The Annals of Probability
1974, Vol. 2, No. 1, 14-31

LIMIT THEOREMS FOR SOME STOCHASTIC EVOLUTION MODELS[1]

BY HARRY A. GUESS

University of Rochester

Limit theorems are proved for two stochastic models of molecular evolution in finite populations of fixed size. An additive fitness model is shown to be asymptotically neutral in the sense that the relative fitnesses converge in probability to one and the gene frequency distribution converges to the same limiting distribution as when all mutations are selectively neutral. A multiplicative fitness model is studied and weak convergence theorems are proved for the vector whose components are the fitnesses of individuals in the population.

1. Introduction and summary. We present a mathematical description and analysis of models incorporating the combined effects of genetic drift and natural selection, acting on a finite population of constant size undergoing repeated mutations. The mutations alter the relative reproductive capacities (fitnesses) of the individuals in which they occur. The main genetical motivation is the work of W. F. Bodmer and L. L. Cavalli-Sforza (1972) on mathematical models of molecular evolution in finite populations of fixed size and I. Eshel (1972) on evolution in infinite populations.

Two models are considered. In the additive fitness model, each mutant is assigned a fitness equal to the sum of its parent's fitness and a random fitness change, the sum being suitably truncated so that the resultant fitness remains positive. This is the form of the model introduced by Cavalli-Sforza and Bodmer, who took the fitness changes to be independent identically distributed (i.i.d.) normal random variables and studied the model using computer simulation. When the fitness changes are i.i.d. random variables with nonnegative expected value and finite second moment, this model is shown to be asymptotically neutral in the sense that the relative fitnesses converge in probability to 1, and the gene frequency distribution converges to the same limiting distribution as one gets when all mutations are selectively neutral. The explicit form of the limiting gene frequency distribution for the diffusion approximation to the additive model in the neutral case was first deduced heuristically by W. J. Ewens (1972) and was rigorously established by Karlin and McGregor (1972).

In the multiplicative fitness model, the fitness of each mutant is the product of its parent's fitness and a random fitness multiplier. The fitness multipliers are

---

Received October 30, 1972.

[1] Research supported in part by the National Science Foundation Grant GP 28876 A1, Stanford University. This work forms a part of the author's Ph. D. dissertation at Stanford University.

AMS 1970 *subject classifications.* Primary 60F05; Secondary 92A10.

Key words and phrases. Stochastic model, evolution, natural selection, mutation, Brownian motion, weak convergence.

assumed to be independent identically distributed positive random variables, whose logarithms have finite positive second moments. The change in the fitness distribution from one generation to the next is determined by natural selection, mutation, and random sampling of gametes. By exploiting a regenerative process embedded in the distribution of fitnesses of individuals relative to the population sample average fitness for each generation (the relative fitness process), limit theorems for the behavior over time of the population fitness vector and the vector of log fitnesses are obtained.

2. Description of the evolution models. It is assumed that at a certain locus in a monoecious haploid population of fixed size N, alleles from the infinite sequence $A_1, A_2, \cdots$ can occur. In any generation, a gene will mutate with probability μ $(0 \leqq \mu \leqq 1)$ to form an entirely new allelic type not currently or previously existing in the population.

Let $W_n = (W_{1n}, \cdots, W_{Nn})$ be the random vector whose components are the fitnesses of the N genes in the nth generation, relative to those in the -1st generation. We call $\{W_n\}_{n \geqq -1}$ the *fitness process*. Let $\bar{W}_{na} = (1/N) \sum_{i=1}^{N} W_{in}$ be the sample average fitness in the nth generation, and let $X_n = (1/\bar{W}_{na})W_n$ be the random vector of relative fitnesses in the nth generation. We call $\{X_n\}_{n \geqq -1}$ the *relative fitness process.*

The $(n + 1)$st generation is produced from the nth generation in two stages as follows. Given W_n, we choose, with replacement, N independent samples $W_{\sigma(1)n}, \cdots, W_{\sigma(N)n}$ according to the distribution

$$P(W_{\sigma(i)n} = W_{jn} \mid W_n) = \frac{W_{jn}}{\sum_{k=1}^{N} W_{kn}} = \frac{X_{jn}}{N} .$$

Each individual $W_{\sigma(i)n}$ mutates independently with probability μ to form an entirely new allelic type not currently or previously existing in the population. An individual which does not mutate retains the type and fitness of its parent. A mutant is assigned a new fitness stochastically as described below.

For each realization ω of the additive fitness process, we set

$$\begin{aligned} W_{in+1}(\omega) &= W_{\sigma(i)n}(\omega) + \Delta_{in+1}(\omega) && \text{if} \quad W_{\sigma(i)n}(\omega) + \Delta_{in+1}(\omega) > 0 , \\ &= \gamma > 0 && \text{otherwise} , \end{aligned}$$

where the random variables Δ_{in} are all independent identically distributed i.i.d. random variables and where γ is a fixed positive constant. The constant γ is needed to avoid zero or negative fitnesses which could cause the relative fitnesses to be undefined. In all cases discussed in this paper, the value of $\gamma > 0$ plays no important role in the asymptotic behavior of the process. Cavalli-Sforza and Bodmer took $\gamma = .0001$.

In the multiplicative fitness process, we set

$$W_{in+1} = W_{\sigma(i)n} v_{in+1} ,$$

where the random variables v_{in} are all i.i.d. positive random variables.

Note that in the additive fitness model, a value of $\Delta = 0$ is assigned to a nonmutant, and in the multiplicative fitness model, a value of $v = 1$ is assigned to a nonmutant. We let $\bar{W}_{n+1} = (1/N) \sum_{i=1}^{n} W_{\sigma(i)n}$ be the sample average fitness in the $(n + 1)$st generation after reproduction but prior to mutation. Recall that $\bar{W}_{n+1a}$ designates the sample average fitness in the $(n + 1)$st generation after mutation.

When there is only one allele present after the sampling process by which generation n is produced from generation $(n - 1)$ but prior to the mutation process in generation n, we say that generation n is a monoallelic generation. When generation n is monoallelic we have $W_{\sigma(1)n-1} = \cdots = W_{\sigma(N)n-1} = \bar{W}_n$. As a convenient normalization we take $W_{i,-1} \equiv 1 \equiv X_{i,-1}$ so that generation 0 is a monoallelic generation. Let $T_0 = 0$ and let $T_1, T_2, \cdots$ denote the successive monoallelic generations. Let $N(n) = \max(k : T_k \leqq n)$. Then $T_{N(n)}$ is the time of the last monoallelic generation up through generation n. In the multiplicative model, but *not* in the additive model, $\{T_k\}_{k=0}^{\infty}$ is a discrete renewal process with renewal function $N(\cdot)$. The following elementary lemma applies to both the additive and the multiplicative models.

Lemma 1. *$P(n - T_{N(n)} > m) \leqq (1/r)(1 - r)^{m+1}$ for all $n, m \geqq 1$, where $r = (1/N)^{N-1}$.*

Proof. $P((n + 1)$st generation is monoallelic $| \, n$th generation$) \geqq \sum_{k=1}^{N} (X_{k,n}/N)^N \geqq (1/N)^{N-1} = r > 0$. Hence $P(T_n < \infty) = 1$ for each n, and we have $P(T_n - T_{n-1} > m \mid T_{n-1} = j) \leqq (1 - r)^m$ for all $0 \leqq j \leqq n - 1$, $m \geqq 0$, $n \geqq 1$. Thus

$$\begin{aligned} P(n - T_{N(n)} > m) &= \textstyle\sum_{k=0}^{\infty} P(n - T_k > m;\, T_k \leqq n < T_{k+1}) \\ &= \textstyle\sum_{j=0}^{n-m-1} \sum_{k=0}^{j} P(T_{k+1} > n \mid T_k = j) P(T_k = j) \\ &\leqq \frac{1}{r}(1 - r)^{m+1}, \end{aligned}$$

where empty summations are equal to 0.

3. Additive fitness model.

We prove that when $E(\Delta_{10}) \geqq 0$ and $0 < E(\Delta_{10}^2) < \infty$, the relative fitnesses converge to one in probability and the distribution of the number and frequencies of distinct alleles present in the population converges to the same limiting distribution as when all mutations are selectively neutral. Thus the additive fitness model of Cavalli-Sforza and Bodmer will be seen to be "asymptotically neutral" in the above sense.

Theorem 1. *If $E(\Delta_{10}) \geqq 0$ and $0 < E(\Delta_{10}^2) < \infty$, then, as $n \to \infty$,*

$$\frac{\max_{1 \leq i \leq N} W_{in}}{\min_{1 \leq i \leq N} W_{in}} \to_P 1$$

and

$$\bar{W}_n \to_P \infty .$$

PROOF. Let

$$\begin{aligned} B_n &= \{\omega : [\bar{W}_{T_{N(n)}} - (n - T_{N(n)})(|\min_{1 \leqq i \leqq N, 0 \leqq j \leqq n-1} \Delta_{ij}|)] > 0\} \\ &= \left\{\omega : \left[1 - (n - T_{N(n)}) \left(\frac{|\min_{1 \leqq i \leqq N, 0 \leqq j \leqq n-1} \Delta_{ij}|}{n^{\frac{1}{2}}}\right)\left(\frac{n^{\frac{1}{2}}}{\bar{W}_{T_{N(n)}}}\right)\right] > 0\right\}. \end{aligned}$$

On the set B_n, we have

$$(1) \qquad \left(\frac{\max_{1 \leqq i \leqq N} W_{in}}{\min_{1 \leqq i \leqq N} W_{in}}\right) \leqq \frac{\bar{W}_{T_{N(n)}} + (n - T_{N(n)})(\max_{1 \leqq i \leqq N, 0 \leqq j \leqq n-1} |\Delta_{ij}| + \gamma)}{\bar{W}_{T_{N(n)}} - (n - T_{N(n)})(|\min_{1 \leqq i \leqq N, 0 \leqq j \leqq n-1} \Delta_{ij}|)}$$

$$= \frac{1 + (n - T_{N(n)}) \dfrac{(\max_{1 \leqq i \leqq N, 0 \leqq j \leqq n-1} |\Delta_{ij}| + \gamma)}{n^{\frac{1}{2}}} \left(\dfrac{n^{\frac{1}{2}}}{\bar{W}_{T_{N(n)}}}\right)}{1 - (n - T_{N(n)}) \dfrac{(|\min_{1 \leqq i \leqq N, 0 \leqq j \leqq n-1} \Delta_{ij}|)}{n^{\frac{1}{2}}} \left(\dfrac{n^{\frac{1}{2}}}{\bar{W}_{T_{N(n)}}}\right)}.$$

Since the Δ_{in}'s are i.i.d. with $E(\Delta_{10}^2) < \infty$, it follows from the Borel-Cantelli lemma that

$$(2) \qquad \lim_{n \to \infty} \frac{\max_{1 \leqq i \leqq N, 0 \leqq j \leqq n-1} |\Delta_{ij}|}{n^{\frac{1}{2}}} = 0 \quad \text{a.s.},$$

and

$$(3) \qquad \lim_{n \to \infty} \frac{|\min_{1 \leqq i \leqq N, 0 \leqq j \leqq n-1} \Delta_{ij}|}{n^{\frac{1}{2}}} = 0 \quad \text{a.s.}$$

If we let $U_{i0} = 0$ and $U_{in+1} = [U_{in} + \Delta_{in}]^+$ for $n \geqq 0$, then it is intuitively clear from the definitions of U_{in} and W_{in} that for each $x \geqq 0$,

$$(4) \qquad P(W_{in} \leqq x) \leqq P(U_{in} \leqq x).$$

A proof of (4) is given below in Lemma 2. From standard theorems on random walks [Feller (1966) page 197], we know that U_{in} has the same distribution as $\max(S_0, S_1, \cdots, S_n)$, where $S_0 = 0$ and $S_n = \sum_{k=1}^{n} \Delta_{ik}$. Hence when $E(\Delta_{10}) \geqq 0$ and $0 < E(\Delta_{10}^2) < \infty$, we have [Billingsley (1968) page 72]

$$\lim_{\varepsilon \to 0+} \limsup_{n \to \infty} P\left(\frac{U_{in}}{n^{\frac{1}{2}}} < \varepsilon\right) = 0,$$

which, together with (4), implies

$$(5) \qquad \lim_{\varepsilon \to 0} \limsup_{n \to \infty} P\left(\frac{W_{1n}}{n^{\frac{1}{2}}} < \varepsilon\right) = 0.$$

Now

$$W_{1n} \leqq \bar{W}_{T_{N(n)}} + (n - T_{N(n)})(\max_{1 \leqq i \leqq N, 0 \leqq j \leqq n-1} |\Delta_{ij}| + \gamma),$$

and, by (2) and Lemma 1,

$$(n - T_{N(n)}) \frac{(\max_{1 \leqq i \leqq N, 0 \leqq j \leqq n-1} |\Delta_{ij}| + \gamma)}{n^{\frac{1}{2}}} \to_P 0;$$

so, by (5),

$$\lim_{\varepsilon \to 0+} \limsup_{n \to \infty} P\left(\frac{\bar{W}_{T_{N(n)}}}{n^{\frac{1}{2}}} < \varepsilon\right) = 0$$

and

$$\lim_{M\to\infty} \limsup_{n\to\infty} P\left(\frac{n^{\frac{1}{2}}}{\bar{W}_{T_{N(n)}}} > M\right) = 0 .$$

Hence

$$(n - T_{N(n)}) \frac{(|\min_{1\leqq i\leqq N, 0\leqq j\leqq n-1} \Delta_{ij}|)}{n^{\frac{1}{2}}} \left(\frac{n^{\frac{1}{2}}}{\bar{W}_{T_{N(n)}}}\right) \to_P 0 .$$

This implies that $P(B_n) \to 1$, and the theorem now follows from (1).

It remains to establish (4). If X and Y are two random variables we write $X \subset Y$ whenever $P[X > t] \leqq P[Y > t]$ for all t.

LEMMA 2. *$h(U_n) \subset h(W_n)$ for every non-decreasing Borel measurable $h: R_+{}^N \to R^1$.*

PROOF. Since $U_0 = (0, \cdots, 0)$ and $W_{i0} > 0$ a.s., the assertion holds for $n = 0$. Suppose it holds for n. Let any such h be given, and let $g(y) = P(h(U_{n+1}) \leqq x \mid U_n = y)$. Then g is Borel measurable, and $-g$ is non-decreasing in y; so by the induction hypothesis, $-g(U_n) \subset -g(W_n)$. Hence

$$P(h(U_{n+1}) \leqq x) = E[g(U_n)] \geqq E[g(W_n)] .$$

But $W_{in} > 0$ a.s. for each i, and, as is easily seen,

$$g(y) = P(h(U_{n+1}) \leqq x \mid U_n = y) \geqq P(h(W_{n+1}) \leqq x \mid W_n = y)$$

for $y \geqq 0$. Hence

$$E[g(W_n)] \geqq E[P(h(W_{n+1}) \leqq x \mid W_n)] = P(h(W_{n+1}) \leqq x) .$$

So $h(U_{n+1}) \subset h(W_{n+1})$, and the proof is complete.

Taking $h(x_1, \cdots, x_N) = x_i$ yields $U_{in} \subset W_{in}$, so $P(W_{in} \leqq x) \leqq P(U_{in} \leqq x)$ for all $x \geqq 0$. This establishes (4).

To prove that the distribution of the number and frequencies of alleles present converges to the same limiting distribution as one gets under selective neutrality, we introduce some additional notation.

For the additive fitness process, let $Z_n = (k; n_1, \cdots, n_k)$ be the vector giving the number k of distinct alleles present in the population in the nth generation, together with their frequencies $n_1, \cdots, n_k$, which satisfy $n_1 + \cdots + n_k = N$. Let Y_n be the corresponding vector for a process in which all mutations are selectively neutral ($\Delta_{in} \equiv 0$), and let S be the finite state space of the processes Z_n and Y_n. Since Y_n is an aperiodic irreducible recurrent finite state space Markov chain, it has a limiting distribution. Let Y be a random vector with this limiting distribution. For $s_0, s_1 \in S$ let $P^k_{s_0 s_1} = P(Y_{n+k} = s_1 \mid Y_n = s_0)$ and $\pi_{s_1} = P(Y = s_1) = \lim_{n\to\infty} P^n_{s_0 s_1}$. For each $\delta > 0$, we define

$$A_\delta = \left\{x \in R_+{}^N : \frac{\max_{1\leqq i\leqq N} x_i}{\bar{x}} \leqq 1 + \delta ; \frac{\min_{1\leqq i\leqq N} x_i}{\bar{x}} \geqq 1 - \delta ; \bar{x} > \frac{1}{\delta}\right\} ,$$

where $\bar{x} = (1/N) \sum_{i=1}^N x_i$.

LEMMA 3. *For each* $s_0, s_1 \in S$, *each* $k = 1, 2, \cdots$, *and each* $\varepsilon > 0$, $\exists \delta > 0$ *such that*

$$|P(Z_k = s_1 \mid Z_0 = s_0; W_0) - P^k_{s_0 s_1}| < \varepsilon$$

for all $W_0 \in A_\delta$.

PROOF. The proof is by induction on k. When $k = 1$, and s_0, s_1 are fixed states in S, $P(Z_1 = s_1 \mid Z_0 = s_0; W_0)$ is a polynomial in the relative fitnesses $X_{i0} = W_{i0}/\bar{W}_{0a}$. When $X_{10} = \cdots X_{N0} = 1$, the value of the polynomial is $P^1_{s_0 s_1}$. Hence the assertion holds for $k = 1$, since polynomials are continuous.

The remainder of the proof is carried out using the Chapman-Kolmogorov equations for the homogeneous Markov processes $(Z_n; W_n)$ and Y_n, together with the easily established fact that, given $\delta, \varepsilon > 0$, there exists a δ_1 with $0 < \delta_1 < \delta$ such that $P(W_1 \notin A_\delta \mid Z_0 = s_0; W_0) < \varepsilon$ for all $W_0 \in A_{\delta_1}$.

THEOREM 2.

$$\lim_{n\to\infty} P\{Z_n = s_1 \mid Z_0 = s_0; W_0\} = \pi_{s_1}$$

for all $s_0, s_1 \in S$ *and all* W_0.

PROOF. Let $\varepsilon > 0$ be given, and choose k_0 such that for $k \geqq k_0$, $|P^k_{ss_1} - \pi_{s_1}| < \varepsilon/4$ for all $s \in S$. By Lemma 3, $\exists \delta > 0$ such that

$$|P(Z_{k_0} = s_1 \mid Z_0 = s; W_0) - P^{k_0}_{ss_1}| < \frac{\varepsilon}{4}$$

for all $W_0 \in A_\delta$. Since the process $(Z_n; W_n)$ has a stationary transition function, we have

$$|P(Z_{n+k_0} = s_1 \mid Z_n = s; W_n) - \pi_{s_1}| < \frac{\varepsilon}{2}$$

for all $W_n \in A_\delta$, all $n \geqq 0$. Hence

$$|P(Z_{n+k_0} = s_1 \mid Z_0 = s_0; W_0) - \pi_{s_1}| < \frac{\varepsilon}{2} + P(W_n \notin A_\delta \mid Z_0 = s_0; W_0) .$$

By Theorem 1, $\lim_{n\to\infty} P(W_n \notin A_\delta \mid Z_0 = s_0; W_0) = 0$; hence for all n sufficiently large,

$$|P(Z_{n+k_0} = s_1 \mid Z_0 = s_0; W_0) - \pi_{s_1}| < \varepsilon ,$$

and Theorem 2 follows.

Although, as has just been shown, $(Y_n, X_n) \to_{\mathscr{D}} (Y, l)$, this limiting distribution is not a stationary distribution for the additive process, and in fact the process has no stationary distribution. Neutrality is approached but is never actually attained.

In the additive fitness model, changes in relative fitness become progressively smaller as the population average fitness increases over that of the original type. The foregoing discussion shows that this assumption leads to a degenerate relative fitness distribution as long as $\bar{W}_n$ increases rapidly enough to take care of stochastic fluctuations in the Δ's. If we assume that the fitness changes act multiplicatively rather than additively, then changes in relative fitness are not diminished with increasing population average fitness. In what follows we investigate the multiplicative fitness model more fully.

4. Multiplicative fitness process. Let $(\Omega, \mathscr{B}, P)$ be a fixed probability triple on which is defined a multiplicative fitness process as described above with $0 < E(\log^2 v) < \infty$. It follows from the proof of Lemma 1 that $P(T_{k-1} - T_k > m) \leqq (1 - r)^m$ where $r = (1/N)^{N-1}$. Hence

$$E[(T_{k+1} - T_k)^n] < \infty \qquad \text{for all } n = 1, 2, \cdots, \text{ and all } k . \tag{6}$$

Let $\mathscr{G}_n$ denote the sub-σ-field of $\mathscr{B}$ generated by W_{ik}, $W_{\sigma(i)k}$ and v_{ij} for $1 \leqq i \leqq N$, $0 \leqq k \leqq n - 1$, $0 \leqq j \leqq n - 1$. Then each T_k is optional with respect to the increasing sequence of σ-fields $\mathscr{G}_n$, $n \geqq 1$. At each generation T_k, we have $W_{\sigma(1)T_k-1} = \cdots = W_{\sigma(N)T_k-1} = \bar{W}_{T_k}$, and so X_{T_k} has the same distribution as X_0. Since the fitness of each mutant is a multiple of its parent's fitness, with the fitness multipliers being i.i.d. positive random variables, the relative fitness process starts over at each generation T_k. In other words, the random vectors

$$\mathscr{Y}_k = (T_k - T_{k-1}; X_{T_{k-1}}, \cdots, X_{T_k-1}) , \qquad k \geqq 1 ,$$

are i.i.d. Similarly, the random vectors

$$\mathscr{X}_k = \left(T_k - T_{k-1}; \frac{\bar{W}_{T_{k-1}+1}}{\bar{W}_{T_{k-1}}}, \cdots, \frac{\bar{W}_{T_k}}{\bar{W}_{T_k-1}}\right),$$

are i.i.d.

Proofs of these results are analogous to those given in Chung [(1968) Chapter 8] and Iglehart (1971a).

Lemma 4. *The process $\{X_n\}$ converges weakly to a probability distribution on R_+^N. The limiting distribution is non-degenerate as long as the distribution of the fitness multiplier v is non-degenerate.*

Proof. This is a direct consequence of the discrete renewal theorem [Feller (1957) page 291]. We have for each Borel set A in R_+^N,

$$P(X_n \in A) = P(X_n \in A; T_1 > n) + \textstyle\sum_{j=1}^n P(X_n \in A \mid T_1 = j)P(T_1 = j) ,$$

and, for $1 \leqq j \leqq n$,

$$P(X_n \in A \mid T_1 = j) = P(X_{T_1+(n-j)} \in A \mid T_1 = j) .$$

The event $[T_1 = j]$ is in the σ-field $\mathscr{B}(\mathscr{Y}_1)$ generated by $\mathscr{Y}_1$, and the event $[X_{T_1+(n-j)} \in A]$ is in the σ-field $\mathscr{B}(\mathscr{Y}_2)$ generated by $\mathscr{Y}_2$. Since these σ-fields are independent, we have that, for $1 \leqq j \leqq n$,

$$P(X_{T_1+(n-j)} \in A \mid T_1 = j) = P(X_{T_1+(n-j)} \in A) = P(X_{n-j} \in A) .$$

Hence

$$P(X_n \in A) = P(X_n \in A; T_1 > n) + \textstyle\sum_{j=1}^n P(X_{n-j} \in A)P(T_1 = j) ,$$

and so by the discrete renewal theorem, since $P(T_1 = j) > 0$ for each j,

$$\lim_{n\to\infty} P(X_n \in A) = \frac{1}{E(T_1)} \textstyle\sum_{m=0}^\infty P(X_m \in A; T_1 > m) .$$

It is clear that $\lim_{n\to\infty} P(X_n \in A)$ defines a probability measure on R_+^N and that the distribution is non-degenerate as long as the distribution of v is non-degenerate.

We now define the *stationary relative fitness process* $\{\mathscr{X}_n\}_{n\geqq 0}$ and the *limiting fitness process* $\{\mathscr{W}_n\}_{n\geqq 0}$ as follows.

By enlarging the given probability triple $(\Omega, \mathscr{B}, P)$ if necessary, we may let $\{\mathscr{W}_n\}_{n\geqq 0}$ be a multiplicative fitness process defined on $(\Omega, \mathscr{B}, P)$, with reproduction as specified in Section 2 and with initial distribution

$$P(\mathscr{W}_0 \in A) \equiv \lim_{n\to\infty} P(X_n \in A)$$

for each Borel set $A \subset R_+^N$. We define $\{\mathscr{X}_n\}_{n\geqq 0}$ to be the relative fitness process associated with $\{\mathscr{W}_n\}_{n\geqq 0}$. Since $(1/N)\sum_{i=1}^N X_{in}(\omega) \equiv 1$, we have $P(\bar{\mathscr{W}}_{0a} = 1) = 1$ so $P(\mathscr{X}_0 = \mathscr{W}_0) = 1$, and the distribution of $\mathscr{X}_0$ is also given by

$$P(\mathscr{X}_0 \in A) = \lim_{n\to\infty} P(X_n \in A) .$$

It will be shown below that the process $\{\mathscr{X}_n\}_{n\geqq 0}$ is strictly stationary.

LEMMA 5. *If f: $R_+^N \to R^1$ is nonnegative or bounded and is Borel measurable, then*

$$E(f(\mathscr{X}_0)) = \frac{1}{E(T_1)} E\{\textstyle\sum_{k=1}^{T_1} f(X_{k-1})\} . \tag{7}$$

PROOF. If f is an indicator function of a Borel set $A \subset R_+^N$, then by Lemma 4 and the definition of $\mathscr{X}_0$,

$$\begin{aligned} E(f(\mathscr{X}_0)) = P(\mathscr{X}_0 \in A) &= \frac{1}{E(T_1)} \textstyle\sum_{k=0}^{\infty} P(X_k \in A;\ T_1 > k) \\ &= \frac{1}{E(T_1)} \textstyle\sum_{k=0}^{\infty} \int_{\{T_1 > k\}} f(X_k)\, dP \\ &= \frac{1}{E(T_1)} \int_\Omega \textstyle\sum_{k=0}^{T_1-1} f(X_k)\, dP = \frac{1}{E(T_1)} E\{\textstyle\sum_{k=1}^{T_1} f(X_{k-1})\} . \end{aligned}$$

By linearity, the result applies to any nonnegative Borel measurable simple function. If f is any nonnegative Borel measurable function, then there is a sequence of nonnegative Borel measurable simple functions $\{f_n\}$ such that $f_n(x) \uparrow f(x)$ $\forall x \in R_+^N$. Applying the preceding result to each f_n and using the monotone convergence theorem on both sides yields (7).

COROLLARY 1. *If f: $R_+^N \to R^1$ is bounded and Borel measurable, then* $\lim_{n\to\infty} 1/n \sum_{k=1}^n f(X_{k-1}) = E(f(\mathscr{X}_0))$ a.s.

LEMMA 6. *The process $\{\mathscr{X}_0\}$ is strictly stationary.*

PROOF. Since $\{X_n\}$ is a Markov process with stationary transition function,

$$P(X_{m+n+k} \in A_k, \cdots, X_{m+n} \in A_0 \mid X_m = \cdot) = f(\cdot)$$

is a nonnegative Borel measurable function F_+^N not dependent on m, for each fixed n and fixed Borel sets $A_0, \cdots, A_k$ in R_+^N. So by Corollary 1,

$$\begin{aligned} &\lim_{r\to\infty} \frac{1}{r} \textstyle\sum_{m=1}^r P(X_{m+n+k} \in A_k, X_{m+n} \in A_0 \mid X_m) \\ &\qquad = E(f(\mathscr{X}_0)) = P(\mathscr{X}_{n+k} \in A_k, \cdots, \mathscr{X}_n \in A_0) \quad \text{a.s.} \end{aligned} \tag{8}$$

Hence

$$\lim_{r\to\infty}\frac{1}{r}\sum_{m=1}^{r} P(X_{m+n+k}\in A_k,\cdots,X_{m+n}\in A_0)$$

$$= E\left\{\lim_{r\to\infty}\frac{1}{r}\sum_{m=1}^{r} P(X_{m+n+k}\in A_k,\cdots,X_{m+n}\in A_0\,|\,X_m)\right\}$$

$$= P(\mathscr{X}_{n+k}\in A_k,\cdots,\mathscr{X}_n\in A_0)\ .$$

But

$$\lim_{r\to\infty}\frac{1}{r}\sum_{m=1}^{r} P(X_{m+n+k}\in A_k,\cdots,X_{m+n}\in A_0)$$

$$= \lim_{r\to\infty}\frac{1}{r}\sum_{m=1}^{r} E\{P(X_{m+n+k}\in A_k,\cdots,X_{m+n}\in A_0\,|\,X_{m+n})\}$$

$$= E\left\{\lim_{r\to\infty}\frac{1}{r}\sum_{j=n+1}^{n+r} P(X_{k+j}\in A_k,\cdots,X_j\in A_0\,|\,X_j)\right\}$$

$$= E\left\{\lim_{r\to\infty}\left[\left(\frac{n+r}{r}\right)\frac{1}{n+r}\sum_{j=1}^{n+r} P(X_{k+j}\in A_k,\cdots,X_j\in A_0\,|\,X_j)\right.\right.$$

$$\left.\left.-\frac{1}{r}\sum_{j=1}^{n} P(X_{k+j}\in A_k,\cdots,X_j\in A_0\,|\,X_j)\right]\right\}$$

$$= P(\mathscr{X}_k\in A_k,\cdots,\mathscr{X}_0\in A_0)$$

by the same argument as was used to obtain (8).

Thus for all n, $(\mathscr{X}_{n+k},\cdots,\mathscr{X}_n)$ has the same distribution as $(\mathscr{X}_k,\cdots,\mathscr{X}_0)$, so the process is strictly stationary.

The following theorem provides a method by which the limiting relative fitness distribution can be computed in terms of the distribution of the fitness multiplier v. Let $v_1,\cdots,v_N$ be N independent samples with the distribution v and let $\bar{v}=(1/N)\sum_{i=1}^{N} v_i$. Let x, y and z denote projection on $R_+{}^{N-1}$ of points in the unit simplex in $R_+{}^N$. Thus $x=(x_1,\cdots,x_{N-1})$ where $0\leqq x_i\leqq 1$, $x_1+\cdots+x_{N-1}\leqq 1$, and $x_N=1-(x_1+\cdots+x_{N-1})$. We set

$$F(x) = P(\mathscr{X}_{i0}/N\leqq x_i,\ i=1,\cdots,N-1)$$

$$K_1(x) = P(v_i/N\bar{v}\leqq x_i,\ i=1,\cdots,N-1)$$

$$K_2{}^1(x,y) = \sum_\sigma\left\{P\left[v_i y_{\sigma(i)}\Big/\sum_{j=1}^{N} v_j y_{\sigma(j)}\leqq x_i,\ i=1,\cdots,N-1\right]\prod_{j=1}^{N} y_{\sigma(j)}\right\}$$

and

$$K_2{}^{n+1}(x,y) = \int K_2{}^n(x,z)K_2{}^1(dz,y)\ ,$$

where the summation of the definition of $K_2{}^1$ extends over all N^N-N possible N-tuples $\{\sigma(1),\cdots,\sigma(N)\}$ formed by sampling without replacement from $\{1,\cdots,N\}$ such that not all $\sigma(j)$'s are equal.

THEOREM 3.

$$F(x) = \frac{K_1(x)+\sum_{i=1}^{\infty}\int K_2{}^i(x,y)K_1(dy)}{K_1(\mathbf{1})+\sum_{i=1}^{\infty}\int K_2{}^i(\mathbf{1},y)K_1(dy)}\ .$$

PROOF. By Lemma 6,

$$
\begin{aligned}
F(x) &= P(\mathscr{X}_{i1}/N \leqq x_i,\ i = 1, \cdots, N-1) \\
&= P(\mathscr{X}_{i1}/N \leqq x_i,\ i = 1, \cdots, N-1 \text{ and } \sigma(1) = \cdots = \sigma(N)) \\
&\qquad + P(\mathscr{X}_{i1}/N \leqq x_i,\ i = 1, \cdots, N-1 \text{ and not all } \sigma(j)\text{'s are equal}) \qquad (9) \\
&= K_1(x) \int \textstyle\sum_{j=1}^{N} y_j{}^N F(dy) + \int K_2{}^1(x, y)F(dy) \\
&= cK_1(x) + \int K_2{}^1(x, y)F(dy)\,, \qquad \text{where } c = NE[(\mathscr{X}_{10}/N)^N]\,.
\end{aligned}
$$

Also, $K_2(x, y) \leqq 1 - \sum_{i=1}^{N} y_i{}^N \leqq 1 - r$ where $r = (1/N)^{N-1}$, and by induction $K_2{}^n(x, y) \leqq (1 - r)^n$. Hence $\int K_2{}^n(x, y)F(dy) \leqq (1 - r)^n$. It follows that upon iterating (9) and taking a limit we get

$$F(x) = c[K_1(x) + \textstyle\sum_{i=1}^{\infty} K_2{}^i(x, y)K_1(dy)]\,.$$

Since $F(\mathbf{1}) = 1$, $1/c = K_1(\mathbf{1}) + \sum_{i=1}^{\infty} \int K_2{}^i(\mathbf{1}, y)K_1(dy)$ and the theorem follows.

We next establish several elementary results which will be used in proving convergence theorems for the fitness process.

LEMMA 7. $E(|\log \bar{W}_{T_1}|^2) < \infty$.

PROOF.

$$\textstyle\prod_{j=1}^{n} \min_{1 \leqq i \leqq N} v_{ij-1} \leqq \bar{W}_n \leqq \prod_{j=1}^{n} \max_{1 \leqq i \leqq N} v_{ij-1}$$

for each k, and hence for each realization of the process

$$|\log \bar{W}_n| \leqq \textstyle\sum_{j=1}^{n} \max_{1 \leqq i \leqq N} |\log v_{ij-1}|\,.$$

Let

$$V_{j-1} = \max_{1 \leqq i \leqq N} |\log v_{ij-1}|\,, \tag{10}$$

and let $\xi_j = V_{j-1} - E(V_0)$, $S_n = \sum_{j=1}^{n} \xi_j$. The random variables $\{\xi_j\}$ are i.i.d. with $E(\xi_1) = 0$, $E(\xi_1{}^2) < \infty$. We have

$$
\begin{aligned}
E(|\log \bar{W}_{T_1}|^2) &\leqq E(\textstyle\sum_{j=1}^{T_1} V_{j-1})^2 \\
&= E((S_{T_1} + T_1 E((V_0))^2) \qquad (11) \\
&\leqq E(S_{T_1}^2) + 2E(V_0)E^{\frac{1}{2}}(T_1{}^2)E^{\frac{1}{2}}(S_{T_1}^2) + E^2(V_0)E(T_1{}^2)\,.
\end{aligned}
$$

By a well-known martingale argument, $E(S_{T_1}^2) \leqq E(T_1)E(\xi_1{}^2)$ so the right-hand side of (11) is finite.

COROLLARY 2. *$E(\sum_{j=1}^{T_1} V_{j-1})^2 < \infty$ where V_{j-1} is defined by* (10).

COROLLARY 3.

$$E\left(\textstyle\sum_{j=1}^{T_1} \left|\log \dfrac{\bar{W}_j}{\bar{W}_{j-1}}\right|\right) < \infty\,.$$

PROOF.

$$
\begin{aligned}
E(\textstyle\sum_{k=1}^{T_1} |\log \bar{W}_k|) &\leqq E(\textstyle\sum_{k=1}^{T_1} \sum_{j=1}^{k} V_{j-1}) \leqq E(\sum_{k=1}^{T_1} \sum_{j=1}^{T_1} V_{j-1}) \\
&= E(T_1 \textstyle\sum_{j=1}^{T_1} V_{j-1}) \leqq E^{\frac{1}{2}}(T_1{}^2)E^{\frac{1}{2}}[(\sum_{j=1}^{T_1} V_{j-1})^2] < \infty\,.
\end{aligned}
$$

We prove below that the time average of the changes in log average fitness from one generation to the next converges a.s. to the expected change per generation in log average fitness for a population whose relative fitness vector has the stationary distribution $\mathscr{X}_0$. This limiting expected change in log average fitness will be seen to be the sum of $E(\log v)$ and an entropy-like term representing the fitness increase produced by natural selection acting on the stationary relative fitness distribution. A functional central limit theorem is proved for the log fitness process. Additional functional limit theorems follow readily.

The functional central limit theorem is similar to those obtained for waiting times in queueing processes [cf. Iglehart (1971a)], partial sums or functionals of a positive recurrent Markov chain [cf. Chung (1960) and Freedman (1967)], and renewal processes [cf. Billingsley (1968)]. The Markov chain results and the results below depend crucially on the Doeblin (1938) decomposition. The functional limit theorems are for processes such as $\log W_{[nt]}$, which are random functions in the space D^N of all right continuous R^N-valued functions on [0, 1] having left limits. Ordinary central limit theorems in R^N for processes such as $\log W_n$ can be obtained from Smith (1955).

THEOREM 4.

$$E\left[\log\left(\frac{\bar{\mathscr{W}}_{k+1}}{\bar{\mathscr{W}}_k}\right)\right] = \frac{1}{E(T_1)}\, E(\log \bar{W}_{T_1}) \qquad \textit{for all} \quad k \geqq 1\,.$$

PROOF. For any $M > 0$.

$$E\left(\left[\log^+\left(\frac{\bar{W}_k}{\bar{W}_{k-1}}\right)\right] \wedge M \,\middle|\, X_{k-2} = \cdot\right) = f_M(\cdot)$$

is a bounded nonnegative Borel measurable function on ${R_+}^N$ to ${R_+}^1$, which is independent of k.

Hence, by Corollary 1,

$$\begin{aligned}
&\lim_{n\to\infty} \frac{1}{n} \textstyle\sum_{k=1}^n E\left(\left[\log^+\left(\frac{\bar{W}_k}{\bar{W}_{k-1}}\right)\right] \wedge M\right)\\
(12)\qquad &= \lim_{n\to\infty} \frac{1}{n} \textstyle\sum_{k=1}^n E\left\{E\left(\left[\log^+\left(\frac{\bar{W}_k}{\bar{W}_{k-1}}\right)\right] \wedge M \,\middle|\, X_{k-2}\right)\right\}\\
&= E\left\{\lim_{n\to\infty} \frac{1}{n} \textstyle\sum_{k=1}^n f_M(X_{k-2})\right\}\\
&= E(f_M(\mathscr{X}_0)) = E\left(\left[\log^+\left(\frac{\bar{\mathscr{W}}_2}{\bar{\mathscr{W}}_1}\right)\right] \wedge M\right).
\end{aligned}$$

But we can also see that

$$\begin{aligned}
\textstyle\sum_{k=1}^n \left(\left[\log^+\left(\frac{\bar{W}_k}{\bar{W}_{k-1}}\right)\right] \wedge M\right) &= \textstyle\sum_{k=1}^{N(n)} \sum_{j=T_{k-1}+1}^{T_k} \left(\left[\log^+\left(\frac{\bar{W}_j}{\bar{W}_{j-1}}\right)\right] \wedge M\right)\\
&\quad + \textstyle\sum_{j=T_{N(n)}+1}^{n} \left(\left[\log^+\left(\frac{\bar{W}_j}{\bar{W}_{j-1}}\right)\right] \wedge M\right).
\end{aligned}$$

By the discussion at the beginning of this section, the random variables

$$Q_k = \textstyle\sum_{j=T_{k-1}+1}^{T_k} \left(\left[\log^+\left(\frac{\bar{W}_j}{\bar{W}_{j-1}}\right)\right] \wedge M\right)$$

are i.i.d. Since

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=T_{N(n)}+1}^{n} \left(\left[\log^+\left(\frac{\bar{W}_j}{\bar{W}_{j-1}}\right)\right] \wedge M\right)$$
$$\leqq \lim_{n\to\infty} \left(\frac{1}{n}\right) [\max_{1\leqq k\leqq n} (T_{k+1} - T_k)] \cdot M = 0\,,$$

it follows in the usual way that

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \left(\left[\log^+ \frac{\bar{W}_k}{\bar{W}_{k-1}}\right] \wedge M\right) = \frac{1}{E(T_1)} E(Q_1)\,; \tag{13}$$

thus, taking expectations in (13)

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} E\left(\left[\log^+\left(\frac{\bar{W}_k}{\bar{W}_{k-1}}\right)\right] \wedge M\right) \tag{14}$$
$$= \frac{1}{E(T_1)} E(Q_1) = \frac{1}{E(T_1)} E\left(\sum_{j=1}^{T_1} \log^+\left(\frac{\bar{W}_j}{\bar{W}_{j-1}}\right) \wedge M\right).$$

Comparing (12) and (14) gives

$$E\left(\left[\log^+\left(\frac{\mathscr{W}_2}{\mathscr{W}_1}\right)\right] \wedge M\right) = \frac{1}{E(T_1)} E\left(\sum_{k=1}^{T_1} \left(\left[\log^+\left(\frac{\bar{W}_k}{\bar{W}_{k-1}}\right)\right] \wedge M\right)\right)$$

for each $M > 0$, and hence

$$E\left(\log^+\left(\frac{\mathscr{W}_2}{\mathscr{W}_1}\right)\right) = \frac{1}{E(T_1)} E\left(\sum_{k=1}^{T_1} \log^+\left(\frac{\bar{W}_k}{\bar{W}_{k-1}}\right)\right).$$

The same argument implies

$$E\left(\log^-\left(\frac{\mathscr{W}_2}{\mathscr{W}_1}\right)\right) = \frac{1}{E(T_1)} E\left(\sum_{k=1}^{T_1} \log^-\left(\frac{\bar{W}_k}{\bar{W}_{k-1}}\right)\right),$$

and, by Corollary 3, all of these quantities are finite. Hence

$$E\left(\log\left(\frac{\mathscr{W}_2}{\mathscr{W}_1}\right)\right) = \frac{1}{E(T_1)} E\left(\sum_{j=1}^{T_1} \log\left(\frac{\bar{W}_j}{\bar{W}_{j-1}}\right)\right) = \frac{1}{E(T_1)} E(\log \bar{W}_{T_1})\,.$$

It is easy to see that $E(\log \bar{W}_n/\bar{W}_{n-1} \mid \mathscr{G}_{n-1})$ is a function of X_{n-2} alone. So since the process $\{\mathscr{X}_n\}$ is stationary,

$$E\left(\log\left(\frac{\mathscr{W}_{n+1}}{\mathscr{W}_n}\right)\right) = E\left(\log\left(\frac{\mathscr{W}_2}{\mathscr{W}_1}\right)\right)$$

for all $n \geqq 1$.

Corollary 4.

$$\lim_{n\to\infty} \frac{1}{n} \log \bar{W}_n =_{\text{a.s.}} \lim_{n\to\infty} \frac{1}{n} E(\log \bar{W}_n) = E\left(\log\left(\frac{\mathscr{W}_2}{\mathscr{W}_1}\right)\right).$$

Proof. The process $\log \bar{W}_n$ is a cumulative process in the sense of Smith (1955) and $E|\log \bar{W}_{T_1}| < \infty$ by Lemma 7. So by Smith (1955) $\lim_{n\to\infty} n^{-1} \log \bar{W}_n =_{\text{a.s.}} \lim_{n\to\infty} n^{-1} E(\log \bar{W}_n) = (1/E(T_1))E(\log \bar{W}_{T_1})$ and the result now follows from Theorem 4.

COROLLARY 5.

$$E\left(\log\left(\frac{\mathscr{W}_2}{\mathscr{W}_1}\right)\right) = E(\log v) + \int \psi(x)F(dx)$$

where

$$\psi(x) = \sum_{i=1}^{N} x_i \log x_i - \frac{1}{N}\sum_{i=1}^{N} \log x_i ,$$

and where F *is given by Theorem* 3. *For any* $x = (x_1, \cdots, x_N)$ *with* $0 < x_i$ *and* $x_1 + \cdots + x_N = 1$, $\psi(x) \geqq 0$ *with equality only if* $x = (1/N, \cdots, 1/N)$.

PROOF. Let $W_{\alpha(1)n}$ denote the fitness of a randomly-chosen individual in the nth generation. (For each n, α has the distribution $P(\alpha(i) = j \mid X_{1n}, \cdots, X_{Nn}) = 1/N$, independently for all $i, j = 1, \cdots, N$.) We have

$$\begin{aligned} &E\{\log W_{\alpha(1)n+1} - \log W_{\alpha(1)n} \mid \mathscr{G}_n\} \\ (15)\qquad &= \sum_{j=1}^{N} \frac{W_{jn}}{N\bar{W}_{na}} \log W_{jn} + E(\log v) - \frac{1}{N}\sum_{j=1}^{N} \log W_{jn} \\ &= E(\log v) + \sum_{j=1}^{N} \frac{W_{jn}}{N\bar{W}_{na}} \log\left(\frac{W_{jn}}{N\bar{W}_{na}}\right) - \frac{1}{N}\sum_{j=1}^{N} \log\left(\frac{W_{jn}}{N\bar{W}_{na}}\right) \\ &= E(\log v) + \psi\left(\frac{X_n}{N}\right). \end{aligned}$$

Upon replacing X by $\mathscr{X}$ and W by $\mathscr{W}$, we get

$$E\left[\log\left(\frac{\mathscr{W}_{\alpha(1)1}}{\mathscr{W}_{\alpha(1)0}}\right)\right] = E(\log v) + E\left[\xi\left(\frac{\mathscr{X}_0}{N}\right)\right].$$

It is easy to see that $\psi(x) \geqq 0$ for all x as in the statement of the Corollary, with equality only if $x = (1/N, \cdots, 1/N)$.

It follows as in Corollary 4 that

$$\lim_{n\to\infty} \frac{1}{n} E(\log W_{\alpha(1)n}) = \lim_{n\to\infty} \frac{1}{n} E(\log \bar{W}_n) = E(\log \mathscr{W}_2/\mathscr{W}_1) .$$

Equation (15) implies

$$\frac{1}{N} E(\log W_{\alpha(1)n}) = E(\log v) + \frac{1}{n} E\left[\sum_{j=1}^{n} \psi\left(\frac{X_{j-1}}{N}\right)\right] + \frac{1}{n} E(\log W_{\alpha(1)0}) .$$

By Fatou's lemma,

$$\begin{aligned} &E(\log \mathscr{W}_2/\mathscr{W}_1) - E(\log v) \\ &= \lim_{n\to\infty} E\frac{1}{n}\left(\sum_{j=1}^{n} \psi\left(\frac{X_{j-1}}{N}\right)\right) \\ (16)\qquad &\geqq E\left\{\liminf_{n\to\infty} \frac{1}{n}\sum_{j=1}^{n} \psi\left(\frac{X_{j-1}}{N}\right)\right\} \\ &\geqq E\left\{\liminf_{n\to\infty} \frac{1}{n}\sum_{k=1}^{N(n)} \sum_{j=T_{k-1}+1}^{T_k} \psi\left(\frac{X_{j-1}}{N}\right)\right\} \\ &= \frac{1}{E(T_1)} E\left(\sum_{j=1}^{T_1} \psi\left(\frac{X_{j-1}}{N}\right)\right) \geqq 0 . \end{aligned}$$

Since $E(\log \bar{\mathscr{W}}_2/\bar{\mathscr{W}}_1)$ and $E(\log v)$ are finite, (16) implies that the i.i.d. random variables $Q_k{}^1 = \sum_{j=T_{k-1}+1}^{T_k} \psi(X_{j-1}/N)$ have finite expected value. Hence

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} \psi\left(\frac{X_{j-1}}{N}\right) = \frac{1}{E(T_1)} E\left(\sum_{j=1}^{T_1} \psi\left(\frac{X_{j-1}}{N}\right)\right) = E\left(\psi\left(\frac{\mathscr{X}_0}{N}\right)\right).$$

The sequence $\{1/n \sum_{j=1}^{n} \psi(X_{j-1}/N)\}$ is uniformly integrable because

$$0 \leqq \frac{1}{n} \sum_{j=1}^{n} \psi\left(\frac{X_{j-1}}{N}\right) \leqq \frac{1}{n} \sum_{j=1}^{N(n)} Q_k{}^1 + \frac{1}{n} \max_{1\leqq k\leqq n+1} Q_k{}^1 \leqq \frac{2}{n} \sum_{k=1}^{n+1} Q_k{}^1 .$$

So

$$\lim_{n\to\infty} E\left(\frac{1}{n} \sum_{j=1}^{n} \psi\left(\frac{X_{j-1}}{N}\right)\right) = \lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} \psi\left(\frac{X_{j-1}}{N}\right) = E\left(\psi\left(\frac{\mathscr{X}_0}{N}\right)\right)$$

and hence

$$E(\log \bar{\mathscr{W}}_2/\bar{\mathscr{W}}_1) = E(\log v) + \int \psi(x)F(dx) .$$

This completes the proof.

The entropy-like quantity $E(\psi(\mathscr{X}_0/N))$ is strictly positive and represents the effect of natural selection acting on the stationary relative fitness distribution.

In preparation for proving functional limit theorems for the $W_{[nt]}$-process, we introduce some additional definitions.

As in Billingsley (1968) let C be the space of all continuous real-valued functions on $[0, 1]$, let D be the space of all real-valued functions on $[0, 1]$ that are right continuous and have left limits, and let d_0 be the Skorohod metric on D. We denote by D^N the product of N copies of D and set $d_0{}^N(x, y) = \max_{1\leqq i\leqq N} \{d_0(x_i, y_i)\}$ and $\rho_N(x, y) = \max_{1\leqq i\leqq N} \{\rho(x_i, y_i)\}$ for $x, y \in D^N$. If X is any random function in $D[0, 1]$, then $\mathbf{1}X = (X, \cdots, X)$ is a random function in D^N, all of whose values lie on the diagonal $\Delta = \{(x_1, \cdots, x_n) \in D^N : x_1 = x_2 = \cdots = x_N\}$.

We let $\xi(\cdot)$ denote a random function in D whose distribution in Wiener measure. Then $\mathbf{1}\xi(\cdot)$ is a random function in D^N and the stochastic process $\{\mathbf{1}\xi(t) : 0 \leqq t \leqq 1\}$ is standard one-dimensional Brownian motion along the diagonal in R^N. For the construction and properties of Wiener measure on D, the reader is referred to Billingsley (1968).

THEOREM 5.

$$\frac{\log W_{[n\cdot]} - \mathbf{1}E(\log \bar{\mathscr{W}}_2/\bar{\mathscr{W}}_1)n\cdot}{(n\sigma^2)^{\frac{1}{2}}} \rightarrow_{\mathscr{D}} \mathbf{1}\xi(\cdot)$$

in D^N, *where* $\sigma^2 = (1/E(T_1))\sigma^2(\log \bar{W}_{T_1} - E(\log \bar{\mathscr{W}}_2/\bar{\mathscr{W}}_1)T_1)$.

PROOF. Evidently, $\log W_{[n\cdot]}$ is a random function in D^N and

$$\mathbf{1}E(\log \bar{\mathscr{W}}_2/\bar{\mathscr{W}}_1)n\cdot \in D^N ,$$

where $[nt]$ denotes the greatest integer less than or equal to nt. Using Theorem

4, we have, for all $0 \leqq t \leqq 1$,

$$\log W_{[nt]} - \mathbf{1}E(\log \mathscr{W}_2/\mathscr{W}_1)nt = \mathbf{1}\left(\log \bar{W}_{T_{N([nt])}} - \left[E\left(\frac{\log \bar{W}_{T_1}}{E(T_1)}\right)\right]T_{N([nt])}\right)$$
$$+ \mathbf{1}\left(\left[E\left(\frac{\log \bar{W}_{T_1}}{E(T_1)}\right)\right](T_{N([nt])} - nt)\right)$$
$$+ \log W_{[nt]} - \mathbf{1}\log \bar{W}_{T_{N([nt])}} .$$

The random variables Y_k defined for $k \geqq 1$ by

$$Y_k = \log\left(\frac{\bar{W}_{T_k}}{\bar{W}_{T_{k-1}}}\right) - \left(E\left(\frac{\log \bar{W}_{T_1}}{E(T_1)}\right)\right)(T_k - T_{k-1})$$

are i.i.d. with $E(Y_1) = 0$ and $E(Y_1^2) < \infty$ by (6) and Lemma 7. As long as v is non-degenerate, Y_1 is also non-degenerate, so $\sigma^2(Y_1) > 0$. Since

$$\bar{W}_{T_{N(n)}} \prod_{j=T_{N(n)}+1}^{n+1} (\min_{1\leqq i\leqq N} v_{ij-1}) \leqq W_{in} \leqq \bar{W}_{T_{N(n)}} \prod_{j=T_{N(n)}+1}^{n+1} (\max_{1\leqq i\leqq N} v_{ij-1})$$

for $1 \leqq i \leqq N$, we have

$$|\log W_{i,[nt]} - \log \bar{W}_{T_{N([nt])}}| \leqq \sum_{j=T_{N([nt])}+1}^{[nt]+1} V_{j-1} \leqq Z_{N([nt])+1}$$
$$\leqq \max_{1\leqq k\leqq n+1} Z_k ,$$

where V_{j-1} is defined by (10) and

$$Z_k = \sum_{j=T_{k-1}+1}^{T_k} V_{j-1} .$$

Hence

$$\rho_N(\log W_{([n\cdot])}, \mathbf{1}\log \bar{W}_{T_{N([n\cdot])}}) \leqq \max_{1\leqq k\leqq n+1} Z_k .$$

Also,

$$|nt - T_{N([nt])}| \leqq 1 + \max_{0\leqq j\leqq n} (T_{j+1} - T_j) ,$$

so

$$\rho_N(1(T_{N([n\cdot])} - n\cdot), 0) \leqq 1 + \max_{0\leqq j\leqq n} (T_{j+1} - T_j) .$$

It follows that

$$\rho_N\left(\frac{\log W_{[n\cdot]} - \mathbf{1}E(\log \bar{\mathscr{W}}_2/\bar{\mathscr{W}}_1)n\cdot}{(n\sigma^2(Y_1)/E(T_1))^{\frac{1}{2}}}, \frac{1}{(n\sigma^2(Y_1)/E(T_1))^{\frac{1}{2}}} \sum_{k=1}^{N([n\cdot])} Y_k\right)$$
$$\leqq \rho\,\frac{|E(\log \bar{W}_{T_1})E(T_1)|}{(\sigma^2(Y_1)/E(T_1))^{\frac{1}{2}}}\left(\frac{1 + \max_{0\leqq j\leqq n}(T_{j+1} - T_j)}{n^{\frac{1}{2}}}\right) \tag{17}$$
$$+ \left(\frac{E(T_1)}{\sigma^2(Y_1)}\right)^{\frac{1}{2}}\left(\frac{\max_{1\leqq j\leqq n+1} Z_j}{\bar{n}}\right) .$$

Since $\{Z_j\}_{j=1}^{\infty}$ are i.i.d., and $\{T_{j+1} - T_j\}_{j=0}^{\infty}$ are i.i.d. with $E((T_1 - T_0)^2) < \infty$ and $E(Z_1^2) < \infty$ by (6) and Corollary 2, respectively, the right-hand side of (17) converges to 0 a.s. By Billingsley [(1968) pages 144–145],

$$(n\sigma^2(Y_1)/E(T_1))^{-\frac{1}{2}} \sum_{k=1}^{N([n\cdot])} Y_k \rightarrow_{\mathscr{D}} \xi(\cdot)$$

in D, so, by Billingsley [(1968) Theorem 5.1] together with the fact that the mapping $x \rightarrow (x, \cdots, x)$ from D to D^N is continuous.

$$(n\sigma^2(Y_1)/E(T_1))^{-\frac{1}{2}} \sum_{k=1}^{N([n\cdot])} Y_k \rightarrow_{\mathscr{D}} \mathbf{1}\xi(\cdot)$$

in D^N.

Since $d_0^N(x, y) \leqq \rho_N(x, y)$ the theorem now follows from (17) and Billingsley [(1968) Theorem 4.1].

Limit theorems for functionals of the $W_{[nt]}$ process can be obtained from the above theorem by the methods of Billingsley [(1968) Sections 10 and 11]. We state two of the more interesting ones below.

COROLLARY 6.

$$(W_{[n\cdot]})^{1/(n\sigma^2)^{\frac{1}{2}}} \exp\{-(n^{\frac{1}{2}}E(\log \bar{\mathscr{W}}_2/\bar{\mathscr{W}}_1)/\sigma)\cdot\} \rightarrow_{\mathscr{D}} \mathbf{1}e^{\xi(\cdot)}, \tag{18}$$

$$\lim_{n\to\infty} P\left\{m\left[0 < s \leqq n: \frac{1}{s}f(s) > E(\log \bar{\mathscr{W}}_2/\bar{\mathscr{W}}_1)\right] \leqq nx\right\} \tag{19}$$

$$= \frac{2}{\pi}\arcsin x^{\frac{1}{2}},$$

where $f(\cdot)$ *is any of the functions*

$$\max_{1\leqq i\leqq N} \log W_{i,[\cdot]}, \qquad \min_{1\leqq i\leqq N} \log W_{i,[\cdot]}, \qquad \textit{or} \quad \log \bar{W}_{[\cdot]},$$

and m *is Lebesgue measure on* R^1.

Using a result of Strassen (1964), Iglehart (1971a) proved a functional law of the iterated logarithm in D for random sums of i.i.d. random variables. As Iglehart observed, the ordinary law of the iterated logarithm in R^1 follows readily from the version in D. The arguments used to prove Theorem 5 imply that the analogous results hold for the process $\log W_{[n\cdot]}$. We will state these as Corollary 7 after first introducing some additional terminology.

For each $\delta > 0$, let $K(\delta)$ denote the set of absolutely continuous functions $x \in \mathscr{C}$, such that $x(0) = 0$ and

$$\int_0^1 \left(\frac{d}{dt}\{x(t)\}\right)^2 dt \leqq \delta^2 .$$

Let $\mathbf{1}K(\delta) = \{(x, x, \cdots, x) \in D^N : x \in K(\delta)\}$. We call a sequence $\{x_n\}$ in a complete separable metric space S *relatively compact* in S if every subsequence contains a convergent subsequence or, equivalently, if $\overline{\{x_n\}}_{n\geqq 0}$ is compact.

COROLLARY 7. (a) *With probability one, the sequence*

$$\left\{\frac{\log W_{[n\cdot]} - \mathbf{1}nE(\log \bar{\mathscr{W}}_2/\bar{\mathscr{W}}_1)\cdot}{(2\sigma^2 n \log\log n)^{\frac{1}{2}}}\right\}_{n\geqq 3}$$

is relatively compact in D^N, *and its set of limit points is* $\mathbf{1}K(1)$.

(b) *With probability one, the sequence*

$$\left\{\frac{\log W_n - \mathbf{1}nE(\log \bar{\mathscr{W}}_2/\bar{\mathscr{W}}_1)}{(2\sigma^2 n \log\log n)^{\frac{1}{2}}}\right\}_{n\geqq 3}$$

is relatively compact in R^N, *and its set of limit points is the set* $\{(s, s, \cdots, s) \in R^N : -1 \leqq s \leqq 1\}$.

5. Concluding remarks. The above results continue to hold when the generation times are random, occurring in accordance with a renewal process independent of the fitness process, and having inter-event times with finite second moments. Similar theorems can be proved for a Moran model with non-overlapping generations by conditioning on when the population becomes monoallelic. By introducing two kinds of fitnesses, viability and fecundity, the fitness of each individual becomes a two-dimensional vector. Using results in Iglehart (1968), one can prove weak convergence theorems for these fitness vectors.

All of these results can be viewed collectively as a corollary to the various existing forms of exponential growth laws [cf. (18)]. Namely, when the size of a population is constrained to be constant, or even to stay in a bounded range $0 < N_1 \leqq N \leqq N_2 < \infty$, while natural selection and mutation act stochastically to cause repeated multiplicative changes in fitness, the N-dimensional log fitness process behaves asymptotically like a one-dimensional random walk. This behavior is illustrated quantitatively by (19), which says that the population minimum fitness and the population maximum fitness both spend the same amount of time above the line whose slope is the asymptotic mean rate of change in fitness.

Although from one generation to the next there is a large amount of independence inherent in the reproduction process, this independence is entirely eliminated in the normalized asymptotic distribution of the log fitness process, whose mass is concentrated on the diagonal. This pronounced tendency for individuals in the population to evolve together as a unit rather than individually is a concomitant of the strong recurrence of monoallelic generations. These occur frequently enough that accumulated differences in fitness among individuals within the population remain small, but not so small as to disappear entirely as in the additive model.

Acknowledgment. I would like to thank Professor James L. McGregor, my thesis adviser, for suggesting the area of investigation and providing continual help and encouragement; Professor Donald L. Iglehart for introducing me to weak convergence theory; and Professor L. L. Cavalli-Sforza for numerous stimulating discussions concerning the evolutionary behavior of natural populations.

REFERENCES

Billingsley, P. (1968). *Convergence of Probability Measures.* Wiley, New York.

Bodmer, W. F. and Cavalli-Sforza, L. L. (1972). Variation in fitness and molecular evolution. *Proc. Sixth Berkeley Symp. Math. Statist. Prob.* Univ. of California Press, Berkeley.

Chung, K. L. (1960). *Markov Chains with Stationary Transition Probabilities.* Springer, Berlin.

Chung, K. L. (1968). *A Course in Probability Theory.* Harcourt, Brace and World, New York.

Doeblin, W. (1938). Sur deux problemes de M. Kolmogorov concernant les chains denombrables. *Bull. Soc. Math. France* **66** 210-220.

Eshel, I. (1972). Evolution processes with continuity of types. *Adv. Appl. Probability* **4** 475-507.

Ewens, W. J. (1972). The sampling theory of selectively neutral alleles. *Theor. Pop. Bio.* **3** 87-112.

Feller, W. (1957). *An Introduction to Probability Theory and its Applications* **1** 2nd ed. Wiley, New York.

Feller, W. (1966). *An Introduction to Probability Theory and its Applications* **2** 1st ed. Wiley, New York.

FREEDMAN, D. (1967). Some invariance principles for functionals of a Markov chain. *Ann. Math. Statist.* **38** 1-7.

IGLEHART, D. (1968). Weak convergence of probability measures on product spaces with application to sums of random vectors. Stanford Univ. OR Tech. Report 109.

IGLEHART, D. (1971a). Functional limit theorems for the queue GI/G/1 in light traffic. *Adv. Appl. Prob.* **3** 269-281.

IGLEHART, D. (1971b). Multiple channel queues in heavy traffic, IV: Law of the iterated logarithm. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, **17** 168-180.

KARLIN, S. and MCGREGOR, J. L. (1972). Addendum to a paper of Ewens. *Theor. Pop. Bio.* **3** 113-116.

SMITH, W. (1955). Regenerative stochastic processes. *Proc. Roy. Soc. London, Ser. A.* **232** 6-31.

STRASSEN, V. (1964). An invariance principle for the law of the iterated logarithm. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **3** 211-226.

DEPARTMENT OF MATHEMATICS
MATHEMATICAL SCIENCES BUILDING
THE UNIVERSITY OF ROCHESTER
ROCHESTER, NEW YORK 14627