) < 0$ of the theorem can be satisfied, so that a polymorphism is maintained, even though $E(w_{1n}) > 1$ and $E(w_{2n}) < 1$, a situation which would produce fixation of A_1 in the absence of environmental variation. Karlin and Lieberman (1974) indicate that convergence of $\mathscr{L}(p_n)$ is a very general phenomenon, even for dependent fitness processes $\{w_n, n \geqq 0\}$, but they give no proofs.

2. Two lemmas

Let X_n be the Markov process $X_n = \ln(p_n/q_n)$, with state space $R = (-\infty, \infty)$. We begin the proof with two lemmas.

Lemma 1. *For any λ with $\mid \lambda \mid$ sufficiently small, and any initial state $x \in R$, the sequence $\{E_x(\exp(\lambda X_n)), n \geqq 0\}$ is bounded.*

This implies that $\{\mathscr{L}(X_n \mid X_0 = x), n \geqq 0\}$ is tight, and

$$E_x(\liminf \exp(\lambda X_n)) < \infty,$$

by Fatou's lemma. (We write $\liminf a_n$ instead of $\liminf_{n \to \infty} a_n$.) Consequently,

$$\liminf X_n < \infty \tag{2}$$

and

$$\limsup X_n > -\infty \tag{3}$$

almost surely (a.s.). Thus we can see already that fixation of either allele has probability 0. The proof of Lemma 1 for $\lambda > 0$ depends only on (1) and $E(\ln w_{1n}) < 0$, hence (2) holds under these conditions. Lamperti's ((1960), Theorem 3.1) methods can be adapted to show that (2) holds even when $E(\ln w_{1n}) = 0$, provided that $\operatorname{var}(\ln w_{1n}) > 0$ and $|\ln w_{in}| \leqq K$ for some constant K and $i = 1, 2$.

For $x \in \bar{R} = [-\infty, \infty]$, let

$$\delta(x) = \delta(x, w) = \ln\left(\frac{pw_1 + q}{qw_2 + p}\right), \tag{4}$$

where $x = \ln(p/q)$. Then

$$\Delta X_n = X_{n+1} - X_n = \delta(X_n, w_n),$$

$$P(\Delta X_n \in B \mid X_n = x) = \Pi(\delta(x) \in B),$$

and

$$E(f(\Delta X_n) \mid X_n = x) = \int f(\delta(x, w))\Pi(dw) \tag{5}$$

for $x \in R$.

Proof of Lemma 1. For fixed w, $\delta(x, w)$ is monotonic in x, with $\delta(-\infty, w) = -\ln w_2$ and $\delta(\infty) = \ln w_1$, hence

$$|\delta(x, w)| \leqq \max\{|\ln w_1|, |\ln w_2|\}. \tag{6}$$

In view of (5),

$$E(e^{\varepsilon|\Delta X_n|} \mid X_n = x) \leqq K, \tag{7}$$

where $K = \int e^{\varepsilon|\ln w_1|}\Pi(dw) + \int e^{\varepsilon|\ln w_2|}\Pi(dw)$ is finite by (1). Also (5), (6) and the dominated convergence theorem imply

$$E(\Delta X_n \mid X_n = x) \to a \tag{8}$$

as $x \to \infty$, where $a = \int \ln w_1 \Pi(dw) < 0$.

The remainder of the proof for $\lambda > 0$, which we now present, depends only on (7) and (8), hence does not require $\int \ln w_2 \Pi(dw) < 0$. The proof for $\lambda < 0$ is similar, and does not require $\int \ln w_1 \Pi(dw) < 0$.

Taylor expansion yields $e^y = 1 + y + 2^{-1}y^2e^{by}$, where $0 \leqq b \leqq 1$, hence, for $\lambda < \varepsilon$,

$$\begin{aligned} E(e^{\lambda\Delta X_n} \mid X_n = x) &\leqq 1 + \lambda E(\Delta X_n \mid X_n = x) \\ &\quad + 2^{-1}\lambda^2 E((\Delta X_n)^2 e^{\lambda|\Delta X_n|} \mid X_n = x). \end{aligned} \tag{9}$$

In view of (8), there is a constant, k, such that $E(\Delta X_n \mid X_n = x) \leqq 2^{-1}a$ for $x \geqq k$. Moreover,

$$2^{-1}(\Delta X_n)^2 \leqq (\varepsilon - \lambda)^{-2} e^{(\varepsilon - \lambda)|\Delta X_n|},$$

so it follows from (7) that the last term on the right in (9) is at most $\lambda^2(\varepsilon-\lambda)^{-2}K$. Therefore

$$
\begin{aligned}
E(e^{\lambda \Delta X_n} \mid X_n = x) &\leqq 1 + 2^{-1} a\lambda + \lambda^2(\varepsilon-\lambda)^{-2}K \\
&= \alpha_\lambda
\end{aligned} \tag{10}
$$

for $x \geqq k$. Since $a<0$, there is an $\varepsilon' \in (0, \varepsilon]$ such that $0<\alpha_\lambda<1$ for $0<\lambda<\varepsilon'$.

It follows from (10) that

$$E(e^{\lambda X_{n+1}} \mid X_n = x) \leqq \alpha_\lambda e^{\lambda x}$$

for $x \geqq k$. As a consequence of (7),

$$
\begin{aligned}
E(e^{\lambda X_{n+1}} \mid X_n = x) &\leqq K e^{\lambda x} \\
&\leqq K e^{\lambda k}
\end{aligned}
$$

for $x \leqq k$, hence

$$E(e^{\lambda X_{n+1}} \mid X_n = x) \leqq \alpha_\lambda e^{\lambda x} + K e^{\lambda k}$$

for all $x \in R$. Therefore

$$E_x(e^{\lambda X_{n+1}}) \leqq \alpha_\lambda E_x(e^{\lambda X_n}) + K e^{\lambda k},$$

and thus, by induction,

$$E_x(e^{\lambda X_n}) \leqq \alpha_\lambda^n e^{\lambda x} + (1-\alpha_\lambda)^{-1} K e^{\lambda k}.$$

This completes the proof.

For $x \in \bar{R}$, let

$$P^+(x) = \Pi(\delta(x) > 0) \quad \text{and} \quad P^-(x) = \Pi(\delta(x) < 0).$$

Lemma 2. *If* $c \in (-\infty, \infty]$, *and* $P^+(x) > 0$ *for all* $x \in (-\infty, c)$, *then* $\limsup X_n \geqq c$ *a.s. If* $c \in [-\infty, \infty)$, *and* $P^-(x) > 0$ *for all* $x \in (c, \infty)$, *then* $\liminf X_n \leqq c$ *a.s.*

The first statement of Lemma 2 can be rephrased as follows. If $P(X_{n+1} > x \mid X_n = x) > 0$ for all $x \in (-\infty, c)$, then $P(X_n > x \text{ i.o.}) = 1$ for all $x \in (-\infty, c)$.

Proof. We begin by showing that, for any $\gamma \in R$ and $y \in \bar{R}$,

$$\liminf\nolimits_{x \to y} \Pi(\delta(x) > \gamma) \geqq \Pi(\delta(y) > \gamma), \tag{11}$$

i.e., $\Pi(\delta(x) > \gamma)$ is a lower semicontinuous function of $x \in \bar{R}$. Let $\chi(u) = 1$ if $u > \gamma$ and $\chi(u) = 0$ if $u \leqq \gamma$. Since $\delta(x, w)$ is continuous in x for each w, χ is continuous except at γ, and $\chi(\gamma) = 0$, it follows that $\chi(\delta(x, w))$ is lower semicontinuous in x. But

$$\Pi(\delta(x) > \gamma) = \int \chi(\delta(x, w))\Pi(dw),$$

so (11) follows from Fatou's lemma.

The two statements of Lemma 2 are symmetric, so we consider only the first. Let $y \in (-\infty, c)$ be given. Since $P^+(y) > 0$, there is a $\gamma = \gamma_y > 0$ such that $b = \Pi(\delta(y) > \gamma) > 0$. Since $\Pi(\delta(x) > \gamma)$ is a lower semicontinuous function of x, there is an $\eta = \eta_y \in (0, \gamma/2)$ such that $\Pi(\delta(x) > \gamma) \geqq b/2$ if $x \in I_y = (y - \eta, y + \eta)$. Now $X_n \in I_y$ and $\Delta X_n > \gamma$ imply $X_{n+1} > y + \eta$, so

$$P(X_{n+1} > y + \eta \mid X_n) \geqq b/2$$

for $X_n \in I_y$. It follows (see Theorem 9.5.2 of Chung (1974)) that

$$[X_n \in I_y \text{ i.o.}] - [X_n > y + \eta \text{ i.o.}]$$

has probability 0. But lim sup $X_n \in I_y$ implies this event, so

$$P(\limsup X_n \in I_y) = 0.$$

According to the Lindelöf theorem, there is a countable set, S, such that $\bigcup_{y \in S} I_y \supset (-\infty, c)$. Hence

$$P(-\infty < \limsup X_n < c) = 0.$$

In view of (3), the proof of Lemma 2 is complete.

The remainder of the proof of the ergodic theorem involves consideration of two cases.

Case 1. $P^+(x) > 0$ for all $x \in R$, or $P^-(x) > 0$ for all $x \in R$.
Case 2 is the residual case.

In view of Lemma 2, $\limsup X_n = \infty$ a.s. or $\liminf X_n = -\infty$ a.s. in Case 1. This case is treated in Section 3. The remaining sections treat Case 2. It will emerge that, in this case, both lim sup X_n and lim inf X_n are finite a.s.

3. Case 1

Since the two possibilities under this case are symmetrical, it suffices to consider the first, $P^+(x) > 0$ for all $x \in R$. By Lemma 2, lim sup $X_n = \infty$ a.s. Consequently, for any $y \in R$, the random variable N defined by

$$N = \min\{n \geqq 0: X_n \geqq y\}$$

is finite a.s.

Let U be the transition operator of the process $\{X_n\}$,

$$Uf(x) = E(f(X_{n+1}) \mid X_n = x)$$
$$(12) \qquad = \int f(x + \delta(x, w))\Pi(dw),$$

for f bounded and measurable on R. Its iterates, U^k, satisfy

$$U^k f(X_n) = E(f(X_{n+k}) \mid \mathcal{F}_n)$$

a.s., where $\mathcal{F}_n$ is the σ-field generated by $X_0, \cdots, X_n$. Differentiating (4) with respect to $x \in R$, we obtain the important equality

$$(13) \qquad \delta'(x) = \frac{pq(w_1 w_2 - 1)}{(pw_1 + q)(qw_2 + p)}.$$

Hence

$$(14) \qquad \delta'(x) > -\frac{pq}{(pw_1 + q)(qw_2 + p)}$$
$$> -1.$$

Consequently, $x + \delta(x, w)$ is an increasing function of x for every w. It follows from (12) that *U maps non-decreasing functions into non-decreasing functions, as does U^k.*

Suppose, henceforth, that f is bounded, non-negative and non-decreasing. Since $f \geqq 0$,

$$E(f(X_n)) \geqq E(f(X_n), N \leqq n)$$

$(E(Y, A) = E(YI_A)$, where I_A is the indicator of $A)$

$$= \sum_{j=0}^{n} E(f(X_n), N = j)$$

$$= \sum_{j=0}^{n} E(E(f(X_n), N = j \mid \mathcal{F}_j))$$

$$= \sum_{j=0}^{n} E(E(f(X_n) \mid \mathcal{F}_j), N = j)$$

$$= \sum_{j=0}^{n} E(U^{n-j} f(X_j), N = j)$$

$$= \sum_{j=0}^{n} E(U^{n-N} f(X_N), N = j)$$

$$= E(U^{n-N} f(X_N), N \leqq n)$$

$$\geqq E(U^{n-N} f(y), N \leqq n),$$

since $X_N \geqq y$. Fatou's lemma and $N < \infty$ a.s. then yield

(15) $$\liminf E(f(X_n)) \geqq \liminf U^n f(y).$$

This holds for every $y \in R$. Note that the quantity on the right is a non-decreasing function of y. Let

$$U^\infty f = \lim_{y\to\infty} \liminf_{n\to\infty} U^n f(y).$$

It follows from (15) that

(16) $$\liminf E(f(X_n)) \geqq U^\infty f.$$

This is valid for any $\mathscr{L}(X_0)$, and $U^\infty f$ does not depend on $\mathscr{L}(X_0)$.

For any $n \geqq k \geqq 0$,

$$\begin{aligned} E(f(X_n)) &= E(U^k f(X_{n-k})) \\ &= E(U^k f(X_{n-k}),\ X_{n-k} \leqq y) \\ &\quad + E(U^k f(X_{n-k}),\ X_{n-k} > y) \\ &\leqq U^k f(y) + \|f\| M(y), \end{aligned}$$

where $\|f\| = \sup_{x\in R} |f(x)|$ and

$$M(y) = \sup_{j\geqq 0} P(X_j > y).$$

It follows that

$$\limsup E(f(X_n)) \leqq U^k f(y) + \|f\| M(y)$$

for all $k \geqq 0$, so

(17) $$\limsup E(f(X_n)) \leqq \liminf U^n f(y) + \|f\| M(y).$$

Suppose now that $X_0 = x$ a.s., for some constant x. Letting $y \to \infty$ in (17) and noting that $M(y) \to 0$ as a consequence of Lemma 1, we obtain

$$\limsup E_x(f(X_n)) \leqq U^\infty f.$$

In combination with (16), this yields

(18) $$\lim_{n\to\infty} E_x(f(X_n)) = U^\infty f$$

for all $x \in R$ and all bounded, non-negative, non-decreasing f.

Since $\{\mathscr{L}(X_n \mid X_0 = 0), n \geqq 0\}$ is tight, there is a subsequence that converges to a probability ν. By (18), with $x = 0$, $U^\infty f = \int f\, d\nu$ if f is bounded, continuous, non-negative, and non-decreasing. If probabilities ν_1 and ν_2 satisfy $\int f\, d\nu_1 = \int f\, d\nu_2$ for all such functions, then $\nu_1 = \nu_2$. Hence (18) and tightness of $\{\mathscr{L}(X_n \mid X_0 = x), n \geqq 0\}$ imply that $X_n \Rightarrow \nu$ for all initial states x (Breiman (1968), Corollary 8.16), hence for all initial distributions $\mathscr{L}(X_0)$.

If f is bounded and continuous, Uf is too, hence

$$E(f(X_{n+1})) = E(Uf(X_n))$$
$$\to \int Uf\,d\nu$$

as $n\to\infty$. But $E(f(X_{n+1}))\to\int f\,d\nu$, so $\int Uf\,d\nu = \int f\,d\nu$ for all bounded continuous f. This implies that ν is stationary. If ν' is any stationary distribution, take $\mathscr{L}(X_0)=\nu'$ and note that $\nu'=\mathscr{L}(X_n)\Rightarrow\nu$, as $n\to\infty$, so $\nu'=\nu$. Thus ν is the unique stationary distribution of $\{X_n, n\geqq 0\}$. Our results for X_n translate immediately into comparable results for p_n.

4. Case 2, preliminaries

Note first that $P^+(-\infty)=\Pi(-\ln w_2>0)>0$, since $E(\ln w_2)<0$. Since P^+ is lower semicontinuous, $P^+(x)>0$ for x in some neighbourhood of $-\infty$. On the other hand, in Case 2, $P^+(x)=0$ for some $x\in R$, hence

$$a^+=\sup\{y\colon P^+(x)>0 \text{ for all } x<y\}$$

is finite, as is

$$a^-=\inf\{y\colon P^-(x)>0 \text{ for all } x>y\}.$$

We say that A implies B a.s. if $P(A-B)=0$.

Lemma 3. *$P^+(a^+)=0$, $(-\infty,a^+)$ is stochastically closed, and $X_0<a^+$ implies* $\limsup X_n = a^+$ *a.s. Similarly, $P^-(a^-)=0$, (a^-,∞) is stochastically closed, and $X_0>a^-$ implies* $\liminf X_n=a^-$ a.s. *Finally, $a^-\leqq a^+$.*

Proof. There is a sequence x_j such that $x_j\geqq a^+$, $P^+(x_j)=0$, and $x_j\to a^+$ as $j\to\infty$. But P^+ is lower semicontinuous, so $P^+(a^+)=0$, or $\delta(a^+,w)\leqq 0$ for almost all w. Thus $a^+ + \delta(a^+,w)\leqq a^+$ a.s. But $x+\delta(x,w)$ is a strictly increasing function of x, so, for almost all w, $x+\delta(x,w)<a^+$ for all $x<a^+$. Hence $P(X_{n+1}<a^+\mid X_n=x)=1$ for $x<a^+$, and $(-\infty,a^+)$ is stochastically closed. It follows that $X_0<a^+$ implies $\limsup X_n\leqq a^+$ a.s. But $P^+(x)>0$ for $x<a^+$, so $\limsup X_n\geqq a^+$ a.s. by Lemma 2. Thus $X_0<a^+$ implies $\limsup X_n = a^+$ a.s. The assertions concerning a^- hold by symmetry.

Since $P^+(a^+)=0$ and $P^-(a^-)=0$, we have $\delta(a^+,w)\leqq 0$ and $\delta(a^-,w)\geqq 0$ for almost all w. Suppose that $a^+<a^-$. In view of (13), we must have $w_1w_2\geqq 1$ a.s., hence $\ln w_1+\ln w_2\geqq 0$ a.s., and $E(\ln w_1)+E(\ln w_2)\geqq 0$. This, however, is inconsistent with our assumption that both of these expectations are strictly negative. Consequently $a^+\geqq a^-$. This concludes the proof of Lemma 3.

We shall distinguish two subcases of Case 2.

Case 2a. $a^+ > a^-$

Case 2b. $a^+ = a^-$

These are considered in the two subsequent sections. Let $I = [a^-, a^+]$. It follows from Lemma 3 that $(-\infty, a^+]$ and $[a^-, \infty)$ are stochastically closed, hence their intersection, I, is too. Moreover, in Case 2a, the lemma implies that $X_n \in I$ for some n, thus for all sufficiently large n, a.s. In Case 2b, I reduces to a single point, a, and we shall see that $\lim_{n\to\infty} X_n = a$ a.s. Hence, in either case, the limiting distribution of X_n is concentrated on I.

5. Case 2a

Since $\delta(a^-) \geqq 0$ and $\delta(a^+) \leqq 0$ a.s., (13) shows that $w_1 w_2 \leqq 1$ a.s. Moreover $w_1 w_2 = 1$ a.s. is inconsistent with $E(\ln w_i) < 0$, $i = 1, 2$, so $\Pi(\delta'(x) < 0) = \Pi(w_1 w_2 < 1) > 0$. Since $-1 < \delta'(x) \leqq 0$, we have $0 < 1 + \delta'(x) \leqq 1$. If f' is bounded, then

$$(d/dx) f(x + \delta(x)) = f'(x + \delta(x))(1 + \delta'(x))$$

is bounded, so differentiation and integration can be interchanged in (12) yielding

$$(Uf)'(x) = \int f'(x + \delta(x))(1 + \delta'(x))\, d\Pi.$$

Let $\|f\| = \sup_{x \in I} |f(x)|$, where $I = [a^-, a^+]$. As noted previously, I is stochastically closed, hence

$$|(Uf)'(x)| \leqq \|f'\| \left(1 + \int \delta'(x)\, d\Pi\right)$$

for $x \in I$, and

$$\|(Uf)'\| \leqq \alpha \|f'\|,$$

where

$$\alpha = \max_{x \in I} \left(1 + \int \delta'(x)\, d\Pi\right).$$

By induction,

$$\|(U^n f)'\| \leqq \alpha^n \|f'\|.$$

But $\int \delta'(x)\, d\Pi$ is continuous and strictly negative on the compact interval I, so $\alpha < 1$, and $\|(U^n f)'\| \to 0$ as $n \to \infty$. Now $\sup_{x \in I} Uf(x) \leqq \sup_{x \in I} f(x)$, so, by induction, $s_n = \sup_{x \in I} U^n f(x)$ is a non-increasing sequence. Similarly, $i_n = \inf_{x \in I} U^n f(x)$ is non-decreasing. But

$$s_n - i_n \leqq (a^+ - a^-)\|(U^n f)'\|,$$

so s_n and i_n converge to the same limit, which we denote $U^\infty f$. Since, for each $x \in I$, $U^n f(x)$ and $U^\infty f$ are in $[i_n, s_n]$, we see that $\| U^n f - U^\infty f \| \to 0$ as $n \to \infty$.

Suppose now that $\mathscr{L}(X_0)$ is unrestricted. By Lemma 3, $X_n \in I$ for some n a.s., so $N = \min\{n: X_n \in I\}$ is finite a.s. As in our treatment of Case 1,

$$E(f(X_n)) = E(U^{n-N} f(X_N), N \leqq n) + E(f(X_n), N > n).$$

Since $X_N \in I$, the first term on the right converges to $U^\infty f$ as $n \to \infty$. The second term converges to 0. Hence $E(f(X_n)) \to U^\infty f$, for any $\mathscr{L}(X_0)$ and bounded function f with bounded derivative. The assertions of the theorem follow easily from this.

6. Case 2b

It suffices to show that $X_n \to a$ a.s., where a is the common value of a^+ and a^-. The asymptotic behavior of p_n in this case was described by Karlin and Lieberman (1974).

We can, without loss of generality, assume that $X_0 = x_0$ a.s. and $x_0 < a$. Let

$$Y_n = Y_n(\alpha) = -(a - X_n)^\alpha,$$

where $0 < \alpha < 1$ will be determined subsequently. According to Lemma 3, lim sup $Y_n = 0$ a.s. We shall show that

$$E(\Delta Y_n(\alpha) \mid Y_n(\alpha) = y) \geqq 0, \tag{19}$$

if $y < 0$ is sufficiently close to 0. Lamperti's (1960) Theorem 2.2 then yields $\lim Y_n = 0$ or $\lim X_n = a$ a.s. (Lamperti's assumption that the process is non-negative is unnecessary.) Thus it remains only to establish (19).

If $Y_n = y = -(a - x)^\alpha$, then

$$\Delta Y_n / |y| = 1 - (1 + \Delta X_n (x - a)^{-1})^\alpha.$$

(The quantity $1 + \Delta X_n (x - a)^{-1}$ is a.s. positive.) Hence it is sufficient to show that

$$\int (1 + \delta(x)(x - a)^{-1})^\alpha d\Pi \leqq 1 \tag{20}$$

for $x < a$ sufficiently large. Since $P^+(a) = 0$ and $P^-(a) = 0$, we have $\delta(a, w) = 0$ for almost all w. Hence

$$\begin{aligned} \delta(x)(x - a)^{-1} &\leqq \sup_{y \in R} \delta'(y) \\ &\leqq \max\{w_1 w_2 - 1, 0\} \end{aligned}$$

by (13), and the integrand in (20) is bounded by $\exp(\alpha|\ln w_1 w_2|)$, which is integrable for $\alpha \leqq \varepsilon/2$. Thus the dominated convergence theorem implies that the integral in (20) converges to $g(\alpha) = \int (1+\delta'(a))^\alpha d\Pi$ as $x \uparrow a$. Consequently, we need only show that $g(\alpha) < 1$ for some α. Granting, for the moment, that

$$g(\alpha)^{1/\alpha} \to \exp\left[\int \ln(1+\delta'(a))\, d\Pi\right] \tag{21}$$

as $\alpha \downarrow 0$, it remains only to prove that

$$d = \int \ln(1+\delta'(a))\, d\Pi < 0. \tag{22}$$

Let $\zeta = \zeta(w) = \ln(1+\delta'(a, w))$. To establish (21), we must show that $(g(\alpha) - 1)\alpha^{-1} \to \int \zeta\, d\Pi$ as $\alpha \downarrow 0$, or

$$\lim_{\alpha \downarrow 0} \int (e^{\alpha\zeta} - 1)\alpha^{-1}\, d\Pi = \int \zeta\, d\Pi. \tag{23}$$

The integrand in (23) equals $\zeta \int_0^1 e^{u\alpha\zeta} du$, which clearly decreases to ζ as $\alpha \downarrow 0$. Moreover, the integral in (23) exists for $\alpha \leqq \varepsilon/2$. Hence (23) follows from the monotone convergence theorem.

We shall now complete the proof by establishing (22). In view of (4), $\delta(a) = 0$ a.s. implies that $(w_1 - 1)p = (w_2 - 1)q$ a.s. (where $a = \ln(p/q)$). Let this quantity be denoted z, so that $w_1 = 1 + (z/p)$ and $w_2 = 1 + (z/q)$ a.s. Substituting into (13), we see that

$$\delta'(a) = \frac{z}{z+1}$$

or

$$\delta'(a) = \frac{w_1 - 1}{w_1 + r} = \frac{w_2 - 1}{w_2 + s} \tag{24}$$

a.s., where $r = q/p$ and $s = p/q$. If $u > 0$ and x is sufficiently large that $(e^x - 1)(e^x + u)^{-1} > -1$, let

$$f_u(x) = \ln[1 + (e^x - 1)(e^x + u)^{-1}].$$

By (24),

$$d = \int f_r(\ln w_1)\, d\Pi = \int f_s(\ln w_2)\, d\Pi$$

($\ln w_1$ and $\ln w_2$ belong a.s. to the domains of definition of f_r and f_s, respectively, as a consequence of (14)). It is easy to show that $f_u'(x) > 0$ and $f_u''(x) < 0$ if $u \leqq 1$, i.e., f_u is increasing and concave on its domain if $u \leqq 1$. Hence, if $r = q/p \leqq 1$,

$$d = \int f_r(\ln w_1)\,d\Pi < f_r\left(\int \ln w_1\,d\Pi\right)$$

by Jensen's inequality,

$$< f_r(0)$$

$$= 0.$$

A similar calculation using f_s shows that $d < 0$ if $s = p/q \leqq 1$. Thus (22) is valid in any case, and the proof of the ergodic theorem is complete.

References

BREIMAN, L. (1968) *Probability*. Addison-Wesley, Reading, Massachusetts.

CHUNG, K. L. (1974) *A Course in Probability Theory*, 2nd ed. Academic Press, New York.

EWENS, W. J. (1969) *Population Genetics*. Methuen, London.

GILLESPIE, J. (1973) Polymorphism in random environments. *Theor. Pop. Biol.* **4,** 193–195.

KARLIN, S. AND LIEBERMAN, U. (1974) Random temporal variation in selection intensities: case of large population size. *Theor. Pop. Biol.* **6,** 355–382.

LAMPERTI, J. (1960) Criteria for recurrence or transience of stochastic processes, I. *J. Math. Anal. Appl.* **1,** 314–330.