Published in final edited form as:

Stoch Anal Appl. 2010 ; 28(6): 1040–1060. doi:10.1080/07362994.2010.515486.

Limiting Distributions for Multitype Branching Processes

Andrei Y. Yakovlev^{†,1,*} and Nikolay M. Yanev^{2,*}

University of Rochester and Institute of Mathematics, Sofia

Abstract

In this paper the asymptotic behavior of multitype Markov branching processes with discrete or continuous time is investigated in the positive regular and nonsingular case when both the initial number of ancestors and the time tend to infinity. Some limiting distributions are obtained as well as multivariate asymptotic normality is proved. The paper considers also the relative frequencies of distinct types of individuals which is motivated by applications in the field of cell biology. We obtained non-random limits for the frequencies and multivariate asymptotic normality when the initial number of ancestors is large and the time of observation increases to infinity. In fact this paper continues the investigations of Yakovlev and Yanev [32] where the time was fixed. The new obtained limiting results are of special interest for cell kinetics studies where the relative frequencies but not the absolute cell counts are accessible to measurement.

Keywords

multitype branching processes; relative frequencies; limit theorems; multivariate asymptotic normality

1. Introduction

This paper deals with multitype Markov branching processes with discrete or continuous time assuming a large number of ancestors. The asymptotic behavior in the positive regular and nonsingular case is investigated when both the number of ancestors N and the time t tend to infinity.

Recall that the terminology “branching processes” was first introduced by Kolmogorov and his coauthors [20, 21] considering multitype branching processes in the Markov case, which have received much attention in the literature on stochastic processes. For a further development of the theory of branching processes and their applications in biology we refer the reader to several books [1, 10, 12, 16, 18, 25, 28, 29]. Remember that the main problems of the theory of branching processes are focused on the asymptotic behavior of the probabilistic characteristics of the processes when $t \rightarrow \infty$ assuming that the process begins from one ancestor.

Note that the multitype branching processes are considered in the present paper also from a new perspective: modeling of the relative frequencies of different types of individuals

[†]Professor Andrei Yakovlev died suddenly on February 27, 2008, when our work was in preparation. This paper is a tribute to his stimulating ideas and the friendship that we shared in our collaboration.

^{*}The paper is supported by NIH/NINDS grant NS39511, NIH/NCI R01 grant CA134839 and NIH grant N01-AI-050020.

¹Department of Biostatistics and Computational Biology, University of Rochester, 601 Elmwood Avenue, Box 630, Rochester, New York 14642, U.S.A. http://www.urmc.rochester.edu/smd/biostat/Andrei_Yakovlev.pdf, <http://www.biology-direct.com/content/3/1/10>

²Department of Probability and Statistics, Institute of Mathematics and Informatics, Bulgarian Academy of Sciences, 8 G.Bonchev, Sofia 1113, Bulgaria, yanev@math.bas.bg, Nikolay_Yanev@urmc.rochester.edu

Mathematics subject classifications. Primary: 60J80, 60J85; secondary: 62P10, 92D25.

(instead of the usual counts of individuals) as functions of time. A short biological background and motivation is provided in *Section 2* and the utility of relative frequencies (fractions, proportions) of distinct cell types in the analysis of biological studies of proliferation and differentiation of cells is considered. The need for such characteristics of cell kinetics arises in experimental situations where absolute cell counts are not accessible to measurement. In fact, the present article continues the investigations in this direction started in [32], where we considered the case with a large number of ancestors but the time was fixed. In this situation a limiting multivariate normal distribution was obtained (see *Theorem A* in *Section 5*) and a relevant statistical inference was developed. Now a new problem is investigated when both N and t increase to infinity and we would like to point out that the developed methods and the obtained results are quite different.

Feller [7] was probably the first who considered branching process with a large number of ancestors. For a classical Bienaymé-Galton-Watson (BGW) process he showed a diffusion approximation in the near-critical case. Further results in this direction were obtained by Lamperti [22], Jirina [17], Lindvall [23, 24], and Grimvall [9]. Lamperti [22] derived also some interesting limiting distributions. These results were summarized and discussed by Jagers [16].

Statistical inference for BGW processes with an increasing number of ancestors as well as limiting distributions when N and t tend to infinity were developed by Yanev [33] and Dion and Yanev [3, 4, 5] (see also a review chapter by Yanev [36]).

In the present paper some of the obtained limiting results by Lamperti [22], Yanev [33] and Dion and Yanev [3, 4, 5] are generalized for multitype Markov branching processes not only with discrete but also with continuous time. Note that usually the last two cases are investigated separately while in the present paper they are considered simultaneously. On the other hand, as usually in the theory of branching processes, the multitype generalization is quite different from the one dimensional case.

The paper is organized as follows. *Section 3* introduces the basic notions and preliminary results for the further investigations. In *Section 4*, the probability for extinction are considered (*Theorem 1*) and some limit results are obtained for the relative frequencies (*Theorem 2*) and for the processes (*Theorem 3*). *Section 5* deals with multivariate asymptotic normality of the processes (*Theorem 4*) as well as for the relative frequencies (*Theorem 5*). As a contrast of these results it is given *Theorem A*, where the obtained results in [32] are summarized. Note that in [32] we dealt with sequences of branching vectors while in the present paper we consider branching fields. In fact *Theorem 2* can be interpreted as an analogue of a LLN in the case of branching arrays and *Theorem 4* and *Theorem 5* as analogues of a CLT in the same case. Finally in *Section 6* it is shown that the obtained asymptotic results can be used to develop new statistical approach for estimation the basic parameters in branching processes with possible applications in the cell proliferation kinetics.

2. Biological Background and Motivation

The theory of branching processes has a long history of biological applications. It is worth to point out that the first asymptotic result for branching processes was obtained by Kolmogorov [19] considering some biological problems. Note that the branching processes have been proven especially useful in cell proliferation kinetics with multiple types of cells (see e.g. [29], [31], [35]).

Consider briefly the following example which is described with most details in [32]:

Oligodendrocyte type-2 astrocyte progenitor cells (referred to as O-2A progenitor cells), are known to be precursors of oligodendrocytes in the developing central nervous system. These cells grow in clones giving rise to oligodendrocytes, when they are plated *in vitro* and stimulated to divide by purified cortical astrocytes or by platelet-derived growth factor. Every O-2A progenitor cell can go to differentiation into an oligodendrocyte (with probability p) or it is able to proliferate (with probability $1 - p$).

The first stochastic model of oligodendrocyte development in cell culture was proposed by Yakovlev et al. [30] and developed in [13, 14, 37]. The model structure was defined following a set of assumptions that specified it as a special case of the Bellman-Harris branching process with two types of cells similar to that studied by Jagers [15].

In the above-described example the scheme of observation allows to count the numbers of O-2A progenitor cells and oligodendrocytes in every of N clones but it may not be possible in many other experimental situation. As it is noted in [32], it is technically impossible to count the number of cells of a given type in the blood or bone marrow in animal experiments. Similarly in the case when studying suspension cell cultures which consist of those cell types for which no specific antibodies are available. Note that in these cases it is possible to observe only the proportions of different types of cells. This is routinely practiced in experimental and clinical laboratories but the mathematical models of cell population kinetics are traditionally formulated in terms of cell counts. That is why it is very important to investigate also the relative frequencies in multitype branching stochastic models, which is one of the motivations for the present work.

The asymptotic results obtained in this paper and in [32] give a new direction toward statistical inference and applications of branching processes in cell proliferation kinetics.

3. Multitype Markov Branching Processes with Discrete or Continuous Time

Usually branching processes with discrete or continuous time are investigated separately but in this paper we will try to treat them together. Throughout of this paper we will consider a multitype Markov branching process $\mathbf{Z}(t) = (Z_1(t), Z_2(t), \dots, Z_d(t))$, where $Z_k(t)$ denotes the number of cells (particles, individuals) of type k ($k = 1, 2, \dots, d$) at time $t \in \mathbf{T}$. The time may be discrete $\mathbf{T} = \mathbf{N}_0 = \{0, 1, 2, \dots\}$ or continuous $\mathbf{T} = \mathbf{R}^+ = [0, \infty)$, but the process is assumed to be positive regular and nonsingular. For most details we refer to [1, 12, 25, 28].

Further on we will use the following notions:

$$\delta_i = (\delta_{i1}, \dots, \delta_{ij}, \dots, \delta_{id}), \delta_{ij} = 1 \text{ for } i=j, \delta_{ij} = 0 \text{ for } i \neq j, i, j = 1, \dots, d;$$

$$\mathbf{S} = (\mathbf{s}_1, \dots, \mathbf{s}_d), |\mathbf{s}| \leq \mathbf{1} = (\mathbf{1}, \dots, \mathbf{1}), \alpha = (\alpha_1, \dots, \alpha_d), \mathbf{S}^\alpha = \prod_{k=1}^d \mathbf{S}_k^{\alpha_k}.$$

Introduce the probability generating functions (p.g.f.)

$$F_i(t, \mathbf{s}) = \mathbb{E}\{\mathbf{s}^{\mathbf{Z}(t)} | \mathbf{Z}(0) = \mathbf{e}_i\}, i = 1, 2, \dots, d, \mathbf{F}(t, \mathbf{s}) = (F_1(t, \mathbf{s}), \dots, F_d(t, \mathbf{s})).$$

Remember that in the Markov case

$$\mathbf{F}(\mathbf{t} + \tau; \mathbf{s}) = \mathbf{F}(\mathbf{t}; \mathbf{F}(\tau; \mathbf{s})), \quad t, \tau \in \mathbf{T}, \quad (1)$$

which has the following stochastic analog

$$\mathbf{Z}(t + \tau) = \sum_{i=1}^d \sum_{j=1}^{Z_i(t)} \mathbf{Z}^{(i)}(\tau; j),$$

where the vectors $\{\mathbf{Z}^{(i)}(\tau; j)\}_{j=1}^{Z_i(t)}$ are *i.i.d.* as $\mathbf{Z}^{(i)}(\tau)$ and $\mathbf{Z}^{(i)}(\tau)$ means $\mathbf{Z}(\tau)$ with $\mathbf{Z}(0) = \mathbf{e}_i$.

In the discrete-time case we will use the offspring p.g.f.

$$h_i(\mathbf{s}) = \mathbf{F}_i(\mathbf{1}; \mathbf{s}) = \sum_{\alpha \in \mathbf{R}^d} \mathbf{p}_i(\alpha) \mathbf{s}^\alpha, \mathbf{h}_i(\mathbf{1}) = \mathbf{1}, i = 1, 2, \dots, d; \mathbf{h}(\mathbf{s}) = (h_1(\mathbf{s}), \dots, h_d(\mathbf{s})).$$

Then by (1) one has

$$\mathbf{F}(\mathbf{t} + \mathbf{1}; \mathbf{s}) = \mathbf{F}(\mathbf{t}; \mathbf{h}(\mathbf{s})) = \mathbf{h}(\mathbf{F}(\mathbf{t}; \mathbf{s})), \mathbf{t} = \mathbf{0}, \mathbf{1}, \mathbf{2}, \dots, \quad (2)$$

which means that $\mathbf{F}(\mathbf{t}; \mathbf{s})$ is the $\mathbf{t}$ -th iterate of $\mathbf{h}(\mathbf{s})$.

In the continuous-time case as $t \rightarrow 0$

$$\mathbf{F}(t; \mathbf{s}) = \mathbf{s} + \mathbf{f}(\mathbf{s})t + o(t) \text{ uniformly for } |\mathbf{s}| \leq \mathbf{1},$$

where $\mathbf{f}(\mathbf{s}) = (f_1(\mathbf{s}), \dots, f_d(\mathbf{s}))$ is the vector of the infinitesimal generating functions

$$f_i(\mathbf{s}) = \sum_{\alpha \in \mathbf{R}^d} \mathbf{r}_i(\alpha) \mathbf{s}^\alpha, \mathbf{f}_i(\mathbf{1}) = \mathbf{0}, i = 1, 2, \dots, d.$$

Then instead (2) the Kolmogorov backward and forward equations hold

$$\frac{\partial \mathbf{F}(\mathbf{t}; \mathbf{s})}{\partial t} = \mathbf{f}(\mathbf{F}(\mathbf{t}; \mathbf{s})), \quad \frac{\partial \mathbf{F}(\mathbf{t}; \mathbf{s})}{\partial \mathbf{t}} = \sum_{i=1}^d \mathbf{f}_i(\mathbf{s}) \frac{\partial \mathbf{F}(\mathbf{t}; \mathbf{s})}{\partial \mathbf{s}_i}, \quad \mathbf{F}(0; \mathbf{s}) = \mathbf{s}.$$

It is well-known that all characteristics of the process can be obtained from the p.g.f. In particular, the following notation will be used:

$$A_{ij}(t) = \mathbb{E}\{Z_j(t) \mid \mathbf{Z}(0) = \delta_i\} = \frac{\partial}{\partial s_j} F_i(t; \mathbf{s})|_{\mathbf{s}=\mathbf{1}},$$

$$A_{ij} = \frac{\partial h_i(\mathbf{s})}{\partial s_j}|_{\mathbf{s}=\mathbf{1}}, \quad a_{ij} = \frac{\partial f_i(\mathbf{s})}{\partial s_j}|_{\mathbf{s}=\mathbf{1}}, \quad i, j = 1, 2, \dots, d.$$

Introduce the matrices $\mathbf{A}(t) = \|A_{ij}(t)\|$, $\mathbf{A} = \|A_{ij}\|$ and $\mathbf{a} = a_{ij}$. Then

$$\mathbf{A}(t) = \mathbf{A}^t, \quad t \in \mathbf{N}_0; \quad \mathbf{A}(t) = \mathbf{e}^{at} = \sum_{n=0}^{\infty} \mathbf{a}^n t^n / n!, \quad t \in \mathbf{R}^+.$$

Similarly one can obtain

$$B_{jk}^{(i)}(t) = \mathbb{E} \{Z_j(t) [Z_k(t) - \delta_{jk}] | \mathbf{Z}(0) = \delta_i\} = \frac{\partial^2}{\partial s_j \partial s_k} F_i(t; \mathbf{s}) |_{\mathbf{s}=\mathbf{1}},$$

$$B_{jk}^{(i)} = \frac{\partial^2 h_i(\mathbf{s})}{\partial s_j \partial s_k} |_{\mathbf{s}=\mathbf{1}}, \quad b_{jk}^{(i)} = \frac{\partial^2 f_i(\mathbf{s})}{\partial s_j \partial s_k} |_{\mathbf{s}=\mathbf{1}}; \quad i, j, k = 1, 2, \dots, d.$$

Further on we will consider the positive regular and nonsingular case and then we will denote the Perron-Frobenius eigenvalue by R for the matrix $\mathbf{A}$ and by r for the matrix $\mathbf{a}$. In both cases we will use one and same notion for the corresponding right $\mathbf{u} = (u_1, \dots, u_d)$ and left $\mathbf{v} = (v_1, \dots, v_d)$ eigenvectors. It is well-known that they can be chosen positive and

normalized such that $\sum_{i=1}^d u_i = 1$, $\sum_{i=1}^d v_i = 1$. Then it is well-known that as t

$$A_{ij}(t) \sim u_i v_j R^t, \quad t \in \mathbf{N}_0; \quad A_{ij}(t) \sim u_i v_j \mathbf{e}^{rt}, \quad t \in \mathbf{R}^+; \quad (3)$$

Note that the moment structure of multitype Galton-Watson process is investigated in Quine [27] where some linear recurrence relations are derived.

We will need also the following notation

$$B(t) = \sum_{i,j,k=1}^d v_i B_{jk}^{(i)}(t) u_j u_k, \quad B = \sum_{i,j,k=1}^d v_i B_{jk}^{(i)} u_j u_k, \quad b = \sum_{i,j,k=1}^d v_i b_{jk}^{(i)} u_j u_k. \quad (4)$$

The main purpose of this paper is the investigation of the asymptotic behavior of the processes with a large number of ancestors. For simplicity (and without any essential restriction) we will assume that $\mathbf{Z}(0) = (N, 0, \dots, 0)$ and $Z_1(0) = N$. In this case we will use the notation $\mathbf{Z}(t, N) = (Z_1(t, N), Z_2(t, N), \dots, Z_d(t, N))$, $t \in \mathbf{T}$, and we will consider the situation when both N and t go to infinity simultaneously. In general, N and t are free parameters, but further on one can assume that $N = N(t)$ and $N(t) \rightarrow \infty$ as $t \rightarrow \infty$.

By the branching property it follows that

$$Z_i(t; N) = \sum_{k=1}^N Z_{i,k}(t), \quad i = 1, 2, \dots, d, \quad (5)$$

where $\{Z_{i,k}(t)\}_{k=1}^N$ are i.i.d. copies of the branching process $Z_i(t)$ with $\mathbf{Z}(0) = \mathbf{1}$.

In fact, (5) is equivalent to the following relation

$$F_N(t; \mathbf{s}) = \mathbb{E}\{\mathbf{s}^{\mathbf{Z}(t)} \mid \mathbf{Z}(0) = (N, 0, \dots, 0)\} = F^N(t; \mathbf{s}), \quad (6)$$

where $F(t; \mathbf{s}) = F_1(t; \mathbf{s}) = \mathbb{E}\{\mathbf{s}^{\mathbf{Z}(t)} \mid \mathbf{Z}(0) = \delta_1\}$.

Introduce the total number of cells at the moment t as

$$U(t) = \sum_{k=1}^d Z_k(t), \quad U(0) = Z_1(0) = 1, \quad t \in \mathbf{T}. \quad (7)$$

Then by (5) and (7) it follows that

$$U(t; N) = \sum_{i=1}^d Z_i(t; N) = \sum_{k=1}^N U^{(k)}(t), \quad (8)$$

where $U^{(k)}(t) = \sum_{i=1}^d Z_{i,k}(t)$, $k \geq 1$, are i.i.d. as $U(t)$.

The relative frequencies (fractions, proportions) of types can be defined on the non-extinction set $\{U(t, N) > 0\}$ as follows:

$$\Delta_i(t; N) = Z_i(t; N) / U(t; N), \quad i = 1, 2, \dots, d, \quad (9)$$

with the obvious condition

$$\sum_{k=1}^d \Delta_k(t; N) = 1. \quad (10)$$

The investigation of the relative frequencies is very important for the applications (especially in the cell biology) because there are a lot of situations when it is not possible to observe the numbers of individuals but only their relative proportions.

In what follows, we will also need the following deterministic proportions:

$$p_i(t) = A_{1i}(t) / M(t), \quad i = 1, 2, \dots, d; \quad t \in \mathbf{T}, \quad (11)$$

where

$$M(t) = \mathbb{E}\{U(t)\} = \sum_{j=1}^d A_{1j}(t). \quad (12)$$

4. Probability of Extinction and Limiting Distributions

In this section we will consider critical processes with $R = 1$, $0 < B < 1$, or $r = 0$, $0 < b < 1$, and non-critical processes ($R < 1$ or $r < 0$) with the classical “ $X \log X$ condition”.

Theorem 1

Let $q(t, N) = \Pr\{U(t, N) = 0\}$ and $N, t \rightarrow \infty$.

- i. If $\{R > 1 \text{ or } r > 0\} \setminus \{R = 1 \text{ or } r = 0 \text{ and } N/t \rightarrow 0\} \setminus \{R < 1, NR^t \rightarrow 0 \text{ or } r < 0, Ne^{rt} \rightarrow 0\}$ then $q(t, N) \rightarrow 0$ (non-extinction);
- ii. If $\{R = 1 \text{ or } r = 0 \text{ and } N/t \rightarrow 0\} \setminus \{R < 1, NR^t \rightarrow 0 \text{ or } r < 0, Ne^{rt} \rightarrow 0\}$ then $q(t, N) \rightarrow 1$ (extinction);
- iii. If $0 < C < \infty$ and $\{R = 1 \text{ or } r = 0 \text{ and } N/t \rightarrow C\} \setminus \{R < 1, NR^t \rightarrow C \text{ or } r < 0, Ne^{rt} \rightarrow C\}$ then $q(t, N) \rightarrow q, 0 < q < 1$ (extinction with probability q or non-extinction with probability $1 - q$).

Remark 1

The constants in the critical and subcritical cases as well as for the discrete or continuous processes can be different but in (iii) and further on we will use for simplicity one and the same letter (without any indices).

Proof

Note that by (6) one has $q(t, N) = F_N(t, \mathbf{0}) = q^N(t)$, where

$$q(t) = \Pr\{U(t) = 0\} = F(t, \mathbf{0}), \mathbf{0} = (0, 0, \dots, 0),$$

is the extinction probability and $q(t) < 1$ for every fixed t .

Therefore $\log q(t, N) = N \log[1 - Q(t)]$, where $Q(t) = 1 - q(t)$.

1) Let $R < 1$ or $r < 0$. Then $Q(t) \sim Ku_1 R^t$ or $Q(t) \sim Ku_1 e^{rt}$ as $t \rightarrow \infty$ where K are some positive constants.

Therefore

$$\log q(t; N) \sim -Ku_1 NR^t \text{ or } \log q(t; N) \sim -Ku_1 Ne^{rt}.$$

Hence

$$\begin{aligned} q(t; N) &\rightarrow 0 \text{ if } NR^t \rightarrow \infty (Ne^{rt} \rightarrow \infty), \\ q(t; N) &\rightarrow 1 \text{ if } NR^t \rightarrow 0 (Ne^{rt} \rightarrow 0) \text{ and} \\ q(t; N) &\rightarrow q \text{ if } NR^t \rightarrow C (Ne^{rt} \rightarrow C), 0 < C < \infty, \text{ where } 0 < q = e^{-Ku_1 C} < 1. \end{aligned}$$

2) Let $R = 1$ or $r = 0$. In this case $Q(t) \sim 2u_1/Bt$ or $Q(t) \sim 2u_1/bt$ as $t \rightarrow \infty$.

Then

$$\log q(t; N) \sim -2u_1 N/Bt \text{ or } \log q(t; N) \sim -2u_1 N/bt.$$

Therefore one obtains

$$\begin{aligned}
 q(t;N) &\rightarrow 0 \text{ if } N/t \rightarrow \infty, \\
 q(t;N) &\rightarrow 1 \text{ if } N/t \rightarrow 0 \text{ and} \\
 q(t;N) &\rightarrow q \text{ if } N/t \rightarrow C, 0 < C < \infty, \text{ where } q = e^{-2u_1 C/B} \text{ or } q = e^{-2u_1 C/b} \text{ and } 0 < q < 1.
 \end{aligned}$$

3) Let $R > 1$ or $r > 0$. In this case $q(t) \rightarrow q$ as $t \rightarrow \infty$. Therefore $q(t, N) = q^N(t) \rightarrow 0$ as $N \rightarrow \infty$ uniformly for $0 < t < \infty$.

Theorem 2

Let $p_k = v_k/V, k = 1, \dots, d, V = \sum_{k=1}^d v_k$, and $N, t \rightarrow \infty$.

- i. If $\{R < 1, NR^t \rightarrow r < 0, Ne^{rt} \rightarrow 0\} \cup \{R = 1 \text{ or } r = 0 \text{ and } N/t \rightarrow C\}$ then $q_k(t, N) \rightarrow p_k$ in probability, $k = 1, 2, \dots, d$;
- ii. If $\{R > 1 \text{ or } r > 0\}$ then $q_k(t, N) \rightarrow 0$ a.s., $k = 1, 2, \dots, d$.

Proof

Remember that by (6)

$$\mathbb{E}\{e^{-\lambda Z_k(t;N)}\} = F_{(k)}^N(t; e^{-\lambda}), \quad k = 1, 2, \dots, d. \quad (13)$$

where $F_{(k)}(t; s_k) = F(t; 1, \dots, 1, s_k, 1, \dots, 1)$.

1) Let first consider the subcritical case $R < 1$ with $NR^t \rightarrow r < 0$ ($t \in \mathbb{N}_0$) or $r < 0$ with $Ne^{rt} \rightarrow 0$ ($t \in \mathbb{R}^+$).

Denote by $q_{(k)}(t, N; \lambda)$, $\lambda > 0$, anyone of the following Laplace transforms:

$$\mathbb{E}\{e^{-\lambda Z_k(t;N)/NR^t}\} = F_{(k)}^N(t; e^{-\lambda/NR^t}), \quad t \in \mathbb{N}_0, \text{ or}$$

$$\mathbb{E}\{e^{-\lambda Z_k(t;N)/Ne^{rt}}\} = F_{(k)}^N(t; e^{-\lambda/Ne^{rt}}), \quad t \in \mathbb{R}^+.$$

Consider first the discrete-time case for which it is known that (see Athreya and Ney [1972, Ch. V. 3], Mode [1971, Ch. 1.11] or Sevastyanov [1971, Ch. VI. 2])

$$1 - F_{(k)}(t; 0) \sim Ku_1 R^t, \quad 0 < K < \infty,$$

$$\left\{ \lim_{N \rightarrow \infty} [1 - F_{(k)}(t; s)] / [1 - F_{(k)}(t; 0)] = 1 - F_{(k)}^*(s), \quad A_k^* = \frac{\partial}{\partial s_k} F_{(k)}^*(s) |_{s=1} = v_k/K. \right. \quad (14)$$

Since

$$\log q_{(k)}(t, N; \lambda) = N \log F_{(k)}(t; e^{-\lambda/NR^t}) \sim -N[1 - F_{(k)}(t; e^{-\lambda/NR^t})]$$

then applying (14) one obtains

$$\log \beta_{(k)}(t, N; \lambda) \sim -Ku_1 NR^t \{1 - F_{(k)}^* (e^{-\lambda/NR^t})\}. \quad (15)$$

On the other hand

$$\lim\{[1 - F_{(k)}^* (t; e^{-\lambda/NR^t})]/[1 - e^{-\lambda/NR^t}]\} = A_k^* \text{ and } 1 - e^{-\lambda/NR^t} \sim \lambda/NR^t.$$

Therefore by (15) one has $\log \beta_{(k)}(t, N; \lambda) \sim -u_1 v_k$ and by the continuity theorem for Laplace transforms as $NR^t \rightarrow \infty$ one obtains

$$Z_k(t; N)/NR^t \rightarrow u_1 v_k \text{ in probability, } k=1, \dots, d. \quad (16)$$

Then from (8) and (16) it follows

$$U(t; N)/NR^t \rightarrow u_1 V \text{ in probability, } V = \sum_{k=1}^d v_k. \quad (17)$$

Finally from (9), (16) and (17) it is not difficult to see that in the case $R < 1$, $NR^t \rightarrow \infty$

$$\Delta_k(t; N) \rightarrow p_k = v_k/V \text{ in probability, } k=1, 2, \dots, d. \quad (18)$$

In the continuous-time case the following relations hold (similarly to (14))

$$1 - F_{(k)}(t; 0) \sim Ku_1 e^{rt}, 0 < K < \infty,$$

$$\{ \lim\{[1 - F_{(k)}(t; s)]/[1 - F_{(k)}(t; 0)]\} = 1 - F_{(k)}^*(s), \sum_{i=1}^d f_i(s) \frac{\partial}{\partial s_i} F_{(k)}^*(s) = -r [1 - F_{(k)}^*(s)] \}. \quad (19)$$

By differentiating the last relation in (19) and putting $s = \mathbf{1}$ one obtains

$$rA_k^* = \sum_{i=1}^d A_i^* a_{ik}, k=1, \dots, d.$$

Therefore by the Frobenius Theorem it follows that $A_k^* = C_1 v_k$ for some constant C_1 . Now using a similar arguments as in the proof of the discrete case (see Theorem 6.2.4 in Sevastyanov, 1971) it is not difficult to obtain that in fact $C_1 = 1/K$. On the other hand, by (3) and the first relation from (19) one has

$$\mathbb{E}\{Z_k(t) | Z_k(t) > 0, Z(0) = \delta_1 = A_{1k}(t)/[1 - F_{(k)}(t; 0)]\} \rightarrow v_k/K = A_k^*.$$

Hence applying (19) one obtains

$$\log \beta_{(k)}(t, N; \lambda) = N \log F_{(k)}^N(t; e^{-\lambda/Ne^{rt}}) \sim -N\{1 - F_{(k)}^N(t; e^{-\lambda/Ne^{rt}})\}$$

and similarly to the discrete-time case

$$\log \beta_{(k)}(t, N; \lambda) \sim -Ku_1 Ne^{rt}\{1 - F_{(k)}^*(e^{-\lambda/Ne^{rt}})\}, \quad (20)$$

where $\lim\{[1 - F_{(k)}^*(t; e^{-\lambda/Ne^{rt}})]/[1 - e^{-\lambda/Ne^{rt}}]\} = A_k^* = v_k/K$ and $1 - e^{-\lambda/Ne^{rt}} \sim \lambda/Ne^{rt}$.

Finally from (20) one obtains that $\log \beta_{(k)}(t, N; \lambda) \sim -u_1 v_k$ and by the continuity theorem as $Ne^{rt} \rightarrow \infty$ one has

$$Z_k(t; N)/Ne^{rt} \rightarrow u_1 v_k \text{ in probability, } k = 1, \dots, d. \quad (21)$$

Hence by (8) and (21) it follows

$$U(t; N)/Ne^{rt} \rightarrow u_1 V \text{ in probability, } V = \sum_{k=1}^d v_k. \quad (22)$$

Then from (9), (21) and (22) it is not difficult to see that in the case $r < 0$, $Ne^{rt} \rightarrow \infty$

$$\Delta_k(t; N) \rightarrow p_k = v_k/V \text{ in probability, } k = 1, 2, \dots, d. \quad (23)$$

2) Consider now the critical case ($R = 1$ or $r = 0$) with $N/t \rightarrow \infty$.

Using (13) one can introduce the following Laplace transform

$$\gamma_{(k)}(t, N; \lambda) = \mathbb{E}\{e^{-\lambda Z_k(t; N)/N}\} = F_{(k)}^N(t; e^{-\lambda/N}), \quad t \in \mathbf{T}. \quad (24)$$

Then $\log \gamma_{(k)}(t, N; \lambda) = N \log F_{(k)}^N(t; e^{-\lambda/N}) \sim N\{1 - F_{(k)}^N(t; e^{-\lambda/N})\}$.

As $t \rightarrow \infty$ in the discrete-time case one has (see Athreya and Ney [1972, Ch. V. 5], Mode [1971, Ch. 1.10] and Sevastyanov [1971, Ch. VI. 3])

$$1 - F(t; \mathbf{s}) \sim \{\mathbf{u}_1 \sum_{j=1}^d v_j(1 - s_j)\} / \{1 + \frac{B}{2} t \sum_{i=1}^d v_i(1 - s_i)\} \quad (25)$$

uniformly for $\mathbf{0} \leq \mathbf{s} \leq \mathbf{1}$, $\mathbf{s} \neq \mathbf{1}$, where B is defined in (4).

Therefore by (13) and (23) one obtains

$$\log \gamma_{(k)}(t, N; \lambda) \sim -Nu_1 v_k(1 - e^{-\lambda/N}) / \{1 + \frac{B}{2} t v_k(1 - e^{-\lambda/N})\} \sim -u_1 v_k \lambda / (1 + \frac{B}{2} v_k \lambda \frac{t}{N}). \quad (26)$$

Since $\gamma_{(k)}(t, N; \lambda) \sim e^{-u_1 v_k}$ as $N/t \rightarrow \infty$ then by the continuity theorem it follows that

$$Z_k(t; N)/N \rightarrow u_1 v_k \text{ in probability, } k=1, \dots, d. \quad (27)$$

Hence by (8) and (27) one obtains

$$U(t; N)/N \rightarrow u_1 V \text{ in probability, } V = \sum_{k=1}^d v_k. \quad (28)$$

Finally by (9), (27) and (28) it follows that in the case $R < 1$, NR^t

$$\Delta_k(t; N) \rightarrow p_k = v_k/V \text{ in probability, } k=1, 2, \dots, d. \quad (29)$$

Note that some results obtained for discrete-time Markov branching processes can be transferred to the continuous-time case using the method of embedded branching processes. In fact, for $F(t, s)$, $t \in \mathbf{R}^+$, and $n = 2^{-n}$, $n = 1, 2, \dots$, one can consider the p.g.f. of the embedded process $\tilde{F}(t; s) = F(\mathbf{k} \varepsilon_n; s) = \mathbf{F}_n(\mathbf{k}; s)$, where $t \in \{k \varepsilon_n, k = 0, 1, 2, \dots\}$. In this case $R_n = e^{r \varepsilon_n}$ and for the critical processes $B(\varepsilon_n) = b \varepsilon_n (1 + o(1))$ as $n \rightarrow \infty$. Therefore the asymptotic relations (25) and (26) will be valid also in the continuous-time case where B can be replaced by b from (4). Hence (27) – (29) hold.

3) In the supercritical case ($R > 1$ or $r > 0$) it is well-known (see Athreya and Ney [1972, Ch. V. 6 & 7], Mode [1971, Ch. 1.8 and 1.9] or Sevastyanov [1971, Ch. VI. 4]) that for $W_k(t) = Z_k(t)/R^t$, $t \in \mathbf{N}_0$, or $W_k(t) = Z_k(t)/e^{rt}$, $t \in \mathbf{R}^+$, the following relations hold as t

$$W_k(t) \rightarrow v_k W \text{ a.s., } k=1, \dots, d; \mathbb{E}\{W | \mathbf{Z}(0) = \delta_1\} = \mathbf{u}_1. \quad (30)$$

Let $\tilde{Z}_k(t; N) = Z_k(t; N)/NR^t$, $t \in \mathbf{N}_0$, or $\tilde{Z}_k(t; N) = Z_k(t; N)/Ne^{rt}$, $t \in \mathbf{R}^+$. Then by (5) one has

$$\eta_k(t; N) = \frac{1}{N} \sum_{j=1}^N W_{k,j}(t), \text{ where } \{W_{k,j}(t)\}_{j=1}^N \text{ are i.i.d. as } W_k(t), k=1, \dots, d.$$

On the other hand, as N, t

$$\eta_k(t; N) = v_k W + \frac{1}{N} \sum_{j=1}^N \zeta_{k,j}(t) \rightarrow v_k W \text{ a.s., } k=1, \dots, d,$$

where $\zeta_{k,j}(t) = W_{k,j}(t) - v_k W$ and by (30) one has that $\frac{1}{N} \sum_{j=1}^N \zeta_{k,j}(t) \rightarrow 0$ a.s.

Therefore by (8) one obtains that $U(t; N)/NR^t$, $t \in \mathbf{N}_0$, or $U(t; N)/Ne^{rt}$, $t \in \mathbf{R}^+$, converges a.s.

to VW as $N, t \rightarrow \infty$, $V = \sum_{k=1}^d v_k$. Then the statement (ii) follows.

Remark 2

In other words, the fractions $\tilde{Z}_i(t; N)$, $i=1, \dots, d$, are consistent estimators for p_i under the conditions of Theorem 2. By virtue of the fact that $0 \leq \tilde{Z}_i(t; N) \leq 1$ and the dominated convergence theorem (DCT), it follows that $\mathbb{E}\{\tilde{Z}_i(t; N)\}$ converges to p_i implying that $\tilde{Z}_i$

(t, N) is an asymptotically unbiased estimator for p_i , $i = 1, 2, \dots, d$. Similarly, one has that $\text{IE}\{\Delta_i^2(t; N)\} \rightarrow p_i^2$ by the DCT and therefore $\text{Var}\{\Delta_i(t; N)\} \rightarrow 0$.

Note also that for each i the defined in (11) quantity $p_i(t)$ may be interpreted as a probability for a randomly chosen cell attime t to be of the type i and by (3), (11) and (12) it is not

difficult to see that $p_i(t) \rightarrow p_i$ as $t \rightarrow \infty$, $\sum_{i=1}^d p_i = 1$.

On the other hand, if t is fixed and $N \rightarrow \infty$ then from (5) and (8) applying the law of large numbers (LLN) one obtains

$$\Delta_i(t; N) = \left[\frac{1}{N} \sum_{k=1}^N Z_{i,k}(t) \right] / \left[\frac{1}{N} \sum_{k=1}^N U^{(k)}(t) \right] \rightarrow p_i(t) \text{ a.s., } i = 1, \dots, d.$$

Put it another way, the frequencies $\Delta_i(t, N)$, $i = 1, \dots, d$, are strongly consistent estimators for $p_i(t)$ when considered as functions of the initial number of ancestors N (but the time t is fixed). Moreover, in this case the frequencies are also asymptotically normal (see *Theorem A* in the next section and *Remark 3*).

Theorem 3

Let $N, t \rightarrow \infty$.

(i) If $\{R < 1, NR^t \rightarrow 0 \text{ or } r < 0, Ne^{rt} \rightarrow 0\} \cup \{R = 1 \text{ or } r = 0 \text{ and } N/t \rightarrow 0\}$ then $Z_k(t, N) \rightarrow 0$ in probability, $k = 1, 2, \dots, d$.

(ii) If $\{R < 1, NR^t \rightarrow C \text{ or } r < 0, Ne^{rt} \rightarrow C \text{ and } 0 < C < \infty\} \cup \{R = 1 \text{ or } r = 0 \text{ and } N/t \rightarrow C, 0 < C < \infty\}$ then $Z_k(t; N) \xrightarrow{d} \chi_k$, $k = 1, 2, \dots, d$, and $U(t; N) \xrightarrow{d} \chi$, where

$$\varphi_k(\lambda) = \text{IE}\{e^{-\lambda\chi_k}\} = \exp\{-u_1 CK[1 - F_{(k)}^*(e^{-\lambda/C})]\},$$

$$\varphi(\lambda) = \text{IE}\{e^{-\lambda\chi}\} = \exp\{-u_1 CK[1 - F^*(e^{-\lambda/C})]\}.$$

The p.g.f. $F_{(k)}^*(s)$ and $F^*(s)$ are well defined in the discrete-time case by (14) or by (19) in the continuous-time case.

(iii) If $\{R = 1 \text{ or } r = 0 \text{ and } N/t \rightarrow C, 0 < C < \infty\}$ then $Z_k(t, N)/N \xrightarrow{d} \varrho_k$, $k = 1, 2, \dots, d$, and $U(t, N)/N \xrightarrow{d} \varrho$, where

$$\psi_k(\lambda) = \text{IE}\{e^{-\lambda\varrho_k}\} = \exp\{-\alpha[1 - \beta_k/(\beta_k + \lambda)]\},$$

$$\psi(\lambda) = \text{IE}\{e^{-\lambda\varrho}\} = \exp\{-\alpha[1 - \beta/(\beta + \lambda)]\}$$

$$\text{and } \alpha = 2u_1/C, \beta_k = 2/CBv_k \text{ or } \beta_k = 2/Cbv_k, \beta^{-1} = \sum_{k=1}^d \beta_k^{-1}.$$

Proof

(i) In this case for the subcritical processes by (15) and (20) it is not difficult to obtain that $\log_{(k)}(t, N) \rightarrow 0$ and for the critical processes $\log_{(k)}(t, N) \rightarrow 0$ by (26), which means that $\mathbb{IE}\{e^{-Z_k(t, N)}\} \rightarrow 1$, showing the statement by the continuity theorem.

(ii) This case follows immediately from (15) and (20). Note that limiting distributions are

Compound Poisson, i.e. $\chi_k = \sum_{j=1}^n \xi_k(j)$, where ξ_k has a Poisson distribution with parameter $\lambda_k = u_1 CK$ and $\xi_k(j), j = 1, \dots$, are *i.i.d.* with a Laplace transform $\mathbb{IE}\{e^{-\lambda \xi_k(j)}\} = F_{(k)}^*(e^{-\lambda/C})$. Similarly for ϱ_k .

(iii) The statement follows directly from (26) (with a remark that in the continuous-time case one can use b instead B). The limiting distributions are also Compound Poisson but now

$\varrho_k = \sum_{j=1}^n \xi_k(j)$ with $Po(\lambda_k)$, $\lambda_k = 2u_1/C$, and $\{\xi_k(j)\}$ are *i.i.d.* and exponentially distributed with parameter $\lambda_k = 2/CB$ or $\lambda_k = 2/Cb$. Similarly for $\varrho_k = \sum_{j=1}^n \xi_k(j)$ where $\{\xi_k(j)\}$ are *i.i.d.* and $V = \sum_{k=1}^d v_k$. $\chi_k \sim \text{Exp}(\lambda_k)$, $\lambda_k = 2/BKV$ or $\lambda_k = 2/bKV$.

5. Multivariate Asymptotic Normality

In this section we will assume that the processes have finite second moments.

Consider $\mathbf{X}(t, N) = (X_1(t, N), \dots, X_d(t, N))$, $t \in \mathbf{T}$, where for some norming functions $D_k(t)$

$$X_k(t; N) = [Z_k(t; N) - A_{1k}(t; N)] / [D_k(t) \sqrt{N}] = \frac{1}{\sqrt{N}} \sum_{j=1}^N [Z_{k,j}(t) - A_{1k}(t)] / D_k(t), \quad (31)$$

because of (5) and the fact that $A_{1k}(t, N) = \mathbb{IE}\{Z_k(t, N)\} = N A_{1k}(t)$, $k = 1, \dots, d$.

Theorem 4

Assume $N, t \rightarrow \infty$.

(i) If $\{R < 1, NR^t \rightarrow \infty, \text{ or } r < 0, Ne^{rt} \rightarrow \infty\}$ and $D_k(t) \equiv D(t) \sim \{\sqrt{R^t} \text{ or } \sqrt{e^{rt}}\}$ for $k = 1, \dots, d$, then

$$\mathbf{X}(t; \mathbf{N}) \xrightarrow{d} \xi = (\xi_1, \dots, \xi_d), \quad (32)$$

and $\xi = (\xi_1, \dots, \xi_d)$ has a multivariate normal distribution with $\mathbb{IE} \xi_i = 0$ and a covariance matrix $C = C_{jk}$, where

$$C_{jk} = \lim_{N \rightarrow \infty} B_{jk}^{(1)}(t) / R^t, \quad t \in \mathbf{N}_0 \quad \text{or} \quad C_{jk} = \lim_{N \rightarrow \infty} B_{jk}^{(1)}(t) / e^{rt}, \quad t \in \mathbf{R}^+.$$

(ii) If $\{R = 1 \text{ or } r = 0 \text{ with } N/t \rightarrow 0\}$ and $D_k(t) \sim \{v_k \sqrt{u_1 B t} \text{ or } v_k \sqrt{u_1 b t}\}$, $k = 1, \dots, d$, then

$$\mathbf{X}(t; \mathbf{N}) \xrightarrow{d} \boldsymbol{\eta} = (\eta_1, \dots, \eta_d), \quad (33)$$

where $\eta_1 = \dots = \eta_d \text{ a.s.}, \eta_1 \sim N(0, 1)$.

(iii) If $R > 1 \text{ or } r > 0$ and $D_k(t) \sim \{v_k R^t \text{ or } v_k e^{rt}\}$, $k = 1, \dots, d$, then

$$\mathbf{X}(t; \mathbf{N}) \xrightarrow{d} \boldsymbol{\zeta} = (\zeta_1, \dots, \zeta_d), \quad (34)$$

where $\eta_1 = \dots = \eta_d \text{ a.s.}, \eta_1 \sim N(0, \sigma^2)$ and $\sigma^2 = \text{IE}\{\mathbf{W}^2 \mid \mathbf{Z}(0) = \delta_1\}$ is defined in (30).

Proof

Using (31) one can introduce the multivariate characteristic functions

$$\Psi_{t,N}(\alpha) = \text{IE}\{\mathbf{e}^{i\alpha \cdot \mathbf{X}^T(t;N)}\} = [\text{IE}\{\mathbf{e}^{i \sum_{k=1}^d \alpha_k \tilde{Z}_k(t) / \sqrt{N}} \mid \mathbf{Z}(0) = \delta_1\}]^N,$$

where $\tilde{Z}_k(t) = [Z_k(t) - A_{1k}(t)] / D_k(t)$, $k = 1, \dots, d$.

Further on we will use the following representation

$$\begin{aligned} \Psi_{t,N}(\alpha) &= \mathbf{e}^{-i\sqrt{N} \sum_{k=1}^d \alpha_k A_{1k}(t) / D_k(t)} [\text{IE}\{\mathbf{e}^{i \sum_{k=1}^d \alpha_k \tilde{Z}_k(t) / \sqrt{N}} \mid \mathbf{Z}(0) = \delta_1\}]^N \\ &= \mathbf{e}^{-i\sqrt{N} \sum_{k=1}^d \alpha_k A_{1k}(t) / D_k(t)} F^N(t; \mathbf{e}^{i\alpha_1 / D_1(t) \sqrt{N}}, \dots, \mathbf{e}^{i\alpha_d / D_d(t) \sqrt{N}}), \end{aligned}$$

where $F(t, \mathbf{s})$ is introduced in (6).

Hence,

$$\log \Psi_{t,N}(\alpha) = -i\sqrt{N} \sum_{k=1}^d \alpha_k A_{1k}(t) / D_k(t) + N \log F(t; \mathbf{e}^{i\alpha_1 / D_1(t) \sqrt{N}}, \dots, \mathbf{e}^{i\alpha_d / D_d(t) \sqrt{N}}). \quad (35)$$

(i) Note that in this case $D(t) \sqrt{N} \sim \sqrt{N} R^t$ (or $\sqrt{N} e^{rt}$) and $D(t) \sqrt{N} \rightarrow \infty$. Therefore

$$\log \Psi_{t,N}(\alpha) \sim -i\sqrt{N} \sum_{k=1}^d \alpha_k A_{1k}(t) / D(t) - N_{\varphi t,N}(\alpha), \quad (36)$$

where

$$\varphi_{t,N}(\alpha) = 1 - F(t; \mathbf{e}^{i\alpha_1 / D(t) \sqrt{N}}, \dots, \mathbf{e}^{i\alpha_d / D(t) \sqrt{N}}). \quad (37)$$

Hence developing (37) it is not difficult to see that

$$\begin{aligned}\varphi_{t,N}(\alpha) \sim & - \sum_{k=1}^d (1 - e^{i\alpha_k / D(t) \sqrt{N}}) A_{1k}(t) \\ & - \frac{1}{2} \sum_{k,l=1}^d (1 - e^{i\alpha_k / D(t) \sqrt{N}}) (1 - e^{i\alpha_l / D(t) \sqrt{N}}) B_{jk}^{(1)}(t) .\end{aligned}$$

Then the following asymptotic result hold

$$\varphi_{t,N}(\alpha) \sim - (i / \sqrt{N}) \sum_{k=1}^d \alpha_k A_{1k}(t) / D(t) + \frac{1}{2N} \sum_{k,l=1}^d \alpha_k \alpha_l B_{kl}^{(1)}(t) / D^2(t) . \quad (38)$$

It is well-known that in the subcritical case

$$B_{jk}^{(i)}(t) \sim C_{jk}^{(i)} R^t, \quad t \in \mathbf{N}_0; \quad B_{jk}^{(i)}(t) \sim C_{jk}^{(i)} e^{rt}, \quad t \in \mathbf{R}^+. \quad (39)$$

Further on we will use the notation $C_{jk} = C_{jk}^{(1)}$ and the corresponding matrix $C = (C_{jk})$.

Therefore from (36) – (39) one obtains $\log \Psi_{t,N}(\alpha) \rightarrow -\frac{1}{2} \sum_{k,l=1}^d \alpha_k \alpha_l C_{kl}$, which proves (32) by the continuity theorem.

(ii) Note that in this case $D_k(t) \sqrt{N} \sim v_k \sqrt{u_1 B N t}$ (or $v_k \sqrt{u_1 b N t}$) and therefore (36) and (37) hold. Further on we will consider only the discrete-time case because the continuous-time case is similar as it was pointed out in the proof of *Theorem 2*. Then by (25) and (37) it is not difficult to obtain that

$$\begin{aligned}\varphi_{t,N}(\alpha) \sim & \{u_1 \sum_{j=1}^d v_j (1 - e^{i\alpha_j / D_j(t) \sqrt{N}})\} / \{1 + \frac{B}{2} t \sum_{l=1}^d v_l (1 - e^{i\alpha_l / D_l(t) \sqrt{N}})\} \\ \sim & -(iu_1 / \sqrt{N}) \sum_{j=1}^d v_j \alpha_j / D_j(t) / [1 - (iBt/2 \sqrt{N}) \sum_{l=1}^d v_l \alpha_l / D_l(t)] .\end{aligned}$$

From here using (36) and the fact that $A_{1k}(t) \sim u_{1k}$ one has

$$\log \Psi_{t,N}(\alpha) \sim -iu_1 \sqrt{N} \sum_{k=1}^d \frac{\alpha_k v_k}{D_k(t)} + iu_1 \sqrt{N} \sum_{j=1}^d \frac{v_j \alpha_j}{D_j(t)} / [1 - i \frac{Bt}{2 \sqrt{N}} \sum_{l=1}^d \frac{v_l \alpha_l}{D_l(t)}] = - (u_1 B t / 2) \{ \sum_{k=1}^d \frac{\alpha_k v_k}{D_k(t)} \}^2 / [1 - i \frac{Bt}{2 \sqrt{N}} \sum_{l=1}^d \frac{v_l \alpha_l}{D_l(t)}] .$$

Since $D_k(t) \sim v_k \sqrt{u_1 B t}$ then

$$\log \Psi_{t,N}(\alpha) \sim -\frac{1}{2} \left\{ \sum_{k=1}^d \alpha_k \right\}^2 / \left[1 - \frac{i\sqrt{B}}{2\sqrt{u_1}} \sqrt{\frac{t}{N}} \sum_{l=1}^d \alpha_l \right].$$

Finally $\Psi_{t,N}(\alpha) \rightarrow e^{-\frac{1}{2}(\sum_{k=1}^d \alpha_k)^2}$ which proves (33) by the continuity theorem.

(iii) In this case, using (30) and the references cited there, one has as

$$t \rightarrow \infty \sum_{k=1}^d \alpha_k W_k(t) \rightarrow W \sum_{k=1}^d \alpha_k v_k, \text{ a.s. and } L_2.$$

Therefore in the discrete-time case as $t \rightarrow \infty$ the following convergence holds

$$F(t; \mathbf{e}^{i\alpha_1/R^t}, \dots, \mathbf{e}^{i\alpha_d/R^t}) = \mathbb{E}\{\mathbf{e}^{i \sum_{k=1}^d \alpha_k Z_k(t)/R^t} | \mathbf{Z}(0) = \delta_1\}$$

$$\rightarrow \mathbb{E}\{\mathbf{e}^{iW \sum_{k=1}^d \alpha_k v_k} | \mathbf{Z}(0) = \delta_1\} = \phi_1\left(\sum_{k=1}^d \alpha_k v_k\right).$$

Note that a similar relation is fulfilled also in the continuous-time case where only R^t has to be replaced by $\mathbf{e}^{t\mathbf{r}}$.

On the other hand, for the discrete-time processes the following equations hold

$$\phi_k(Rx) = F_k(\phi_1(x), \dots, \phi_d(x)), k=1, 2, \dots, d, \quad (40)$$

where $\phi_k(x) = \mathbb{E}\{\mathbf{e}^{ixW} | \mathbf{Z}(0) = \delta_k\}$, $x \in \mathbf{R}^1$, and $A_k^* = \mathbb{E}\{W | \mathbf{Z}(0) = \delta_k\} = \mathbf{u}_k$.

Therefore as $x \rightarrow 0$

$$1 - \phi_1(x) \sim -iu_1 x + \tau^2 x^2 / 2, \quad (41)$$

where τ^2 is defined in (34).

Note that in this case the relations (35) – (37) hold and by (37) and (41) one has

$$\varphi_{t,N}(\alpha) \sim 1 - \phi_1\left(\sum_{k=1}^d \alpha_k / \sqrt{N}\right) \sim -iu_1 \sum_{k=1}^d \alpha_k / \sqrt{N} + (\tau^2 / 2N) \left\{ \sum_{k=1}^d \alpha_k \right\}^2.$$

Sine $A_{1k}(t)/Dk(t) \rightarrow u_1$ then by (36) we obtain

$$\log \Psi_{t,N}(\alpha) \sim -iu_1 \sqrt{N} \sum_{k=1}^d \alpha_k - N[-iu_1 \sum_{k=1}^d \alpha_k / \sqrt{N} + (\tau^2/2N) \{ \sum_{k=1}^d \alpha_k \}^2] = -(\tau^2/2) \{ \sum_{k=1}^d \alpha_k \}^2,$$

which proves (34) in the discrete-time case (applying the continuity theorem).

In the continuous-time case instead (40) the following system of differential equations holds

$$\frac{d}{dx} \phi_k(x) = f_k(\phi_1(x), \dots, \phi_d(x)) / r x, \quad k=1, 2, \dots, d. \quad (42)$$

From (42) it is not difficult to obtain that $rA_k^* = \sum_{j=1}^d a_{kj}A_j^*, \quad k=1, 2, \dots, d.$

Therefore by the Frobenius Theorem it follows that $A_k^* = Cu_k$ for some constant C and it is not difficult to prove that $C=1$. Indeed, since $A_{ik}(t) \sim u_i k e^{rt}$ then

$$\mathbb{E}\{W_k(t)/v_k | \mathbf{Z}(0) = \delta_1\} = A_{1k}(t)/v_k e^{rt} \rightarrow u_1 = A_1^* = \{W | \mathbf{Z}(0) = \delta_1\}.$$

Hence (41) holds and the rest of the proof is similar as in the discrete-time case.

As it was pointed out in *Remark 2*, the relative frequencies $f_i(t, N), i=1, 2, \dots, d$, introduced by (9), are strongly consistent and asymptotically unbiased estimators for $p_i(t), i=1, 2, \dots, d$, (defined by (11) and (12)) as $N \rightarrow \infty$ for any fixed t . As one can see by the following theorem (proved in [32]), in this case the frequencies have also asymptotic multivariate normal distributions. Note that the theorem is valid for any kind of branching processes with discrete or continuous time (Markov or non Markov, reducible or irreducible) assuming only the usual independence of the individual evolutions.

Introduce the following notions:

$$\begin{aligned} \sigma_i^2(t) &= \text{Var}\{Z_i(t) | \mathbf{Z}(0) = \delta_1\}, \quad i=1, 2, \dots, d, \\ \mathbf{R}^{(d)}(t) &= \|r_{ij}(t)\|, \quad r_{ij}(t) = \text{Cor}\{Z_i(t), Z_j(t) | \mathbf{Z}(0) = \delta_1\}, \quad i, j=1, 2, \dots, d, \\ \mathbf{C}^{(d)}(t) &= \|c_{ij}(t)\|, \quad c_{ii}(t) = \sigma_i(t)(1 - p_i(t)), \quad c_{ij}(t) = -\sigma_i(t)p_j(t) \text{ for } i \neq j; \quad i, j=1, 2, \dots, d. \end{aligned}$$

Theorem A

([32]). Let t be fixed and $N \rightarrow \infty$. Then for the r .

$$V_i(t; N) = [Z_i(t; N) - A_{1i}(t; N)] / [\sigma_i(t) \sqrt{N}] \quad \text{and} \quad W_i(t; N) = M(t) \sqrt{N} [\Delta_i(t; N) - p_i(t)],$$

$i=1, 2, \dots, d$, the following statements are valid:

$$(i) (V_1(t; N), \dots, V_d(t; N)) \xrightarrow{d} (X_1(t), \dots, X_d(t)),$$

where the random vector $\mathbf{X}^{(d)}(t) = (X_1(t), \dots, X_d(t))$ has a joint normal distribution with

$$\mathbb{E}\{X_i(t)\} = 0, \text{Var}\{X_i(t)\} = 1 \text{ and } \text{Cov}\{X_i(t), X_j(t)\} = r_{ij}(t).$$

(ii) $W_i(t; N) \xrightarrow{d} Y_i(t)$, where $Y_i(t)$ is a normally distributed r. . with

$$\mathbb{E}\{Y_i(t)\} = 0 \text{ and } \text{Var}\{Y_i(t)\} = \sum_{j,k=1}^d r_{jk}(t) c_{ji}(t) c_{ki}(t), \quad i = 1, 2, \dots, d;$$

(iii) For every $k = 2, 3, \dots, d-1$

$$(W_1(t; N), \dots, W_k(t; N)) \xrightarrow{d} (Y_1(t), \dots, Y_k(t)),$$

and $\mathbf{Y}^{(k)}(t) = (Y_1(t), \dots, Y_k(t))$ has a multivariate normal distribution with a covariance matrix which can be calculated as follows:

$$\mathbf{D}^{(k)}(t) = \|\text{Cov}\{Y_i(t), Y_j(t)\}\| = [\mathbf{C}_{d \times k}(t)]^T \mathbf{R}^{(d)}(t) \mathbf{C}_{d \times k}(t),$$

where $[\mathbf{C}_{d \times k}(t)]^T = [c_{jk}(t)]$, $j = 1, 2, \dots, k$; $i = 1, 2, \dots, d$, is the corresponding transposed matrix of $[k \times d]$ dimensions.

Remark 3

While condition (10) implies that the fractions $X_i(t; N)$, $i = 1, 2, \dots, d$, are linear dependent, there exist $d-1$ joint normal distributions of lower dimensions that are asymptotically non-degenerate.

In a contrast of *Theorem A* the following results hold as both the number of the ancestors and the time increase to infinity.

Theorem 5

Assume $N, t \rightarrow \infty$ and Condition (i) of Theorem 4. Then

(i) For every $k = 1, 2, \dots, d$

$$\Phi_k(t; N) = u_1 V \sqrt{ND(t)} [\Delta_k(t; N) - p_k(t)] \xrightarrow{d} Y_k,$$

where Y_k is a normally distributed r. . with $\mathbb{E}\{Y_k\} = 0$ and

$$\text{Var}\{Y_k\} = C_{kk} - 2pk_{j=1}^d C_{kj} + p_{ki,j=1}^{2d} C_{ij}.$$

(ii) For every $k = 2, 3, \dots, d-1$ and every subset $(n_1, n_2, \dots, n_k)$ with nonrecurring elements from the set $\{1, 2, \dots, d\}$ the following joint distributions hold

$$(\Phi_{n_1}(t; N), \dots, \Phi_{n_k}(t; N)) \xrightarrow{d} (Y_{n_1}, \dots, Y_{n_k}),$$

where the random variables $(Y_{n1}, \dots, Y_{nk})$ have a multivariate normal distribution with

$$\text{Cov}(Y_i, Y_j) = C_{ij} - p_{ik=1}^d C_{kj} - p_{jl=1}^d C_{li} + p_i p_{jk,l=1}^d C_{kl}. \quad (43)$$

Proof

Using (31) it is not difficult to obtain the following relations

$$\Phi_k(t; N) = \frac{u_1 V N D^2(t)}{U(t; N)} \{ [1 - p_k(t)] X_k(t; N) - p_k(t) \sum_{j \neq k}^d X_j(t; N) \}, \quad k = 1, 2, \dots, d. \quad (44)$$

Note first that from (17) and (22) one obtains $\frac{N D^2(t) u_1 V}{U(t; N)} \rightarrow 1$ in probability. On the other hand, by (3), (11) and (12) it is not difficult to see that $p_k(t) \rightarrow p_k$, $k = 1, 2, \dots, d$. Finally applying *Theorem 4*, (32), one has $X_j(t; N) \xrightarrow{d} \xi_j$, where $(\xi_1, \dots, \xi_d)$ have a multivariate normal distribution with zero means and a covariance matrix $C = C_{jk}$. Therefore by (44) it follows that

$$\Phi_k(t; N) \xrightarrow{d} Y_k = (1 - p_k) \xi_k - p_k \sum_{j \neq k}^d \xi_j = \xi_k - p_k \sum_{j=1}^d \xi_j, \quad k = 1, 2, \dots, d, \quad (45)$$

which obviously proves the statement (i).

The statement (ii) follows immediately by (44) and (45) applying *Theorem 4*, (i), where one has to check only that $\text{IE}\{Y_i Y_j\}$ is well presented by (43).

Remark 4

Surprisingly in the critical and supercritical cases there are not analogs of *Theorem 5*. We will consider only the discrete-time case, because the continuous-time case is quite similar.

Indeed, under the conditions of *Theorem 4*, (ii), it is not difficult to see that

$$\sqrt{N u_1 / t B p_k^2} [\Delta_k(t; N) - p_k(t)] \xrightarrow{d} Y_k = (1 - p_k) \eta_k - \sum_{j \neq k} p_j \eta_j, \quad k = 1, 2, \dots, d,$$

where $\eta_1 = \dots = \eta_d$ a.s., $\eta_1 \sim N(0, 1)$. Since $\sum_{j=1}^d p_j = 1$ then $Y_k = 0$ a.s., $k = 1, 2, \dots, d$.

Similarly, under the conditions of *Theorem 4*, (iii),

$$\sqrt{N / p_k^2} [\Delta_k(t; N) - p_k(t)] \xrightarrow{d} W^{-1} Y_k \quad \text{and} \quad (U(t; N) / \nu_k R^t \sqrt{N}) [\Delta_k(t; N) - p_k(t)] \xrightarrow{d} Y_k,$$

where

$$Y_k = (1 - p_k)S_k - \sum_{j \neq k}^d p_j S_j, \quad S_1 = \dots = S_d \text{ a.s., } S_1 \in N(0, \tau^2), \quad \tau^2 = \text{IE}\{W^2 \mid Z(0) = \delta_1\}.$$

Therefore $Y_k = 0$ a.s., $k = 1, 2, \dots, d$.

Remark 5

It is interesting to compare *Theorem A* with *Theorem 4* and *5* (see also *Remark 4*). While *Theorem A* is valid for any kind of branching processes and does not depend of the criticality, then *Theorem 4* and *5* are proved only in the Markov case and depend essentially by the usual trichotomy of the processes as well as by the relationship between N and t . *Theorem 4* and *5* complete *Theorem A* in the case when t is large enough (but the obtained results are quite different).

6. Concluding Remarks

As it was pointed out in the Introduction and in *Section 2*, there are experimental situations in the cell biology where analyzing the relative frequencies, $\lambda(t, N)$, of cell types rather than the total cell counts $Z_\lambda(t, N)$, $i = 1, 2, \dots, d$, may be quite advantageous. Should this be the case, the property of asymptotic normality given by *Theorem A* ((ii) and (iii)) and *Theorem 5* could be useful in developing the needed statistical inference of model parameters from experimental data. On the other hand, if the process can be observed directly then one can apply *Theorem 4* or *Theorem A*, (i).

In particular, the following observation process is directly relevant to quantitative studies of proliferation, differentiation, and death of cells. Suppose that the process under study begins

with $N = \sum_{k=1}^n N_k$ cells of type T_1 and the values of N_k are all large, i.e., $N_0 = \min\{N_1, N_2, \dots, N_n\}$. The descendants of the first N_1 ancestors are examined only once at time t_1 to determine the observations of $Z_\lambda(t_1; N_1)$ or $\lambda(t_1; N_1) = 1, 2, \dots, d$, whereupon the observation process is discontinued (i.e. the cells under examination are destroyed). At the next moment $t_2 \geq t_1$, the process $Z_\lambda(t_2; N_2)$ or the fractions $\lambda(t_2; N_2)$, $i = 1, 2, \dots, d$, related to the descendants of the second N_2 ancestors are observed, and so on. This procedure results in n independent observations of the form:

$$\mathbf{z}_k = \mathbf{Z}(t_k; N_k) = (Z_1(t_k; N_k), \dots, Z_d(t_k; N_k)) \text{ or } \mathbf{z}_k = (\lambda(t_k; N_k)) = (\lambda_1(t_k; N_k), \dots, \lambda_d(t_k; N_k)), \quad k = 1, 2, \dots, n; \quad t_1 \leq t_2 \leq \dots \leq t_n,$$

where each vector $\mathbf{z}_k$ is asymptotically normal in accordance with *Theorems 4, 5* or *A*.

Denoting the corresponding contribution to the log-likelihood function by $L_k(\mathbf{z}_k; t_k, N_k)$, the overall log-likelihood is given by

$$\Lambda_n(\zeta_1, \zeta_2, \dots, \zeta_n) = \sum_{k=1}^n L_k(\zeta_k; t_k, N_k). \quad (46)$$

The log-likelihood (45) depends on the offspring parameters only, which are of primary interest in applications and especially in cell kinetics studies. Finally the parameters can be estimated from the data on the process or the relative frequencies by maximizing the log-likelihood (46). Some applications are given in [32].

Lamperti [22] obtained interesting limiting distributions for supercritical BGW processes with infinite offspring variance. It is an open problem to consider multitype processes with infinite variance when N and t tend to infinity.

Jagers [15] was probably the first to consider relative frequencies (proportions, fractions) of cells within the framework of multitype branching processes. He studied asymptotic (as $t \rightarrow \infty$) properties of a reducible age-dependent branching process with two types of cells and proved convergence of their relative frequencies to non-random limits in mean square and almost surely on the non-extinction set. The usefulness of such frequencies in cell cycle analysis was further demonstrated by Mode [26] considering a four-type irreducible age-dependent branching process. Mode built his cell cycle analysis on a model of multitype positively regular age-dependent branching process. In the supercritical case, he proved that $\lim_{t \rightarrow \infty} \delta_k(t) = \delta_k$ a.s. as $t \rightarrow \infty$, providing the population does not become extinct. It should be noted that the constants δ_k , $k = 1, 2, \dots, d$, depend only on the offspring characteristics. In

fact, $\delta_k = v_k / \sum_{j=1}^d v_j$, where $v_k = \eta_k(1 - G_k^*(\alpha))$, $\eta = (\eta_1, \eta_2, \dots, \eta_d)$ is a left eigenvector of the matrix $\mathbf{H}(\alpha) = \parallel \mathbf{G}_k^*(\alpha) \mathbf{A} \parallel$ with a Malthusian parameter α , while $G_k^*(\lambda)$ is the Laplace-Stieltjes transform of the life-span distribution $G_k(t)$ for the type $k(k=1, 2, \dots, d)$. In his monograph, Mode [25] also considered the utility of fractions and reported a similar result for the BGW process.

The results obtained by Jagers and Mode for some age-dependent branching processes suggest that it would be interesting to investigate the asymptotic behavior of the fractions (t, N) as well as the processes $\mathbf{Z}(t, N)$ in the non Markov case when both parameters N and t tend to infinity simultaneously. It is anticipated that such asymptotic properties will depend on a specific branching model and its reducibility. Another problem is the investigation (in the same situation) of the the most complicated models of branching processes such as considered in [2, 6, 34].

Acknowledgments

This paper was prepared while N. Yanev was a visiting professor in the Department of Bio-statistics and Computational Biology, University of Rochester, and he is grateful for hospitality and inspiring communications with his colleagues.

References

1. Athreya, KB.; Ney, PE. Branching Processes. Springer; Berlin: 1972.
2. Del Puerto IM, Yanev NM. Stationary distributions for branching processes with multi-type random control functions. J Applied Statistical Science. 2008; 16(1):91–102.
3. Dion JP, Yanev NM. Statistical inference for branching processes with an increasing random number of ancestors. J Statistical Planing and Inference. 1994; 39:329–359.
4. Dion JP, Yanev NM. CLT for martingales in BGWR branching processes with some statistical applications. Mathematical Methods in Statistics. 1995; 4:344–358.
5. Dion JP, Yanev NM. Limit theorems and estimation theory for branching processes with an increasing random number of ancestors. J Applied Probability. 1997; 34:309–327.
6. Drmota M, Louchard G, Yanev NM. Analysis of a Recurrence Related to Critical Nonhomogeneous Branching Processes. Stochastic Analysis and Applications. 2006; 24(1):37–59.
7. Feller, W. Diffusion processes in genetics. Proceedings Second Berkeley Symp Math Stat Prob; Univ of California Press; 1951. p. 227–246.
8. Feller, W. An Introduction to Probability and Its Applications. 2. Vol. 2. Wiley; New York: 1971.
9. Grimvall A. On the convergence of sequences of branching processes. Annals of Probability. 1974; 2:1027–1045.

10. Guttorp, P. Statistical inference for Branching Processes. Wiley; New York: 1991.
11. Haccou, P.; Jagers, P.; Vatutin, VA. Branching Processes: Variation, Growth and Extinction of Populations. Cambridge University Press; Cambridge: 2005.
12. Harris, TE. Branching Processes. Springer; New York: 1963.
13. Hyrien O, Mayer-Proschel M, Noble M, Yakovlev AY. Estimating the life-span of oligodendrocytes from clonal data on their development in cell culture. *Mathematical Biosciences*. 2005; 193:255–274. [PubMed: 15748733]
14. Hyrien O, Mayer-Proschel M, Noble M, Yakovlev AY. A stochastic model to analyze clonal data on multi-type cell populations. *Biometrics*. 2005; 61:199–207. [PubMed: 15737094]
15. Jagers P. The proportions of individuals of different kinds in two-type populations. A branching process problem arising in biology. *J Applied Probability*. 1969; 6:249–260.
16. Jagers, P. Branching Processes with Biological Applications. Wiley; London: 1975.
17. Ji in a P. On Feller's branching diffusion processes. *Ĝasopis P St Math*. 1969; 94:84–90.
18. Kimmel, M.; Axelrod, DE. Branching Processes in Biology. Springer; New York: 2002.
19. Kolmogorov AN. Zur Lösung einer biologischen Aufgabe. *Proceedings Tomsk State Univ*. 1938; 2:1–6.
20. Kolmogorov AN, Dmitriev NA. Branching random processes. *Dokl Akad Nauk (Proceedings Acad Sci USSR)*. 1947; 56:7–10. (in Russian).
21. Kolmogorov AN, Sevastyanov BA. Calculation of final probabilities of branching random processes. *Dokl Akad Nauk (Proceedings Acad Sci USSR)*. 1947; 56:783–786. (in Russian).
22. Lamperti, J. Limiting distributions for branching processes. 1967; *Proceedings Fifth Berkeley Symp Math Stat Prob; Univ of California Press*; p. 225-241.
23. Lindvall T. Weak convergence of probability measures and random functions in the function space $D[0, \infty)$. *J Applied Probability*. 1973; 10:109–121.
24. Lindvall T. Limit theorems for some functionals of certain Galton-Watson branching processes. *Advances of Applied Probability*. 1974; 6:309–321.
25. Mode, CJ. Multitype Branching Processes: Theory and Applications. Elsevier; New York: 1971.
26. Mode CJ. Multitype age-dependent branching processes and cell cycle analysis. *Mathematical Biosciences*. 1971; 10:177–190.
27. Quine MP. A note of the moment structure of the multitype Galton-Watson process. *Biometrika*. 1970; 57:219–222.
28. Sevastyanov, BA. Branching Processes. Nauka; Moscow: 1971. (in Russian)
29. Yakovlev, AY.; Yanev, NM. Transient Processes in Cell Proliferation Kinetics, *Lecture Notes in Biomathematics*. Vol. 82. Springer; Berlin: 1989.
30. Yakovlev AY, Boucher K, Mayer-Proschel M, Noble M. Quantative insight into proliferation and differentiation of oligodendrocyte type 2 astrocyte progenitor cells in vitro. *Proceeding National Acad Sci USA*. 1998; 95:14164–14167.
31. Yakovlev AY, Yanev NM. Branching stochastic processes with immigration in analysis of renewing cell populations. *Mathematical Biosciences*. 2006; 203:37–63. [PubMed: 16904129]
32. Yakovlev AY, Yanev NM. Relative frequencies in multitype branching processes. *Annals of Applied Probability*. 2009; 19(1):1–14.
33. Yanev NM. On the statistics of branching processes. *Theory of Probability and Applications*. 1975; 20:612–622.
34. Yanev NM, Mitov KV. Critical branching processes with nonhomogeneous migration. *Annals of Probability*. 1985; 13:923–933.
35. Yanev N, Jordan CT, Catlin S, Yakovlev A. Two-type Markov branching processes with immigration as a model of leukemia cell kinetics. *Proceedings Bulg Acad Sci*. 2005; 58:1025–1032.
36. Yanev, NM. Statistical Inference for Branching Processes. In: Ahsanullah, M.; Yanev, GP., editors. *Records and Branching Processes*. NOVA Science Publishers; 2008. p. 143-168.

37. Zorin AV, Yakovlev AY, Mayer-Proschel M, Noble M. Estimation problems associated with stochastic modeling of proliferation and differentiation of O-2A progenitor cells in vitro. *Mathematical Biosciences*. 2000; 67:109–121. [PubMed: 10998484]