Statistics and Probability Letters 98 (2015) 54–58

Contents lists available at ScienceDirect

Statistics and Probability Letters

journal homepage: www.elsevier.com/locate/stapro

Markov limit of line of decent types in a multitype supercritical branching process

Jyy-I Hong *, K.B. Athreya

Iowa State University, United States

ARTICLE INFO

Article history:
Received 31 July 2014
Received in revised form 22 October 2014
Accepted 13 November 2014
Available online 12 December 2014

MSC:
primary 60J80
secondary 60G50

Keywords:
Branching processes
Multitype
Supercritical
Markov

ABSTRACT

In a multitype (d types) supercritical positively regular Galton–Watson branching process, let $\{X_n, X_{n-1}, \dots, X_0\}$ denote the types of a randomly chosen (i.e., uniform distribution) individual from the nth generation and this individual's n ancestors. It is shown here that this sequence converges in distribution to a Markov chain $\{Y_0, Y_1, \dots\}$ with transition probability matrix $(p_{ij})_{1\le i,j\le d}$ and having the stationary distribution. We also consider the critical case conditioned on non-extinction.

1. Introduction

Let $\mathbf{Z}_n = (Z_{n,1}, Z_{n,2}, \dots, Z_{n,d})$ be the population vector in the nth generation, $n = 0, 1, 2, \dots,$ where $Z_{n,i}$ is the number of individuals of type i in the nth generation. We assume that each individual of type i, $i = 1, 2, \dots, d$, lives a unit of time and, upon death, produces children of all types according to the offspring distribution $\left\{p^{(i)}(\mathbf{j}) \equiv p^{(i)}(j_1, j_2, \dots, j_d)\right\}_{\mathbf{j}\in\mathbb{N}^d}$ and independently of other individuals, where $p^{(i)}(j_1, j_2, \dots, j_d)$ is the probability that a type i parent produces j_1 children of type $1, j_2$ children of type $2, \dots, j_d$ children of type d.

Let $m_{ij} = E(Z_{1,j}|\mathbf{Z}_0 = \mathbf{e}_i)$ be the expected number of type j offspring of a single type i individual in one generation for any $i, j = 1, 2, \dots, d$. Then,

$$\mathbf{M} \equiv \{m_{ij} : i, j = 1, 2, \dots, d\} \tag{1.1}$$

is called the offspring mean matrix.

In a discrete-time multi-type positively regular Galton–Watson branching process, by the Perron–Frobenius theorem, the matrix $\mathbf{M}$ has a maximal eigenvalue ρ and has associated strictly positive normalized right and left eigenvectors $\mathbf{u} = (u_1, u_2, \dots, u_d)$ and $\mathbf{v} = (v_1, v_2, \dots, v_d)$ such that

$$\mathbf{u}\cdot\mathbf{v} = 1 \quad \text{and} \quad \mathbf{u}\cdot\mathbf{1} = 1. \tag{1.2}$$

The maximal eigenvalue ρ of the offspring mean matrix $\mathbf{M}$ plays a crucial role. The process is called a supercritical, critical or subcritical branching process according as $1 < \rho < \infty$, $\rho = 1$ or $\rho < 1$, respectively (see Athreya and Ney, 2004 for the details).

* Corresponding author.
E-mail addresses: hongjyyi@gmail.com (J. Hong), kbathreya@gmail.com (K.B. Athreya).

http://dx.doi.org/10.1016/j.spl.2014.11.013

Now, we consider Galton–Watson branching process with a finite offspring mean matrix **M** whose maximal eigenvalue $1 < \rho < \infty$ and with no extinction. Then choose an individual at random, i.e., uniform distribution, from the nth generation and denote its type X_n. Let X_{n-1} be the type of its parent, X_{n-2} the type of its grandparent, etc., down to X_0 being the type of the ancestor in the generation 0. We show that, for any integer k, $(X_n, X_{n-1}, \dots, X_{n-k})$ converges in distribution to $(Y_0, Y_1, \dots, Y_k)$ where $\{Y_n\}_{n\geq 0}$ is a Markov chain with a unique stationary distribution.

The work of Jagers and Nerman (1996) considers a similar process in a more general setting. However, their principal assumption is that the process has been evolving for an infinite amount of time and is already in a stable state. In this paper, we consider the case when the population has evolving up to n generations and prove a limit result about the types of the ancestors of a random chosen individual as $n \to \infty$. Thus, the work reported here is related to but different from that in Jagers and Nerman (1996).

2. Main results

The first result is for the supercritical case. Without lose of generality, we assume that each individual in this supercritical process produces at least one offspring with probability 1 upon death, that is, $P(\mathbf{Z}_1 = \mathbf{0}|\mathbf{Z}_0 = \mathbf{e}_i) = 0$ for all $i = 1, 2, \dots, d$. Thus, the probability of extinction is 0.

Theorem 2.1. *Let $1 < \rho < \infty$, $|\mathbf{Z}_0| = 1$, $E\big(\|Z_1\| \log \|Z_1\| \big| \mathbf{Z}_0 = \mathbf{e}_i\big) < \infty$ for any $i = 1, 2, \dots, d$. Then, for any integer $k \geq 0$, there exists a random vector $(Y_0, Y_1, \dots, Y_k)$ such that*

$$\big(X_n, X_{n-1}, \dots, X_{n-k}\big) \xrightarrow{d} \big(Y_0, Y_1, \dots, Y_k\big) \quad \text{as } n \to \infty,$$

and, for any $i_0, i_1, \dots, i_k \in \{1, 2, \dots, d\}$,

$$P\big(Y_0 = i_0, Y_1 = i_1, \dots, Y_k = i_k\big) = \frac{v_{i_k} m_{i_k i_{k-1}} m_{i_{k-1} i_{k-2}} \cdots m_{i_1 i_0}}{(\mathbf{1}\cdot\mathbf{v})\rho^k}.$$

Moreover, $\{Y_n\}_{n\geq 0}$ is a Markov chain with the state space $\{1, 2, \dots, d\}$,

(a) *initial distribution $\boldsymbol{\lambda}_0 \equiv \big(\lambda_0(1), \lambda_0(2), \dots, \lambda_0(d)\big)$ where*

$$\lambda_0(i) = \frac{v_i}{\mathbf{1}\cdot\mathbf{v}} \quad \text{for any } i = 1, 2, \dots, d,$$

(b) *transition probability $\mathbf{P} \equiv \big(p_{ij} : i, j = 1, 2, \dots, d\big)$, where*

$$p_{ij} = \frac{v_j m_{ji}}{v_i \rho} \quad \text{for any } n = 0, 1, 2, \dots,$$

(c) *and a unique stationary distribution $\boldsymbol{\pi} \equiv \big(\pi_1, \pi_2 \dots, \pi_d\big)$ where*

$$\pi_i = \frac{u_i v_i}{\mathbf{u}\cdot\mathbf{v}} \quad \text{for any } i = 1, 2, \dots, d.$$

A similar result also holds for the critical case conditioned on non-extinction:

Theorem 2.2. *Let $\rho = 1$, $|\mathbf{Z}_0| = 1$ and $E\|Z_1\|^2 < \infty$. Then, for any integer $k \geq 0$, there exists a random vector $(Y_0, Y_1, \dots, Y_k)$ such that*

$$\big(X_n, X_{n-1}, \dots, X_{n-k}\big)\Big| |\mathbf{Z}_n| > 0 \xrightarrow{d} \big(Y_0, Y_1, \dots, Y_k\big) \quad \text{as } n \to \infty,$$

and, for any $i_0, i_1, \dots, i_k \in \{1, 2, \dots, d\}$,

$$P\big(Y_0 = i_0, Y_1 = i_1, \dots, Y_k = i_k\big) = \frac{v_{i_k} m_{i_k i_{k-1}} m_{i_{k-1} i_{k-2}} \cdots m_{i_1 i_0}}{(\mathbf{1}\cdot\mathbf{v})}.$$

Moreover, $\{Y_n\}_{n\geq 0}$ is a Markov chain with the state space $\{1, 2, \dots, d\}$,

(a) *initial distribution $\boldsymbol{\lambda}_0 \equiv \big(\lambda_0(1), \lambda_0(2), \dots, \lambda_0(d)\big)$ where*

$$\lambda_0(i) = \frac{v_i}{\mathbf{1}\cdot\mathbf{v}} \quad \text{for any } i = 1, 2, \dots, d,$$

(b) *transition probability $\mathbf{P} \equiv \big(p_{ij} : i, j = 1, 2, \dots, d\big)$, where*

$$p_{ij} = \frac{v_j m_{ji}}{v_i} \quad \text{for any } n = 0, 1, 2, \dots,$$

(c) *and a unique stationary distribution $\boldsymbol{\pi} \equiv \big(\pi_1, \pi_2 \dots, \pi_d\big)$ where*

$$\pi_i = \frac{u_i v_i}{\mathbf{u}\cdot\mathbf{v}} \quad \text{for any } i = 1, 2, \dots, d.$$

Remark 2.1. Georgii and Baake (2003) also investigated the ancestral types of a typical individual for a supercritical multitype Markov branching process under a weaker condition on the moment of Z_1. The result in their paper can be derived from Theorem 2.1 (c) using the law of large numbers and irreducibility. On the other hand, Theorem 2.1 (c) also follows from the result of Georgii and Baake and a result for Markov chain which asserts that for any irreducible chain the limit of the marginal distribution is also a stationary distribution.

3. Proofs of main results

We will prove Theorem 2.1 using the principle of mathematical induction.

For $k = 0$, since, in the supercritical case under the assumptions in the theorem, it is known (Kesten and Stigum, 1966) that $\frac{\mathbf{Z}_n}{\rho^n} \to \mathbf{v}W$ with probability 1 as $n \to \infty$ where W is a random variable with $P(0 < W < \infty) = 1$, we have, by the bounded convergence theorem, as $n \to \infty$,

$$P\big(X_n = i_0\big) = E\left(\frac{Z_{n,i_0}}{|\mathbf{Z}_n|}\right) = E\left(\frac{Z_{n,i_0}/\rho^n}{|\mathbf{Z}_n|/\rho^n}\right) \to E\left(\frac{v_{i_0}W}{(\mathbf{1}\cdot\mathbf{v})W}\right) = \frac{v_{i_0}}{\mathbf{1}\cdot\mathbf{v}} \equiv \lambda_0(i_0).$$

Also, $\sum_{i=1}^d \lambda_0(i) = \sum_{i=1}^d \frac{v_i}{\mathbf{1}\cdot\mathbf{v}} = 1$, i.e., $\big\{\lambda_0(i) : i = 1, 2, \ldots, d\big\}$ is a proper probability distribution and hence there exists a random variable Y_0 with $P(Y_0 = i) = \lambda_0(i)$ for $i = 1, 2, \ldots, d$ such that $X_n \xrightarrow{d} Y_0$ as $n \to \infty$.

Next, we prove that the theorem holds for $k = 1$.

Let $\boldsymbol{\xi}_{n,j}^{(i)} = \big(\xi_{n,j}^{(i)1}, \xi_{n,j}^{(i)2}, \ldots, \xi_{n,j}^{(i)d}\big)$ be the vector of offsprings of the jth individual of type i in the nth generation. For any fixed $i_0, i_1 = 1, 2, \ldots, d$, $\left\{\xi_{n,j}^{(i_1)i_0}\right\}_{j\geq 1, n\geq 1}$ are i.i.d random variables with $E\Big(\xi_{n,j}^{(i_1)i_0}\Big) = m_{i_1 i_0} < \infty$.

By the assumption of non-extinction, $Z_{n,i_1} \to \infty$ with probability 1, and then by the strong law of large numbers, as $n \to \infty$,

$$\frac{1}{Z_{n,i_1}} \sum_{j=1}^{Z_{n,i_1}} \boldsymbol{\xi}_{n,j}^{(i_1)i_0} \to m_{i_1 i_0} \quad \text{with probability 1.}$$

So, by the bounded convergence theorem,

$$\begin{aligned} P\big(X_{n-1} = i_1 \big| X_n = i_0\big) &= E\left(\frac{\sum_{j=1}^{Z_{n-1,i_1}} \xi_{n-1,j}^{(i_1)i_0}}{Z_{n,i_0}}\right) = E\left(\frac{1}{Z_{n-1,i_1}} \sum_{j=1}^{Z_{n-1,i_1}} \boldsymbol{\xi}_{n-1,j}^{(i_1)i_0} \frac{Z_{n-1,i_1}/\rho^{n-1}}{Z_{n,i_0}/\rho^n} \frac{1}{\rho}\right) \\ &\to E\left(m_{i_1 i_0} \frac{v_{i_1}W}{v_{i_0}W} \frac{1}{\rho}\right) = \frac{v_{i_1} m_{i_1 i_0}}{\rho v_{i_0}} \quad \text{as } n \to \infty. \end{aligned}$$

Hence, as $n \to \infty$,

$$\begin{aligned} P\big(X_n = i_0, X_{n-1} = i_1\big) &= P\big(X_{n-1} = i_1 \big| X_n = i_0\big) P\big(X_n = i_0\big) \\ &\to \frac{v_{i_1} m_{i_1 i_0}}{(\mathbf{1}\cdot\mathbf{v})\rho} \equiv \lambda_1(i_0, i_1) \end{aligned}$$

and

$$\sum_{i=1}^d \sum_{j=1}^d \lambda_1(i,j) = \sum_{i=1}^d \frac{1}{(\mathbf{1}\cdot\mathbf{v})\rho} \left(\sum_{j=1}^d v_j m_{ji}\right) = \sum_{i=1}^d \frac{\rho v_i}{(\mathbf{1}\cdot\mathbf{v})\rho} = \sum_{i=1}^d \lambda_0(i) = 1$$

since $\mathbf{v}$ is the left eigenvector of $\mathbf{M}$ associated with the eigenvalue ρ.

So, $\big\{\lambda_1(i,j) : i, j = 1, 2, \ldots, d\big\}$ is a proper probability distribution with one marginal distribution λ_0. Thus, there exists a random variable Y_1 such that $P\big(Y_0 = i, Y_1 = j\big) = \lambda_1(i,j)$ for $i, j = 1, 2, \ldots, d$ and $\big(X_n, X_{n-1}\big) \xrightarrow{d} \big(Y_0, Y_1\big)$ as $n \to \infty$.

Assume that there exist random variables $Y_0, Y_1, \ldots, Y_k$ such that

$$P\big(Y_0 = i_0, Y_1 = i_1, \ldots, Y_k = i_k\big) = \frac{v_{i_k} m_{i_k i_{k-1}} \cdots m_{i_1 i_0}}{(\mathbf{1}\cdot\mathbf{v})\rho^k} \equiv \lambda_k(i_0, i_1, \ldots, i_k)$$

and, as $n \to \infty$,

$$\big(X_n, X_{n-1}, \ldots, X_{n-k}\big) \xrightarrow{d} \big(Y_0, Y_1, \ldots, Y_k\big).$$

Then

$$P\big(X_{n-(k+1)}=i_{k+1},X_{n-k}=i_k,\ldots,X_{n-1}=i_1\big|X_n=i_0\big)=E\left(\frac{\sum_{j_{k+1}=1}^{Z_{n-(k+1),i_{k+1}}}\sum_{j_k=1}^{\xi^{(i_{k+1})i_k}_{n-(k+1),j_{k+1}}}\cdots\sum_{j_1=1}^{\xi^{(i_2)i_1}_{n-2,j_2}}\xi^{(i_1)i_0}_{n-1,j_1}}{Z_{n,i_0}}\right)$$

$$=E\Bigg(\frac{1}{Z_{n-(k+1),i_{k+1}}}\frac{1}{\xi_{n-(k+1),j_{k+1}}^{(i_{k+1})i_k}}\cdots\frac{1}{\xi^{(i_2)i_1}_{n-2,j_2}}\sum_{j_{k+1}=1}^{Z_{n-(k+1),i_{k+1}}}\sum_{j_k=1}^{\xi^{(i_{k+1})i_k}_{n-(k+1),j_{k+1}}}$$

$$\cdots\sum_{j_1=1}^{\xi^{(i_2)i_1}_{n-2,j_2}}\left(\xi^{(i_{k+1})i_k}_{n-(k+1),j_{k+1}}\cdots\xi^{(i_2)i_1}_{n-2,j_2}\xi^{(i_1)i_0}_{n-1,j_1}\right)\frac{Z_{n-(k+1),i_{k+1}}/\rho^{n-(k+1)}}{Z_{n,i_0}/\rho^n}\frac{1}{\rho^{k+1}}\Bigg)$$

and, again by the strong law of large numbers and the bounded convergence theorem, we have that, as $n\to\infty$,

$$P\big(X_{n-(k+1)}=i_{k+1},X_{n-k}=i_k,\ldots,X_{n-1}=i_1\big|X_n=i_0\big)\to\frac{v_{i_{k+1}}m_{i_{k+1}i_k}m_{i_ki_{k-1}}\cdots m_{i_1i_0}}{v_{i_0}\rho^{k+1}}.$$

Hence, as $n\to\infty$,

$$\begin{aligned}P\big(X_n=i_0,X_{n-1}=i_1,\ldots,X_{n-(k+1)}=i_{k+1}\big)&=P\big(X_{n-(k+1)}=i_{k+1},X_{n-k}=i_k,\ldots,X_{n-1}=i_1\big|X_n=i_0\big)P\big(X_n=i_0\big)\\&\to\frac{v_{i_{k+1}}m_{i_{k+1}i_k}m_{i_ki_{k-1}}\cdots m_{i_1i_0}}{(\mathbf{1}\cdot\mathbf{v})\rho^{k+1}}\equiv\lambda_{k+1}(i_0,i_1,\ldots,i_k)\end{aligned}$$

and

$$\sum_{i_0=1}^d\sum_{i_1=1}^d\cdots\sum_{i_{k+1}=1}^d\lambda_{k+1}(i_1,i_1,\ldots,i_{k+1})=\sum_{i_0=1}^d\sum_{i_1=1}^d\cdots\sum_{i_k=1}^d\lambda_k(i_1,i_1,\ldots,i_k)=1.$$

So, there exists a random variable Y_{k+1} such that

$$\begin{aligned}P\big(Y_0=i_0,Y_1=i_1,\ldots,Y_k=i_k,Y_{k+1}=i_{k+1}\big)&=\lambda_{k+1}(i_0,i_1,\ldots,i_k,i_{k+1})\\&=\frac{v_{i_{k+1}}m_{i_{k+1}i_k}\cdots m_{i_1i_0}}{(\mathbf{1}\cdot\mathbf{v})\rho^{k+1}}\end{aligned}$$

and $\big(X_n,X_{n-1},\ldots,X_{n-k},X_{n-(k+1)}\big)\xrightarrow{d}\big(Y_0,Y_1,\ldots,Y_k,Y_{k+1}\big)$ as $n\to\infty$.

By the principle of the mathematical induction, we have proved the existence of the random variables $Y_0,Y_1,\ldots,Y_k$ for any integer $k\ge 0$ and have found their joint distribution.

Now, we prove the Markov property of $\{Y_n\}_{n\ge0}$.

For any $n\ge1$ and any $i,j,i_0,\ldots,i_{n-1}\in\{1,2,\ldots,d\}$, we have

$$\begin{aligned}P\big(Y_{n+1}=j\big|Y_n=i,Y_{n-1}=i_{n-1},\ldots,Y_0=i_0\big)&=\frac{P\big(Y_{n+1}=j,Y_n=i,Y_{n-1}=i_{n-1},\ldots,Y_0=i_0\big)}{P\big(Y_n=i,Y_{n-1}=i_{n-1},\ldots,Y_0=i_0\big)}\\&=\frac{v_jm_{ji}m_{ii_{n-1}}\cdots m_{i_1i_0}/(\mathbf{1}\cdot\mathbf{v})\rho^{n+1}}{v_im_{ii_{n-1}}\cdots m_{i_1i_0}/(\mathbf{1}\cdot\mathbf{v})\rho^{n}}\\&=\frac{v_jm_{ji}}{v_i\rho}\equiv p_{ij}.\end{aligned}$$

So, the conditional probability distribution of the future state of the chain $\{Y_n\}_{n\ge0}$, given the present state and the past states, only depends on the present state. Therefore, $\{Y_n\}_{n\ge0}$ is a Markov chain with the state space $\{1,2,\ldots,d\}$ such that

(a) the initial distribution $\boldsymbol{\lambda}_0\equiv\big(\lambda_0(1),\lambda_0(2),\ldots,\lambda_0(d)\big)$ where

$$\lambda_0(i)=\frac{v_i}{\mathbf{1}\cdot\mathbf{v}}\quad\text{for any } i=1,2,\ldots,d,$$

(b) the transition probability $\mathbf{P}\equiv\big(p_{ij}:i,j=1,2,\ldots,d\big)$, where

$$p_{ij}=\frac{v_jm_{ji}}{v_i\rho}\quad\text{for any } n=0,1,2,\ldots$$

It remains to show that the Markov chain $\{\tilde{Y}_n\}_{n\ge0}$ has a stationary distribution $\boldsymbol{\pi}\equiv\big(\pi_1,\pi_2\ldots,\pi_d\big)$ where

$$\pi_i=\frac{u_iv_i}{\mathbf{u}\cdot\mathbf{v}}\quad\text{for any } i=1,2,\ldots,d.$$

Since $\mathbf{u} > \mathbf{0}$ and $\mathbf{v} > \mathbf{0}$, $\pi_i = \frac{u_i v_i}{\mathbf{u}\cdot\mathbf{v}} > 0$. Also,

$$\sum_{i=1}^{d} \pi_i = \sum_{i=1}^{d} \frac{u_i v_i}{\mathbf{u}\cdot\mathbf{v}} = \frac{\mathbf{u}\cdot\mathbf{v}}{\mathbf{u}\cdot\mathbf{v}} = 1.$$

So, $\boldsymbol{\pi} \equiv \left(\pi_1, \pi_2 \ldots, \pi_d\right)$ is a probability distribution.

Moreover, since $\mathbf{u}$ is a right eigenvector of $\mathbf{M}$ associated with the eigenvalue ρ, for any $j = 1, 2, \ldots, d$,

$$\begin{aligned}\sum_{i=1}^{d} \pi_i p_{ij} &= \sum_{i=1}^{d} \frac{u_i v_i}{\mathbf{u}\cdot\mathbf{v}} \frac{v_j m_{ji}}{v_i \rho} \\ &= \frac{v_j}{\rho(\mathbf{u}\cdot\mathbf{v})} \sum_{i=1}^{d} m_{ji} u_i = \frac{v_j}{\rho(\mathbf{u}\cdot\mathbf{v})} \rho u_j = \frac{v_j u_j}{\mathbf{u}\cdot\mathbf{v}} = \pi_j\end{aligned}$$

and hence π is a stationary distribution of the transition probability $\mathbf{P}$.

Therefore, the proof is complete.

Remark 3.1. A similar proof can be adopted to prove the result in Theorem 2.2 for the critical case.

References

Athreya, K.B., Ney, P., 2004. Branching Processes. Dover, New York.

Georgii, H.-O., Baake, E., 2003. Supercritical multitype branching processes: the ancestral types of typical individuals. Adv. Appl. Probab. 35, 1090–1110.

Jagers, P., Nerman, O., 1996. The Asymptotic Composition of Supercritical Multitype Beanching Populations. In: Springer Lecture Notes in Mathematics, vol. 1626. pp. 40–54.

Kesten, H., Stigum, B.P., 1966. A limit theorem for multidimensional Galton–Watson processes. Ann. Math. Statist. 37, 1211–1223.