Mathematical Statistics
Stockholm University

Genealogy for supercritical branching processes

Andreas Nordvall Lagerås
Anders Martin-Löf

Research Report 2006:2

ISSN 1650-0377

Postal address:

Mathematical Statistics
Dept. of Mathematics
Stockholm University
SE-106 91 Stockholm
Sweden

Internet:

http://www.math.su.se/matstat

Mathematical Statistics
Stockholm University
Research Report **2006:2**,
http://www.math.su.se/matstat

Genealogy for supercritical branching processes

Andreas Nordvall Lagerås*
Anders Martin-Löf

March 2006

Abstract

We study so-called immortal branching processes, i.e. branching processes where each individual upon death is replaced by at least one new individual. All supercritical branching processes, conditioned to explode, contain a sub-tree of individuals who have an infinite line of the descent. Those individuals form an immortal branching process.

By studying the genealogy of an immortal branching process we can conclude that the distribution of the number of individuals at any given time, minus the ancestor, has a compound geometric distribution. The result implies that also the limiting distribution of a properly scaled supercritical branching process has a compound geometric distribution.

Marginal distributions for the size of a branching process are generally hard to find, but we find an explicit expression for the marginal distribution for a class of branching processes that have recently appeared in the theory of coalescent processes and continuous stable random trees. The limiting distribution can be expressed in terms of the Fox H-function, and in special cases by the Meijer G-function.

Keywords: Supercritical branching process, Compound geometric distribution, Self-decomposability, Fox H-function, Meijer G-function.

MSC Primary: 60J80
MSC Secondary: 60E07, 60F05, 33C60

*Postal address: Mathematical Statistics, Stockholm University, SE-106 91, Stockholm, Sweden. E-mail: andreas@math.su.se

1 Introduction

A branching process is a Markov process in continuous time. Heuristically it can be considered as the number of particles in a population where the particles behave independent of each other, live for exponentially distributed periods of time and at death give birth to new particles according to some distribution on the non-negative integers. Usually the population is assumed to start with one particle. The name branching process is appropriate since one can describe the evolution of the population by drawing a family tree in which the lifetime of each particle corresponds to a particular branch.

We shall mainly consider branching processes where each particle give birth to at least two new particles. We call such processes immortal branching processes as they can be seen as the number of particles in a population where no particle dies, but instead repeatedly gives birth to new offspring after exponential periods of time, independently of each other.

Our main result is that the number of individuals in an immortal branching processes have compound geometric distributions. As a corollary we obtain that the limiting distribution of properly scaled supercritical branching processes are compound geometric.

2 Branching processes

We can describe the dynamics of a branching process $Z = \{Z_t\}_{t\geq 0}$, where the number of new particles at each birth has the distribution $\{p_k\}_{k\geq 0}$, as follows (Athreya and Ney [1]): The process starts with $Z_0 = 1$. If the process is in state i at any time, it continues there for an amount of time which is exponentially distributed with parameter $i\mu$, where μ is called the intensity of the process, and then jumps into state $j \geq i-1$ with probability p_{j-i+1}. It then stays in state j for an exponentially distributed time with parameter $j\mu$ and jumps to state $k \geq j-1$ with probability p_{k-j+1}, etc.

Let

$$f(s) = \sum_{k=0}^{\infty} p_k s^k$$

be the generating function of the distribution $\{p_k\}$. A necessary and sufficient condition for the process described above not to explode in finite time almost surely, that is $P(Z_t < \infty) = 1$, is that

$$\int_{1-\epsilon}^{1} \frac{ds}{f(s) - s} \tag{1}$$

diverges for every $\epsilon > 0$, see Harris [5]. A sufficient condition for this to hold is $f'(1) < \infty$.

The Kolmogorov backward equation for the generating function $F(s,t) = E[s^{Z_t}]$ is

$$\frac{\partial}{\partial t} F(s,t) = \mu(f(F(s,t)) - F(s,t)) \tag{2}$$

and the Kolmogorov forward equation is

$$\frac{\partial}{\partial t} F(s,t) = \mu(f(s) - s) \frac{\partial}{\partial s} F(s,t). \tag{3}$$

From this equation we see that there is no loss of generality to assume that $p_1 = 0$ since if Z is a branching process with $p_1 > 0$ and intensity μ, then it is equally distributed as the branching process Z^* with intensity $\mu^* = \mu(1-p_1)$ and generating function for the offspring

$$f^*(s) = \frac{f(s) - p_1 s}{1 - p_1} = \sum_{k=0}^{\infty} p_k^* s^k,$$

where $p_1^* = 0$ and $p_k^* = p_k/(1-p_1)$ for $k = 0, 2, 3 \ldots$. Thus we can, and will, assume that $p_1 = 0$.

As $t \to \infty$, the branching process almost surely either dies out, $Z_t \to 0$ or explodes, $Z_t \to \infty$. What behaviour we will have only depends on the expected number of offspring of each particle. If $m = f'(1) \leq 1$, then the process dies out almost surely. If $m > 1$ then there is a positive probability $1 - q$ of explosion and the process is called supercritical. It is easily shown that q is the smallest non-negative root of $q = f(q)$.

In the case of a supercritical branching process we can scale it in time to obtain a non-trivial random variable in the limit.

Proposition 1 With notation as above, let $\lambda = \mu(m-1)$, with $m > 1$. There exists a non-negative random variable W such that

$$e^{-\lambda t} Z_t \to W \quad a.s. \tag{4}$$

as $t \to \infty$. Furthermore $W \not\equiv 0$ if and only if

$$\sum_{k=2}^{\infty} p_k k \log k < \infty \tag{5}$$

For proof, see [1].

We shall in particular study branching processes with $p_0 = 0$. We call such processes immortal branching processes. As mentioned above these processes can be interpreted as the size of a population of immortal particles who independently of each other give birth to new particles at exponential times. We introduce the auxiliary probability distribution $\{q_k\}_{k\geq 1}$, where $q_k = p_{k+1}$, and we interpret this distribution as the distribution of new offspring at each birth. Let $g(s) = f(s)/s$ be the generating function of $\{q_k\}$. The backward equation can thus be written

$$\frac{\partial}{\partial t}F(s,t) = \mu F(s,t)(g(F(s,t)) - 1). \tag{6}$$

Since the immortal branching process is nondecreasing we have $P(Z_t \to \infty) = 1$.

There is a connection between supercritical processes conditioned on exploding, and immortal branching processes. Consider the supercritical branching process Z_t, conditioned on exploding. At any given time this process will have individuals of two types, those who will have an infinite line of descent and those who will not. Let $\widetilde{Z}_t$ be the number of the former individuals at time t. $\widetilde{Z}_t$ is an immortal branching process. To understand this we derive the generating function of the distribution of the offspring $\widetilde{f}(s)$. Each child of an individual in $\widetilde{Z}_t$ will have an infinite line of descent with probability $1-q$ independently of each other. Conditional on k children, the number of children with infinite line of descent thus has a binomial distribution with parameters k and $1-q$, conditioned on being non-zero, since the parent has an infinite line of descent, and thus at least one child. So conditional on k children the generating function is

$$\frac{((1-q)s+q)^k - q^k}{1-q}$$

The unconditional generating function is therefore

$$\begin{aligned}\widetilde{f}(s) &= \sum_{k=1}^{n} p_k \frac{((1-q)s+q)^k - q^k}{1-q} \\ &= \frac{f((1-q)s-q) - f(q)}{1-q} \\ &= \frac{f((1-q)s-q) - q}{1-q}\end{aligned}$$

The exact relation between the limiting distribution of proposition 1 for an exploding branching process Z_t and the associated immortal process $\widetilde{Z}_t$ is given by the following result.

Proposition 2 The proportion $\widetilde{Z}_t/Z_t$ will, conditional on Z_t exploding, converge almost surely to $(1-q)$ as $t \to \infty$, where q is the extinction probability of the supercritical branching process.

See [1] sections I.12 and III.7 for proof. Because of this result, it is sufficient to study immortal branching processes for understanding the limiting behaviour of supercritical branching processes conditioned on exploding.

If Z_t is an immortal branching process, then $Z_t \geq 1$. We will often consider the distribution of $Z_t - 1$ with generating function $F(s,t)/s$, since the results are a little easier to state for the latter process.

3 Compound distributions

We recall some results about compound distributions. All random variables in this section are assumed to be non-negative. More details can be found in [8]. A random variable X is compound-N if

$$X \stackrel{d}{=} \sum_{i=1}^{N} Y_i$$

where N is a random variable with distribution on $\mathbb{N}_0$, the non-negative integers, and $Y_1, Y_2, \ldots$ are independent and identically distributed. Let $g_N(s) = Es^N$, and $L_Y(\theta) = Ee^{-\theta Y_1}$ be the generating function of N and the Laplace-Stieltjes transform of the distribution of Y_1 respectively. Then the Laplace-Stieltjes transform of the distribution of X is

$$L_X(\theta) = E[e^{-\theta X}] = E[E[e^{-\theta X}|N]] = E[L_Y(s)^N] = g_N(L_Y(\theta))$$

If Y_1 has a distribution on $\mathbb{N}_0$ we can write the generating function $g_X(s) = g_N(g_Y(s))$. Some examples of discrete compound distributions are the compound Poisson with generating function

$$F(s) = \exp(\nu(G(s) - 1)) \tag{7}$$

and the compound geometric distribution with generating function

$$F(s) = \frac{1-p}{1-pH(s)} \tag{8}$$

Since

$$\begin{aligned} \frac{1-p}{1-pH(s)} &= \exp\left(-\log(1-p)\left(\frac{\log(1-pH(s))}{\log(1-p)} - 1\right)\right) \\ &= \exp(-\log(1-p)(g_N(H(s)) - 1))), \end{aligned}$$

where $g_N(s)$ is the generating function of the logarithmic distribution, we see that all compound geometric random variables are also compound Poisson. Note that if X has generating function $F(s)$ then $\widetilde{F}(s) = (F(s)-F(0))/(1-F(0))$ is the generating function of $X|X>0$, with $\widetilde{F}(0)=0$. By rescaling ν and p, respectively, we can always choose $G(0)=H(0)=0$ in (7) and (8).

The probability distribution of the random variable X is said to be infinitely divisible if there for all positive integers n exist some i.i.d. random variables $X_1,\ldots,X_n$ such that $X \stackrel{d}{=} X_1+\cdots+X_n$, where $\stackrel{d}{=}$ denotes equality in distribution. This is equivalent to $L_X(\theta) = L_{X_1}(\theta)^n$, where $L_{X_1}(\theta)$ is the Laplace-Stieltjes transform of the distribution of X_1.

Example 1 All compound Poisson random variables are infinitely divisible since $F(s) = \exp(\nu(G(s)-1)) = \exp(\frac{\nu}{n}(G(s)-1))^n$, where $\exp(\frac{\nu}{n}(G(s)-1))$ is a Laplace-Stieltjes transform of a distribution function.

The compound Poisson distributions are important in part due to the following facts, a proof of which can be found in [8].

Proposition 3 All infinitely divisible distributions with $F(0)>0$ are compound Poisson, and all infinitely divisible distributions can be obtained as the weak limit of compound Poisson distributions. Furthermore, all weak limits of infinitely divisible distributions are infinitely divisible.

Example 2 The negative binomial distribution with parameters $\alpha>0$ and $0<p<1$ has generating function

$$F(s) = \left(\frac{1-p}{1-ps}\right)^{\alpha} = \left(\left(\frac{1-p}{1-ps}\right)^{\alpha/n}\right)^{n}$$

So we see that it is infinitely divisible for all α. Now we can investigate if it is also compound geometric.

$$F(s) = \frac{1-q}{1-qH(s)},$$

with $H(0)=0$ implies that $q = 1-(1-p)^{\alpha}$ and

$$H(s) = \frac{1-(1-ps)^{\alpha}}{1-(1-p)^{\alpha}} = \frac{1}{1-(1-p)^{\alpha}} \sum_{k=1} (-1)^{k+1} \binom{\alpha}{k} (ps)^k \qquad (9)$$

and $p_k = -\binom{\alpha}{k}(-p)^k/(1-(1-p)^{\alpha})$ for $k \geq 1$ is a probability distribution if and only if $0<\alpha\leq 1$. Thus the negative binomial distribution is compound geometric if and only if $\alpha \leq 1$.

For non-integer-valued N we define X, a compound-N variable, as follows: Let Y be infinitely divisible. Thus $L_Y(\theta)^a$ is a Laplace-Stieltjes transform for all $a > 0$. We define X by its Laplace-Stieltjes transform $L_X(\theta) = E[L_Y(\theta)^N] = E[\exp(N \log L_Y(\theta))] = L_N(-\log L_Y(\theta))$. If N is infinitely divisible, so is X.

Example 3 The compound exponential distribution has Laplace-Stieltjes transform

$$L_X(\theta) = \frac{\lambda}{\lambda - \log L_Y(\theta)} = \frac{1}{1 - \log L_Y(\theta)^{1/\lambda}}$$

which we obtain with $N \sim \text{Exp}(\lambda)$. Note that we can always choose $\lambda = 1$ by changing Y. The exponential distribution is infinitely divisible since

$$L_N(\theta) = \frac{\lambda}{\lambda + \theta} = \left(\left(\frac{\lambda}{\lambda + \theta}\right)^{1/n}\right)^n,$$

and $(\lambda/(\lambda + \theta))^{1/n}$ is the Laplace-Stieltjes transform of the gamma distribution with parameters $\frac{1}{n}$ and λ, all compound exponential distributions are infinitely divisible.

The following result is an analogue of proposition 3 for compound exponential and compound geometric distributions.

Proposition 4 All compound exponential distributions with $L_X(0) > 0$ are compound geometric and all compound exponential distributions can be obtained as the weak limit of compound geometric distributions. Furthermore, all weak limits of compound exponential distributions are compound exponential.

The latter two statements can be proved, for example, by using the continuity theorem for Laplace-Stieltjes transforms. The first statement is easily checked: Since $0 < P(X = 0) = \lim_{\theta\to\infty} L_X(\theta) = \lim_{\theta\to\infty} 1/(1 - \log L_Y(\theta))$, we obtain $\lim_{\theta\to\infty} L_Y(\theta) = P(Y = 0) > 0$. Since Y is infinitely divisible, we have, by proposition 3, that Y is compound Poisson so $L_Y(\theta) = \exp(\nu(L_V(\theta) - 1))$ for some random variable V. We get

$$L_X(\theta) = \frac{1}{1 - \log L_Y(\theta)} = \frac{1}{1 - \nu(L_V(\theta) - 1)} = \frac{1 - \frac{\nu}{1+\nu}}{1 - \frac{\nu}{1+\nu} L_V(\theta)}.$$

We will also need the following result.

Lemma 1 If X conditional on $X > 0$ is compound exponential, then X is compound geometric.

Proof Let $p = P(X > 0)$. We need to prove that

$$L_{X|X>0}(\theta) = \frac{1}{1 - \log L_Y(\theta)}$$

for some infinitely divisible Y implies that

$$L_X(\theta) = \frac{1 - p'}{1 - p' L_V(\theta)}$$

for some random variable V and some $0 < p' < 1$. $P(X = 0) = L_X(0) = 1-p$ implies that $p' = p$

$$\begin{aligned} L_X(\theta) &= (1-p)L_{X|X=0}(\theta) + pL_{X|X>0}(\theta) \\ &= (1-p) + \frac{p}{1 - \log L_Y(\theta)} \\ &= \frac{1-p}{1 - p\frac{1/(1-p)}{1/(1-p) - \log L_Y(\theta)}} \\ &= \frac{1-p}{1 - pL_V(\theta)}, \end{aligned}$$

where

$$L_V(\theta) = \frac{1/(1-p)}{1/(1-p) - \log L_Y(\theta)}$$

clearly is a Laplace-Stieltjes transform of a probability distribution (a compound exponential one).

Example 4 The gamma distribution with parameters $\alpha > 0$ and $\lambda > 0$ has Laplace-Stieltjes transform

$$L(\theta) = \left(\frac{\lambda}{\lambda + \theta}\right)^{\alpha} = \left(\left(\frac{\lambda}{\lambda + \theta}\right)^{\alpha/n}\right)^{n}$$

which is infinitely divisible for all α. Let N be negative binomial with parameters α and $e^{-c\lambda}$, where $\lambda > 0$ and $c > 0$, then cN is compound geometric by example 2 if and only if $\alpha \le 1$.

$$\lim_{c \to 0} L_{cN}(\theta) = \lim_{c \to 0}\left(\frac{1 - e^{-c\lambda}}{1 - e^{-c(\lambda+\theta)}}\right)^{\alpha} = \left(\frac{\lambda}{\lambda + \theta}\right)^{\alpha}$$

Thus cN converges in distribution to X and X is by proposition 4 therefore compound exponential if $\alpha \le 1$. That the gamma distribution is compound exponential only if $\alpha \le 1$ is shown in example III.5.4 in [8].

4 Infinite divisibility

Theorem 1 All immortal branching processes are infinitely divisible.

Proof Let $\{Z_t\}$ be an immortal branching process with intensity μ and new offspring distributed as Y, that is $P(Y = k) = q_k$. Fix a positive integer n. Let $\{X_t^{(n,1)}\}, \ldots, \{X_t^{(n,n)}\}$ be n i.i.d. copies of an immortal branching process with intensity μ/n and new offspring distributed as nY. Consider the process $\{X_t\}$ defined by $X_t = \sum_{i=1}^n X_t^{(n,i)}$. Note that the value of $\{X_t\}$ is divisible by n at all times. $\{X_t\}$ jumps from state kn with intensity $kn \cdot \mu/n = k\mu$ and when it jumps from state kn it will make a jump with a length distributed as nY. Clearly $\{X_t\} \stackrel{d}{=} \{nZ_t\}$, and thus we can write $Z_t \stackrel{d}{=} \sum_{i=1}^n Z_t^{(n,i)}$ where $Z_t^{(n,i)} = X_t^{(n,i)}/n$ for $i = 1, \ldots, n$. Since n is arbitrary, $\{Z_t\}$ is infinitely divisible.

For fixed t, the distribution of $Z_t - 1$ is thus compound Poisson. This can be proved with the Kolmogorov backward equation. From (6) we have

$$
\begin{aligned}
\frac{\partial}{\partial t} \log \left(\frac{F(s,t)}{s} \right) &= \mu(g(F(s,t)) - 1) \\
\frac{F(s,t)}{s} &= \exp\left(\mu \left(\int_0^t g(F(s,r))dr - t \right) \right) \\
&= \exp\left(\mu t \left(\frac{1}{t} \int_0^t g(F(s,r))dr - 1 \right) \right) \qquad (10)
\end{aligned}
$$

which is the generating function of a compound Poisson distribution since $\frac{1}{t}\int_0^t g(F(s,r))dr$ is a generating function. This distribution has an interesting probabilistic interpretation. Consider the first individual in the process. He will give births according to a Poisson process with intensity μ. Thus the number of births will be Poisson distributed with parameter μt. At each birth the number of children will have the generating function $g(s)$ and if a birth occurred at time $t - r$ each child in that litter will be the ancestor of a number of individuals with generating function $F(s,r)$. Thus the total contribution of a birth at time $t - r$ to the final size will have generating function $g(F(s,r))$. We finally note that each time of birth for the ancestor will be uniformly distributed over $(0,t)$ if we disregard the order of the times.

Corollary 1 Let $\{Z_t\}$ be a supercritical branching process and W the limiting random variable of proposition 1. Assume (5) holds. Then W conditional on $W > 0$ has an infinitely divisible distribution.

Proof Let $\widehat{W}$ have the distribution of $W|W > 0$, and let $\{\widetilde{Z}_t\}$ be the branching process whose individuals are those individuals of $\{Z_t\}$ that have an infinite line of descent. As noted earlier, $\{\widetilde{Z}_t\}$ is an immortal branching process. Since $\widetilde{Z}_t$ is infinitely divisible for all t, so is $e^{-\lambda t}\widetilde{Z}_t$, which by proposition 3 implies that $\widetilde{W} = \lim_{t\to\infty} e^{-\lambda t}\widetilde{Z}_t$ also is infinitely divisible. Now $\widehat{W} = \widetilde{W}/(1-q)$ where $q = P(W = 0)$, and thus even $\widehat{W} \stackrel{d}{=} (W|W > 0)$ is infinitely divisible.

Remark 1 Earlier proofs of the infinite divisibility of an immortal branching process and its limiting random variable can be found in [4]. Our proof is different, and our interpretation of the result also prepares us for proving further results, such as theorem 2.

5 The Yule process

The easiest example of an immortal branching process is the so called Yule process, with $p_2 = 1$, or equivalently $q_1 = 1$. The Yule process is one of the few branching processes whose marginal distribution can be found explicitly. One way of finding the distribution is by solving the Kolmogorov backward equation, but a more direct approach provides the joint distribution of the number of individuals in the process at time t, and their times of birth. This additional information will be useful later on. Let $0 = \tau_{(0)} < \tau_{(1)} < \dots$ be the times of birth in the Yule process and let $(t_1, \dots, t_{n+1}) \in (0,t)^{n+1}$ and let $(t_{(1)}, \dots, t_{(n+1)})$ be the ordered sample of $(t_1, \dots, t_{n+1})$, and set $t_{(0)} = 0$. Recall that $\tau_{(k+1)} - \tau_{(k)} \sim \mathrm{Exp}(k\mu)$ for $k \geq 1$.

$$
\begin{aligned}
P(\tau_{(1)} \in dt_{(1)}, \dots, \tau_{(n)} \in dt_{(n)}, \tau_{(n+1)} > t) = \\
&= \prod_{k=1}^{n} P(\tau_{(k)} \in dt_{(k)} | \tau_{(k-1)} = t_{(k-1)}) \cdot P(\tau_{(n+1)} > t | \tau_{(n)} = t_{(n)}) \\
&= \prod_{k=1}^{n} k\mu e^{-k\mu(t_{(k)} - t_{(k-1)})} \cdot e^{-(n+1)\mu(t - t_{(n)})} dt_{(1)} \dots dt_{(n)} \\
&= n!\mu^n e^{-\mu((n+1)t - \sum_{k=1}^{n} t_{(k)})} dt_{(1)} \dots dt_{(n)} \\
&= n! e^{-\mu t} \prod_{k=1}^{n} \mu e^{-\mu(t - t_{(k)})} dt_{(1)} \dots dt_{(n)} \\
&= e^{-\mu t}(1 - e^{-\mu t})^n n! \prod_{k=1}^{n} \frac{\mu e^{-\mu(t - t_{(k)})}}{1 - e^{-\mu t}} dt_{(1)} \dots dt_{(n)} \qquad (11) \\
&= e^{-\mu t}(1 - e^{-\mu t})^n \prod_{k=1}^{n} \frac{\mu e^{-\mu(t - t_k)}}{1 - e^{-\mu t}} dt_1 \dots dt_n \qquad (12)
\end{aligned}
$$

Since $P(Z_t - 1 = n) = P(\tau_{(n)} \leq t < \tau_{(n+1)})$ we see from (11) or (12) that $Z_t - 1$ has a geometric distribution with parameter $1 - e^{-\mu t}$. Furthermore, we see from (11) that the joint distribution of the times of birth have the density

$$n! \prod_{k=1}^{n} \frac{\mu e^{-\mu(t-t_{(k)})}}{1-e^{-\mu t}}$$

conditional on $Z_t - 1 = n$. From (12) we see that this is the distribution of the ordered sample of the i.i.d. random variables $\tau_1, \ldots, \tau_n$ with probability density

$$h(r) = P(\tau_k \in dr)/dr = \frac{\mu e^{-\mu(t-r)}}{1-e^{-\mu t}} \tag{13}$$

We note that $\tau_k \stackrel{d}{=} t - \epsilon$ conditional on $\epsilon < t$ for an $\epsilon \sim \text{Exp}(\mu)$. The generating function for the distribution of $Z_t - 1$ is

$$\frac{F(s,t)}{s} = \frac{e^{-\mu t}}{1-(1-e^{-\mu t})s}, \tag{14}$$

which we also could have found by solving the Kolmogorov backward equation.

More generally we can also find the distribution when $q_k = 1$ for $k \geq 2$. Here we have

$$\frac{F(s,t)}{s} = \left(\frac{e^{-\mu k t}}{1-(1-e^{-\mu k t})s^k}\right)^{1/k},$$

so $Z_t - 1$ has the distribution of a multiple of a negative binomial random variable. We can also find the distribution of the limiting variable W of proposition 1. We note that $\lim_{t\to\infty} e^{-\lambda t} Z_t = \lim_{t\to\infty} e^{-\lambda t}(Z_t - 1)$ and that $\lambda = \mu(m-1) = \mu k$. The Laplace-Stieltjes transform of the distribution W is given by

$$\begin{aligned} L_W(\theta) &= \lim_{t\to\infty} E[\exp(-\theta e^{-\mu k t}(Z_t - 1))] \\ &= \lim_{t\to\infty} \left(\frac{e^{-\mu k t}}{1-(1-e^{-\mu k t})\exp(-k\theta e^{-\mu k t})}\right)^{1/k} \\ &= \lim_{r\to 0} \left(\frac{r}{1-(1-r)e^{-rk\theta}}\right)^{1/k} \\ &= \left(\frac{1}{1+k\theta}\right)^{1/k}, \end{aligned}$$

so the distribution is gamma, and, in particular, the distribution is exponential for the limit of the Yule process.

6 Further distributional properties

The results of the previous section are in concordance with theorem 1 and corollary 1, but we note that the distributions obtained are not only infinitely divisible but also compound geometric, and compound exponential in the limit. This is in fact true for all immortal branching processes.

Theorem 2 All immortal branching processes have compound geometric distributions.

Proof We make the ansatz (compare with (14))

$$\frac{F(s,t)}{s} = \frac{e^{-\mu t}}{1-(1-e^{-\mu t})H(s,t)},$$

where $H(s,t)$ is a generating function. We can rewrite the Kolmogorov backward equation (6) as

$$\frac{\partial}{\partial t}F(s,t) = \mu F(s,t)(F(s,t)K(s,t)-1), \tag{15}$$

where $K(s,t) = g(F(s,t))/F(s,t)$. We note that $K(s,t)$ is a generating function.

$$\begin{aligned}\frac{\partial}{\partial t}F(s,t) &= \frac{-\mu s e^{-\mu t}}{1-(1-e^{-\mu t})H(s,t)}+\\ &+\frac{s e^{-\mu t}\left(\mu e^{-\mu t}H(s,t)+(1-e^{-\mu t})\frac{\partial}{\partial t}H(s,t)\right)}{(1-(1-e^{-\mu t})H(s,t))^2}\\ &= \mu F(s,t)\left(\frac{F(s,t)}{s}\left(H(s,t)+\frac{e^{\mu t}-1}{\mu}\frac{\partial}{\partial t}H(s,t)\right)-1\right)\end{aligned}$$

Comparing with (15) we get

$$\begin{aligned}H(s,t)+\frac{e^{\mu t}-1}{\mu}\frac{\partial}{\partial t}H(s,t) &= sK(s,t)\\ \mu e^{-\mu t}H(s,t)+(1-e^{-\mu t})\frac{\partial}{\partial t}H(s,t) &= \mu s e^{-\mu t}K(s,t)\\ \frac{\partial}{\partial t}((1-e^{-\mu t})H(s,t)) &= \mu s e^{-\mu t}K(s,t)\\ H(s,t) &= \frac{\mu s}{1-e^{-\mu t}}\int_0^t e^{-\mu u}K(s,u)du\\ \{r=t-u\} &= s\int_0^t K(s,t-r)h(r)dr,\end{aligned}$$

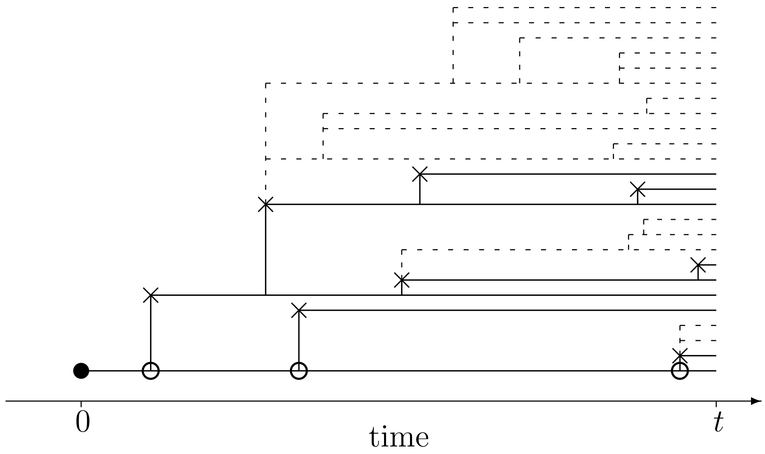

Figure 1: A realisation of an immortal branching process.

which really is a generating function since $K(s,t)$ is a generating function and $h(r)$ is a probability distribution.

We can give a nice probabilistic interpretation of this result if we compare the general immortal branching processes with the simple Yule process. Imagine that the ancestor has a certain title that he passes on to only one child each time he gives birth. The children themselves also pass on the title, if they have any, to only one of their children at each birth.

Now we can consider the group of individuals that have the specific title at any given time t. This group will form a Yule process $\{\widehat{Z}_t\}$ with intensity μ since they give birth to only one child with the title each time they give birth. Thus $\widehat{Z}_t - 1$ has a geometric distribution with parameter $1 - e^{-\mu t}$.

The rest of the population originates from siblings of individuals in the Yule process $\{\widehat{Z}_t\}$. The number of siblings to an individual who was given the title have generating function $g(s)/s$, and each of those siblings will produce a subtree of its own with total size having generating function $F(s, t-r)$ if that sibling is born at time r. Thus the size of all subtrees of siblings of a given titlebearer, born at time r, has generating function $K(s, t-r)$. Finally, we know from (13) that the unconditional distribution of a time of birth in the Yule process $\widehat{Z}_t$ has distribution $h(r)$. See figure 1 for an illustration.

In figure 1, each line corresponds to an individual. The circles, $\circ$, denote the times when the ancestor gives birth, thus they form a realisation of a Poisson process. The $\times$'s denote the times of births of individuals with the same title as the ancestor. The times of the $\times$'s have the distribution $h(r)$.

Theorem 2 also provides information about the limiting random variable W of proposition 1.

Corollary 2 Let $\{Z_t\}$ be a supercritical branching process and W the limiting random variable of proposition 1. Assume (5) holds. Then W has a compound geometric distribution, and W conditional on $W > 0$ has a compound exponential distribution

Proof If we carry through the proof of corollary 1 with the information that $\widetilde{Z}_t$ is not only infinitely divisible, but also compound geometric, we obtain, using proposition 4, the result that $W|W > 0$ has a compound exponential distribution. By lemma 1, this implies that W has a compound geometric distribution.

7 Another branching distribution

So far we have only seen quite trivial immortal branching processes with $q_k = p_{k+1} = 1$ for some $k \geq 1$. There is another class of immortal branching processes whose marginal distribution can be found explicitly. Let

$$p_k = \frac{(\gamma+1)\Gamma(k-1-\gamma)}{k!\Gamma(1-\gamma)} = (\gamma+1)\frac{(1-\gamma)(2-\gamma)\cdots(k-2-\gamma)}{k!},$$

for $0 < \gamma < 1$ and $k \geq 2$. For $\gamma = 1$ we set $p_2 = 1$, and for $\gamma = 0$ we set $p_k = 1/(k(k-1))$. The generating function is

$$f(s) = \begin{cases} (1-s)\frac{(1-s)^\gamma - 1}{\gamma} + s & \text{for } 0 < \gamma \leq 1 \\ (1-s)\log(1-s) + s & \text{for } \gamma = 0. \end{cases} \tag{16}$$

We note that $\gamma = 1$ is the ordinary Yule case, so this family of branching processes can be seen as a generalization of the Yule process.

In the case $\gamma = 0$ the expected number of offspring at each birth is infinite, $m = f'(1) = \infty$, so proposition 1 does not hold and there is no limiting random variable W. Nonetheless, the integral in (1) diverges, so the branching process does not explode in finite time almost surely. Let us solve the Kolmogorov backward equation when $\gamma = 0$. Let $\bar{F}(s,t) = 1 - F(s,t)$. From (2) and (16) we get $\frac{\partial}{\partial t}\log\bar{F}(s,t) = -\mu\log\bar{F}(s,t)$, so $F(s,t) = 1 - \exp(e^{-\mu t}k(s))$ for some function $k(s)$. The boundary condition $F(s,0) = s$ gives us $k(s) = \log(1-s)$ so

$$F(s,t) = 1 - (1-s)^{e^{-\mu t}}$$

Note that this is the distribution of (9) when $\alpha = e^{-\mu t}$ and $p \to 1$.

A branching process with offspring distribution $\{p_k\}$ with $0 < \gamma < 1$ has appeared in the theory of coalescent processes and continuous (stable) random trees, [3]. We once again solve the Kolmogorov backward equation: With the same notation as above and $0 < \gamma \leq 1$,

$$\begin{aligned}
\frac{\partial}{\partial t}\bar{F}(s,t) &= -\frac{\mu}{\gamma}\bar{F}(s,t)\left(\bar{F}(s,t)^{\gamma}-1\right) \\
\left(\frac{-\gamma\bar{F}(s,t)^{-\gamma-1}}{1-\bar{F}(s,t)^{-\gamma}}\right)\frac{\partial}{\partial t}\bar{F}(s,t) &= -\mu \\
\log(1-\bar{F}(s,t)^{-\gamma}) &= -\mu t + k(s) \\
F(s,t) &= 1 - \frac{1}{(1-e^{-\mu t+k(s)})^{1/\gamma}}
\end{aligned}$$

where $k(s)$ again is some function of s. Using $F(s,0) = s$, we get

$$F(s,t) = 1 - \frac{1}{(1-e^{-\mu t}+e^{-\mu t}(1-s)^{-\gamma})^{1/\gamma}}$$

We see that this result is in concordance with (14) for $\gamma = 1$.

The expected number of new offspring at each birth is $m-1 = f'(1)-1 = 1/\gamma$. The limiting random variable W of (4) has Laplace-Stieltjes transform

$$\begin{aligned}
L_W(\theta) &= \lim_{t\to\infty} E[\exp(-\theta e^{-\mu t/\gamma}Z_t)] \\
&= \lim_{t\to\infty} F(\exp(-\theta e^{-\mu t/\gamma}),t) \\
&= \lim_{r\to 0} F(e^{-\theta r}, -\frac{\gamma}{\mu}\log r) \\
&= \lim_{r\to 0}\left(1 - \frac{1}{(1-r^{\gamma}+r^{\gamma}(1-e^{-\theta r})^{-\gamma})^{1/\gamma}}\right) \\
&= 1 - \lim_{r\to 0}\frac{1}{(1-r^{\gamma}+r^{\gamma}(\theta r+o(r^2))^{-\gamma})^{1/\gamma}} \\
&= 1 - \lim_{r\to 0}\frac{1}{(1-r^{\gamma}+(\theta+o(r))^{-\gamma})^{1/\gamma}} \\
&= 1 - \frac{1}{(1+\theta^{-\gamma})^{1/\gamma}}
\end{aligned}$$

Let $F_\gamma(x)$ be the distribution function of W, and $f_\gamma(x)$ its probability density. Then $1-F_\gamma(x)$ has Laplace transform

$$\begin{aligned}
\mathcal{L}(1-F_\gamma(x)) &= \frac{1}{\theta} - \mathcal{L}(F_\gamma(x)) = \frac{1}{\theta} - \frac{1}{\theta}\mathcal{L}(f_\gamma(x)) \\
&= \frac{1}{(1+\theta^{\gamma})^{1/\gamma}} = \theta^{-1}(1+\theta^{-\gamma})^{-1/\gamma} = \sum_{k=0}^{\infty}\binom{-1/\gamma}{k}\theta^{-\gamma k-1}
\end{aligned}$$

By inverting this series term by term we find that

$$1 - F_\gamma(x) = \sum_{k=0}^{\infty} \binom{-1/\gamma}{k} \frac{x^{\gamma k}}{\Gamma(1+\gamma k)} = \frac{1}{\Gamma(\frac{1}{\gamma})} \sum_{k=0}^{\infty} \frac{\Gamma(\frac{1}{\gamma}+k)}{\Gamma(1+\gamma k)} \frac{(-x^\gamma)^k}{k!}$$

We note that $F_1(x) = 1 - e^{-x}$ as expected. When $0 < \gamma < 1$, this series is a special case of several different special functions. With the notation of [6], equation (1.7.8),

$$1 - F_\gamma(x) = \frac{1}{\Gamma(\frac{1}{\gamma})} {}_1\Psi_1 \left[\begin{matrix} (\frac{1}{\gamma}, 1); \\ (1, \gamma); \end{matrix} - x^\gamma \right] = \frac{1}{\Gamma(\frac{1}{\gamma})} H^{1,1}_{1,2} \left[x^\gamma \middle| \begin{matrix} (1-\frac{1}{\gamma}, 1) & \\ (0,1) & (0,\gamma) \end{matrix} \right],$$

where Ψ is a certain generalization of the hypergeometric function and H is the Fox H-function. For rational values of γ, $1 - F_\gamma(x)$ can be expressed with the Meijer G-function, which has the advantage of being implemented in software packages such as `Mathematica` and `Maple`. For example, if $\gamma = \frac{p}{q}$ with $p < q$ being two relatively prime positive integers, then, by equations (8.3.2.7) and (8.3.2.22) in [7],

$$1 - F_{\frac{p}{q}}(x) = \frac{q^{q/p}}{\Gamma(\frac{q}{p})\sqrt{p}(\sqrt{2\pi})^{2q-p-1}} G^{q,q}_{q,p+q} \left(\left(\frac{px}{q^2}\right)^p \middle| \begin{matrix} \frac{1}{q}-\frac{1}{p}, \frac{2}{q}-\frac{1}{p}, \dots, 1-\frac{1}{p} \\ 0, \frac{1}{q}, \dots, \frac{q-1}{q}, 0, \frac{1}{p}, \dots, \frac{p-1}{p} \end{matrix} \right)$$

8 Self-decomposability

A random variable X is called self-decomposable if for all $0 < \beta < 1$ there are random variables $X_{(\beta)}$ independent of X such that $X \stackrel{d}{=} \beta X + X_{(\beta)}$. This is equivalent to $L_X(\theta) = L_X(\beta\theta)L_\beta(\theta)$, with $L_X(\theta)$ and $L_\beta(\theta)$ being the Laplace-Stieltjes transforms of the distributions of X and $X_{(\beta)}$ respectively. It is known that the limiting random variable W in (4) conditioned on $W > 0$ is self-decomposable, see [2]. Self-decomposability implies infinite divisibility but not that a distribution is compound exponential, and vice versa there are compound exponential distributions that are not self-decomposable.

Example 5 The gamma distribution with parameters α and λ is self-decomposable since

$$\begin{aligned} L_X(\theta) = \left(\frac{\lambda}{\lambda+\theta}\right)^\alpha &= \left(\frac{\lambda}{\lambda+\beta\theta}\right)^\alpha \left(\frac{\lambda+\beta\theta}{\lambda+\theta}\right)^\alpha \\ &= L_X(\beta\theta) \left(\frac{\beta}{1-(1-\beta)\frac{\lambda/\beta}{\lambda/\beta+\theta}}\right)^\alpha \\ &= L_X(\beta\theta)L_\beta(\theta), \end{aligned}$$

with $L_\beta(\theta) = g_N(L_Y(\theta))$, where N is negative binomial with parameters α and $1-\beta$, and Y is exponential with parameter λ/β. According to example 4, the gamma distribution is compound exponential if and only if $\alpha \leq 1$, so for $\alpha > 1$ we have an example of a self-decomposable distribution which is not compound exponential.

Example 6 Let X be compound exponential with Laplace-Stieltjes transform $L_X(\theta) = 1/(1 + \log(1+\theta))$. Consider $L_\beta(\theta) = L_X(\theta)/L_X(\beta\theta) = (1+\log(1+\beta\theta))/(1+\log(1+\theta))$ with $0 < \beta < 1$. If $L_\beta(\theta)$ were a Laplace-Stieltjes transform of a (positive) random variable $X_{(\beta)}$ it would be decreasing in θ. But $1 = L_\beta(0) \geq L_\beta(\theta) \to 1$ as $\theta \to \infty$, so $L_\beta(\theta)$ is not decreasing for all values of θ. Therefore, X is a compound exponential random variable that is not self-decomposable.

Since we have shown that immortal branching processes are infinitely divisible and compound geometric even before the limit, one might wonder if the distribution also is self-decomposable for finite t or at least easily is shown to be self-decomposable in the limit.

First we note that no discrete non-trivial random variable X can be self-decomposable since βX is not integer-valued. There is a property similar to self-decomposability that is called discrete self-decomposability that at first sight might seem appropriate to investigate. Define the binomial thinning operator $\odot$, by $\beta \odot n \sim \mathrm{Bin}(n, \beta)$. A random variable with distribution on $\mathbb{N}_0$ is called discrete self-decomposable if for all $0 < \beta < 1$ there are random variables $X_{(\beta)}$ independent of X such that $X \stackrel{d}{=} \beta \odot X + X_{(\beta)}$.

Branching processes are in general *not* discretely self-decomposable for finite t:

Example 7 Consider the branching process with $q_2 = 1$. At each birth, the number of individuals increases with two, so $Z_t - 1$ is even, but $\beta \odot (Z_t - 1)$ has both positive probability of being odd and of being even, so $Z_t - 1$ is therefore not discretely self-decomposable.

References

[1] Athreya, Krishna B. and Ney, Peter E. (1972) *Branching processes* Springer-Verlag, New York.

[2] Biggins, J. D. and Shanbhag, D. N. (1981) Some divisibility problems in branching processes *Math. Proc. Camb. Phil. Soc.,* **90,** 321–330

[3] Berestycki, J., Berestycki, N. and Schweinsberg, J. (2006) Beta-coalescents and continuous stable random trees Preprint: arXiv:math.PR/0602113 v1, 7 Feb 2006

[4] Bühler, W. J. (1968) Some results on the behaviour of branching processes *Th. Prob. Appl.,* **13** 52–64

[5] Harris, Theodore E. (1963) *The Theory of Branching Processes* Springer-Verlag, Berlin.

[6] Mathai, A.M. and Saxena, R.K. (1978) *The H-function with Applications in Statistics and Other Disciplines* Wiley Eastern Limited, New Delhi.

[7] Prudnikov, A.P., Brychkov, J.A. and Marichev, O.I. (1990) *Integrals and Series, vol. 3, More Special Functions* Gordon and Breach, New York.

[8] Steutel, Fred W. and van Harn, Klaas (2004) *Infinite divisibility of probability distributions on the real line* Marcel Dekker, New York.