Serik Sagitov *Chalmers University and University of Gothenburg*

Linear-fractional age-dependent branching processes

(work in progress, supported by the

Swedish Research Council grant 621-2010-5623)

General linear-fractional generating functions

Notation: vectors $\mathbf{x} = (x_1, x_2, \ldots)$ of possibly infinite dimension

$$\mathbf{0} = (0, 0, \ldots),\ \mathbf{1} = (1, 1, \ldots),\ \mathbf{e}_i = (1_{\{i=1\}}, 1_{\{i=2\}}, \ldots),\ x = \mathbf{x}\mathbf{1}^{\mathrm{tr}}$$

where $\mathbf{x}\mathbf{y}^{\mathrm{tr}} = x_1 y_1 + x_2 y_2 + \ldots$.

A general LF pgf for a random vector $\mathbf{Z} = (Z_1, Z_2, \ldots)$

$$E(\mathbf{s}^{\mathbf{Z}}) := E(s_1^{Z_1} s_2^{Z_2} \ldots)$$

has the representation

$$E(\mathbf{s}^{\mathbf{Z}}) = r_0 + \frac{t_0 \sum_{i=1}^{\infty} r_i s_i}{1 - \sum_{i=1}^{\infty} t_i s_i},$$

where all r_i and t_i are non-negative and

$$\sum_{i=0}^{\infty} r_i = \sum_{i=0}^{\infty} t_i = 1.$$

This implies

$$\mathbf{Z} = (\mathbf{Z}' + \mathbf{Z}'')1_{\{\mathbf{Z}\neq\mathbf{0}\}},$$

where $P(\mathbf{Z} = \mathbf{0}) = r_0$, random vectors $\mathbf{Z}', \mathbf{Z}''$ are independent with

$$P(\mathbf{Z}' = \mathbf{e}_i) = r_i/(1 - r_0),\ i \geq 1,$$

and $\mathbf{Z}''$ having non-negative integer valued components

$$P(\mathbf{Z}'' = \mathbf{k}) = \binom{k}{k_1, k_2, \ldots} t_0 t_1^{k_1} t_2^{k_2} \ldots$$

Simulation of $\mathbf{Z}$ using two dice with countably many sides each:

1. the $\mathbf{r}-$die gives the type of the 1st particle, if any,
2. the $\mathbf{t}-$die generates the Geom (t_0) number of the remaining particles as well as their types.

Proof of the two dice representation. Since,

$$E(\mathbf{s}^{\mathbf{Z}'}) = \frac{1}{1-r_0}\sum_{i=1}^{\infty} r_i s_i$$

and

$$\begin{aligned} E(\mathbf{s}^{\mathbf{Z}''}) &= t_0 \sum_{\mathbf{k}} \binom{k}{k_1, k_2, \ldots} (t_1 s_1)^{k_1} (t_2 s_2)^{k_2} \ldots \\ &= \frac{t_0}{1 - \sum_{i=1}^{\infty} t_i s_i}, \end{aligned}$$

the decomposition $\mathbf{Z} = (\mathbf{Z}' + \mathbf{Z}'')1_{\{\mathbf{Z} \neq \mathbf{0}\}}$ is confirmed by

$$\begin{aligned} E(\mathbf{s}^{\mathbf{Z}}) &= r_0 + (1-r_0)E(\mathbf{s}^{\mathbf{Z}'})E(\mathbf{s}^{\mathbf{Z}''}) \\ &= r_0 + \frac{t_0 \sum_{i=1}^{\infty} r_i s_i}{1 - \sum_{i=1}^{\infty} t_i s_i}. \end{aligned}$$

With this formula for a LF pgf as a starting point, a **naive attempt** to introduce a branching process with countably many types would define the reproduction law of type i particles by pgfs

$$E(\mathbf{s}^{\mathbf{Z}^{(1)}}|\mathbf{Z}^{(0)} = \mathbf{e}_i) = r_{i0} + \frac{t_{i0}\sum_{j=1}^{\infty} r_{ij}s_j}{1 - \sum_{j=1}^{\infty} t_{ij}s_j}$$

allowing for most general dependence on the mother's type.

However, as it was shown in the finite-dimensional case [6], for the convolutions of the LF functions to be again LF, it is necessary that the parameters $t_j = t_{ij}$ are independent of the mother type i.

Next we will justify this restriction on a unique $\mathbf{t}-$die using a simple argument referring to the contour process of the random tree generated by the LF reproduction.

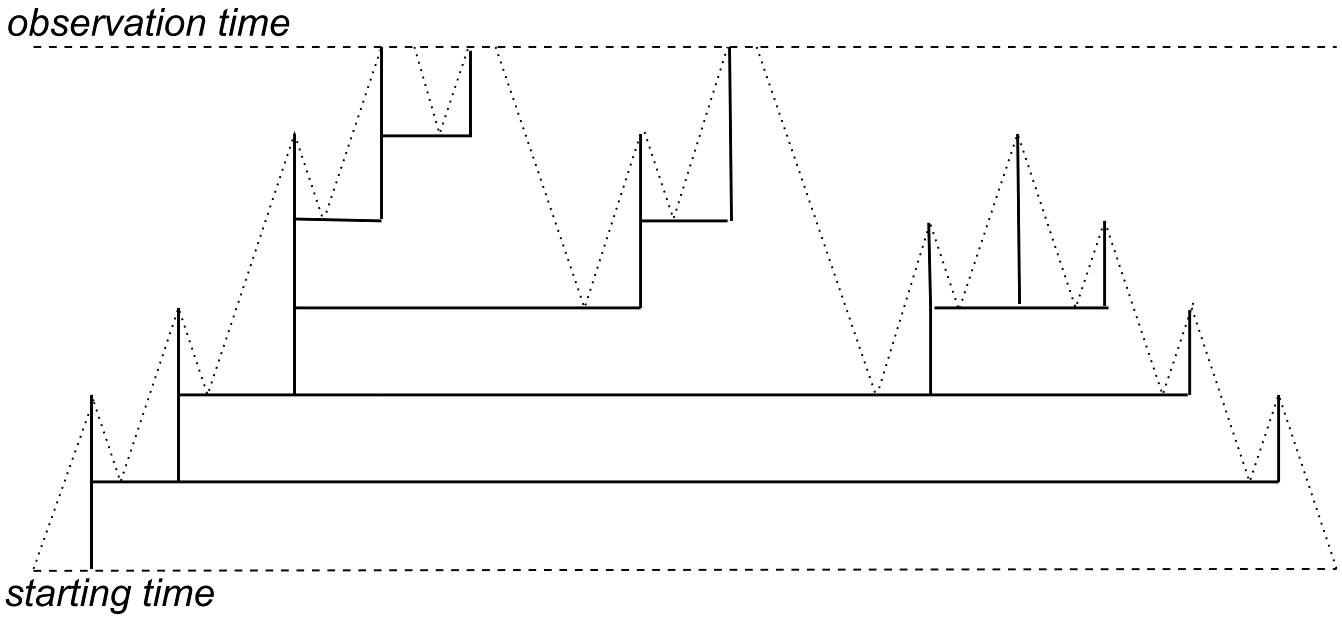

Figure 1: A tree (thick line) and the corresponding contour profile (dashed line).

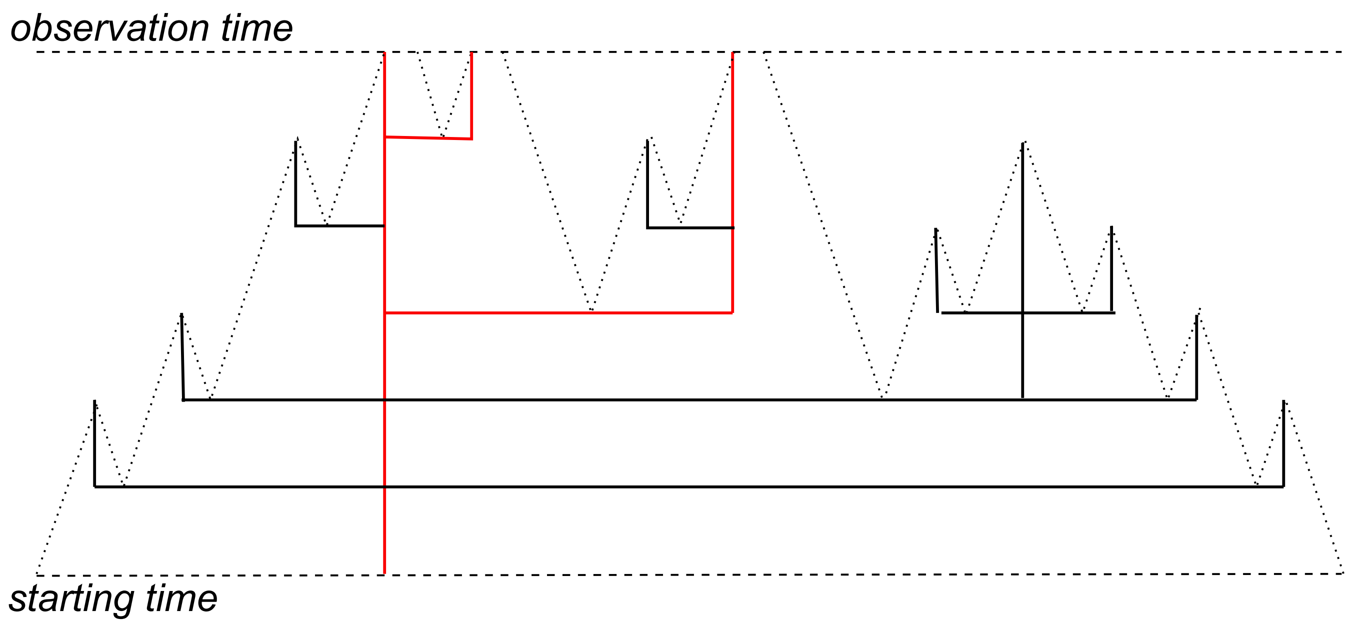

Figure 2: An alternative tree representation for the same contour profile as in Figure 1.

Contour process of a LF-multitype Galton-Watson tree

The contour process of a LF Galton-Watson tree is an alternating random walk following the following rules:

if the previous step was up of type i, move up with mark j with probability r_{ij} and move down with probability r_{i0}

after a move down go down with probability t_0 or up with mark j with probability t_j

The corresponding ARW as before moves **down** Geom$(1 - t_0)$ number of steps plus one step up marked j with prob $g_j = \frac{t_j}{1-t_0}$.

Clearly, the $\mathbf{t}-$die should not depend on the particle type to ensure that for a given level, the number of downwards excursions is distributed geometrically.

LF pgf for BPs with countably many types

We now define a LF branching process with countably many types

$$\mathbf{Z}^{(n)} = (Z_1^{(n)}, Z_2^{(n)}, \ldots)$$

in terms of a sequence of pgfs for $i \geq 1$

$$E(\mathbf{s}^{\mathbf{Z}^{(1)}} | \mathbf{Z}^{(0)} = \mathbf{e}_i) = r_{i0} + \frac{t_0 \sum_{j=1}^{\infty} r_{ij} s_j}{1 - \sum_{j=1}^{\infty} t_j s_j}$$

characterized by a sub-stochastic matrix $\mathbf{R} = (r_{ij})_{i,j=1}^{\infty}$ with

$$r_{ij} \geq 0,\ \ r_{i0} = 1 - \sum_{i=1}^{\infty} r_{ij} \in [0, 1)$$

and a vector $\mathbf{t} = (t_1, t_2, \ldots)$ with

$$t_i \geq 0,\ \ t_0 = 1 - \sum_{i=1}^{\infty} t_i \in (0, 1).$$

LF pgf for BPs with countably many types

The matrix of mean offspring numbers $\mathbf{M} = (m_{ij})_{i,j=1}^{\infty}$:

$$m_{ij} = E(Z_j^{(1)} | \mathbf{Z}^{(0)} = \mathbf{e}_i)$$

is easily computed

$$m_{ij} = r_{ij} + (1 - r_{i0}) t_j t_0^{-1}.$$

It follows, that the total mean offspring numbers are

$$m_i := m_{i1} + m_{i2} + \cdots = \frac{1 - r_{i0}}{t_0}$$

From $\mathbf{M}$ and $\mathbf{t}$ it is easy to recover $\mathbf{R}$:

$$r_{ij} = m_{ij} - m_i t_j, \; r_{i0} = 1 - m_i t_0.$$

LF PGF FOR BPS WITH COUNTABLY MANY TYPES

The expected sizes of the n-th generation $m_{ij}^{(n)} = E(Z_j^{(n)} | \mathbf{Z}^{(0)} = \mathbf{e}_i)$:

$$(m_{ij}^{(n)})_{i,j=1}^{\infty} = \mathbf{M}^n$$

For the process stems from a single newborn individual put

$$M_n = E_{\text{newborn}}(Z^{(n)} - 1 | Z^{(n)} \geq 1)$$

Looking at the spinal picture Fig 2, we see that given $Z^{(n)} \geq 1$, $Z^{(n)} = 1 + \sum$(contributions stemming from the spinal lineage)

$$\boxed{M_n = t_0^{-1}\mathbf{t}(\mathbf{I} + \mathbf{M} + \cdots + \mathbf{M}^{n-1})\mathbf{1}^{\text{tr}}}$$

Similarly, we obtain a key expression for the parameters $t_j^{(n)}$

$$(t_0^{(n)})^{-1}\mathbf{t}^{(n)} = t_0^{-1}\mathbf{t}(\mathbf{I} + \mathbf{M} + \cdots + \mathbf{M}^{n-1})$$

implying $t_0^{(n)} = \frac{1}{1+M_n}$.

The n-th generation's LF pgfs:

$$\boxed{E(\mathbf{s}^{\mathbf{Z}^{(n)}}|\mathbf{Z}^{(0)} = \mathbf{e}_i) = r_{i0}^{(n)} + \frac{t_0^{(n)} \sum_{j=1}^{\infty} r_{ij}^{(n)} s_j}{1-\sum_{j=1}^{\infty} t_j^{(n)} s_j}}$$

As in the case $n = 1$, we can put

$$m_i^{(n)} = \sum_{j=1}^{\infty} m_{ij}^{(n)}$$

and write

$$r_{i0}^{(n)} = 1 - m_i^{(n)} t_0^{(n)} \qquad \boxed{P(\mathbf{Z}^{(n)} = \mathbf{0}|\mathbf{Z}^{(0)} = \mathbf{e}_i) = r_{i0}^{(n)}}$$

$$r_{ij}^{(n)} = m_{ij}^{(n)} - m_i^{(n)} t_j^{(n)} \qquad \boxed{P(\mathbf{Z}^{(n)} \neq \mathbf{0}|\mathbf{Z}^{(0)} = \mathbf{e}_i) = \frac{m_i^{(n)}}{1+M_n}}$$

Perron-Frobenius theorem for countable matrices

If the number of types h is finite and $\mathbf{M}$ is irreducible and aperiodic, then according to the PFT there exist $0 \le \rho_0 < \rho < \infty$ such that

$$m_{ij}^{(n)} = \rho^n u_i v_j + O(\rho_0^n),\ \ i,j = 1,\ldots,h.$$

Here ρ is the max eigenvalue of $\mathbf{M}$ which is positive and single:

$$\mathbf{M}\mathbf{u}^{\mathrm{tr}} = \rho\mathbf{u}^{\mathrm{tr}},\ \ \mathbf{v}\mathbf{M} = \rho\mathbf{v},\ \ \mathbf{v}\mathbf{u}^{\mathrm{tr}} = 1,\ \ \mathbf{v}\mathbf{1}^{\mathrm{tr}} = 1.$$

Two positive eigenvectors:

$\mathbf{v}$ describes the asymptotic balance among types,
$\mathbf{u}$ compares productivity of different types: $m_i^{(n)} \sim \rho^n u_i$.

This immediately implies the classical weak convergence theorems for the the sub- ($\rho < 1$), super- ($\rho > 1$), and critical ($\rho = 1$) LF-cases. However, the key characteristics ρ, u_i, v_i are not explicitly expressed in terms of r_{ij}, t_i.

Perron-Frobenius theorem for countable matrices

PF-theory for countable matrices is given in Chapter 6 in Seneta's book [16]. Its application to the LF-case is less straightforward.

Clearly all powers $\mathbf{M}^n$ are element-wise finite. Assume again that $\mathbf{M}$ is irreducible and aperiodic. Due to Theorem 6.1 the power series

$$T_{ij}(z) = \sum_{n=0}^{\infty} m_{ij}^{(n)} z^n, \ i,j = 1,2,\ldots$$

all have common convergence radius $0 \le R < \infty$, called the convergence parameter of the matrix $\mathbf{M}$.

Furthermore, one of the two alternatives holds:

R-transient case: $\sum_{n=0}^{\infty} m_{ii}^{(n)} R^n < \infty, \ i = 1,2,\ldots$

R-recurrent case: $\sum_{n=0}^{\infty} m_{ii}^{(n)} R^n = \infty, \ i = 1,2,\ldots$

Perron-Frobenius theorem for countable matrices

According to Theorem 6.2 (and a remark afterwards) in the R-recurrent case there exist unique up to constant multipliers positive vectors $\mathbf{u}$ and $\mathbf{v}$

$$R\mathbf{M}\mathbf{u}^{\mathrm{tr}} = \mathbf{u}^{\mathrm{tr}}, \ R\mathbf{v}\mathbf{M} = \mathbf{v}$$

Form a *stochastic* matrix $Rv_j m_{ji}/v_i$ and apply results from Ch. 5.

The R-recurrent case is divided in two sub-cases:

R-null, if $\mathbf{v}\mathbf{u}^{\mathrm{tr}} = \infty$, and
R-positive, if $\mathbf{v}\mathbf{u}^{\mathrm{tr}} < \infty$.

In the R-null case (and clearly also in the R-transient case)

$$R^n m_{ij}^{(n)} \to 0, \ i, j = 1, 2, \ldots$$

In the R-positive (see Theorem 6.5) one can scale $\mathbf{v}\mathbf{u}^{\mathrm{tr}} = 1$ and

$$R^n m_{ij}^{(n)} \to u_i v_j, \;\; i, j = 1, 2, \ldots$$

The R-positive case is similar to the finite-dimensional irreducible aperiodic case with $\rho = 1/R$.

One of the few known results for BP with countably many types is Theorem 1 in [12]: in the supercritical case $R < 1$ given that

$$\sum_{i=1}^{\infty} v_i E((\mathbf{Z}_1 \mathbf{u}^{\mathrm{tr}})^2 | \mathbf{Z}_0 = \mathbf{e}_i) < \infty,$$

for any $\mathbf{w}$ such that $\mathbf{w} \le c\mathbf{u}$, convergence

$$R^n \mathbf{Z}_n \mathbf{w}^{\mathrm{tr}} \to \mathbf{v}\mathbf{w}^{\mathrm{tr}} Y$$

holds in mean square, where $Y \ge 0$ has a finite second moment.

In the framework of multitype LF processes we can define a discrete time BP with overlapping generations, where **individuals** can give birth during their lifespan and also at the moment of death.

Let

$$\mathbf{R} = \begin{pmatrix} t_1 & t_2 & t_3 & \dots & t_i & t_{i+1} & \dots \\ 1 & 0 & 0 & \dots & 0 & 0 & \dots \\ 0 & 1 & 0 & \dots & 0 & 0 & \dots \\ \vdots & \vdots & \vdots & \dots & \vdots & \vdots & \dots \\ 0 & 0 & 0 & \dots & 0 & 1 & \dots \\ \vdots & \vdots & \vdots & \dots & \vdots & \vdots & \dots \end{pmatrix}$$

Here

type i particles = individuals whose remaining age is i years,

life length L of a newborn individual: $P(L = k) = g_k = \frac{t_k}{1-t_0}$.

The special form of the mean offspring numbers

$$\mathbf{M} = \begin{pmatrix} t_1 + \frac{1-t_0}{t_0} t_1 & t_2 + \frac{1-t_0}{t_0} t_2 & t_3 + \frac{1-t_0}{t_0} t_3 & \dots \\ 1 + \frac{t_1}{t_0} & \frac{t_2}{t_0} & \frac{t_3}{t_0} & \dots \\ \frac{t_1}{t_0} & 1 + \frac{t_2}{t_0} & \frac{t_3}{t_0} & \dots \\ \frac{t_1}{t_0} & \frac{t_2}{t_0} & 1 + \frac{t_3}{t_0} & \dots \\ \vdots & \vdots & \vdots & \dots \end{pmatrix} = m\mathbf{G} + \mathbf{D}$$

where $m = (1 - t_0)/t_0$ is the average litter size and

$$\mathbf{G} = \begin{pmatrix} g_1 & g_2 & g_3 & \dots \\ g_1 & g_2 & g_3 & \dots \\ g_1 & g_2 & g_3 & \dots \\ g_1 & g_2 & g_3 & \dots \\ \vdots & \vdots & \vdots & \dots \end{pmatrix}, \quad \mathbf{D} = \begin{pmatrix} 0 & 0 & 0 & \dots \\ 1 & 0 & 0 & \dots \\ 0 & 1 & 0 & \dots \\ 0 & 0 & 1 & \dots \\ \vdots & \vdots & \vdots & \dots \end{pmatrix}.$$

We are able to compute $\mathbf{M}^n$ in terms of the representation

$$\mathbf{M}^n = \sum_{k=0}^{n} m^k \Phi_{k,n}$$

where $\Phi_{0,n} = \mathbf{D}^n$, $\Phi_{n,n} = \mathbf{G}$ and

$$\Phi_{k,n} = \mathbf{G}\Phi_{k-1,n-1} + \mathbf{D}\Phi_{k,n-1},\ 1 \le k \le n-1.$$

Introduce the generating function of the tail probabilities

$\boxed{f(s) = \sum_{k=1}^{\infty} q_k s^k}$ where $q_k = P(L \ge k) = \sum_{i=k}^{\infty} g_i$.

If R be a positive solution of $mf(R) = 1$, then for all $s \in [0, R)$

$$\sum_{n=0}^{\infty} s^n \mathbf{a}\mathbf{M}^n \mathbf{b}^{\mathrm{tr}} = \frac{\psi_a(s)\psi_b(s)}{(1 - mf(s))f(s)}$$

$$\psi_a(s) = \sum_{i=1}^{\infty} a_i(1 + s + \ldots + s^{i-1}), \quad \psi_b(s) = \sum_{i=1}^{\infty} s^i \sum_{k=1}^{\infty} q_{i+k-1} b_k$$

In particular, with $(\mathbf{a} = \mathbf{e}_i, \mathbf{b} = \mathbf{1})$ and $(\mathbf{a} = \mathbf{g}, \mathbf{b} = \mathbf{1})$ we get

$$\sum_{n=0}^{\infty} m_i^{(n)} s^n = \frac{1 + s + \ldots + s^{i-1}}{1 - mf(s)}$$

$$\sum_{n=0}^{\infty} \sum_{i=1}^{\infty} g_i m_i^{(n)} s^n = \frac{f(s)}{s(1 - mf(s))}$$

The asymptotics of $M_n = t_0^{-1} \sum_{k=0}^{n-1} \sum_{i=1}^{\infty} t_i m_i^{(k)}$ will follow from

$$t_0^{-1} \sum_{n=0}^{\infty} \sum_{i=1}^{\infty} t_i m_i^{(n)} s^n = \frac{mf(s)}{s(1 - mf(s))}$$

The mean life length $E(L)$, finite or infinite, is

$$\lambda = \sum_{k=1}^{\infty} k g_k = \sum_{i=1}^{\infty} q_i.$$

The mean total offspring number per individual is $\mu = m\lambda$

In view of

$$\sum_{n=0}^{\infty} m_1^{(n)} s^n = \frac{1}{1 - mf(s)}$$

the function $m_1^{(n)}$ can be treated as a renewal function so that

$$m_1^{(n)} \sim \rho^n \beta^{-1}, \; n \to \infty,$$

where ρ is defined by $mf(\rho^{-1}) = 1$ and $\beta = m \sum_{i=1}^{\infty} i q_i \rho^{1-i}$.

Rearranging $mf(\rho^{-1}) = 1$ as

$$\sum_{i=1}^{\infty} mq_i \rho^{-i} = 1$$

and referring to the so called Malthusian parameter $\alpha = \ln \rho$ we obtain the familiar formula defining α for a CMJ process (cf [5]):

$$\sum_{i=1}^{\infty} m_i e^{-\alpha i} = 1$$

where $m_i = mq_i$ is the mean litter number of a individual at age i.

$\beta =$ the average age at childbearing.

If $\mu = 1$, then $m = \lambda^{-1}$, $\alpha = 0$ and $\beta = \lambda^{-1} \sum_{i=1}^{\infty} iq_i$.

In the critical case

$$\sum_{n=0}^{\infty} m_1^{(n)} s^n = \frac{1}{1 - \lambda^{-1} f(s)}$$

and if furthermore $E(L^2) < \infty$, then $\beta = \frac{E(L^2)+\lambda}{2\lambda}$ is finite.

A critical LF CMJ process with $\beta = \infty$ is an R-null Markov chain

Assume that for a slowly varying at zero function $\mathcal{L}$

$$E(s^L) = 1 + \lambda(s-1) + (1-s)^{1+\alpha}\mathcal{L}(1-s),\ \ \alpha \in (0,1]$$

It follows,

$$1 - \lambda^{-1} f(s) \sim (1-s)^{\alpha}\mathcal{L}(1-s), \quad s \to 1-$$

A Tauberian theorem and the monotone density theorem imply

$$M_n \sim (\Gamma(1+\alpha))^{-1} n^{\alpha} \mathcal{L}^{-1}(n^{-1}), \quad n \to \infty$$

$$m_1^{(n)} \sim \alpha(\Gamma(1+\alpha))^{-1} n^{\alpha} \mathcal{L}^{-1}(n^{-1}), \; n \to \infty$$

In view of

$$P(\mathbf{Z}^{(n)} \neq \mathbf{0} | \mathbf{Z}^{(0)} = \mathbf{e}_i) = \tfrac{m_i^{(n)}}{1+M_n},$$

we derive

$$P(\mathbf{Z}^{(n)} \neq \mathbf{0} | \mathbf{Z}^{(0)} = \mathbf{e}_1) \sim \tfrac{\alpha}{n}, \; n \to \infty$$

and furthermore, due to $m_i^{(n)} = m_1^{(n-1)} + \ldots + m_1^{(n-i)}$

$$\boxed{P(\mathbf{Z}^{(n)} \neq \mathbf{0} | \mathbf{Z}^{(0)} = \mathbf{e}_i) \sim \tfrac{\alpha i}{n}, \; n \to \infty}$$

A version of the last result was first noticed by Borovkov and Vatutin (cf Th 3 in [3]) in their paper on the maxima in critical Galton-Watson processes.

Their proof is based on the Pakes paper [14] devoted to the maximum of a left-continuous random walk. The latter can be interpreted as the contour process for our LF CMJ process.

Final remarks

The general LF branching process with countably many types can be treated as a LF CMJ process with the life length having a phase-type distribution (see for example Asmussen [1]).

Making births very rare events we can approximate the LF CMJ processes by the continuous time CMJ processes with Poisson reproduction, recently studied by Lambert [10] and coauthors [11].

References

[1] ASMUSSEN, S. (2003) *Applied probability and queues*, Springer, New York.

[2] ATHREYA, K. AND NEY, P. (1972) *Branching processes*, John Wiley & Sons, London-New York-Sydney.

[3] BOROVKOV, K. AND VATUTIN, V. (1996) On distribution tails and expectations of maxima in critical branching processes. *J. Appl. Prob.* 33, 614–622

[4] JAGERS P. (1975) *Branching processes with Biological Applications*, John Wiley & Sons, London-New York-Sydney.

[5] JAGERS, P. AND SAGITOV, S. (2008) General branching processes in discrete time as random trees. *Bernoulli* 14, 949–962.

[6] JOFFE A. AND G. LETAC Multitype linear fractional branching processes, J. Appl. Probab. 43 (4) (2006), pp. 1091–1106.

[7] Hoppe FM. (1997) Coupling and the Non-degeneracy of the Limit in Some Plasmid Reproduction Models. Theor Popul Biol. 52(1):27-31.

[8] Kesten, H. (1989) Supercritical branching processes with countably many types and the sizes of random cantor sets. In T.W. Anderson, K.B. Athreya and D. Iglehart (eds), Probability,Statistics and Mathematics Papers in Honor of Samuel Karlin, pp. 108-121. New York:Academic Press.

[9] Kimmel, M. and Axelrod, D. (2002) *Branching Processes in Biology*, Springer, New York.

[10] Lambert, A. (2010) The contour of splitting trees is a Levy process. Ann. Probab. 38, 348–395.

[11] Lambert, A., Simatos, F., and Zwart, B. (2011) Scaling limits via excursion theory: Interplay between Crump-Mode-Jagers branching processes and Processor-Sharing queues. Preprint: arXiv:1102.5620v1

[12] Moy,S.-T. C. (1967) Extensions of a limit theorem of Everett Ulam and Harrison multi-type branching processes to a branching process with countably many types. Ann. Math. Statist., 38. 992-999.

References

[13] Nee, S., May, R. M., and Harvey, P. H. (1994) The reconstructed evolutionary process. Philos. Trans. Roy. Soc. London, B. Biol. Sci. 344, 1309, 305–311.

[14] Pakes, A. (1978) On the maximum and absorption time of a left-continuous random walk. *J. Appl. Prob.* 15, 292–299.

[15] Pollak, E. (1974) Survival probabilities and extinction times for some multitype branching processes. Adv. Appl. Prob. 6, 446–462.

[16] Seneta, E. (2006). *Non-negative matrices and Markov chains*, Springer Series in Statistics No. 21, Springer, New-York.

[17] Seneta E, Tavare S. (1983) Some stochastic models for plasmid copy number, Theoretical Population Biology, 23, 241–256.

THANK YOU FOR THE ATTENTION!