Branching processes and random trees

Svante Janson

Sto-Upp
Stockholm, 28 May, 2018

Part I – Galton–Watson trees

Let ξ be a random variable with $\xi \in \mathbb{N} := \{0, 1, 2, \dots\}$. Let $p_k := \mathbb{P}(\xi = k)$.

A *Galton–Watson process* starts with a single individual; every individual gets a number of children; these are independent copies of ξ. ξ (or its distribution (p_k)) is called the *offspring distribution*.

A Galton–Watson process generates a random rooted tree $\mathcal{T}$ (finite or infinite), with the initial individual as the root. $\mathcal{T}$ is called a *Galton–Watson tree*.

Fundamental theorem: $\mathcal{T}$ is a.s. finite $\iff \mathbb{E}\,\xi \leq 1$.

The case $\mathbb{E}\,\xi$ is called *critical*.

A *conditional Galton–Watson tree* with n nodes is the random tree $\mathcal{T}$ conditioned on $|\mathcal{T}| = n$. (Denoted $\mathcal{T}_n$.)

Remarks.

1. Conditioned Galton–Watson trees are (essentially) the same as *simply generated trees*, as defined by combinatorists (introduced by Meir and Moon, 1978).

2. We obtain the same $\mathcal{T}_n$ if we replace the offspring distribution ξ by a conjugate distribution $\tilde{\xi}$, i.e. with

$$\mathbb{P}(\tilde{\xi} = k) = ca^k \, \mathbb{P}(\xi = k)$$

for some constants $a, c > 0$.

3. Typically (but not in some exceptional cases, causing condensation) we can therefore assume $\mathbb{E}\,\xi = 1$, a *critical* Galton–Watson tree.

This turns out to be the natural choice of ξ.

Many kinds of random trees occuring in various applications can be seen as conditioned Galton–Watson trees. Some examples, all critical ($\mathbb{E}\,\xi = 1$):

Example A *Geometric distribution* Ge(1/2), $p_k = 2^{-k-1}$, yields uniformly random *ordered trees* (*plane trees*).

Example A *Poisson distribution* Po(1), $p_k = e^{-1}/k!$, yields uniformly random *labelled trees*.

Example A *Binary distribution* Bi(2, 1/2), $w_k = \binom{2}{k}\frac{1}{4}$, yields uniformly random *binary trees*.

Critical Galton–Watson trees form a nice and natural setting, with many known results (possibly with extra assumptions).

Sometimes, but not always, $\sigma^2 := \operatorname{Var} \xi < \infty$ has to be assumed.

Local limit close to the root

Theorem

$\mathcal{T}_n \xrightarrow{\mathrm{d}} \widehat{\mathcal{T}}$ as $n \to \infty$, where $\widehat{\mathcal{T}}$ is an infinite modified Galton–Watson tree (see below).

The limit (in distribution) in the theorem is for a topology where convergence means convergence of outdegree for any fixed node; it thus really means local convergence close to the root.

(It is for this purpose convenient to regard the trees as subtrees of the infinite Ulam–Harris tree.)

Kennedy (1975), Aldous & Pitman (1998), Kolchin (1984), Jonsson & Stefánsson (2011), et al + J

The infinite limit tree

The infinite limit tree $\widehat{\mathcal{T}}$ has nodes of two types, *normal* and *special*. The root is special.

Normal nodes have offspring (outdegree) as copies of ξ.
Special nodes have offspring as copies of $\widehat{\xi}$, where

$$\mathbb{P}(\widehat{\xi} = k) := k\pi_k, \qquad k = 0, 1, 2, \dots$$

(This is a probability distribution because $\mathbb{E}\,\xi = 1$. It is the *size-biased* distribution of ξ.)

When a special node gets children, one of its children (selected uniformly at random) is special.
All other children are normal.

This is the same as the *size-biased Galton–Watson tree* defined by Kesten.

The spine

The limit tree $\widehat{\mathcal{T}}$ can also be described as follows:

The special nodes form an infinite path from the root; we call this path the *spine* of $\widehat{\mathcal{T}}$.

Each outdegree in $\widehat{\mathcal{T}}$ is finite, so the tree is infinite but locally finite.

The spine

The limit tree $\widehat{\mathcal{T}}$ can also be described as follows:

The special nodes form an infinite path from the root; we call this path the *spine* of $\widehat{\mathcal{T}}$.

Each outdegree in $\widehat{\mathcal{T}}$ is finite, so the tree is infinite but locally finite.

Alternative construction: Start with the spine (an infinite path from the root). At each node in the spine attach further branches; the number of branches at each node in the spine is a copy of $\widehat{\xi} - 1$ and each branch is a copy of the Galton–Watson tree $\mathcal{T}$ with offspring distributed as ξ; furthermore, at a node where k new branches are attached, the number of them attached to the left of the spine is uniformly distributed on $\{0, \dots, k\}$.

Since the critical Galton–Watson tree $\mathcal{T}$ is a.s. finite, it follows that $\widehat{\mathcal{T}}$ a.s. has exactly one infinite path from the root, viz. the spine.

Local limit close to the boundary

Given a tree T, let T_v be the fringe tree at v, i.e., the subtree rooted at v, and let T^* be the fringe tree at a uniformly random node v.

Theorem

Let $\mathcal{T}_n^*$, be the random fringe tree of $\mathcal{T}_n$. Then, as $n \to \infty$, $\mathcal{T}_n^*$ converges in distribution to the (unconditioned) Galton–Watson tree $\mathcal{T}$ with offspring distribution ξ, i.e., for any fixed (finite) tree T,

$$\mathbb{P}(\mathcal{T}_n^* = T) \to \mathbb{P}(\mathcal{T} = T).$$

Explicit in Aldous (1991), referring to Kolchin (1986).

Extended fringe trees

Even more generally:

Define the extended fringe tree T^{**} by adding also the mother of v, with its descendents, the grandmother, and so on, i.e., by considering T "shifted" with centre at the random node v.

Theorem

The extended fringe tree $\mathcal{T}_n^{**}$ converges to a random tree $\hat{\hat{\mathcal{T}}}$ constructed as follows:

Add an infinite spine backwards from the root of $\widehat{\mathcal{T}}$; let each node in the spine be special (with a $\widehat{\xi}$ offspring distribution), and add independent forward Galton–Watson trees $\mathcal{T}$ to all their children.

(Implicit in Jagers and Nerman.)

Quenched version

Let $n_T(\mathcal{T}_n)$ be the number of fringe subtrees of $\mathcal{T}_n$ that are isomorphic to T.

Theorem

Assume $\mu := \mathbb{E}\,\xi = 1$ *and* $\operatorname{Var}\xi < \infty$.

(i). *For any fixed tree* T,

$$\frac{n_T(\mathcal{T}_n)}{n} = \mathbb{P}(\mathcal{T}_n^* = T \mid \mathcal{T}_n) \xrightarrow{\text{p}} \mathbb{P}(\mathcal{T} = T).$$

(ii).

$$\frac{n_T(\mathcal{T}_n) - n\,\mathbb{P}(\mathcal{T} = T)}{\sqrt{n}} \xrightarrow{\text{d}} N(0, \gamma^2)$$

for some $\gamma^2 = \gamma_T^2 < \infty$.

General subtrees

Let $S(T)$ be the number of arbitrary (non-fringe) subtrees of T.

Theorem

Suppose that $0 < \operatorname{Var} \xi < \infty$.

(i). *There exist constants* $\mu, \sigma^2 > 0$ *such that*

$$\frac{\log S(\mathcal{T}_n) - \mu n}{\sqrt{n}} \xrightarrow{\mathrm{d}} N(0, \sigma^2).$$

(ii). *If* ξ *has an exponential moment, i.e.* $\mathbb{E}\, e^{t\xi} < \infty$ *for some* $t > 0$*, then, assuming a technical condition,*

$$\mathbb{E}\, S(\mathcal{T}_n)^m \sim \gamma'_m \tau_m^n$$

for some constants $\gamma_m > 0$ *and* $1 < \tau_1 < \tau_2 < \dots$.

Cai and Janson (2018)

Global limit

The global shape of a conditioned Galton–Watson tree with finite offspring variance is asymptotically given by a Brownian excursion $B_{ex}(t)$.

The typical distance to the root is of order $\sqrt{n}$, so we scale distances by this factor.

Aldous (1990).

Part II – general CMJ branching processes

A Crump–Mode–Jagers process is a branching process in continuous time, where each individual has a random number N of children (with $0 \le N \le \infty$), born at times when the individual itself has ages $\xi_1 \le \xi_2 \dots$; these are also random (and may be dependent in any way). (Technically, best seen as a point process.)

Different individuals have i.i.d. life stories.

Let $\mathcal{T}_\infty$ be the complete family tree of the process, starting with a single individual born at time 0, and let $\mathcal{T}_t$ be the subtree of individuals born up to time t.

We are interested in cases when $\mathcal{T}_\infty$ is infinite but each $\mathcal{T}_t$ a.s. is finite. Thus assume $\mathbb{E}\, N > 1$ (supercritical case) and assume for simplicity $N \ge 1$.

Let $Z_t := |\mathcal{T}_t|$, the number of individuals at time t.

Assume some technical conditions.

Then there exists $\alpha > 0$ (the Malthusian parameter), such that

$$e^{-\alpha t} Z_t \xrightarrow{\text{a.s.}} W$$

for some random variable $W > 0$.

(Crump, Mode, Jagers, Nerman, et al)

Define $\tau_n := \inf\{t : Z_t \geq n\}$ and $T_n := \mathcal{T}_{\tau(n)}$.

Thus T_n has n nodes (if birth times are a.s. distinct).

Fringe trees

Theorem

(i). *(Annealed version.) The random fringe tree T_n^* converges in distribution to the random tree $\overline{\mathcal{T}} = \mathcal{T}_\tau$, where $\tau \sim \mathrm{Exp}(\alpha)$ is a random time, independent of $\mathcal{T}$.*

(ii). *(Quenched version.) For every finite tree T,*

$$\mathbb{P}(T_n^* = T \mid T_n) = \frac{|\{v : T_{n;v} = T\}|}{|T_n|} \xrightarrow{\text{a.s.}} \mathbb{P}(\overline{\mathcal{T}} = T).$$

Fringe trees

Theorem

(i). *(Annealed version.) The random fringe tree* T_n^* *converges in distribution to the random tree* $\overline{\mathcal{T}} = \mathcal{T}_\tau$*, where* $\tau \sim \mathrm{Exp}(\alpha)$ *is a random time, independent of* $\mathcal{T}$*.*

(ii). *(Quenched version.) For every finite tree* T*,*

$$\mathbb{P}(T_n^* = T \mid T_n) = \frac{|\{v : T_{n;v} = T\}|}{|T_n|} \xrightarrow{\text{a.s.}} \mathbb{P}(\overline{\mathcal{T}} = T).$$

Central limit theorem?

Extended fringe trees

Define a sin-tree $\tilde{T}$ as follows:

- Start with a copy of the branching process, starting with o born at time 0.
- Give o an infinite line of ancestors, $o^{(1)}, o^{(2)}, \dots$, each having a modified life history where one child is distinguished, and called *heir*, and the probability is weighted by a factor $e^{-\alpha\xi}$, where ξ is the time the heir is born.
- Let the heir of $o^{(i)}$ be $o^{(i-1)}$. This defines the (negative) birth times of the ancestors. Let all other children of the ancestors start new copies of $\mathcal{T}$.

Theorem

(i). *(Annealed.) The extended fringe tree of T_n converges in distribution to $\tilde{T}$.*

(ii). *(Quenched.) This holds also conditioned on T_n, a.s.*

Random recursive tree

Example

Children born with independent Exp(1) waiting times, i.e., according to a Poisson process with rate 1. The branching process is the Yule process.

T_n is the random recursive tree. The next node is added as a child to a uniformly chosen node.

General preferential attachment tree

Example

Let $(w_k)_0^\infty$ be a sequence of weights with $w_k \geq 0$ and $w_0 > 0$.

Grow a tree by choosing the mother of each new node randomly with probability proportional to w_d where d is the outdegree (number of existing children).

T_n where the waiting time for child k is $\text{Exp}(w_{k-1})$.

General preferential attachment tree

Example

Let $(w_k)_0^\infty$ be a sequence of weights with $w_k \geq 0$ and $w_0 > 0$.

Grow a tree by choosing the mother of each new node randomly with probability proportional to w_d where d is the outdegree (number of existing children).

T_n where the waiting time for child k is $\text{Exp}(w_{k-1})$.

Standard case: $w_k = k + 1$.

General preferential attachment tree

Example

Let $(w_k)_0^\infty$ be a sequence of weights with $w_k \geq 0$ and $w_0 > 0$.

Grow a tree by choosing the mother of each new node randomly with probability proportional to w_d where d is the outdegree (number of existing children).

T_n where the waiting time for child k is $\mathrm{Exp}(w_{k-1})$.

Standard case: $w_k = k + 1$.

Linear case: $w_k = \chi k + \rho$.

Binary search tree

Example

Each individual gets two children, one left and one right; each after an Exp(1) time (independent).

m-ary search tree (with external nodes)

Example

$m \geq 2$ fixed.

A newborn has 0 "keys". It get $m-1$ keys after independent waiting times $Y_1, \dots, Y_{m-1}$ with $Y_i \sim \text{Exp}(i)$. When the last key arrives, m children are born.

The m-ary search tree T_n is defined with a fixed number n keys, while the number of nodes is random. The theory extends to this case too, using the notion of *random characteristic*.

Fragmentation trees

Example

Start with an object of mass $x_0 > 0$; break it ito b pieces with masses $V_1 x_0, \dots, V_b x_0$, where $(V_1, \dots, V_b)$ is a random vector with $V_i \geq 0$ and $\sum_i V_i = 1$. Continue recursively with each piece of mass $\geq x_1$, using a new copy of $(V_1, \dots, V_b)$ each time.

Regard the fragments of masses $\geq x_1$ seen during the process as nodes in the *fragmentation tree*.

CMJ process: An individual has b children, born at times $\xi_1, \dots, \xi_b$ with $\xi_i := -\log V_i$.

The fragmentation tree is the tree $\mathcal{T}_{\log(x_0/x_1)}$.