Reduced critical processes for small populations

Vladimir Vatutin (Steklov Mathematical Institute)

Minzhi Liu (Bejing Normal University)

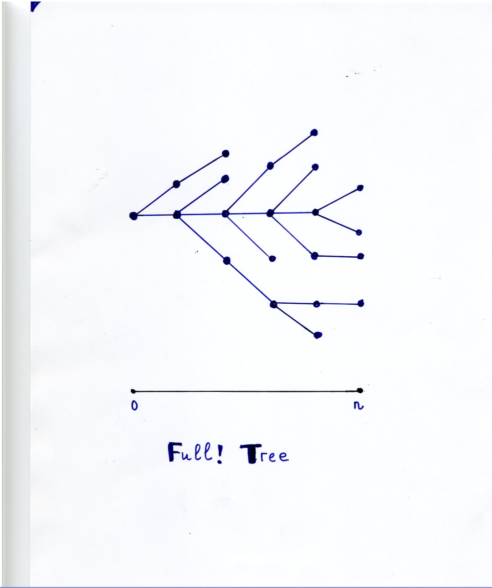
Badajoz, April, 2018

Notation:

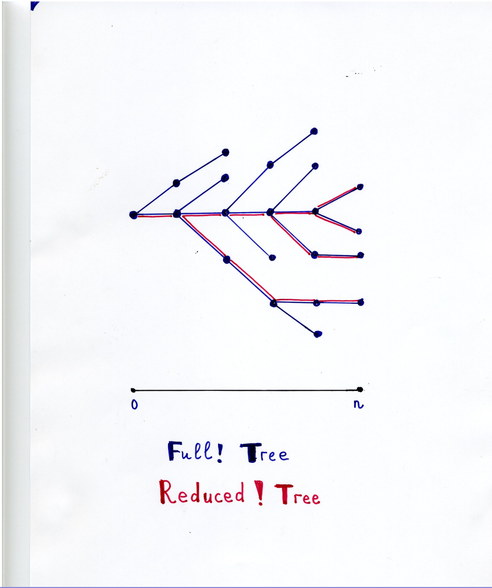
- Let $\{Z(n), n \geq 0\}$ be a Galton-Watson branching process with $Z(0) = 1$;
- $$f(s) = \mathbf{E}s^{\xi} = \sum_{k=0}^{\infty} f_k s^k$$

  be the offspring generating function of the process;
- $Z(m, n)$ be the number of particles in the process at moment $m \leq n$ having a positive number of descendants at moment n.
- The process $\{Z(m, n), 0 \leq m \leq n\}$ is called a reduced process.

Genealogical tree for a branching process



Genealogical tree for the respective reduced process



- **Fleischmann and Prehn (1974)** - the subcritical reduced processes.
- **Zubkov (1975)** - the distance to the most recent common ancestor (MRCA) for the supercritical Galton–Watson processes and for the critical processes with possibly infinite variance of the offspring size.
- **Fleischmann and Siegmund-Schultze (1977)** - a functional conditional limit theorem on the convergence of the reduced critical Galton–Watson branching process to the Yule processes.

- Different questions related to the problem of the distribution of the MRCA for the k particles selected at random among the $Z(n) \geq k$ particles existing in the population at moment n were considered, for instance, by **K.Athreya, O'Connell, Durrett, Lambert, S.C. Harris, Johnston, Roberts, Moelle, Sagitov**.
- Properties of the GW processes given that the total amount of particles in the process is fixed or belongs to some sets were considered by **Dwass, Kolchin... and many others, including Abraham and Delmas.**

Abraham and Delmas (2014): the structure of the tree given that the number of particles in generation n is fixed.

- A Kesten tree $\mathcal{T}^*$ is an infinite random tree with vertices labeled by tuples $V_1 V_2 \cdots V_n, n \in \mathbb{N}$ (the size-biased tree) whose distribution is as follows.
- There exists a **unique infinite random sequence** $(V_l; l \in \mathbb{N})$ of positive integers such that, for every $h \in \mathbb{N}$, $V_1 V_2 \cdots V_h \in \mathcal{T}^*$, with the convention that $V_1 V_2 \cdots V_h = \oslash$ if $h = 0$.

The joint distribution of $(V_l; l \in \mathbb{N})$ and $\mathcal{T}^*$ is determined recursively as follows:

- for each $h \in \mathbb{N}$, given $V_1 V_2 \cdots V_h$ and $\mathcal{T}_h^*$ (restriction of $\mathcal{T}^*$ to the first h generations) we have:
- The number of children of all particles of generation h are independent and distributed according to

$$f(s) = \mathbf{E} s^{\xi} = \sum_{j=0}^{\infty} f_j s^j$$

if $V_1 V_2 \cdots V_h \notin \mathcal{T}^*$ and according to

$$f'(s) = \mathbf{E}\left[\xi s^{\xi - 1}\right]$$

if $V_1 V_2 \cdots V_h \in \mathcal{T}^*$.

- Given also the total number M_h of children of all particles of the $h-$th generation, the integer V_{h+1} is uniformly distributed on the set of integers $1, 2, ..., M_h$.

Let $\mathcal{T}$ be the genealogical tree of an ordinary GW process.

Theorem

Assume that $f_0 + f_1 < 1$ *and* $\mathbf{E}\xi = 1$. *Let* $\{\alpha_n, n = 1, 2, ...\}$ *be a sequence of positive integers tending to infinity such that, for any* $j = 0, 1, ...$

$$\lim_{n\to\infty} \frac{\mathbf{P}\left(Z(n-j) = \alpha_n\right)}{\mathbf{P}\left(Z(n) = \alpha_n\right)} = 1.$$

Then

$$\mathcal{L}\left(\mathcal{T}\ |Z(n) = \alpha_n\right) \to \mathcal{L}\left(\mathcal{T}^*\right).$$

Our result is, in a sense a commentary to this paper and to the results due to **Fleischmann+Prehn and Zubkov** and concerns the reduced trees.
We know that if

$$\mathbf{E}\xi = 1, \quad 2B := Var\xi \in (0, \infty),$$

then

$$Q(n) := \mathbf{P}(Z(n) > 0) \sim \frac{1}{Bn} \quad \text{as} \quad n \to \infty$$

and, for any $y \geq 0$

$$\lim_{n\to\infty} \mathbf{P}\left(\frac{Z(n)}{Bn} \leq y | Z(n) > 0\right) = 1 - e^{-y}.$$

In addition **(Fleischmann+Prehn)**, for any fixed $t \in [0, 1)$ and all $s \in [0, 1]$

$$\lim_{n\to\infty} \mathbf{E}\left[s^{Z(nt,n)} | Z(n) > 0\right] = s\frac{1-t}{1-ts}.$$

We study the asymptotic properties of the reduced process when the condition $\{Z(n) > 0\}$ is replaced either

- by the assumption that $\{0 < Z(n) \leq B\varphi(n)\}$ for a function $\varphi(n) = o(n)$ as $n \to \infty$
- or by the assumption that $\{0 < Z(n) \leq aBn\}$ for some $a > 0$.

Theorem

If g.c.d. $\{k : f_k > 0\} = 1$,

$$\mathbf{E}\xi = 1, \quad 2B := Var\xi \in (0, \infty),$$

and $\varphi(n) \to \infty$ *in such a way that* $\varphi(n) = o(n)$, *then for any* $x \in (0, \infty)$

$$\lim_{n\to\infty} \mathbf{E}\left[s^{Z(n-x\varphi(n),n)} \middle| 0 < Z(n) \le B\varphi(n) \right] = sx\frac{1 - e^{-(1-s)/x}}{1-s}.$$

Let

$$\beta(n) := \max\left(0 \leq m < n : Z(m,n) = 1\right)$$

be the birth moment of the MRCA of all particles existing in the population at moment n and let $d(n) := n - \beta(n)$ be the distance from the point of observation n to the birth moment of the MRCA.

Corollary

Under the basic conditions

$$\lim_{n\to\infty} \mathbf{P}\left(d(n) \leq x\varphi(n) | 0 < Z(n) \leq B\varphi(n)\right) = x\left(1 - e^{-1/x}\right).$$

Proof of the corollary. Let

$$\mathcal{H}(n) := \{0 < Z(n) \leq B\varphi(n)\}.$$

Then

$$\begin{aligned}
&\lim_{n\to\infty} \mathbf{P}\left(d(n) \leq x\varphi(n) | \mathcal{H}(n)\right) \\
&\quad = \lim_{n\to\infty} \mathbf{P}\left(Z(n - x\varphi(n), n) = 1 | \mathcal{H}(n)\right) \\
&\qquad = coeff_s \left[sx \frac{1 - e^{-(1-s)/x}}{1-s} \right] = x\left(1 - e^{-1/x}\right).
\end{aligned}$$

Note that

$$\lim_{x\to\infty} x\left(1 - e^{-1/x}\right) = 1 \text{ and } \lim_{x\to 0} x\left(1 - e^{-1/x}\right) = 0.$$

Theorem

If g.c.d. $\{k : f_k > 0\} = 1$ *and the basic conditions are valid, then, for any fixed* $t \in [0, 1)$ *and any* $a > 0$

$$\lim_{n\to\infty} \mathbf{E}\left[s^{Z(nt,n)}\middle|0 < Z(n) \le aBn\right] = s\frac{1-t}{1-ts}\frac{1-e^{-(1-ts)a/(1-t)}}{1-e^{-a}}.$$

Corollary

Under the basic conditions

$$\lim_{n\to\infty} \mathbf{P}\left(d(n) \le nt|0 < Z(n) \le aBn\right) = t\frac{1-e^{-a/t}}{1-e^{-a}}.$$

Observe that **(Zubkov)**, for $0 < t < 1$

$$\lim_{n\to\infty} \mathbf{P}\left(d(n) \le nt|0 < Z(n)\right) = t.$$

Proof of the corollary

As before:

$$\lim_{n\to\infty} \mathbf{P}\left(d(n) \le tn | 0 < Z(n) \le aBn\right)$$
$$= \lim_{n\to\infty} \mathbf{P}\left(Z(n(1-t), n) = 1 | 0 < Z(n) \le aBn\right)$$
$$= coeff_s \left[s \frac{t}{1-(1-t)\,s} \frac{1-e^{-(1-(1-t)s)a/t}}{1-e^{-a}} \right] = t\frac{1-e^{-a/t}}{1-e^{-a}}.$$

Basic tool:

Let

$$\mathcal{H}(n) := \{0 < Z(n) \leq B\varphi(n)\}.$$

We have

$$\mathbf{P}\left(Z(n - x\varphi(n), n) = j | \mathcal{H}(n)\right)$$

$$= \frac{\mathbf{P}\left(Z(n - x\varphi(n), n) = j\right) \times \mathbf{P}\left(\mathcal{H}(n) | Z(n - x\varphi(n), n) = j\right)}{\mathbf{P}\left(\mathcal{H}(n)\right)}.$$

Nagaev, Wachtel (2006): if the basic conditions are valid and $k, n \to \infty$ in such a way that the ratio k/n remains bounded then

$$\lim_{n\to\infty} n^2 B^2 \left(1 + \frac{1}{Bn}\right)^{k+1} \mathbf{P}\left(Z(n) = k | Z(0) = 1\right) = 1.$$

Therefore, given $k/n \to 0$

$$\begin{aligned} \mathbf{P}\left(\mathcal{H}(n) | Z(0) = 1\right) &= \sum_{1\le k\le B\varphi(n)} \mathbf{P}\left(Z(n) = k | Z(0) = 1\right) \\ &\sim \frac{1}{n^2 B^2} \sum_{1\le k\le B\varphi(n)} 1 \sim \frac{\varphi(n)}{n^2 B}. \end{aligned}$$

Denote $f_n(s)$ the nth iteration of $f(s)$ with itself. Then

$$\mathbf{P}\left(Z(n - x\phi(n), n) = j\right) = \sum_{k=j}^{\infty} \mathbf{P}\left(Z(n - x\phi(n)) = k; Z(n - x\phi(n), n) = j\right)$$

$$= \sum_{k=j}^{\infty} \mathbf{P}\left(Z(n - x\phi(n)) = k\right) C_k^j f_{x\phi(n)}^{k-j}(0) \left(1 - f_{x\phi(n)}(0)\right)^j$$

$$= \frac{\left(1 - f_{x\phi(n)}(0)\right)^j}{j!} f_{n-x\phi(n)}^{(j)}(f_{x\phi(n)}(0)).$$

Now

$$\lim_{n\to\infty} n^2 \left[f_{n+1}(0) - f_n(0)\right] = \frac{1}{B}.$$

We consider for $\lambda > 0$ the function

$$f_m(f^{\lambda}_{x\varphi(n)}(0)) = f_m(e^{\lambda \log f_{x\varphi(n)}(0)})$$

and find r such that

$$1 - f_{r+1}(0) < 1 - f^{\lambda}_{x\varphi(n)}(0) \leq 1 - f_r(0).$$

We know that

$$1 - f^{\lambda}_{x\varphi(n)}(0) \sim \lambda\left(1 - f_{x\varphi(n)}(0)\right) \sim \frac{\lambda}{Bx\varphi(n)}.$$

Hence we get

$$r \sim \frac{x\varphi(n)}{\lambda} = o(n) \text{ as } n \to \infty.$$

Hence we get

$$r \sim \frac{x\varphi(n)}{\lambda} = o(n) \text{ as } n \to \infty.$$

Then for $n - m = x\varphi(n)$

$$\begin{aligned}
& \lim_{n\to\infty} \frac{n^2}{x\varphi(n)} \left[f_m(f_r(0)) - f_m(0)\right] \\
= & \lim_{n\to\infty} \frac{1}{x\varphi(n)} \sum_{k=0}^{r-1} n^2 \left[f_m(f_{k+1}(0)) - f_m(f_k(0))\right] \\
= & \lim_{n\to\infty} \frac{1}{x\varphi(n)} \sum_{k=0}^{r-1} \frac{n^2}{(m+k)^2} (m+k)^2 \left[f_{m+k+1}(0) - f_{m+k}(0)\right] \\
= & \frac{1}{B} \lim_{n\to\infty} \frac{1}{x\varphi(n)} \sum_{k=0}^{r-1} 1 = \frac{1}{B}\frac{1}{\lambda}.
\end{aligned}$$

Thus,

$$\lim_{n\to\infty} \frac{n^2}{x\varphi(n)} \left[f_m(e^{\lambda \log f_{x\varphi(n)}(0)}) - f_m(0) \right] = \frac{1}{B}\frac{1}{\lambda}, \quad \lambda > 0.$$

Clearly, the prelimiting and limiting functions are analytical in the complex semi-plane $\text{Re } \lambda > 0$.

Therefore, the derivatives of any order of the prelimiting functions converge to the respective derivatives of the limiting function for each λ with $\text{Re } \lambda > 0$.

Thus, for each $j \geq 1$

$$\lim_{n\to\infty} \frac{Bn^2}{x\varphi(n)} \frac{d^j}{d\lambda^j} \left[f_m(e^{\lambda \log f_{x\varphi(n)}(0)}) \right] = (-1)^j \frac{j!}{\lambda^{j+1}}.$$

Set $I_j := i_1 + \cdots + i_j$ and

$$\mathcal{D}(j) := \{(i_1, ..., i_j) : 1 \cdot i_1 + 2 \cdot i_2 + \cdots + j i_j = j\},$$

By Faà di Bruno's formula we have

$$\begin{aligned}
&\frac{d^j}{d\lambda^j}\left[f_m(e^{\lambda \log f_{x\varphi(n)}(0)})\right] \\
&= \sum_{\mathcal{D}(j)} \frac{j!}{i_1! \cdots i_j!} f_m^{(I_j)}(e^{\lambda \log f_{x\varphi(n)}(0)}) \prod_{r=1}^{j} \left(\left(\frac{e^{\lambda \log f_{x\varphi(n)}(0)}}{r!}\right)^{(r)}\right)^{i_r} \\
&= \sum_{\mathcal{D}(j)} \frac{j!}{i_1! \cdots i_j!} f_m^{(I_j)}(e^{\lambda \log f_{x\varphi(n)}(0)}) e^{\lambda I_j \log f_{x\varphi(n)}(0)} \prod_{r=1}^{j} \frac{(\log f_{x\varphi(n)}(0))^{r i_r}}{(r!)^{i_r}} \\
&= (\log f_{x\varphi(n)}(0))^j \sum_{\mathcal{D}(j)} \frac{j!}{i_1! \cdots i_j!} f_m^{(I_j)}(e^{\lambda \log f_{x\varphi(n)}(0)}) e^{\lambda I_j \log f_{x\varphi(n)}(0)} \prod_{r=1}^{j} \left(\frac{1}{r!}\right)^{i_r}.
\end{aligned}$$

One can show by induction for $m = n - x\varphi(n)$ that

$$
\begin{aligned}
(-1)^j\, j! &\sim \frac{Bn^2}{x\varphi(n)} \frac{d^j}{d\lambda^j} \left[f_m(e^{\lambda \log f_{x\varphi(n)}(0)}) \right] |_{\lambda=1} \\
\sim \quad & (-1)^j \sum_{\mathcal{D}(j)} \frac{j!}{i_1! i_2! \cdots i_j!} \frac{B^2 n^2}{(Bx\varphi(n))^{j+1}} f_m^{(I_j)}(f_{x\varphi(n)}(0)) \prod_{r=1}^{j} \frac{1}{(r!)^{i_r}} \\
\sim \quad & (-1)^j \frac{j!}{j! 0! \cdots 0!} \frac{B^2 n^2}{(Bx\varphi(n))^{j+1}} f_m^{(j)}(f_{x\varphi(n)}(0)) \\
= \quad & (-1)^j \frac{B^2 n^2}{(Bx\varphi(n))^{j+1}} f_m^{(j)}(f_{x\varphi(n)}(0)).
\end{aligned}
$$

This representation and previous results give

$$\begin{aligned}\mathbf{P}\left(Z(n-x\varphi(n),n)=j\right) &= \frac{\left(1-f_{x\varphi(n)}(0)\right)^{j}}{j!} f^{(j)}_{n-x\varphi(n)}\left(f_{x\varphi(n)}(0)\right) \\ &\sim \frac{1}{j!\,(xB\varphi(n))^{j}} \frac{j!\,(xB\varphi(n))^{j+1}}{B^2n^2} \sim \frac{x\varphi(n)}{Bn^2}.\end{aligned}$$

Let now $Z_1^*(m), \ldots, Z_j^*(m)$ be i.i.d. random variables distributed as $\{Z(m) | Z(m) > 0\}$, and let $\eta_1, \ldots, \eta_j$ be i.i.d. random variables having exponential distribution with parameter 1. Then

$$\begin{aligned}
& \lim_{n\to\infty} \mathbf{P}\left(0 < Z(n) \leq B\phi(n) | Z(n - x\varphi(n), n) = j\right) \\
= & \lim_{n\to\infty} \mathbf{P}\left(Z_1^*(x\varphi(n)) + \cdots + Z_j^*(x\varphi(n)) \leq B\varphi(n)\right) \\
= & \lim_{n\to\infty} \mathbf{P}\left(\frac{Z_1^*(x\varphi(n))}{Bx\varphi(n)} + \cdots + \frac{Z_j^*(x\varphi(n))}{Bx\varphi(n)} \leq \frac{1}{x}\right) \\
= & \mathbf{P}\left(\eta_1 + \cdots + \eta_j \leq \frac{1}{x}\right) = \frac{1}{(j-1)!}\int_0^{1/x} z^{j-1} e^{-z} dz.
\end{aligned}$$

As a result

$$\begin{aligned}
\lim_{n\to\infty} \mathbf{E}\left[s^{Z(n-x\varphi(n),n)}|\mathcal{H}(n)\right] &= \sum_{j=1}^{\infty} \lim_{n\to\infty} \mathbf{P}\left(Z(n-x\varphi(n),n)=j|\mathcal{H}(n)\right)s^j \\
&= \sum_{j=1}^{\infty} \frac{x}{(j-1)!}\int_0^{1/x} s^j z^{j-1} e^{-z} dz \\
&= xs\int_0^{1/x} e^{(s-1)z} dz = \frac{xs}{1-s}\left(1-e^{-(1-s)/x}\right).
\end{aligned}$$

Generalizations:

- **infinite variance for the offspring number**??? :

  there is NO local limit theorem for $Z(n) = j$ for large j (!)

- Age-dependent processes, that is the processes with $G(t) = \mathbf{P}(\tau \leq t)$ being the life-length distribution of the particles of the process and $\mu = \mathbf{E}\tau$.

It is known (**V., 1976, 1979**) that if $\mathbf{E}\xi = 1, \quad 2B := Var\xi \in (0, \infty)$, and

$$1 - G(t) = \mathbf{P}(\tau > t) \sim \frac{C}{t^{\gamma}}$$

then, for $\gamma \in (0, 2)$

$$\lim_{t \to \infty} \mathbf{E}\left[s^{Z(t)} | Z(t) > 0\right] = 1 - \sqrt{1 - s},$$

for $\gamma > 2$

$$\lim_{t \to \infty} \mathbf{P}\left(\frac{\mu}{Bt} Z(t) \leq x \,|\, Z(t) > 0\right) = 1 - e^{-x},$$

If $\gamma = 2$ then
a) for all $s \in [0, 1)$

$$\lim_{t\to\infty} \mathbf{E}\left[s^{Z(t)} | Z(t) > 0\right] = 1 - \frac{\mu + \sqrt{\mu^2 + 4C(1-s)}}{\mu + \sqrt{\mu^2 + 4C}};$$

b) for any $x > 0$

$$\lim_{t\to\infty} \mathbf{P}\left(\frac{\mu}{Bt} Z(t) \leq x \,|\, Z(t) > 0\right) = 1 - r + r(1 - e^{-x}).$$

with

$$r = \frac{2\mu}{\mu + \sqrt{\mu^2 + 4C}}.$$

R. Abraham and J.-F. Delmas. (2014) Local limits of conditioned Galton-Watson trees: the infinite spine case. *Elec. J. of Probab.*, 19:1–19,.

R. Abraham and J.-F. Delmas. (2015) An introduction to Galton-Watson trees and their local limits. arXiv:1506.05571.

R. Abraham, A. Bouaziz, J.-F. Delmas. (2015) Local limits of Galton-Watson trees conditioned on the number of protected nodes. arXiv:1509.02350.

Athreya, K. B. (2012) Coalescence in the recent past in rapidly growing populations. *Stochastic Processes and their Applications*, **122**, 3757–3766.

Athreya, K. B. (2012) Coalescence in critical and subcritical Galton-Watson branching processes. *Journal of Applied Probability*, **49**, 627–638.

Durrett, R. (1978) The genealogy of critical branching processes. *Stochastic Processes and their Applications*, **8**, 101–116.

Fleischmann, K., Prehn, U. (1974) Ein Grenzfersatz für subkritische Verzweigungsprozesse mit eindlich vielen Typen von Teilchen. *Math. Nachr.*, **64**, 233–241.

Fleischmann, K., Siegmund-Schultze, R. (1977) The structure of reduced critical Galton-Watson processes. *Math. Nachr.*, **79**, 233–241.

Harris, S. C., Johnston, S. G. G., and Roberts, M. I. (2017) The coalescent structure of continuous-time Galton-Watson trees. https://arxiv.org/pdf/1703.00299.pdf

Johnston, S. G. G. (2017) Coalescence in supercritical and subcritical continuous-time Galton-Watson trees. https://arxiv.org/pdf/1709.008500v1.pdf

Lambert, A. (2003) Coalescence times for the branching process. *Advances in Applied Probability*, **35**, 1071–1089.

Lambert, A. (2016) Probabilistic models for the subtrees of life. https://arxiv.org/abs/1603.03705

Le, V. (2014) Coalescence times for the Bienaymé-Galton-Watson process. *Journal of Applied Probability,* **51**, 209–218.

Nagaev, S. V. and Vakhtel, V. I. (2006) On the local limit theorem for a critical Galton–Watson process. *Theory Probab. Appl.*, **50**, 400–419.

O'Connell, N. (1995) The genealogy of branching processes and the age of our most recent common ancestor. *Advances in Applied Probability*, **27**, 418–442.

Vatutin, V. A. and D'yakonova, E. E. (2015) Decomposable branching processes with a fixed extinction moment. *Proc. Steklov Inst. Math.*, **290**, 103–124.

Zubkov, A. M. (1975) Limit distributions of the distance to the nearest common ancestor. *Theory Probab. Appl.*, **20**, 602–612.