The Annals of Probability
1989, Vol. 17, No. 3, 1118–1123

ON THE GROWTH OF THE MULTITYPE SUPERCRITICAL BRANCHING PROCESS IN A RANDOM ENVIRONMENT

BY HARRY COHN

University of Melbourne

Let $\{\mathbf{Z}_n\}$ be a multitype branching process in a random environment (MBPRE) which grows to infinity with positive probability for almost all environmental sequences. Under some conditions involving the first two moments of the environmental sequence, it is shown that dividing the $\{\mathbf{Z}_n\}$ components by their environment-conditioned expectations yields a sequence convergent in L^2 to a random vector with equal components.

1. Introduction. Let $\{\mathbf{Z}_n\} = \{(Z_n^{(1)}, \dots, Z_n^{(p)})\}$ be a p-type branching process where $Z_n^{(i)}$ stands for the nth generation size of type i particles with $i = 1, \dots, p$. We shall assume that $\{\mathbf{Z}_n\}$ is a branching process in a random environment of the type introduced by Athreya and Karlin [2]. In this model each particle of type i of the nth generation yields p-type offspring with probability distributions $\{\zeta_n(i)\}$, where $\zeta_n(i) = (\zeta_n^{(1)}(i), \dots, \zeta_n^{(p)}(i))$. We shall say that $\zeta_n = \{(\zeta_n^{(1)}, \dots, \zeta_n^{(p)})\}$ is the environment of the nth generation, and $\zeta = (\zeta_0, \dots, \zeta_n, \dots)$ is the environmental sequence of the process. We assume ζ to be random, which accounts for the qualifier "random environment" used in MBPRE. According to the Athreya–Karlin model, when conditioned on ζ, the $\{\mathbf{Z}_n\}$ becomes a Markov branching process in a varying environment, that is, a branching process with independent lines of descent and generation-dependent offspring distributions. An earlier model assuming a random and independent environment was considered by Smith and Wilkinson [12]. For the axiomatic setup of the MBPRE see Athreya and Karlin [2] and Tanny [13]. Athreya and Karlin [2] deals with extinction criteria for single- and multitype settings. Tanny [13] is concerned with extinction theorems as well as aspects of growth of MBPRE in the supercritical case. Limit results for suitably normed single-type supercritical branching processes in random environment are given in Athreya and Karlin [3] and Tanny [14]. No results of this kind seem to be known in the multitype case which is attempted in this article. It seems that the approaches available for the Galton–Watson process (Harris [7], Athreya and Ney [4] and Asmussen and Hering [1]) which are crucially dependent on the Perron–Frobenius theory break down in the random environment setting. It will appear that the martingale-subsequence approach of Cohn [5] is adaptable to this case. Important ingredients in the proof will be provided by the Furstenberg–Kesten a.s. convergence result [6] and the Coale–Lopez ratio limit theorem [10, 11]. We shall consider random norming for $\{\mathbf{Z}_n\}$ depending on the environment ζ. When conditioned on ζ the norming is expressed by nonrandom vectors which, unlike

Received March 1988; revised July 1988.
AMS 1980 *subject classifications*. Primary 60J80; secondary 60F25.
Key words and phrases. Branching, random environment, multitype, Furstenberg–Kesten theorem, martingale, L^2-convergence.

the Galton–Watson case, do not appear to have asymptotically proportional components. Indeed, in general, the ratio of the conditional expectations of $Z_n^{(i)}$ and $Z_n^{(j)}$ with $i \neq j$ does not seem to converge. Surprisingly, the one-dimensional character of the limit distribution of a Galton–Watson process is preserved for MBPRE.

We shall make the blanket assumption that ζ is stationary and ergodic. As in [2] we write $\{\mathbf{Z}_n(\zeta)\}$ for the branching process conditioned on the environmental sequence ζ. Throughout the article we shall refer to $\{\mathbf{Z}_n(\zeta)\}$ as well as quantities derived from it, but to ease the notation we shall hereafter suppress the qualifier ζ.

We recall that

$$Z_{n+r}^{(j)} = \sum_{i=1}^{p} \sum_{u=1}^{Z_n^{(i)}} Z_{u,i}^{(j)}(n,r), \qquad j = 1,\dots,p, \tag{1}$$

where $Z_{u,i}^{(j)}(n,r)$ is the size of type j offspring at time $n+r$ of the uth type i particle of the nth generation. The random vectors $\{(Z_{u,i}^{(1)}(n,r),\dots,Z_{u,i}^{(p)}(n,r)); u = 1,2,\dots\}$ are i.i.d. given $\mathbf{Z}_n$. Consider the matrices $\{M_n\} = \{m_{i,j}^{(n)}\}$, where $m_{i,j}^{(n)}$ is the expected number of offspring of type i produced by one particle of type j under the environment ζ_n. It is easy to see that if ${}^nM = ({}^{(n)}m_{i,j}) = M_{n-1}M_{n-2}\cdots M_0$, then ${}^{(n)}m_{i,j} = E(Z_n^{(i)}|\mathbf{Z}_0 = \mathbf{e}_j)$, where $\mathbf{e}_j$ is the p-dimensional vector with 1 in the jth place and 0 elsewhere. Let $S_n^{(j)} = (s_{i,k}^{(n)}(j))$, where $s_{i,k}^{(n)}(j) = E(Z_n^{(i)}Z_n^{(k)}|\mathbf{Z}_0 = \mathbf{e}_j)$ and $\hat{S}_n^{(j)} = (\hat{s}_{i,k}^{(n)}(j))$, where

$$\hat{s}_{i,k}^{(n)}(j) = E\left(Z_n^{(i)}Z_n^{(k)}|\mathbf{Z}_{n-1} = \mathbf{e}_j\right) - E\left(Z_n^{(i)}|\mathbf{Z}_{n-1} = \mathbf{e}_j\right)E\left(Z_n^{(k)}|\mathbf{Z}_{n-1} = \mathbf{e}_j\right).$$

Write $\mathbf{\Phi}_{\zeta_n}(\mathbf{s}) = (\Phi_{\zeta_n}^{(1)}(\mathbf{s}),\dots,\Phi_{\zeta_n}^{(p)}(\mathbf{s}))$ for the probability generating function of ζ_n, and consider the conditions

$$\begin{gathered} 0 < C < \partial\Phi_{\zeta_0}^{(i)}(\mathbf{1})/\partial s_i < D < \infty, \\ 0 < \partial^2\Phi_{\zeta_0}^{(i)}(\mathbf{1})/\partial s_j\,\partial s_k < D < \infty \end{gathered} \tag{2}$$

and

$$E\left(\left|\log\left\langle 1 - \mathbf{\Phi}_{\zeta_0}(\mathbf{0}), \mathbf{1}\right\rangle\right|\right) < \infty, \tag{3}$$

where C and D are some constants and $\langle \mathbf{a},\mathbf{b}\rangle = \sum_{i=1}^{p} a_i b_i$ is the scalar product of $\mathbf{a} = (a_1,\dots,a_p)$ and $\mathbf{b} = (b_1,\dots,b_p)$. We shall write $\mathbf{a} < \mathbf{b}$ ($\mathbf{a} \le \mathbf{b}$) if $a_i < b_i$ ($a_i \le b_i$) for $i = 1,\dots,p$. The vectors $\mathbf{0}$ and $\mathbf{1}$ have all their components 0 and 1, respectively. Further, 1_Λ is to denote the indicator of a set Λ and $A \vartriangle B$ is the symmetric difference of the sets A and B. Define the norm of the matrix $M = (m_{i,j})$ by $\|M\| = \max_{1\le i\le n}\sum_{j=1}^{p}|m_{i,j}|$ and denote its transpose by M'. The matrix I will stand for the identity matrix. We assume that

$$\lim_{n\to\infty} n^{-1}\log(\|{}^nM\|) = \lambda > 0. \tag{4}$$

According to [2], (2), (3) and (4) ensure that $P(q(\zeta) < 1) = 1$, where $q(\zeta)$ is the extinction probability under the stronger assumption of an independent environmental sequence ζ. We shall take $\mathbf{Z}_0 = \mathbf{e}_{j_0}$ for an arbitrary fixed j_0, but extension of our result to arbitrary $\mathbf{Z}_0$ with finite second moment is straightforward.

2. An L^2-convergence result. We shall prove the following theorem.

THEOREM. *Assume that* (2), (3) *and* (4) *hold and write* $\mathbf{W}_n = (W_n^{(1)}, \dots, W_n^{(p)})$, *where* $W_n^{(i)} = Z_n^{(i)}/E(Z_n^{(i)}(\zeta))$ *for* $i = 1, \dots, p$. *Then* $\{\mathbf{W}_n\}$ *converges in mean square to a random vector* $\mathbf{W}$ *where* $\mathbf{W} = (W, \dots, W)$ *with* $E(W) = 1$.

PROOF. *Step 1.* First we shall prove that from any subsequence of $\{\mathbf{W}_n\}$ one may extract a further subsequence that converges completely to a random vector whose components are not identically 0. For this it suffices to show that for some finite constant A,

$$E\left(Z_n^{(i)^2}\right)/E^2\left(Z_n^{(i)}\right) < A \quad \text{a.s.}, i = 1, \dots, p;\ n = 1, 2, \dots. \tag{5}$$

To prove (5), we shall need the following lemma.

LEMMA. *Let* ${}^nM^r = M_{n-1}M_{n-2} \dots M_r$ *for* $r = 0, 1, \dots, n-1$ *and* ${}^nM^n = {}^{-1}M = I$. *Then*

$$S_n^{(j)} = \sum_{r=0}^{n} ({}^nM^r)' \left\{ \sum_{k=1}^{p} \hat{S}_r(k)\left({}^{r-1}M\mathbf{e}_j\right)_k \right\} {}^nM^r, \tag{6}$$

where $(\alpha)_k$ *is the kth coordinate of the vector* α.

The proof of (6), which will be omitted, is similar to the classical multitype Galton–Watson process case (see Harris [7] or Jagers [8], pages 88–89).

It is easy to see that (5) will follow if we shall prove that the diagonal entries of (6) divided by the squares of the corresponding entries of the vector ${}^nM\mathbf{e}_j$ are bounded away from ∞. Notice that

$$S_n^{(j)} < D \sum_{r=0}^{n} ({}^nM^r)' \left\{ \sum_{k=1}^{p} \left({}^{r-1}M\mathbf{e}_j\right)_k \right\} {}^nM^r. \tag{7}$$

Write $\delta = D/C$ and ${}^{(s)}m_{i,j}^{(t)}$ for the (i, j) entry of ${}^sM^t$ with $t < s$. According to Lemma 2 of [6],

$$\delta^{-2} \leq {}^{(s)}m_{i_1,j_1}^{(t)}/{}^{(s)}m_{i_2,j_2}^{(t)} \leq \delta^2 \tag{8}$$

for any s, t, i_1, i_2, j_1 and j_2.

Note that by (5) and (7) we shall need to majorate sums of products of three factors divided by $({}^{(n)}m_{i,j})^2$. The first two factors divided by ${}^{(n)}m_{i,j}$, and (8) yield

$$\frac{{}^{(n)}m_{i_1,j_1}^{(r)}\,{}^{(r-1)}m_{i_2,j_2}}{{}^{(n)}m_{i,j}} = \left(\sum_{k,l=1}^{p} \frac{{}^{(n)}m_{i,k}^{(r)} m_{k,l}^{(r)}\,{}^{(r-1)}m_{l,j}}{{}^{(n)}m_{i_1,j_1}^{(r)}\,{}^{(r-1)}m_{i_2,j_2}} \right)^{-1} \leq (p^2C)^{-1}\delta^4. \tag{9}$$

We deal next with the third factor of (7) divided by ${}^{(n)}m_{i,j}$. According to the

Furstenberg–Kesten theorem [6],

$$\lim_{n\to\infty} n^{-1}\log{}^{(n)}m_{i,j} = \lambda \quad \text{a.s. for } i, j = 1,\dots, n. \tag{10}$$

This implies ${}^{(n)}m_{i,j} = \rho^n\delta_{i,j}^{(n)}$, where $\rho = \exp(\lambda)$ and $\{\delta_{i,j}^{(n)}\}$ are some random variables with $\lim_{n\to\infty} n^{-1}\log\delta_{i,j}^{(n)} = 0$ a.s. Thus there must exist some numbers r_0 and ρ_0 with $1 < \rho_0 < \rho$ such that ${}^{(r)}m_{i,j} > \rho_0^r$ a.s. for $r > r_0$ which in conjunction with (8) yields

$$\frac{{}^{(n)}m_{i_1,j_1}^{(r)}}{{}^{(n)}m_{i,j}} = \left(\sum_{k=1}^{p}\frac{{}^{(n)}m_{i,k}^{(r)}{}^{(r)}m_{k,j}}{{}^{(n)}m_{i_1,j_1}^{(r)}}\right)^{-1} \le p^{-1}\delta^2\left(\rho_0^{-1}\right)^r \quad \text{a.s.} \tag{11}$$

It is easy to see that (9) and (11) imply (5). The existence of a weakly convergent subsequence of an arbitrary sequence is a well-known result (see Loève [9], page 181). By another well-known property for moments (see Loève [9], page 186) the limit distribution function of any weakly convergent subsequence must have expectation vector **1** as the limit of the expectation vectors of $\{\mathbf{W}_n\}$. Thus convergence on subsequences to a nonnull limit is complete as stated.

Step 2. We show now that from any subsequence of $\{\mathbf{W}_n\}$ one can extract a weakly convergent subsequence $\{\mathbf{W}_{n_k}\}$ such that $\{\xi_n(\mathbf{x})\}$ with

$$\begin{aligned}\xi_n(\mathbf{x}) &= \lim_{k\to\infty} P\left(\mathbf{W}_{n_k} \le \mathbf{x}|\mathbf{Z}_n\right)\\ &= P\left(\sum_{i=1}^{p}\beta_n^{(i)}\sum_{u=1}^{Z_n^{(i)}}W_{u,i}^{(1)}(n) \le x_1,\dots,\sum_{i=1}^{p}\beta_n^{(i)}\sum_{u=1}^{Z_n^{(i)}}W_{u,i}^{(p)}(n) \le x_p|\mathbf{Z}_n\right)\end{aligned} \tag{12}$$

is a martingale for any continuity point $\mathbf{x} = (x_1,\dots,x_p)$ of the limit distribution of $\{\mathbf{W}_{n_k}\}$, where $\{W_{u,i}^{(j)}(n)\}$ are i.i.d. given $\mathbf{Z}_n$, $E(W_{u,i}^{(j)}(n)) = 1$ and $\{\beta_n^{(i)}\}$ are some positive constants independent of $\{n_k\}$. Indeed, extract a subsequence $\{n_k\}$ such that $\{\mathbf{W}_{n_k}\}$ and $\{Z_{u,i}^{(j)}(n, n_k - n)/E(Z_{u,i}^{(j)}(n, n_k - n));\ i, j = 1,\dots, p;\ n = 1, 2,\dots\}$ converge weakly as $k \to \infty$. This is achievable by a diagonal procedure, since the set of variables considered is countable. Thus $\xi_n(\mathbf{x}) = \lim_{k\to\infty} P(\mathbf{W}_{n_k} \le \mathbf{x}|\mathbf{Z}_n)$ exist for all n and the martingale property is easily checkable. Notice further that $E(Z_n^{(j)}) = {}^{(n_k)}m_{j,j_0}$ and $E(Z_{u,i}^{(j)}(n, n_k - n)) = {}^{(n_k)}m_{j,i}^{(n)}$. Thus by (1),

$$Z_{n_k}^{(j)}/E\left(Z_{n_k}^{(j)}\right) = \sum_{i=1}^{p}\frac{{}^{(n_k)}m_{j,i}^{(n)}}{{}^{(n_k)}m_{j,j_0}}\sum_{u=1}^{Z_n^{(i)}}W_{u,i}^{(j)}(n, n_k - n), \tag{13}$$

where $E(W_{u,i}^{(j)}(n, n_k - n)) = 1$. Applying the Coale–Lopez theorem (or Theorem 3.3 of [11]) to the backwards products of matrices $\{{}^tM^n\}$ yields

$$\lim_{t\to\infty}\frac{{}^{(t)}m_{j,i}^{(n)}}{{}^{(t)}m_{j,l}^{(n)}} =: \alpha_{i,l}^{(n)} > 0.$$

It follows that for any $\{n_k\}$ with $\lim_{k\to\infty} n_k = \infty$,

$$(14)\qquad \begin{aligned} \left(\beta_i^{(n)}\right)^{(-1)} &=: \lim_{k\to\infty} \frac{{}^{(n_k)}m_{j,j_0}}{{}^{(n_k)}m_{j,i}^{(n)}} \\ &= \lim_{k\to\infty} \sum_{l=1}^{n} \frac{{}^{(n_k)}m_{j,l}^{(n)}}{{}^{(n_k)}m_{j,i}^{(n)}} {}^{(n)}m_{l,j_0} = \sum_{l=1}^{p} \alpha_{l,i}^{(n)(n)} m_{l,j_0}. \end{aligned}$$

Using (14) and (13) in (12) on letting $k \to \infty$ yields (12).

Step 3. We shall prove that $\{\mathbf{W}_n\}$ converges in distribution to a random vector with equal components. Indeed, $\{W_{u,i}^{(j)}(n)\}$ have means 1 and, in view of the stationarity of ζ and Step 1, uniformly bounded second moments. In this situation Chebyshev's inequality and the conditional independence of $\{W_{u,i}^{(j)}(n)\}$ on $\mathbf{Z}_n$ lead us to conclude that $\{(Z_n^{(i)})^{-1}\sum_{u=1}^{Z_n^{(i)}} W_{u,i}^{(j)}(n)\}$ converges in probability to 1 as $n \to \infty$ on the set of nonextinction. Since $E(\sum_{i=1}^{n} \beta_n^{(i)} Z_n^{(i)}) = 1$, it follows that $\{\sum_{i=1}^{n} \beta_n^{(i)} \sum_{u=1}^{Z_n^{(i)}} W_{u,i}^{(j)}(n) - \sum_{i=1}^{n} \beta_n^{(i)} Z_n^{(i)}\}$ converges in probability to 0 as $n \to \infty$. But $\{\beta_n^{(i)}\}$ does not depend on $\{n_k\}$. This shows that $\xi(\mathbf{x}) = \lim_{n\to\infty} \xi_n(\mathbf{x})$ a.s. is also independent of $\{n_k\}$. However $E(\xi(\mathbf{x})) = \mathbf{F}(x)$, where $\mathbf{F}$ is the limit distribution of $\{\mathbf{W}_{n_k}\}$, wherefrom we deduce that $\{\mathbf{W}_n\}$ converges weakly to $\mathbf{F}$. Finally, $\{\sum_{i=1}^{p} \beta_n^{(i)} Z_n^{(i)}\}$ is obviously independent of j which in conjunction with (12) proves that $\mathbf{F}$ is the distribution function of a random vector with equal components.

Step 4. We shall show that $\{\mathbf{W}_n\}$ converges in probability. Indeed, according to Steps 2 and 3,

$$(15)\qquad \xi_n(\bar{\mathbf{x}}) = P\left(\sum_{i=1}^{n} \beta_n^{(i)} \sum_{u=1}^{Z_n^{(i)}} W_{u,i}^{(1)}(n) \le x | \mathbf{Z}_n \right)$$

for any continuity point $\bar{\mathbf{x}} = (x, \dots, x)$ of $\mathbf{F}$. It is easy to see that (15) implies

$$(16)\qquad \xi_n(\bar{\mathbf{x}}) \ge P\left(\sum_{i=1}^{n} \beta_n^{(i)} \sum_{u=1}^{[(x-\varepsilon)E(Z_n^{(i)})]} W_{u,i}^{(1)}(n) \le x \right) \quad \text{if } \mathbf{W}_n \le (x - \varepsilon, \dots, x - \varepsilon)$$

for any $\varepsilon > 0$, where $[a]$ is the integral part of a. Taking into account that $\sum_{i=1}^{p} \beta_n^{(i)} E(Z_n^{(i)}) = 1$ as well as the weak law of large numbers in (16) yields $P(\{\xi(\bar{\mathbf{x}}) = 1\}) \ge \mathbf{F}(\bar{\mathbf{x}})$. However, $E(\xi(\bar{\mathbf{x}})) = \mathbf{F}(\bar{\mathbf{x}})$ which implies $P(\{\xi(\bar{\mathbf{x}}) = 1\}) = \mathbf{F}(\bar{\mathbf{x}})$ and $P(\{\xi(\bar{\mathbf{x}}) = 0\}) = 1 - \mathbf{F}(\bar{\mathbf{x}})$. Thus there exists a set $\Lambda_{\bar{\mathbf{x}}}$ such that $\xi(\bar{\mathbf{x}}) = \mathbf{1}_{\Lambda_{\bar{\mathbf{x}}}}$ and now a word-for-word extension of Theorem 3.1(ii) of [5] to the multidimensional case yields $\lim_{n\to\infty} P(\{\mathbf{W}_n \le \bar{\mathbf{x}}\} \vartriangle \Lambda_{\bar{\mathbf{x}}}) = 0$, which is tantamount to convergence in probability for $\{\mathbf{W}_n\}$.

Finally, $\{\mathbf{W}_n\}$ is by (5) L^2 uniformly bounded and L^2-convergence follows. □

Acknowledgment. The author is thankful to the referee for a number of useful suggestions.

REFERENCES

[1] ASMUSSEN, S. and HERING, H. (1983). *Branching Processes*. Birkhäuser, Boston.
[2] ATHREYA, K. B. and KARLIN, S. (1971). On branching processes with random environment. I. Extinction probabilities. *Ann. Math. Statist.* **42** 1499–1520.
[3] ATHREYA, K. B. and KARLIN, S. (1971). On branching processes with random environment. II. Limit theorems. *Ann. Math. Statist.* **42** 1843–1858.
[4] ATHREYA, K. B. and NEY, P. (1972). *Branching Processes*. Springer, New York.
[5] COHN, H. (1985). A martingale approach to supercritical (CMJ) branching processes. *Ann. Probab.* **13** 1179–1191.
[6] FURSTENBERG, H. and KESTEN, H. (1960). Products of random matrices. *Ann. Math. Statist.* **31** 457–469.
[7] HARRIS, T. E. (1963). *The Theory of Branching Processes*. Springer, New York.
[8] JAGERS, P. (1975). *Branching Processes with Biological Applications*. Wiley, New York.
[9] LOÈVE, M. (1977). *Probability Theory* **1**, 4th ed. Springer, New York.
[10] LOPEZ, A. (1961). *Problems in Stable Population Theory*. Office of Population Research, Princeton Univ.
[11] SENETA, E. (1981). *Non-negative Matrices and Markov Chains*. Springer, New York.
[12] SMITH, W. L. and WILKINSON, W. (1969). On branching processes in random environment. *Ann. Math. Statist.* **40** 814–827.
[13] TANNY, D. (1978). Normalizing constants for branching processes in random environments. *Stochastic Process. Appl.* **6** 201–211.
[14] TANNY, D. (1981). On multitype branching processes in a random environment. *Adv. in Appl. Probab.* **13** 464–497.

DEPARTMENT OF STATISTICS
MELBOURNE UNIVERSITY
PARKVILLE, VICTORIA 3052
AUSTRALIA