Pruning procedures on trees

Anita Winter, University of Duisburg-Essen

DAAD Spring School

"Combinatorical stochastic processes and applications"

Vietnam Institute for Advanced Study in Mathematics, Hanoi

March 07th-18th, 2016

UNIVERSITÄT DUISBURG ESSEN

Offen im Denken

Outline of the course

- **Part 1: The discrete picture**

  "The tree-valued Markov chain arising from pruning Galton-Watson trees"

- **Part 2: The continuous picture**

  *"*THE *Continuum Random Tree (CRT) and pruning of continuum trees"*

- **Part 3: Convergence of the discrete to the continuous picture**

  "Leaf sampling weak vague topology and THE *pruning process"*

Outline: Part I
The tree-valued Markov chain arising from pruning Galton-Watson trees

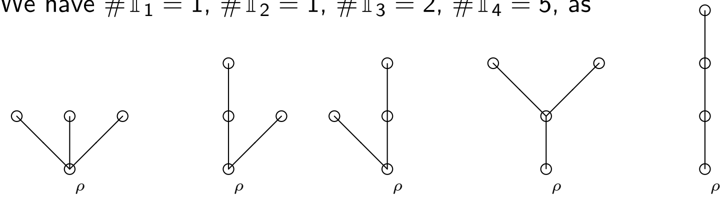
1. Random tree models: the discrete world
   (a) Notation and terminology of trees
   (b) Random trees: link between Galton-Watson and combinatorial trees
      - Galton-Watson trees
      - Galton-Watson trees conditioned on fixed progeny
      - Galton-Watson trees conditioned on number of leaves
2. Pruning Galton-Watson trees
   (a) Edge and node percolation: homogeneous pruning
   (b) Node percolation with degree dependence

Notation and terminology for our trees

- By a **tree** $\mathfrak{t}$ we mean a rooted, labelled tree, i.e., a set of vertices $V = V(\mathfrak{t})$ equipped with a **direct edge** relation $\stackrel{\mathfrak{t}}{\to}$ such that for some **root** $\rho = \rho(\mathfrak{t}) \in V$ there is for each $v \in V$ a unique path from the root to v.
- For $v, w \in \mathfrak{t}$ with $v \stackrel{\mathfrak{t}}{\to} w$, call w a **child** of v and v the **parent** of w.
- $h = h(v, \mathfrak{t})$ is called the **height** of v in the tree $\mathfrak{t}$. The height $h(\mathfrak{t})$ of a tree $\mathfrak{t}$ is the maximal height of a vertex in the tree.
- If a subset $S \subseteq V$ is such that the restriction of $\stackrel{\mathfrak{t}}{\to}$ defines a tree $\mathfrak{s}$, then S or $\mathfrak{s}$ are called a **subtree** of $\mathfrak{t}$.
- Let $\#\mathfrak{t} := \#V(\mathfrak{t})$ denote the **size of the tree**.
- The **number of edges** in $\mathfrak{t}$ equals $\#\mathfrak{t} - 1$.

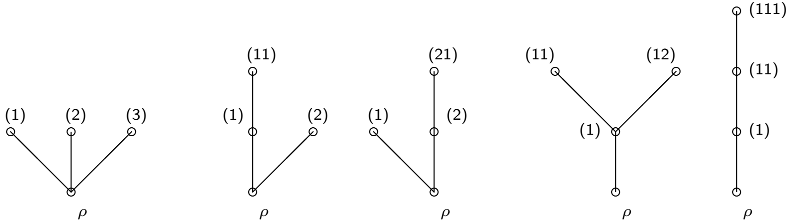
Family trees (=rooted, ordered trees)

- Let $\mathbb{T}_n$ be the set of **all rooted, ordered trees** with n vertices (including the root), i.e., we distinguish between trees when the birth order is not the same.
- We have $\#\mathbb{T}_1 = 1$, $\#\mathbb{T}_2 = 1$, $\#\mathbb{T}_3 = 2$, $\#\mathbb{T}_4 = 5$, as



- We consider **finite trees** $\mathfrak{t}$ as elements in $\mathbb{T} := \bigcup_{n\geq 1} \mathbb{T}_n$.
- For each $\mathfrak{t} \in \mathbb{T}$ and $g = 0, 1, 2, ...$, each vertex at height g corresponds to an individual in the g^{th} **generation** of the family.
- We identify an individual in the g^{th} generation with a sequence of g integers, for instance $(2, 7, 4)$ to indicate a third generation individual who is the 4^{th} child of the 7^{th} child of the 2^{nd} child of the progenitor (root). This generates a **labelling** on trees.

Labelled family trees: illustration



Exercise. Show that the number of rooted, ordered trees equals the **Catalan numbers**, i.e., for all $k = 1, 2, \ldots,$

$$\#\mathbb{T}_k = \frac{1}{k}\binom{2(k-1)}{(k-1)} = 2^{k-1}\frac{1}{k!}(2k-3)!!,$$

where $(2n-1)!! = (2n-1)\cdot(2n-3)\cdot\ldots\cdot 3\cdot 1$. The first numbers are:

$$1,\ 1,\ 2,\ 5,\ 14,\ 42,\ 132,\ 429,\ \ldots$$

Hint. Show that $\#\mathbb{T}_n = \sum_{\ell=1}^{n-1}\#\mathbb{T}_\ell\cdot\#\mathbb{T}_{n-\ell}$, $n\geq 2$, and use this to determine $g(s) := \sum_{n\geq 1} s^n\#\mathbb{T}_n$.

Labelled family trees: ordered, rooted, possibly infinite trees

- To allow for possibly infinite family trees, we consider **trees** $\mathfrak{t}$ as rooted trees for which the vertex set

$$V \subseteq \{\rho\} \cup \bigcup_{g \in \mathbb{N}} \mathbb{N}^g$$

satisfies

(i) *If* $w = (v, j) \in V$ *for some* $v \in \mathbb{N}^g$, $g \geq 1$ *and* $j \in \mathbb{N}$, *then* $w = (v, j') \in V$ *for all* $1 \leq j' \leq j$.

(ii) For all $v \in V$, the **number of v's children is finite**, i.e., $c_v := \#\{j \in \mathbb{N} : (v, j) \in V\} < \infty$.

(iii) *If* $w = (v, j) \in V$ *for some* $v \in \mathbb{N}^g$, $g \geq 1$ *and* $j \in \mathbb{N}$, *then* $v \in V$.

and with the set of **directed edges** $v \stackrel{\mathfrak{t}}{\to} w$ if and only if $w = (v, j)$ for some $v \in \mathbb{N}^g$, $g \geq 1$, and $j \in \mathbb{N}$, or if $v = \rho$ and $w = (j)$ for some $j \in \mathbb{N}$.

- Denote by $\mathbb{T}_\infty$ the set of all possibly infinite family trees.

Convergence of family trees in $\mathbb{T}_\infty$

- The **height** of $\mathfrak{t} \in \mathbb{T}_\infty$ equals the maximal $g \in \mathbb{N}_0$ such that $V(\mathfrak{t}) \cap \mathbb{N}^g \neq \emptyset$.
- For each height $h \in \mathbb{N}_0$ there is a natural **restriction map** $r_h : \mathbb{T}_\infty \to \mathbb{T}^{(h)}$ where $\mathbb{T}^{(h)}$ denotes the set of all finite trees of height at most h. Namely,

$$r_h\mathfrak{t} := \{\rho\} \cup \Big(V(\mathfrak{t}) \cap \big(\bigcup_{g=1}^{h} \mathbb{N}^g\big)\Big).$$

- The restriction maps $(r_h, h \in \mathbb{N})$ satisfy a **projective property**, i.e., $r_h\mathfrak{t} = r_h\big(r_{h+1}\mathfrak{t}\big)$.
- A tree $\mathfrak{t} \in \mathbb{T}_\infty$ can thus be identified with the sequence $(r_h\mathfrak{t};\ h \in \mathbb{N}_0)$.
- We say that **a sequence $(\mathfrak{t}_n)_{n\in\mathbb{N}}$ converges to $\mathfrak{t}$** in $\mathbb{T}_\infty$ if and only if for all $h \in \mathbb{N}_0$, the sequences $(r_h\mathfrak{t}_n)_{n\in\mathbb{N}}$ converges to $r_h\mathfrak{t}$ in $\mathbb{T}^h$ with respect to the discrete topology.

Random family trees

- A **random family tree** $\mathcal{T}$ is a random variable with values in $\mathbb{T}_\infty$.
- Define **convergence of distributions of random trees** by weak convergence of probability measures on $\mathbb{T}_\infty$. That is, for random family trees $\mathcal{T}_n$, $n = 1, 2, \ldots$, we say that $(\mathcal{T}_n)_{n\in\mathbb{N}}$ converges in distribution to $\mathcal{T}$ if for all $h \in \mathbb{N}_0$ and $\mathfrak{t} \in \mathbb{T}^{(h)}$,

$$\mathbb{P}\{r_h\mathcal{T}_n = \mathfrak{t}\} \underset{n\to\infty}{\longrightarrow} \mathbb{P}\{r_h\mathcal{T} = \mathfrak{t}\}.$$

- In this lecture we will mainly focus on the two classes of random trees:
  - **Combinatorial trees.** We choose these trees uniformly in a certain class of trees, e.g., family trees (also called **plane trees**), Cayley trees, binary trees, etc.
  - **Galton-Watson trees.** We construct these trees by choosing the number of "children" of the root, then recursively the number of children of each child, and so on.

  There is a link between several combinatorial trees and Galton-Watson trees conditioned on the progeny.

Galton-Watson trees

Definition (Galton-Watson tree)

Let $p := (p(0), p(1), \ldots)$ be a probability distribution on $\mathbb{N}_0$ with $p(1) < 1$. We call a random tree $\mathcal{G}$ a **Galton-Watson tree** with **offspring distribution** $p(\cdot)$ iff

- the number of children of the root has distribution $p(\cdot)$, and
- for each $h = 1, 2, \ldots$, conditionally given that $r_h\mathcal{G} = \mathfrak{t} \in \mathbb{T}^{(h)}$, the numbers of children $c_v(\mathcal{G})$, $v \in \mathrm{gen}(h, \mathcal{G})$, are i.i.d. w.r.t. $p(\cdot)$.

- For all $\mathfrak{t} \in \mathbb{T}$,
$$\mathbb{P}\{\mathcal{G} = \mathfrak{t}\} = \prod_{v \in V(\mathfrak{t})} p\big(c_v \mathfrak{t}\big). \tag{1}$$
- Let $\mu := \sum_{n \in \mathbb{N}} np(n)$ be the **mean offspring number**, then the following are equivalent:
$$\mu \le 1 \;\Leftrightarrow\; \mathbb{P}\{\#\mathcal{G} < \infty\} = 1 \;\Leftrightarrow\; \mathbb{P}\{\mathrm{height}(\mathcal{G}) \ge h\} \underset{h \to \infty}{\longrightarrow} 0.$$
- Consequently, if $\mu \le 1$, then the distribution of $\mathcal{G}$ is uniquely determined by (1).

- For $\mu > 0$, let $\mathcal{G}_{\mathrm{Pois}(\mu)}$ be a Galton-Watson tree with **Poisson offspring distribution** with mean μ, i.e.,

$$p_\mu(n) := \tfrac{\mu^n}{n!} e^{-\mu}, \quad n = 0, 1, 2, \dots$$

- Denote the distribution of $\mathcal{G}_{\mathrm{Pois}(\mu)}$ by **PGW(μ)**. Notice that for all $\mathfrak{t} \in \mathbb{T}$,

$$\mathbb{P}\{\mathcal{G}_{\mathrm{Pois}(\mu)} = \mathfrak{t}\} = e^{-\mu \#\mathfrak{t}} \mu^{\#\mathfrak{t}-1} \prod_{v \in V(\mathfrak{t})} \tfrac{1}{(c_v \mathfrak{t})!}$$

Use that $\sum_{v \in V(\mathfrak{t})} c_v = \#\mathfrak{t} - 1$.

Example: Binary branching trees

- For $v \in (0,1)$, let $\mathcal{G}_{\mathrm{binary}(p)}$ be a Galton-Watson tree whose offspring distribution satisfies

$$p_v(0) := (1-v), \quad p_v(2) = v.$$

  That is, almost surely any vertex (other than the root) in $\mathcal{G}_{\mathrm{binary}(v)}$ has either degree 1 (= **leaf**) or degree 3 (= inner node). Such tree are called **binary**.

- Notice that if $\mathfrak{t} \in \mathbb{T}$ is **binary, rooted** with $n \geq 2$ leaves (other than the root), then $\#\mathfrak{t} = 2n - 1$. Hence

$$\mathbb{P}\{\mathcal{G}_{\mathrm{binary}(v)} = \mathfrak{t}\} = (1-v)^{\#\mathrm{Lf}(\mathfrak{t})} \cdot v^{(\#\mathrm{Lf}(\mathfrak{t})-1)}.$$

- In the **critical case** $v = \frac{1}{2}$, and

$$\mathbb{P}\{\mathcal{G}_{\mathrm{binary}(v)} = \mathfrak{t}\} = 2^{-\#\mathfrak{t}}.$$

In particular, **all** rooted, binary ordered trees **of the same size** are **equally likely**.

Example: Geometric Galton-Watson trees

- For $u \in (0,1)$, let $\mathcal{G}_{\mathrm{Geom}(u)}$ be a Galton-Watson tree with **geometric offspring distribution** with success parameter u, i.e.,

$$p_u(n) := u \cdot (1-u)^n, \quad n = 0, 1, 2 \ldots.$$

- Denote the distribution of $\mathcal{G}_{\mathrm{Geom}(u)}$ by **Geom**(u). Notice that for all $\mathfrak{t} \in \mathbb{T}$,

$$\mathbb{P}\big\{\mathcal{G}_{\mathrm{Geom}(u)} = \mathfrak{t}\big\} = u^{\#\mathfrak{t}} \cdot \big(1-u\big)^{\#\mathfrak{t}-1}$$

  Use once more that $\sum_{v \in V(\mathfrak{t})} c_v = \#\mathfrak{t} - 1$.

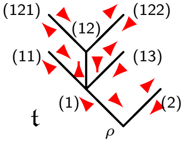
- Specifically, if $u = \frac{1}{2}$,

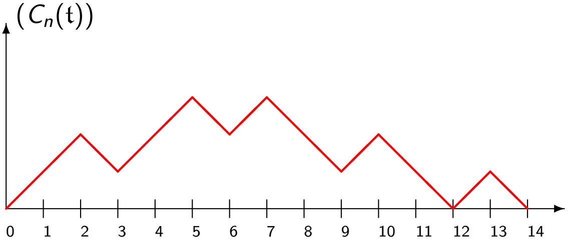
$$\mathbb{P}\big\{\mathcal{G}_{\mathrm{Geom}(\frac{1}{2})} = \mathfrak{t}\big\} = 2^{-(2\#\mathfrak{t}-1)}.$$

In particular, under the law of GW-trees with critical geometric offspring all trees of the same size are **equally likely**.

Coding finite family trees via the contour function

- The **contour function** of a finite rooted, ordered tree $\mathfrak{t}$ is obtained by traversing the tree at speed 1 starting in the root clockwise, and recording the height profile.
- As every edge is traversed exactly twice, if $\mathfrak{t}$ has n edges, then the contour function is a function on $[0, 2n]$.





Contour function representation of a geometric GW-tree

- Characteristic for the geometric distribution (among discrete distributions) is the **memoryless property**, i.e., if G has geometric distribution with success parameter $u \in (0,1)$, then for all $n, m \in \mathbb{N}_0$,
$$\mathbb{P}\big(G = m + n \big| G \geq n\big) = \mathbb{P}\big\{G = m\big\}.$$
- Thus the contour function of geometric GW-trees can be represented by a Markov process.

Lemma

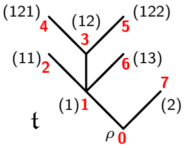
*If $\mathcal{G}$ is **Geom**(u), then the contour process $(\mathcal{C}_n)_{n\in\{0,1,2,\dots,\tau_0\}}$ is a random walk with jump distribution $\mathbb{P}\{\mathcal{C}_k - \mathcal{C}_{k-1} = -1\} = u$ and $\mathbb{P}\{\mathcal{C}_k - \mathcal{C}_{k-1} = 1\} = 1 - u$ stopped one step before it gets negative.*

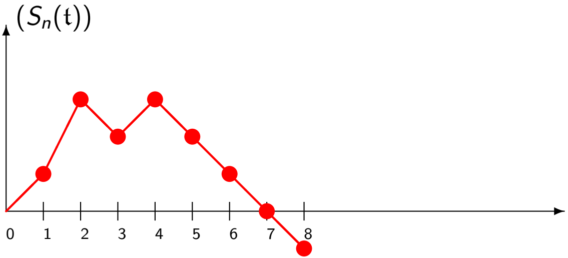
- Notice that for any other offspring distribution, the contour process is **NOT a Markov process**.

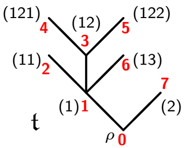
Lukasiewicz walk

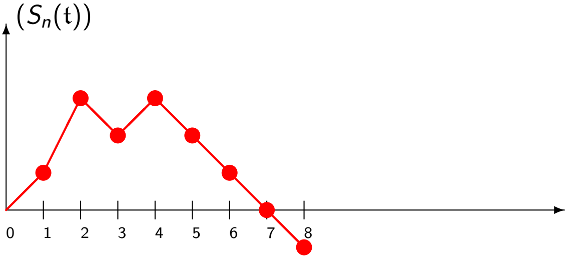
- Enumerate vertices of $\mathfrak{t}$ in lexicographic order, $v_0 := \rho$, $v_1 := (1)$, ..., $v_{\#\mathfrak{t}-1}$.
- Define $S_0 := 0$, and for $0 \leq n \leq \#\mathfrak{t} - 1$,

$$S_{n+1} = S_n + \big(c_{v_n}(\mathfrak{t}) - 1\big).$$









Lemma

If $\mathcal{G}$ is a GW-tree with offspring distribution $p(\cdot)$, then the **Lukasiewicz walk** $(S_n)_{0\le n\le \#\mathcal{G}}$ *is a random walk with jump distribution*

$$\nu(k) = p(k+1), \quad k = -1, 0, \ldots,$$

stopped at its first hitting time of -1. That is, for all $n = 1, 2, \ldots,$

$$\mathbb{P}\{\#\mathcal{G} = n\} = \mathbb{P}\{S_1 \ge 0, \ldots, S_{n-1} \ge 0, S_n = -1\}.$$

Lukasiewicz walk and GW-trees: proof

$S_{n+1} = S_n + \big(c_{v_n}(\mathfrak{t}) - 1\big)$, $c_v(\mathfrak{t})$ **number of children of v in $\mathfrak{t}$**

- For a GW-tree $\mathcal{G}$ with offspring distribution $p(\cdot)$, the $\{c_{v_n}(\mathcal{G}) - 1,\ n = 0, ..., \#\mathcal{G} - 1\}$ have distribution $\nu(k) = p(k+1)$, $k = -1, 0, 1, 2, ...$
- $S_{\#\mathcal{G}} = \sum_{v \in V(\mathcal{G})} c_v(\mathcal{G}) - \#\mathcal{G} = (\#\mathcal{G} - 1) - \#\mathcal{G} = -1.$
- For all $1 \le m \le \#\mathcal{G} - 1$,

$$S_m = \sum_{n=0}^{m-1} \big(c_{v_n}(\mathcal{G}) - 1\big) = \sum_{n=0}^{m-1} c_{v_n}(\mathcal{G}) - m \ge 0,$$

because among all individuals counted in $\sum_{n=0}^{m-1} c_{v_n}(\mathcal{G})$, the individuals v_1, ..., v_m will appear. □

Dwass' observation

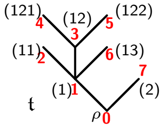
Proposition (Dwass (1962) [4])

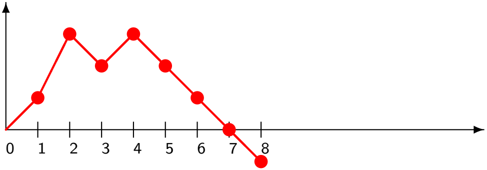
Let $X_1, X_2, \ldots$ i.i.d. with $\mathbb{P}\{X_1 = k\} = p(k+1)$, and $S_n := \sum_{i=1}^{n} X_i$. Then

$$\mathbb{P}\{S_1 \geq 0, S_2 \geq 0, \ldots, S_{n-1} \geq 0, S_n = -1\} = \frac{1}{n}\mathbb{P}\{S_n = -1\}.$$

Sketch of proof: a numerical illustration. Consider all possible **cyclic permutations**:

x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	s_1	s_2	s_3	s_4	s_5	s_6	s_7	s_8
1	2	-1	1	-1	-1	-1	-1	1	3	2	3	2	1	0	-1
2	-1	1	-1	-1	-1	-1	1	2	1	2	1	0	-1	-2	-1
-1	1	-1	-1	-1	-1	1	2	-1	0	-1	-2	-3	-4	-3	-1
1	-1	-1	-1	-1	1	2	-1	1	0	-1	-2	-3	-2	0	-1
-1	-1	-1	-1	1	2	-1	1	-1	-2	-3	-4	-3	-1	- 2	-1
-1	-1	-1	1	2	-1	1	-1	-1	-2	-3	-2	0	-1	0	-1
-1	-1	1	2	-1	1	-1	-1	-1	-2	-1	1	0	1	0	-1
-1	1	2	-1	1	-1	-1	-1	-1	0	2	1	2	1	0	-1





Dwass' observation: sketch of formal proof

Proposition (Dwass (1962) [4])

Let X_1, X_2, ... *be an i.i.d. sequence with* $\mathbb{P}\{X_1 = k\} = p(k+1)$, *and* $S_n := \sum_{i=1}^n X_i$. *Then*

$$\mathbb{P}\{S_1 \geq 0, S_2 \geq 0, ..., S_{n-1} \geq 0, S_n = -1\} = \tfrac{1}{n}\mathbb{P}\{S_n = -1\}.$$

Sketch of a formal proof.

- We consider the n cyclic permutations of a given set of $\{-1, 0, 1, 2, ...\}$-valued numbers x_1, x_2, ..., x_n with $\sum x_i = -1$.
- Denote by $T(\ell)$ the **cyclically permuted** sequence

$$x_\ell, x_{\ell+1}, ..., x_n, x_1, ..., x_{\ell-1}.$$

- Show that $\{T(\ell);\, \ell = 1, ..., n\}$ contains exactly one representative for which all first $n-1$ partial sums are non-negative.
- As each representative has the same probability, the claim follows.

$\square$

Dwass' observation: the cyclic lemma

Lemma

Given $\{-1, 0, 1, 2, ...\}$-valued integers $\{x_i, \ \ i = 1, ..., n\}$ with $\sum_{i=1}^n x_i = -1$, we denote for any $\ell = 1, ..., n$ by $T(\ell)$ the **cyclically permuted** *sequence*

$$x_\ell, x_{\ell+1}, ..., x_n, x_1, ..., x_{\ell-1}.$$

We claim that the set $\{T(\ell), \ \ \ell = 1, ..., n\}$ contains exactly one element for which the minimum of the first $(n-1)$ partial sums in $T(\ell)$ is non-negative.

Proof of existence. We are given $\{x_i, \ \ i = 1, ..., n\}$ with partial sums $s_k := \sum_{i=1}^k x_i$, $s_n = -1$. W.l.o.g. assume that the minimum of the first $(n-1)$ partial sums in $T(1)$ is negative.

- Let $\mu(1)$ denote the first index at which the minimum among $s_1, ..., s_{n-1}$ is attained. That is, $s_\ell - s_{\mu(1)} \geq 1$ for $\ell < \mu(1)$, and $s_{\mu(1)} - s_\ell \geq 0$ for $\ell \geq \mu(1)$.
- It follows that the minimum of the first $(n-1)$ partial sums in $T(\mu(1)+1)$ is non-negative. Indeed, the partial sums of $T(\mu(1)+1)$ are

$$\underbrace{s_{\mu(1)+1} - s_{\mu(1)}}_{\geq 0}, ..., \underbrace{s_n - s_{\mu(1)}}_{\geq 0}, s_n \underbrace{- s_{\mu(1)} + s_1}_{\geq 1}, s_n \underbrace{- s_{\mu(1)} + s_2}_{\geq 1}, ..., s_n - s_{\mu(1)} + s_{\mu(1)}.$$

Dwass' observation: the cyclic lemma

Lemma

Given $\{-1,0,1,2,...\}$-valued integers $\{x_i, \ \ i=1,...,n\}$ with $\sum_{i=1}^n x_i = -1$, we denote for any $\ell = 1,...,n$ by $T(\ell)$ the **cyclically permuted** *sequence*

$$x_\ell, x_{\ell+1}, ..., x_n, x_1, ..., x_{\ell-1}.$$

We claim that the set $\{T(\ell), \ \ \ell = 1,...,n\}$ contains exactly one element for which the minimum of the first $(n-1)$ partial sums in $T(\ell)$ is non-negative.

Proof of uniqueness. We are given $\{x_i, \ \ i=1,...,n\}$ with partial sums $s_k := \sum_{i=1}^k x_i$, $s_n = -1$. W.l.o.g. assume that the minimum of the first $(n-1)$ partial sums in $T(1)$ is non-negative.

- Fix $\ell \in \{2,...,n\}$. Notice that the $n-\ell+1$ partial sum in $T(\ell)$ equals

$$s_n - s_{\ell-1} \leq -1. \qquad \square$$

Consequence of Dwass' observation

Corollary

Let X_1, X_2, ... be an i.i.d. sequence distributed according to the offspring distribution $\nu(k) := p(k+1)$, $k = -1, 0, \ldots$, and $S_n := \sum_{i=1}^{n} X_i$. Then for all $n \in \mathbb{N}$,

$$\mathbb{P}\big\{\#\mathcal{G} = n\big\} = \tfrac{1}{n}\mathbb{P}\big\{S_n = -1\big\}.$$

Equivalently, we also have the following:

Corollary

Let X_1, X_2, ... be an i.i.d. sequence distributed according to the offspring distribution $p(\cdot)$, and $S_n := \sum_{i=1}^{n} X_i$. Then for all $n \in \mathbb{N}$,

$$\mathbb{P}\big\{\#\mathcal{G} = n\big\} = \tfrac{1}{n}\mathbb{P}\big\{S_n = n - 1\big\}.$$

Total progeny of the Poisson-Galton-Watson tree

- Let X_1, X_2, ... be i.i.d. Poisson distributed with mean μ, and $S_n := \sum_{k=1}^n X_k$.
- Then S_n is Poisson distributed with parameter $n\mu$ and we find that

$$\mathbb{P}\{\#\mathcal{G}_{\mathrm{Pois}(\mu)} = n\} = \tfrac{1}{n}\mathbb{P}\{S_n = n-1\} = \frac{(n\mu)^{n-1}}{n!}e^{-n\mu}, \quad n = 1, 2, \ldots$$

- This distribution is called **Borel(μ)-distribution**.

Lemma

If X is Borel(μ)-distributed for $\mu < 1$, then $\mathbb{E}[X] = (1-\mu)^{-1}$.

Proof. Put $\nu(\mu) := \mu e^{-\mu}$. As

$$\mu = \sum_{n\ge 1} \frac{n^{n-1}\mu^n}{n!}e^{-n\mu} = \sum_{n\ge 1} \frac{n^{n-1}\nu^n}{n!},$$

differentiating by ν yields

$$\tfrac{\mathrm{d}\mu}{\mathrm{d}\nu} = \sum_{n\ge 1} n\frac{n^{n-1}\nu^{n-1}}{n!} = e^{\mu}\mathbb{E}[X].$$

Now use that

$$e^{-\mu}\big(\tfrac{\mathrm{d}\nu}{\mathrm{d}\mu}\big)^{-1} = (1-\mu)^{-1}. \qquad \square$$

Total progeny of the binary Galton-Watson tree

s

- Let X_1, X_2, ... be i.i.d. with $\mathbb{P}\{X_1 = 0\} = 1 - v$ and $\mathbb{P}\{X_1 = 2\} = v$, and $S_n := \sum_{k=1}^n X_k$.
- Then $\frac{S_n}{2}$ has binomial distributed with parameters n and v, and we find that

$$\begin{aligned}
&\mathbb{P}\big\{\#\mathcal{G}_{\mathrm{binary}(v)} = 2\ell - 1\big\} \\
&\quad = \tfrac{1}{2\ell-1}\mathbb{P}\big\{S_{2\ell-1} = 2\ell - 2\big\} \\
&\quad = \tfrac{1}{2\ell-1}\binom{2\ell-1}{\ell-1}\big(1-v\big)^{\ell}\cdot v^{\ell-1} \\
&\quad = 2^{\ell-1}\tfrac{1}{(\ell-1)!}(2\ell-3)!!(1-v)\big((1-v)v\big)^{\ell-1}, \;\; \ell = 1, 2, 3, \ldots
\end{aligned}$$

- In particular, if $v = \frac{1}{2}$,

$$\mathbb{P}\big\{\#\mathcal{G}_{\mathrm{binary}(\frac{1}{2})} = 2\ell - 1\big\} = \tfrac{1}{(\ell-1)!}(2\ell-3)!!\cdot 2^{-\ell}.$$

Total progeny of the geometric Galton-Watson tree

- Let $X_1, X_2, \dots$ be i.i.d. geometrically distributed with success parameter $u \in (0,1)$, and $S_n := \sum_{k=1}^n X_k$.
- Then S_n has **negative binomial distribution** with parameters n and u, i.e.,

$$\mathbb{P}\big\{S_n = k\big\} = \tbinom{k+n-1}{k} u^n \cdot \big(1-u\big)^k, \quad k = 0,1,2,\dots$$

We therefore find that for all $n = 1,2,\dots$,

$$\begin{aligned}\mathbb{P}\big\{\#\mathcal{G}_{\mathrm{Geom}(u)} = n\big\} &= \tfrac{1}{n}\mathbb{P}\big\{S_n = n-1\big\} \\ &= \tfrac{1}{n}\tbinom{2(n-1)}{n-1} u^n \cdot \big(1-u\big)^{(n-1)} \\ &= 2^{n-1}(2n-3)!! \cdot \tfrac{u^n(1-u)^{n-1}}{n!},\end{aligned}$$

where

$$(2k-1)!! = (2k-1)\cdot(2k-3)\cdot\dots\cdot 3\cdot 1.$$

- Specifically, if $u = \frac{1}{2}$,

$$\mathbb{P}\big\{\#\mathcal{G}_{\mathrm{Geom}(\frac{1}{2})} = n\big\} = 2^{-n}\tfrac{1}{n!}(2n-3)!!.$$

Total progeny: asymptotic behavior as $n \to \infty$

Apply the **local central limit theorem**.

Theorem (Local CLT)

Let X_1, X_2, ... be i.i.d. with finite second moment and positive variance $\sigma^2 > 0$. Then

$$\sup_{k \in \mathcal{N}} \left| \sigma \sqrt{n} \mathbb{P}\{X_1 + X_2 + \ldots + X_n = k\} - \tfrac{1}{\sqrt{2\pi}} e^{-\frac{k^2}{2\sigma^2 n}} \right| \underset{n \to \infty}{\longrightarrow} 0.$$

The asymptotic behavior for progeny distribution is well-known. Assume that $\mathcal{G}$ is a Galton-Watson tree whose offspring distribution $p(\cdot)$ has finite second moment and positive variance $\sigma^2 > 0$. Let
$d := \text{g.c.d.}\{i \in \mathbb{N} : \ p(i) > 0\}$. Note that GW-tree can only have sizes that are 1 modulo d. Thus if $n_\ell = d\ell + 1$,

$$\mathbb{P}\{\#\mathcal{G} = n\} \sim \tfrac{d}{\sqrt{2\pi\sigma^2 n^3}}.$$

Conditioning on total progeny: assumption

Assume that the offspring distribution $p(\cdot)$ is such that the generating function $g(s) = \sum_{k\geq 0} s^k p(k)$ satisfies

$$\exists\, a > 0 : \quad g(a) = ag'(a), g''(a) < \infty \qquad (2)$$

Note. Assumption (2) is always satisfied if $\mu > 1$. In the case $\mu = 1$ it is satisfied if $\sigma^2 < \infty$. In the case $\mu < 1$ the assumption requires the $p(1), p(2), \ldots$ to decay exponentially.

Lemma (Kennedy (1975), [6])

Let $\mathcal{G}$ be a GW-tree whose offspring generating function satisfies (2). Put $\bar{g}(s) := g(as)/g(a)$ (equivalently, $\bar{p}(n) := \frac{a^n}{g(a)} p(n), \;\; n = 1, 2, \ldots$). Then this offspring distribution is critical and the corresponding Galton-Watson tree satisfies for each $n \in \mathbb{N}$,

$$\mathcal{L}\big(\mathcal{G} \big| \#\mathcal{G} = n\big) = \mathcal{L}\big(\bar{\mathcal{G}} \big| \#\bar{\mathcal{G}} = n\big).$$

In words, a GW tree conditioned on fixed progeny looks always like a critical GW tree conditioned on total progeny.

Conditioning on total progeny: formulation via GW-processes

- Define the new offspring distribution

$$\bar{p}(k) := \tfrac{a^k}{g(a)} p(k), \quad k = 0, 1, 2, \ldots$$

- Let $\{Z_k;\ k = 0, 1, \ldots\}$ and $\{\bar{Z}_k;\ k = 0, 1, \ldots\}$ be the Galton-Watson processes with offspring distributions $p(\cdot)$ and $\bar{p}(\cdot)$, rsespectively. Moreover, put $N := \sum_{i=0}^{\infty} Z_i$ and $\bar{N} := \sum_{i=0}^{\infty} \bar{Z}_i$.

Lemma

For each $n \in \mathbb{N}$, $0 \le k_1, \ldots, k_j \le n$,

$$\mathbb{P}\big\{Z_{k_1} = r_1, \ldots, Z_{k_j} = r_j \big| N = n\big\} = \mathbb{P}\big\{\bar{Z}_{k_1} = r_1, \ldots, \bar{Z}_{k_j} = r_j \big| \bar{N} = n\big\}.$$

Lemma

For each $n \in \mathbb{N}$, $0 \le k_1 < \ldots < k_j \le n$,

$$\mathbb{P}\big\{Z_{k_1} = r_1, \ldots, Z_{k_j} = r_j \big| N = n\big\} = \mathbb{P}\big\{\bar{Z}_{k_1} = r_1, \ldots, \bar{Z}_{k_j} = r_j \big| \bar{N} = n\big\}.$$

Proof. Let $N_k := \sum_{n=0}^{k} Z_n$ and $\bar{N}_k := \sum_{n=0}^{k} \bar{Z}_n$ be the total numbers in the first k generations. Then

$$\begin{aligned}
&\mathbb{P}\big\{Z_{k_1} = r_1, \ldots, Z_{k_j} = r_j \big| N = n\big\} \\
&= \frac{\mathbb{P}\big\{Z_{k_1} = r_1, \ldots, Z_{k_j} = r_j, N = n\big\}}{\mathbb{P}\big\{N = n\big\}} \\
&= \sum_{s=1}^{n} \mathbb{P}\big\{Z_{k_1} = r_1, \ldots, Z_{k_j} = r_j, N_{k_j} = s\big\} \frac{\mathbb{P}\big\{N^1 + \ldots + N^{r_j} = n - s + r_j\big\}}{\mathbb{P}\big\{N = n\big\}},
\end{aligned}$$

where N^1, N^2, ... are i.i.d. with the same distribution as $\#\mathcal{G}_{p(\cdot)}$. The claim follows by exploiting our transformation as follows:

Consequences of the offspring distribution transformation

1. $$\mathbb{P}\{S_n = j\} = \sum_{i_1,\ldots,i_n:\,\sum i_k = j} p(i_1+1)\cdot\ldots\cdot p(i_n+1)$$
$$= \tfrac{g(a)^n}{a^{n+j}}\cdot\mathbb{P}\{\bar{S}_n = j\}.$$

2. Specifically, for $j=-1$, $\mathbb{P}\{N=n\} = \frac{g(a)^n}{a^{n-1}}\mathbb{P}\{\bar{N}=n\}$.

3. As before let N^1, N^2, ... be i.i.d. with the same distribution as $\#\mathcal{G}_{p(\cdot)}$, and $\bar{N}^1$, $\bar{N}^2$, ... be i.i.d. with the same distribution as $\#\mathcal{G}_{\bar{p}(\cdot)}$. Similar as before we conclude that
$$\mathbb{P}\{N^1+\ldots+N^r = k\} = \tfrac{g(a)^k}{a^{r-k}}\mathbb{P}\{\bar{N}^1+\ldots+\bar{N}^r = k\}.$$

4. $$\mathbb{P}\big\{Z_{k_1}=r_1,\ldots,Z_{k_j}=r_j,N_{k_j}=s\big\}$$
$$=\frac{g(a)^{s-r_j}}{a^{s-1}}\mathbb{P}\big\{\bar{Z}_{k_1}=r_1,\ldots,\bar{Z}_{k_j}=r_j,\bar{N}_{k_j}=s\big\}. \qquad \square$$

Example: Binary branching trees

Assume that for some $v \in (0,1)$,

$$p_v(0) := (1 - v) \text{ and } p_v(2) = v.$$

Then $\bar{p}(\cdot)$ is binary as well. By criticality,

$$\bar{p}(0) = \bar{p}(2) = \tfrac{1}{2}.$$

Lemma

Any binary GW-tree conditioned on total progeny n equals in distribution.

Binary GW-trees conditioned on total progeny equals the random rooted, binary ordered trees

Lemma

Denote by $\mathbb{T}_\ell^{(2)}$ *the set of binary, rooted ordered trees with* ℓ *leaves,* $\ell = 1, 2, \ldots$. *Then* $\#\mathbb{T}_\ell^{(2)} = 2^{\ell-1}(2\ell-3)!!\frac{1}{(\ell-1)!}$.

Proof. Let $\mathcal{G}$ denote the binary, rooted GW-tree. For each $\mathfrak{t} \in \mathbb{T}_\ell^{(2)}$, $\ell = 1, 2, \ldots$

$$\begin{aligned}
\mathbb{P}\Big\{\mathcal{G} = \mathfrak{t}\Big|\#\mathcal{G} = 2\ell - 1\Big\} &= \frac{\mathbb{P}\{\mathcal{G}=\mathfrak{t}\}}{\mathbb{P}\{\#\mathcal{G}=2\ell-1\}} \\
&= \frac{(1-v)^\ell v^{\ell-1}}{\frac{2^{\ell-1}}{(\ell-1)!}(2\ell-3)!!(1-v)^\ell v^{\ell-1}} \\
&= \frac{(\ell-1)!}{2^{\ell-1}(2\ell-3)!!}. \qquad \square
\end{aligned}$$

As we have shown before that all critical, binary GW-trees with a fixed number of vertices, (or equivalently, fixed number of leaves) is equally likely, the claim follows.

Example: Geometric Galton-Watson tree

Assume that for some $u \in (0,1)$,

$$p_u(k) := u \cdot \big(1-u\big)^k, \quad k \geq 0.$$

Then

$$\bar{p}_u(k) := u \cdot \big(1-u\big)^k \cdot \tfrac{a^k}{g(a)}, \quad k \geq 0.$$

Thus $\bar{p}_u$ is again geometrically distributed and by criticality,

$$\bar{p}_u(k) := 2^{-(k+1)}, \quad k = 0, 1, 2, \ldots$$

Lemma

Any geometric GW-tree conditioned on total progeny n equals in distribution.

Geometric GW-tree conditioned on total progeny is uniform rooted, ordered tree

Recall that the number of rooted, ordered trees with n vertices equals

$$\#\mathbb{T}_n := 2^{n-1}\tfrac{1}{n!}(2n-3)!!.$$

Proposition

Let $\mathcal{G}$ be the geometric GW-tree with mean offspring 1. Then for all $\mathfrak{t} \in \mathbb{T}_n$, $n \geq 1$,

$$\mathbb{P}\{\mathcal{G} = \mathfrak{t} \big| \#\mathcal{G} = n\} = \big(\#\mathbb{T}_n\big)^{-1}.$$

Proof. For each $\mathfrak{t} \in \mathbb{T}_n$,

$$\begin{aligned}\mathbb{P}\{\mathcal{G} = \mathfrak{t} \big| \#\mathcal{G} = n\} &= \tfrac{\mathbb{P}\{\mathcal{G}=\mathfrak{t}\}}{\mathbb{P}\{\#\mathcal{G}=n\}} \\ &= \frac{n!2^{-(2n-1)}}{2^{-n}(2n-3)!!} \\ &= \frac{n!}{2^{n-1}(2n-3)!!} \qquad \square\end{aligned}$$

Example: Poisson Galton-Watson tree

Assume that for some $\lambda > 0$,

$$p_\lambda(k) := \tfrac{\lambda^k}{k!} e^{-\lambda}, \quad k \geq 0.$$

Then

$$g_\lambda(s) := \sum_{k\geq 0} s^k \tfrac{\lambda^k}{k!} e^{-\lambda} = \exp\big(-\lambda(1-s)\big), \;\; s > 0$$

and

$$s g'_\lambda(s) = g_\lambda(s) \text{ iff } s := \lambda^{-1}, \text{ and } g''_\lambda(\lambda^{-1}) = \lambda^2 \exp\big(1-\lambda\big) < \infty.$$

We find that with $s_0 := \lambda^{-1}$

$$\bar{g}_\lambda(s) = \tfrac{g_\lambda(s_0 s)}{g_\lambda(s_0)} = \exp\big(-(1-s)\big) = g_1(s).$$

Lemma

Any Poisson GW tree conditioned on total progeny n equals in distribution.

Cayley trees: Random rooted, unordered trees

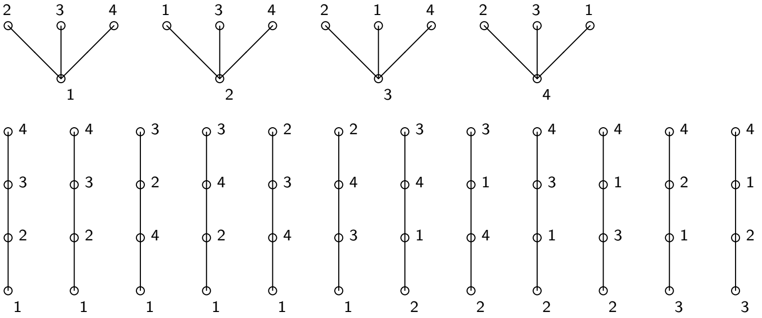
- Consider now a tree as a set of vertices with an edge being an **unordered pair** of vertices.
- For a **labelled** tree with n vertices, the vertices are labelled by $1, 2, ..., n$.
- Labelled trees $\mathfrak{t}$ and $\mathfrak{t}'$ are **isomorphic** iff for each pair (i, j) of labels, (i, j) is an edge in $\mathfrak{t}$ iff it is an edge in $\mathfrak{t}'$.
- Denote by $\mathbf{T}_{[n]}$ the set of **all isomorphy classes** of labelled trees.

Proposition (Cayley's formula)

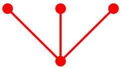
For all $n \geq 2$, $\#\mathbf{T}_{[n]} = n^{n-2}$.

- Two **unlabelled** trees are **isomorphic** iff there exist labellings making them isomorphic as labelled trees.
- Denote by $\widetilde{\mathbf{T}}_{[n]}$ the set of all isomorphy classes of rooted unlabelled trees, and by $\widetilde{\mathfrak{t}}$ the isomorphy class to which $\mathfrak{t}$ belongs.

There are $4^2 = 16$ labelled, unrooted, unordered trees but only 2 unlabelled, unrooted trees with 4 vertices. Thus $\#\mathbf{T}_{[4]} = 4^3 = 48$ and $\#\widetilde{\mathbf{T}}_{[4]} = 8$.



probability 3/4



probability 1/4

Number of rooted, unordered, unlabelled trees

Proposition (Pitman (1997), [8])

There are exactly

$$\frac{n!}{\prod_{v \in V(\mathcal{G})} c_v(\mathfrak{t})!}$$

distinct ways (up to isomorphy) to label a given rooted, unlabelled tree $\mathfrak{t}$ *with* n *vertices.*

PGW conditioned on total progeny equals the uniform unordered tree

Proposition (Aldous (1991), [2])

Let $\mathcal{G}$ be the PGW(1). Then for all $\mathfrak{t} \in \widetilde{\mathbf{T}}_{[n]}$, $n \geq 1$,

$$\mathbb{P}\{\widetilde{\mathcal{G}} = \mathfrak{t} \big| \#\mathcal{G} = n\} = \big(\#\widetilde{\mathbf{T}}_{[n]}\big)^{-1}.$$

Proof. For each $\mathfrak{t} \in \widetilde{\mathbf{T}}_{[n]}$,

$$\begin{aligned}
\mathbb{P}\{\widetilde{\mathcal{G}} = \mathfrak{t} \big| \#\mathcal{G} = n\} &= \frac{\mathbb{P}\big\{\widetilde{\mathcal{G}}=\mathfrak{t}\big\}}{\mathbb{P}\big\{\#\mathcal{G}=n\big\}} \\
&= \frac{n!e^n \prod_{v\in V(\mathcal{G})} p(c_v(\mathfrak{t}))}{n^{n-1}} \\
&= \frac{n!e^n \prod_{v\in V(\mathcal{G})} \frac{e^{-1}}{c_v(\mathfrak{t})!}}{n^{n-1}} \\
&= \frac{\frac{n!}{\prod_{v\in V(\mathcal{G})} c_v(\mathfrak{t})!}}{n^{n-1}} \qquad \square
\end{aligned}$$

Mean number of leaves

Lemma

Let $\mathcal{G}_p$ be a Galton-Watson tree with offspring distribution $p(\cdot)$. Then

$$\mathbb{E}\big[\#\mathrm{Lf}(\mathcal{G}_p)\big] = \frac{p(0)}{1 - \sum_{n\geq 1} np(n)}.$$

Proof. By the branching property, for all $n \geq 1$, $\ell \geq n$,

$$\mathbb{P}\big\{\#\mathrm{Lf}(\mathcal{G}_p) = \ell \big| c_\rho(\mathcal{G}_p) = n\big\} = \mathbb{P}\big\{\sum_{i=1}^{n} L_i = \ell\big\},$$

where $L_1, L_2, \ldots$ are i.i.d. copies of $\#\mathrm{Lf}(\mathcal{G}_p)$. Thus

$$\mathbb{E}\big[\#\mathrm{Lf}(\mathcal{G}_p)\big] = p(0) + \mathbb{E}\big[c_\rho(\mathcal{G}_p)\big]\mathbb{E}\big[\#\mathrm{Lf}(\mathcal{G}_p)\big],$$

which gives $\mathbb{E}\big[\#\mathrm{Lf}(\mathcal{G}_p)\big] = \frac{p(0)}{1-\sum_{n\geq 1} np(n)}$. □

Lemma

Let $\mathcal{G}_p$ be the GW-tree with offspring distribution $p(\cdot)$, and let $\#\mathrm{Lf}(\mathcal{G}_p)$ denote its number of leaves. Then for all $n \geq 0$, there exists a constant $C_p(n)$ such that

$$\mathbb{P}\{\#\mathrm{Lf}(\mathcal{G}_p) = n\} = p^n(0) \cdot C_p(n).$$

Proof. W.l.o.g. assume $p(1) < 1$. If $\mathfrak{t}$ is a rooted ordered family tree **with** m **inner nodes (including the root)** whose offspring numbers are $a_1, a_2, ..., a_m$, then

$$\#\mathrm{Lf}(\mathfrak{t}) = a_1 + a_2 + ... + a_m - m + 1,$$

and thus

$$\mathbb{P}\{\mathcal{G}_p = \mathfrak{t}\} = p(a_1) \cdot p(a_2) \cdot ... \cdot p(a_m) \cdot p^{a_1 + a_2 + ... + a_m - m + 1}(0).$$

Therefore

$$\mathbb{P}\{\#\mathrm{Lf}(\mathcal{G}_p) = n\} = p^n(0) \cdot \sum\nolimits_{\mathfrak{t}, \#\mathrm{Lf}(\mathfrak{t}) = n} p(a_1) \cdot p(a_2) \cdot ... \cdot p(a_m) =: p^n(0) \cdot C_p(n). \quad \square$$

Conditioning on the number of leaves

Proposition (Abraham, Delmas & He (2012), [1])

Let $p(\cdot)$ and $q(\cdot)$ be two offspring distributions. Let $\mathcal{G}_p$ and $\mathcal{G}_q$ be the associated Galton-Watson trees and let $\#\mathrm{Lf}(\mathcal{G}_p)$ and $\#\mathrm{Lf}(\mathcal{G}_q)$ denote their number of leaves. Then for all $n \geq 0$,

$$\mathbb{P}\big(\mathcal{G}_p \in \cdot \big| \#\mathrm{Lf}(\mathcal{G}_p) = n\big) = \mathbb{P}\big(\mathcal{G}_q \in \cdot \big| \#\mathrm{Lf}(\mathcal{G}_q) = n\big)$$

if and only if there exists a $u > 0$ such that for all $k \geq 1$,

$$q(k) = u^{k-1} \cdot p(k).$$

Proof. For all $n \geq 1$ and trees $\mathfrak{t}$ with inner node degrees $(a_1, \ldots, a_m)$ such that $\sum_{i=1}^{m} a_i = n + m - 1$,

$$\mathbb{P}\big(\mathcal{G}_p = \mathfrak{t} \big| \#\mathrm{Lf}(\mathcal{G}_p) = n\big) = \mathbb{P}\big(\mathcal{G}_q = \mathfrak{t} \big| \#\mathrm{Lf}(\mathcal{G}_q) = n\big) \quad \Leftrightarrow \quad \frac{p(a_1)\ldots p(a_m)}{C_p(n)} = \frac{q(a_1)\ldots q(a_m)}{C_q(n)}.$$

If $n = 1$, all trees with 1 leaf are those with one offspring each generation until the last individual dies. Thus for all $k \geq 0$, $C_p(1) = 1/(1 - p(1))$ and $C_q(1) = 1/(1 - q(1))$, and hence

$$p^k(1)\big(1 - p(1)\big) = q^k(1)\big(1 - q(1)\big).$$

We can therefore conclude that $p(1) = q(1)$.

Continuation of Proof. Let $n_0 := \min\{n \geq 2 : \; p(n) > 0\}$, and choose

$$u := \left(\frac{q(n_0)}{p(n_0)}\right)^{1/(n_0-1)}$$

If $p(0) + p(1) + p(n_0) = 1$, $q(k) = u^{k-1} \cdot p(k)$ trivially holds for all $k \geq 1$. On the other hand, for all $n > n_0$ with $p(n_0) > 0$, put $N := 2(n-1)(n_0-1)$. For any tree $\mathfrak{t}$ with $N+1$ leaves, $n-1$ inner nodes with n_0 offspring and $n_0 - 1$ inner nodes with n offspring, we conclude that

$$\frac{p^{n-1}(n_0)p^{n_0-1}(n)}{C_p(N+1)} = \frac{q^{n-1}(n_0)q^{n_0-1}(n)}{C_q(N+1)}.$$

Moreover, for any tree $\mathfrak{t}$ with $N+1$ leaves and $2(n-1)$ inner nodes with n_0 offspring, we conclude that

$$\frac{p^{2(n-1)}(n_0)}{C_p(N+1)} = \frac{q^{2(n-1)}(n_0)}{C_q(N+1)}.$$

Dividing the two latter equations implies that for all $n \geq 1$,

$$q(n) = u^{n-1}p(n).$$

Proof of conditioning on the number of leaves: II

Conversely, lets suppose that for all $n \geq 1$,

$$q(n) = u^{n-1} q(n).$$

Then for all $n \geq 1$ with $C_p(n) \neq 0$, and for every tree $\mathfrak{t}$ with n leaves,

$$\begin{aligned} q(a_1)...q(a_m) &= u^{a_1 - 1} p(a_1) ... u^{a_m - 1} p(a_m) \\ &= u^{n-1} p(a_1)...p(a_m). \end{aligned}$$

Thus $C_q(n) = u^{n-1} C_p(n)$, and therefore

$$\frac{q(a_1)\ldots q(a_m)}{C_q(n)} = \frac{u^{a_1-1}p(a_1)\ldots u^{a_m-1}p(a_m)}{u^{n-1}p(a_1)\ldots p(a_m)} = \frac{p(a_1)\ldots p(a_m)}{C_p(n)}$$

which was shown to be equivalent to

$$\mathbb{P}\big(\mathcal{G}_p \in \cdot \big| \#\mathrm{Lf}(\mathcal{G}_p) = n\big) = \mathbb{P}\big(\mathcal{G}_q \in \cdot \big| \#\mathrm{Lf}(\mathcal{G}_q) = n\big). \quad \square$$

Classical problem: Cutting down trees

Given a rooted tree $(\mathfrak{t}, \rho)$.

1. Remove an edge uniformly at random. This disconnects the tree into two subtrees.
2. Destroy the subtree which does not contain the root.
3. We iterate until the are stuck with a tree without edges. That means, the root is isolated.

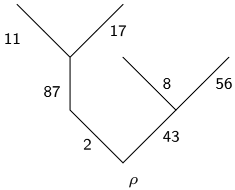
Denote by $N(\mathfrak{t}, \rho)$ the (random) **number of cuts needed to isolate the root**.

Question:

What can we say about the distribution of $N(\mathfrak{t}, \rho)$?

Equivalent formulation in terms of records

- $N(T, \rho)$ appears also when we are consider **records** in a tree.
- Let each edge e have a random value λ_e attached to it, and assume that these values are i.i.d. with a continuous distribution.
- Say that a value λ_e is a **record** if it is the largest value in the path from the root to e.
- Then the number of records equals in distribution $N(T, \rho)$.



- To see this, generate first the values λ_e, and then cut the tree: each time choosing the edge with the largest λ_e among the remaining ones.

Classical record problem

- Take T_n be a path with n edges, from the root to an end.
- Let $N(T_n)$ be the **number of records** on a sequence of n i.i.d. numbers λ_1, ..., λ_n.
- Let A_j be the event that λ_j is a record. Then $\mathbb{P}(A_j) = \frac{1}{j}$, so $\mathbf{1}_{A_j}$ is Bernoulli distributed with success parameter $\frac{1}{j}$. Thus

$$\mathbb{E}\big[N(T_n)\big] = \sum_{i=1}^{n} \frac{1}{j} \sim \ln n.$$

- Moreover, A_1,A_2, ..., A_n are independent and satisfy the **Lyapunov condition**. Hence the **CLT** holds:

$$\frac{N(T_n) - \ln n}{\sqrt{\ln n}} \xrightarrow[n\to\infty]{\text{w}} \mathcal{N}(0,1).$$

Janson's result for Galton-Watson trees

Theorem (Janson (2006), [5])

Let $\mathcal{G}_n$ be the GW-tree with offspring distribution $p(\cdot)$ conditioned to have n vertices. Assume that $p(\cdot)$ is critical, $p(1) < 1$, and $p(\cdot)$ has finite variance σ^2. Then

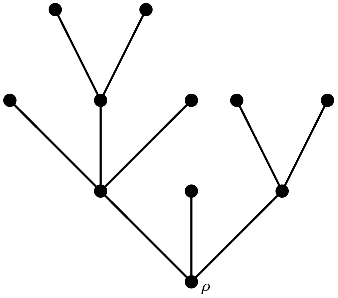
$$\mathbb{P}\left\{N(\mathcal{G}_n, \rho) \geq x\sqrt{n}\sigma\right\} \underset{n\to\infty}{\longrightarrow} e^{-x^2/2}.$$

Proof will be given in Part II. □

Remark. The limit distribution is known as **Rayleigh distribution**.

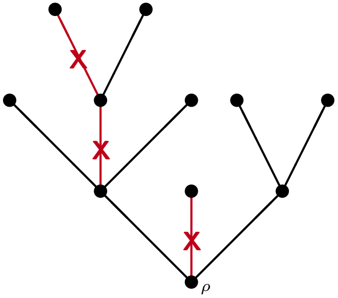
Pruning finite trees: edge percolation

1. Consider a **rooted**, finite **tree** $(\mathfrak{t}, \rho)$



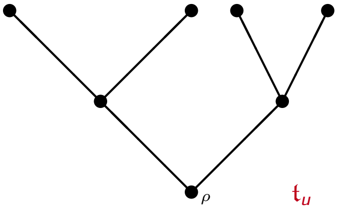
Pruning finite trees: edge percolation

1. Consider a **rooted**, finite **tree** $(\mathfrak{t}, \rho)$
2. Mark edges independently with probability $1 - u$



Pruning finite trees: edge percolation

1. Consider a **rooted**, finite **tree** $(\mathfrak{t}, \rho)$
2. Mark edges independently with probability $1 - u$
3. Call the unmarked component containing ρ the **pruned tree** $\mathfrak{t}_u$



Pruning finite trees: edge percolation

1. Consider a **rooted**, finite **tree** $(\mathfrak{t}, \rho)$
2. Mark edges independently with probability $1 - u$
3. Call the unmarked component containing ρ the **pruned tree** $\mathfrak{t}_u$
4. Couple different pruning procedures such that $\mathfrak{t}_u \subseteq \mathfrak{t}_v$, $u \leq v$, and obtain a non-decreasing process $(\mathfrak{t}_u)_{u\in[0,1]}$

Edge percolation of Galton-Watson trees

1. Consider a GW-tree $\mathcal{G}$ with offspring distribution $p(\cdot)$.

Edge percolation of Galton-Watson trees

1. Consider a GW-tree $\mathcal{G}$ with offspring distribution $p(\cdot)$.
2. Mark edges independently with probability $1 - u$.

Edge percolation of Galton-Watson trees

1. Consider a GW-tree $\mathcal{G}$ with offspring distribution $p(\cdot)$.
2. Mark edges independently with probability $1-u$.

Lemma (Lyons (1992) [7])

The pruned tree $\mathcal{G}_u$ is a GW-tree with offspring generating function

$$g_u(s) = g_1(1-u+us), \quad s \in (0,1).$$

In particular, if $\mathcal{G}$ is PGW(μ) then $\mathcal{G}_u$ is PGW($u\mu$).

3. **Proof.** Given the size of the first generation of $\mathcal{G}$ is Z_1. Then the size $Z_1^{(u)}$ of the first generation of $\mathcal{G}_u$ is distributed as the sum of Z_1 independent Bernoulli variables. Thus for all $s \in (0,1)$,

$$\begin{aligned} g_u(s) &:= \mathbb{E}\big[s^{Z_1^{(u)}}\big] = \mathbb{E}\big[\mathbb{E}\big[s^{Z_1^{(u)}}\big|Z_1\big]\big] \\ &= \mathbb{E}\big[\big((1-u)+su\big)^{Z_1}\big] = g_1\big(1-u+su\big). \end{aligned}$$

If $g_1(s) := e^{-\mu(1-s)}$, then $g_u(s) = e^{-u\mu(1-s)}$. □

Joint distribution of pruned and unpruned GW tree

- For $c, m \geq 0$ and $0 < \alpha < \beta < 1$, denote

$$\bar{P}_{\alpha,\beta}(c;m) := \mathbb{P}\{c_\rho(\mathcal{G}_\beta) - c_\rho(\mathcal{G}_\alpha) = m \big| c_\rho(\mathcal{G}_\alpha) = c\}.$$

- Denote by $p_\alpha(\cdot)$ and $p_\beta(\cdot)$ the offspring laws of the tree pruned with parameter $\alpha, \beta \in [0,1]$. Then obviously,

$$\bar{P}_{\alpha,\beta}(c;m) = \frac{p_\beta(m+c)}{p_\alpha(c)} \binom{m+c}{c} \left(\frac{\alpha}{\beta}\right)^c \left(1 - \frac{\alpha}{\beta}\right)^m.$$

Corollary (Rao & Rubin (1964), [9])

$\bar{P}_{\alpha,\beta}(c;m)$ does not depend on c iff $p_\beta(\cdot)$ is Poisson distributed. That is, $c_\rho(\mathcal{G}_\alpha)$ and $c_\rho(\mathcal{G}_\beta) - c_\rho(\mathcal{G}_\alpha)$ are **independent** *if and only if $\mathcal{G}_\beta$ is a Poisson GW-tree.*

I will leave the **proof** for you as an **exercise**.

Representation of the (un)pruned tree

Proposition (Aldous & Pitman (1998), [3])

Fix $0 < \alpha < \beta < 1$*. Given* $\mathcal{G}_\alpha$*, let* $\{K_\alpha(v),\ v \in V(\mathcal{G}_\alpha)\}$ *be a independent family with*

$$\mathbb{P}\big\{K_\alpha(v) = k\big\} = \bar{P}_{\alpha,\beta}\big(c_v(\mathcal{G}_\alpha), k\big),\quad k = 0, 1, \ldots$$

Moreover, given $\{K_\alpha(v),\ v \in V(\mathcal{G}_\alpha)\}$*, let* $\widetilde{\mathcal{G}}_\beta$ *be defined by random attachments of* $K_\alpha(v)$ *independent copies of* $\mathcal{G}_\beta$ *at vertex* v*. Then*

$$(\mathcal{G}_\alpha, \mathcal{G}_\beta) \stackrel{(d)}{=} (\mathcal{G}_\alpha, \widetilde{\mathcal{G}}_\beta)$$

Sketch of proof.

- Conditionally given the pruned tree $\mathcal{G}_\alpha$, the family $\{c_v(\mathcal{G}_\beta) - c_v(\mathcal{G}_\alpha);\ v \in V(\mathcal{G}_\alpha)\}$ is independent, and thus distributed as the family $\{K_\alpha(v),\ v \in V(\mathcal{G}_\alpha)\}$.
- Each of the children of $v \in \mathcal{G}_\alpha$ in $\mathcal{G}_\beta$ is the root of a subtree of $\mathcal{G}_\beta$ which - identified as a family tree is an independent copy of $\mathcal{G}_\beta$. □

Pruning Poisson GW-trees: the total progeny

Corollary

Fix $0 \le \alpha < \beta < \infty$. *Assume that* $\mathcal{G}_1$ *is a PGW*(λ)*-tree. Given* $\mathcal{G}_\alpha$, *let* $\{N_{\alpha,\beta}(v);\, v \in V(\mathcal{G}_\alpha)\}$ *be an i.i.d. family with Poisson*$((\beta-\alpha)\lambda)$ *distribution, and put*

$$N_{\alpha,\beta} := \sum_{v \in V(\mathcal{G}_\alpha)} N_{\alpha,\beta}(v).$$

Moreover, let $\mathcal{G}_\beta^1, \mathcal{G}_\beta^2, \ldots$ *be independent copies of* $\mathcal{G}_\beta$. *Then*

$$\left(\mathcal{G}_\alpha, \#\mathcal{G}_\beta\right) \stackrel{(d)}{=} \left(\mathcal{G}_\alpha, \#\mathcal{G}_\alpha + \sum_{i=1}^{N_{\alpha,\beta}} \#\mathcal{G}_\beta^i\right).$$

Pruning Poisson GW-trees: a martingale

Proposition (Aldous & Pitman (1998), [3])

Let $\mathcal{G}$ be the PGW(μ) with $\mu < 1$, and $\{\mathcal{G}_u; u \in [0,1]\}$ be the pruned process. Then $(\#\mathcal{G}_u)_{u\in[0,1]}$ is a Markov process, and the process

$$\{(1-\mu u)\#\mathcal{G}_u;\ u \in [0,1]\}$$

is a martingale w.r.t. the filtration generated by $\{\mathcal{G}_u,\ u \in [0,1]\}$.

Proof. Recall that $\mathcal{G}_u$ is PGW($u\mu$), and thus $\mathbb{E}\big[\#\mathcal{G}_u\big] = \big(1-u\mu\big)^{-1}$. Using the representation given before, for $0 \le \alpha < \beta \le 1$,

$$\begin{aligned}\mathbb{E}\big[\#\mathcal{G}_\beta \big| \mathcal{G}_\alpha\big] &= \#\mathcal{G}_\alpha + \#\mathcal{G}_\alpha(\beta-\alpha)\mu\mathbb{E}\big[\#\mathcal{G}_\beta\big] \\ &= \#\mathcal{G}_\alpha + \#\mathcal{G}_\alpha(\beta-\alpha)\mu\big(1-\mu\beta\big)^{-1} = \tfrac{1-\alpha\mu}{1-\beta\mu}\#\mathcal{G}_\alpha.\end{aligned}$$

With the **Markov property** we conclude that

$$\mathbb{E}\big[(1-\beta\mu)\#\mathcal{G}_\beta \big| \{\mathcal{G}_{\alpha'}, \alpha' \in [0,\alpha]\}\big] = \mathbb{E}\big[(1-\beta\mu)\#\mathcal{G}_\beta \big| \mathcal{G}_\alpha\big] = (1-\alpha\mu)\#\mathcal{G}_\alpha. \qquad \square$$

There are also vertex versions for cuttings and records:

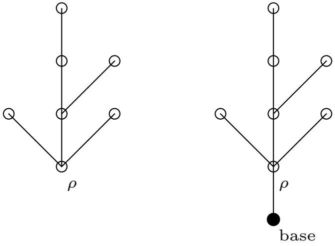
- For cuttings, choose a vertex at random and destroy it together with all its descendants. Continue until the root is chosen and thus the whole tree is destroyed.
- For records, we assign i.i.d. values λ_v (or a random permutation) to the vertices, and define a record as above.

Again, vertex cutting and records are equivalent: Denote by

$N_{\text{vertex}}(\mathfrak{t}, \rho)$

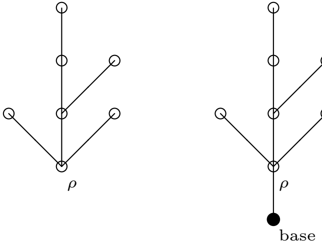
$:= \#$ **number of vertex deletions needed to destroy the tree.**

Given a rooted tree $(\mathfrak{t}, \rho)$ with n vertices. We add a new vertex, called the **base** and link it to the root ρ of $\mathfrak{t}$ by a new edge. This gives a **planted tree** which we denote by $\bar{\mathfrak{t}}$. The set $\bar{E}$ of edges of $\bar{\mathfrak{t}}$ is thus the set E of edges of $\mathfrak{t}$ plus the newly inserted edge.



Duality between rooted tree and planted tree

We consider $\bar{E}$ as a set of vertices, and endow it with a natural tree structure by declaring that e and e' are neighbors if and only if the are adjacent in $\bar{\mathfrak{t}}$. The map $v : \bar{E} \to V(E)$ that associates to an edge e of $\bar{\mathfrak{t}}$ its end point $v(E)$ which is further away from the base is bijective and preserves the tree structure.



Corollary

Any statement expressed in terms of the edges of the planted tree $\bar{\mathfrak{t}}$ can thus be rephrased in terms of the vertices of $\mathfrak{t}$ and vice versa.

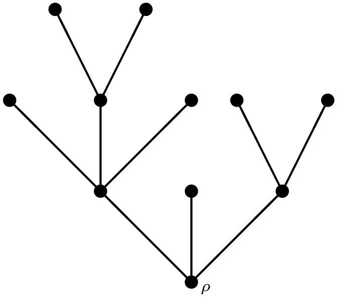
Duality between rooted tree and planted tree

Corollary

The distributions of $N(\mathfrak{t}, \rho)$ *and of* $N_{\mathrm{vertex}}(\mathfrak{t}, \rho)$ *agree.*

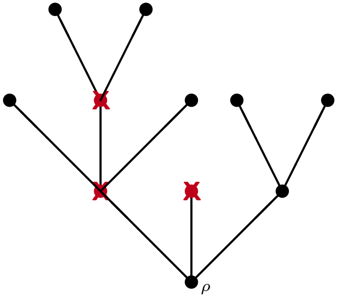
A inhomogeneous pruning

1. Consider a **rooted**, finite **tree** $(\mathfrak{t}, \rho)$



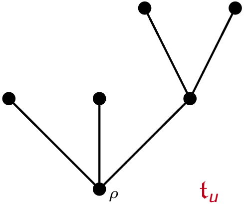
A inhomogeneous pruning

1. Consider a **rooted**, finite **tree** $(\mathfrak{t}, \rho)$
2. Mark vertices independently with probability $1 - u^{\textbf{(\#children-1)}}$



A inhomogeneous pruning

1. Consider a **rooted**, finite **tree** $(\mathfrak{t}, \rho)$
2. Mark vertices independently with probability $1 - u^{(\text{\#children-1})}$
3. Call the unmarked component containing ρ the **pruned tree** $\mathfrak{t}_u$



A inhomogeneous pruning

1. Consider a **rooted**, finite **tree** $(\mathfrak{t}, \rho)$
2. Mark vertices independently with probability $1 - u^{\textbf{(\#children-1)}}$
3. Call the unmarked component containing ρ the **pruned tree** $\mathfrak{t}_u$
4. Couple different such that $\mathfrak{t}_u \subseteq \mathfrak{t}_v$, $u \leq v$, and obtain a non-decreasing process $(\mathfrak{t}_u)_{u \in [0,1]}$

Inhomogeneous pruning of GW-trees

1. Consider a GW-tree $\mathcal{G}$ with offspring distribution $p(\cdot)$.

Inhomogeneous pruning of GW-trees

1. Consider a GW-tree $\mathcal{G}$ with offspring distribution $p(\cdot)$.
2. Mark vertices independently with probability $1 - u^{(\text{\#children-1})}$

Inhomogeneous pruning of GW-trees

1. Consider a GW-tree $\mathcal{G}$ with offspring distribution $p(\cdot)$.
2. Mark vertices independently with probability $1 - u^{\textbf{(\#children-1)}}$

Lemma (Abraham, Delmas & He (2012) [1])

The pruned tree $\mathcal{G}_u$ is a GW-tree with offspring distribution $p_u(\cdot)$:

$$p_u(n) = u^{n-1}p(n), \quad n = 1, 2, \ldots \text{ and } p_u(0) = 1 - \sum_{n \geq 1} p_u(n).$$

Equivalently,

$$g_u(s) = 1 - \frac{g_1(u)}{u} + \frac{g_1(su)}{u}, \quad s \in (0, 1).$$

3. **Proof** follows same lines of argument as in the homogeneous case. I will leave it for you as an **exercise**. □

Representation of the (un)pruned tree

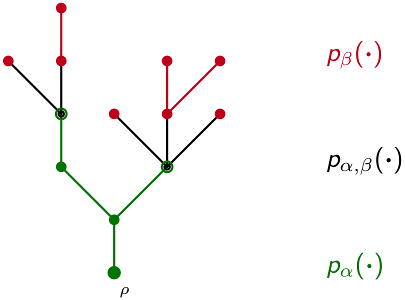
Proposition (Abraham, Delmas & He (2012), [1])

Let $\mathcal{G}$ be a GW-tree with offspring distribution $p(\cdot)$, and $(\mathcal{G}_u)_{u\in[0,1]}$ be the inhomogeneous pruning process. Fix $0 < \alpha < \beta < 1$, and put

$$p_{\alpha,\beta}(k) := \frac{1-\left(\frac{\alpha}{\beta}\right)^{k-1}}{p_\alpha(0)} p_\beta(k), \quad k = 1, 2, \dots \text{ and } p_{\alpha,\beta} = \frac{p_\beta(0)}{p_\alpha(0)}.$$

Define the **modified GW-tree** $\mathcal{G}_{\alpha,\beta}$ *in which the size of the first generation has distribution $p_{\alpha,\beta}$, while these and all subsequent individuals have offspring distribution p_β. If $\widehat{\mathcal{G}}_\beta$ denotes the tree obtained from $\mathcal{G}_\alpha$ by attaching to each of the leaves of $\mathcal{G}_\alpha$ independent copies of $\mathcal{G}_{\alpha,\beta}$. Then*

$$(\mathcal{G}_\alpha, \mathcal{G}_\beta) \stackrel{(d)}{=} (\mathcal{G}_\alpha, \widehat{\mathcal{G}}_\beta)$$



Notice that the **number of leaves process**

$$(\#\mathrm{Lf}(\mathcal{G}_u))_{u\in[0,1]}$$

is a **Markov** process **for all offspring distributions**.

Proof of the representation of the (un-)pruned GW-tree

W.l.o.g. assume that $\mathcal{G}_\beta$ is (sub-)critical. Otherwise argue with $(r_h\mathcal{G}_u)_{u\in[0,1]}$. Fix $0\le\alpha<\beta\le 1$, two trees $\mathfrak{s}$, $\mathfrak{t}$ with $\mathfrak{s}$ being a subtree of $\mathfrak{t}$.

- The definition of $\widehat{\mathcal{G}}_\beta$ readily implies

$$\begin{aligned}\mathbb{P}\{\mathcal{G}_\alpha=\mathfrak{s},\widehat{\mathcal{G}}_\beta=\mathfrak{t}\}&=\mathbb{P}\{\mathcal{G}_\alpha=\mathfrak{s}\}\mathbb{P}\big(\widehat{\mathcal{G}}_\beta=\mathfrak{t}\big|\mathcal{G}_\alpha=\mathfrak{s}\big)\\&=\prod_{v\in V(\mathfrak{s})}p_\alpha(c_v(\mathfrak{s}))\prod_{v\in\mathrm{Lf}(\mathfrak{s})}p_{\alpha,\beta}(c_v(\mathfrak{t}))\prod_{v\in V(\mathfrak{t})\setminus V(\mathfrak{s})}p_\beta(c_v(\mathfrak{t})).\end{aligned}$$

- On the other hand, by the pruning procedure,

$$\begin{aligned}&\mathbb{P}\{\mathcal{G}_\alpha=\mathfrak{s},\mathcal{G}_\beta=\mathfrak{t}\}\\&=\mathbb{P}\{\mathcal{G}_\beta=\mathfrak{t}\}\mathbb{P}\big(\mathcal{G}_\alpha=\mathfrak{t}\big|\mathcal{G}_\beta=\mathfrak{t}\big)\\&=\prod_{v\in V(\mathfrak{t})}p_\beta(c_v(\mathfrak{t}))\prod_{v\in V(\mathfrak{s})\setminus\mathrm{Lf}(\mathfrak{s})}\big(\tfrac{\alpha}{\beta}\big)^{c_v(\mathfrak{t})-1}\prod_{v\in\mathrm{Lf}(\mathfrak{s})\setminus\mathrm{Lf}(\mathfrak{t})}\big(1-\big(\tfrac{\alpha}{\beta}\big)^{c_v(\mathfrak{t})-1}\big)\\&=\prod_{v\in V(\mathfrak{s})\setminus\mathrm{Lf}(\mathfrak{s})}p_\alpha(c_v(\mathfrak{s}))\prod_{v\in\mathrm{Lf}(\mathfrak{s})\setminus\mathrm{Lf}(\mathfrak{t})}\big(1-\big(\tfrac{\alpha}{\beta}\big)^{c_v(\mathfrak{t})-1}\big)\prod_{v\in V(\mathfrak{t})\setminus V(\mathfrak{s})\cup\mathrm{Lf}(\mathfrak{s})}p_\beta(c_v(\mathfrak{t}))\\&=\prod_{v\in V(\mathfrak{s})}p_\alpha(c_v(\mathfrak{s}))\prod_{v\in\mathrm{Lf}(\mathfrak{s})}\frac{p_\beta(c_v(\mathfrak{t}))}{p_\alpha(0)}\big(1-\big(\tfrac{\alpha}{\beta}\big)^{c_v(\mathfrak{t})-1}\mathbf{1}_{\{c_v(\mathfrak{t})>1\}}\big)\prod_{v\in V(\mathfrak{t})\setminus V(\mathfrak{s})}p_\beta(c_v(\mathfrak{t}))\\&=\prod_{v\in V(\mathfrak{s})}p_\alpha(c_v(\mathfrak{s}))\prod_{v\in\mathrm{Lf}(\mathfrak{s})}p_{\alpha,\beta}(c_v(\mathfrak{t}))\prod_{v\in V(\mathfrak{t})\setminus V(\mathfrak{s})}p_\beta(c_v(\mathfrak{t})).\qquad\square\end{aligned}$$

Inhomogeneous pruning of GW-trees: a martingale

Proposition (Abraham, Delmas & He (2012), [1])

Let $\mathcal{G}$ be a GW with offspring distribution $p(\cdot)$, and $(\mathcal{G}_u)_{u\in[0,1]}$ the inhomogeneous pruning. Denote the mean offspring of $\mathcal{G}_u$ by $\mu(u)$. Then

$$\Big\{\frac{1-\mu(u)}{p_u(0)}\cdot \#\mathrm{Lf}(\mathcal{G}_u);\ u\in(0,1]\Big\}$$

is a martingale.

Proof. A simple calculation shows that $p_{\alpha,\beta}$ has mean

$$\mu_{\alpha,\beta}=\frac{\mu(\beta)-\mu(\alpha)}{p_\alpha(0)}.$$

By the representation of $\mathcal{G}_\beta$ given $\mathcal{G}_\alpha$ and the Markov property,

$$\begin{aligned}\mathbb{E}\big[\#\mathrm{Lf}(\mathcal{G}_\beta)\big|\mathcal{G}_\alpha\big] &= \#\mathrm{Lf}(\mathcal{G}_\alpha)\mathbb{E}\big[\#\mathrm{Lf}(\mathcal{G}_{\alpha,\beta})\big]\\ &= \#\mathrm{Lf}(\mathcal{G}_\alpha)\big(p_{\alpha,\beta(0)}+\mu_{\alpha,\beta}\mathbb{E}\big[\#\mathrm{Lf}(\mathcal{G}_\beta)\big]\big)\\ &= \#\mathrm{Lf}(\mathcal{G}_\alpha)\big(p_{\alpha,\beta(0)}+\mu_{\alpha,\beta}\frac{p_\beta(0)}{1-\mu(\beta)}\big)\\ &= \#\mathrm{Lf}(\mathcal{G}_\alpha)\frac{1-\mu(\alpha)}{p_\alpha(0)}\frac{p_\beta(0)}{1-\mu(\beta)}. \qquad \square\end{aligned}$$

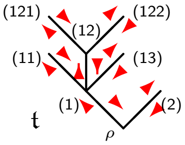
Literature

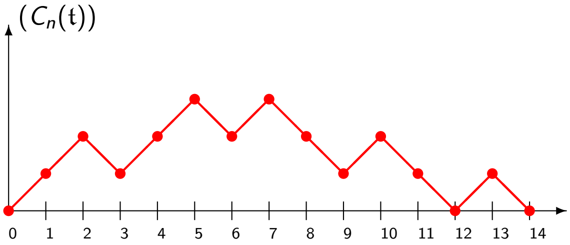
- Romain Abraham, Jean-François Delmas & Hui He (2012). Pruning Galton-Watson trees and tree-valued Markov processes, Annales de l'Institut Henri Poincaré (B).
- David Aldous (1991). The continuum random tree I, Annals of Probability.
- David Aldous & Jim Pitman (1998). Tree-valued Markov chain derived from Galton-Watson trees, Annales de l'Institut Henri Poincaré (B).
- Meyer Dwass (1962). A fluctuation theorem for cyclic random variables, The Annals of Mathematical Statistics.
- Svantje Janson (2006). Random cuttings and records in deterministic and random trees, Random Structures and Algorithms.
- Douglas Kennedy (1975). The Galton-Watson process conditioned on the total progeny, Journal of Applied Probability.
- Russell Lyons (1992). Random walks, capacity and percolation on trees, Annals of Probability.
- Jim Pitman (1997). Enumerations of trees and forests related to branching processes and random walks, Microsurveys in discrete probability.
- Radhakrishna Rao and Herman Rubin (1964). On a characterization of the Poisson distribution, Sankrhya The Indian Journal of Statistics Series (A).

Outline: Part II
THE Continuum Random Tree and continuous pruning procedures

1. Convergence of the Galton-Watson trees
   - Convergence of the contour function
   - Invariance principle via the Lukasiewicz walk
2. Scaling limits
   - The Brownian CRT
   - The Levy tree
3. How many cuts needed to isolate k vertices?
   - How many cuts needed to isolate the root?
   - The cut tree
4. Pruning procedures on continuum trees

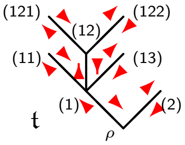
- The **contour function** $(C_n(\mathfrak{t}))_{n=0,1,\ldots,2(\#\mathfrak{t}-1)}$ of a finite rooted, ordered tree $\mathfrak{t}$ was obtained by traversing the tree at speed 1 starting in the root clockwise, and recording the height profile.

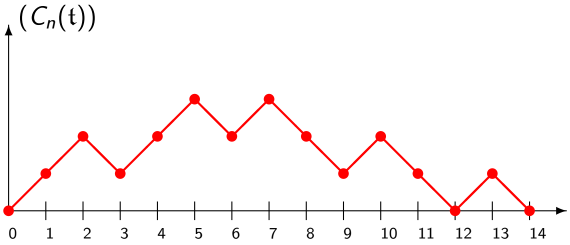




The contour function

- The **contour function** $(C_n(\mathfrak{t}))_{n=0,1,\ldots,2(\#\mathfrak{t}-1)}$ of a finite rooted, ordered tree $\mathfrak{t}$ was obtained by traversing the tree at speed 1 starting in the root clockwise, and recording the height profile.
- Recall that if $\mathcal{G}$ is the GW-tree with geometric offspring distribution, then $(C_n(\mathcal{G}))_{n=0,1,\ldots,2(\#\mathcal{G}-1)}$ has a representation as a **nearest neighbor random walk stopped one time step before it gets negative**.





Proposition

If $\mathcal{G}_n$ is the GW-tree with **geometric offspring** *distribution conditioned to have total progeny n, then*

$$\Big(\tfrac{1}{\sqrt{2n}}\mathcal{C}_{\lfloor 2nt \rfloor}(\mathcal{G}_n)\Big)_{t\in[0,1]} \underset{n\to\infty}{\Longrightarrow} \big(B^{\mathrm{exc}}_t\big)_{t\in[0,1]},$$

where $(B^{\mathrm{exc}}_t)_{t\in[0,1]}$ is the **normalized Brownian excursion**.

Remarks.

- The **normalized Brownian excursion** as the scaling limit is the analogue of standard Brownian motion but conditioned to stay positive for a while, and then come back to zero for the first time at time $t = 1$ (see **Durrett, Iglehart & Miller (1977), [6]**):
  1. Consider a Brownian motion $(B^{\varepsilon}_t)_{t\geq 0}$ starting in $B^{\varepsilon}_0 := \varepsilon > 0$.
  2. Condition $(B^{\varepsilon}_t)_{t\geq 0}$ on the event $\inf\big\{t > 0 : B^{\varepsilon}_t = 0\big\} = 1$.
  3. Let ε tend to zero.
- A more precise construction uses **Ito's excursion theory**.

Aldous' invariance principle

Proposition

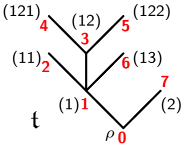
If $\mathcal{G}_n$ is the GW-tree with **general** *critical offspring distribution of finite variance $\sigma^2 > 0$ conditioned to have total progeny n, then*

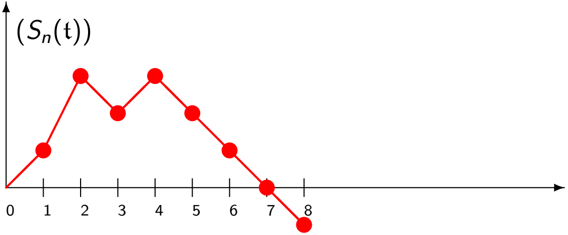
$$\left(\frac{1}{\sqrt{2n}}\mathcal{C}_{\lfloor 2nt \rfloor}(\mathcal{G}_n)\right)_{t\in[0,1]} \underset{n\to\infty}{\Longrightarrow} \left(\frac{\sqrt{2}}{\sigma} B_t^{\mathrm{exc}}\right)_{t\in[0,1]},$$

where $(B_t^{\mathrm{exc}})_{t\in[0,1]}$ is the **normalized Brownian excursion***.*

- This statement agrees with the earlier statement as the critical geometric offspring distribution has variance $\sigma^2 = 2$.
- The proof of the statement follows the line of arguments of the conditioned version of Donsker's theorem if and only if the offspring distribution is geometric.
- For general offspring distributions (finite variance) we could argue by means of the **Lukasiewicz walk**.

- Enumerate the vertices in lexicographic order.
- Define $S_0 := 0$, and for $0 \leq n \leq \#\mathfrak{t} - 1$, $S_{n+1} = S_n + \big(c_{v_n}(\mathfrak{t}) - 1\big)$.





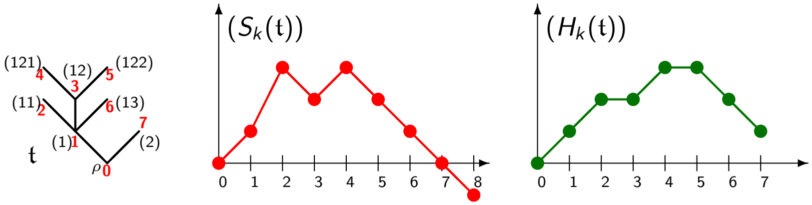
Lemma

If $\mathcal{G}$ is a GW-tree with offspring distribution $p(\cdot)$, then the **Lukasiewicz walk** $(S_n)_{0 \leq n \leq \#\mathcal{G}}$ *is a random walk with jump distribution*

$$\nu(k) = p(k+1), \quad k = -1, 0, \ldots,$$

stopped at its first hitting time of -1.

- We want to link the contour function (which records the height profile while traversing) with the Lukasiewicz walk.
- For that purpose, we traverse the tree in Lukasiewicz's lexicographic order and record the height of a visited vertex.
- The result $(H_k)_{k=0,1,\ldots,\#\mathfrak{t}-1}$ is called the **height function**.



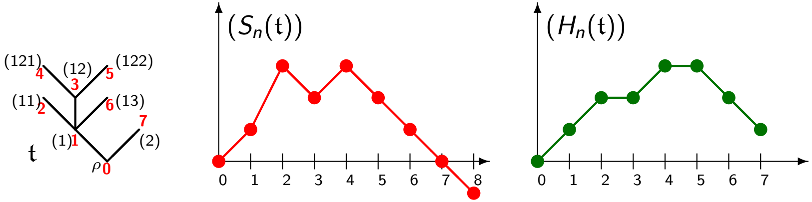
The height function: key formula

- Given the vertex v_k, all vertices in $\mathfrak{t}$ which are on the way from ρ to v_k can be read off the Lukasiewicz walk as

$$\mathcal{H}_k := \Big\{ v_j : 0 \le j < k,\ S_j = \min_{j \le i \le k} S_i \Big\}.$$

- Thus the height H_k of vertex v_k equals

$$H_k := \#\mathcal{H}_k = \#\Big\{ j \in \{0,1,\dots,k-1\} : S_j = \min_{j \le i \le k} S_i \Big\}.$$



For example, $\mathcal{H}_5 := \{0, 1, 3\}$.

Exploiting the Markov property of the Lukasiewicz walk

Proposition (Csaki & Mohanty (1981), [4])

If $\mathcal{G}_n$ is the GW-tree with critical offspring distribution of finite variance $\sigma^2 > 0$ conditioned to have n vertices, then

$$\Big(\frac{1}{\sqrt{n\sigma^2}} S_{\lfloor nt \rfloor}(\mathcal{G}_n)\Big)_{t\in[0,1]} \underset{n\to\infty}{\Longrightarrow} \big(B_t^{\mathrm{exc}}\big)_{t\in[0,1]},$$

where $(B_t^{\mathrm{exc}})_{t\in[0,1]}$ is the **normalized Brownian excursion**.

- The statement is a conditioned version of the classical Donsker's invariance principle.

1. Read off the height function from the Lukasiewicz walk via the **key formula**,

$$H_n := \#\{j \in \{0, 1, ..., n-1\} : S_j = \min_{j \leq i \leq n} S_i\}.$$

Steps in the proof of Aldous' invariance principle

1. Read off the height function from the Lukasiewicz walk via the **key formula**,

$$H_n := \#\{j \in \{0, 1, ..., n-1\} : S_j = \min_{j \le i \le n} S_i\}.$$

2. Show that for critical $p(\cdot)$ the contour process and the height process (up to changing time by a factor of $\frac{1}{2}$) are close, i.e.,

$$n^{-\frac{1}{2}} \sup_{t \in [0,1]} \left| C_{\lfloor 2nt \rfloor}(\mathcal{G}_n) - H_{\lfloor nt \rfloor}(\mathcal{G}_n) \right| \underset{n \to \infty}{\longrightarrow} 0, \quad \text{in probability}.$$

Steps in the proof of Aldous' invariance principle

1. Read off the height function from the Lukasiewicz walk via the **key formula**,

$$H_n := \#\{j \in \{0,1,...,n-1\} : S_j = \min_{j \le i \le n} S_i\}.$$

2. Show that for critical $p(\cdot)$ the contour process and the height process (up to changing time by a factor of $\frac{1}{2}$) are close, i.e.,

$$n^{-\frac{1}{2}} \sup_{t \in [0,1]} \left| C_{\lfloor 2nt \rfloor}(\mathcal{G}_n) - H_{\lfloor nt \rfloor}(\mathcal{G}_n) \right| \underset{n \to \infty}{\longrightarrow} 0, \quad \text{in probability.}$$

3. Show that for critical $p(\cdot)$ with finite variance the Lukasiewicz walk and a multiple of the height function are close, i.e.,

$$n^{-\frac{1}{2}} \sup_{t \in [0,1]} \left| H_{\lfloor nt \rfloor}(\mathcal{G}_n) - \frac{2}{\sigma^2} S_{\lfloor nt \rfloor}(\mathcal{G}_n) \right| \underset{n \to \infty}{\longrightarrow} 0, \quad \text{in probability.}$$

Steps in the proof of Aldous' invariance principle

1. Read off the height function from the Lukasiewicz walk via the **key formula**,

$$H_n := \#\{j \in \{0,1,...,n-1\} : S_j = \min_{j\le i\le n} S_i\}.$$

2. Show that for critical $p(\cdot)$ the contour process and the height process (up to changing time by a factor of $\frac{1}{2}$) are close, i.e.,

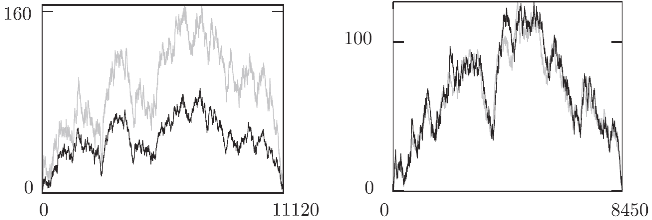
$$n^{-\frac{1}{2}} \sup_{t\in[0,1]} \left|C_{\lfloor 2nt\rfloor}(\mathcal{G}_n) - H_{\lfloor nt\rfloor}(\mathcal{G}_n)\right| \underset{n\to\infty}{\longrightarrow} 0, \text{ in probability.}$$

3. Show that for critical $p(\cdot)$ with finite variance the Lukasiewicz walk and a multiple of the height function are close, i.e.,

$$n^{-\frac{1}{2}} \sup_{t\in[0,1]} \left|H_{\lfloor nt\rfloor}(\mathcal{G}_n) - \tfrac{2}{\sigma^2} S_{\lfloor nt\rfloor}(\mathcal{G}_n)\right| \underset{n\to\infty}{\longrightarrow} 0, \text{ in probability.}$$

4. Apply a conditional version of Donsker's theorem to find that

$$\left(\tfrac{1}{\sqrt{2n}} C_{2nt}\right)_{t\in[0,1]} \approx \tfrac{\sqrt{2}}{\sigma}\left(\tfrac{1}{\sqrt{n\sigma^2}} S_{\lfloor nt\rfloor}\right)_{t\in[0,1]} \underset{n\to\infty}{\Longrightarrow} \tfrac{\sqrt{2}}{\sigma}\left(B_t^{\text{exc}}\right)_{t\in[0,1]}. \qquad \square$$



The paper **Marckert & Mokkadem (2003) ([9])** provides a visual simulation of the joint convergence of the **contour function** and the **Lukasiewicz walk** towards the same Brownian excursion (up to a multiplicative factor):

- The first picture shows a GW-tree of size $n = 5560$ with offspring distribution $p(0) = \frac{13}{18}$, $p(2) = \frac{1}{6}$ and $p(6) = \frac{1}{9}$ (i.e., $\frac{\sigma^2}{2} = \frac{11}{6}$).
- The next picture shows a GW-tree of size $n = 4208$ with offspring distribution $p(0) = \frac{8}{15}$, $p(1) = \frac{4}{15}$, $p(3) = \frac{2}{15}$ and $p(5) = \frac{1}{15}$ (i.e., $\frac{\sigma^2}{2} = \frac{16}{15}$).

The notion of a real tree

Is there a tree associated with the normalized Brownian excursion?

Definition

A complete and separable metric space (T, r) is called a **real tree** iff

1. any two points $a, b \in T$ are joint by a **unique arc**, and
2. this arc is **isometric to a line segment**.

It is a rooted real tree if we distinguish a point $\rho \in T$, called the **root**. $x \in T$ is called a **leaf** or a **a branch point** if $T \setminus \{x\}$ consists of 1 respectively at least 3 connected components.

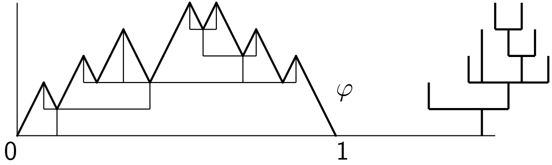
Remarks. A real tree can have

- uncountably infinitely many leaves,
- branch points lying dense in the tree (that is, edge lengths are infinitesimal small).

- A (continuous) **excursion** is a function $\varphi \in C([0,1])$ with $\varphi\big|_{\{0,1\}} = 0$ and $\varphi\big|_{(0,1)} > 0$.
- With every excursion φ we associate a **pseudo-metric on** $[0,1]$**:**

$$r_\varphi(s,t) := \varphi(s) + \varphi(t) - 2 \cdot \inf_{u\in[s,t]} \varphi(u).$$

Fact. $T\big|_\varphi = [0,1]_{/\sim_\varphi}$ is a **compact** real tree with **root** 0.



Definition (THE Continuum Random Tree)

Call the tree "below" $2\cdot$ **Brownian excursion** the **Brownian CRT**.

Measure real trees

- In order to be able to sample points from the **real tree** (T, r) it is often in addition equipped with a **probability measure** μ.
- We refer to μ as the **sampling measure**.

Examples. Assume that T is associated with a continuous excursion φ over $[0, 1]$.

- Equip $T\big|_{\varphi} = [0, 1]_{/\sim_{\varphi}}$ with the (image measure) μ of the Lebesgue measure on $[0, 1]$ under the map which sends $t \in [0, 1]$ to a point in the tree.
- If $\mathfrak{t}$ is finite, then $\#\mathrm{Lf}(\mathfrak{t}) + \#\mathrm{Br}(\mathfrak{t}) < \infty$. Typical choices are
  - the **normalized length measure** μ_{skeleton}, i.e., the normalized length measure on the set $\biguplus_{\ell \in \mathrm{Lf}(\mathfrak{t})}[\rho, \ell]$,
  - the **uniform** distribution μ_{leaf} **on** the set of **leaves**, or
  - the **uniform** distribution μ_{vertex} **on** all **vertices**.

Aldous' CRT

- For $k \geq 2$, we consider **binary trees** with k **leaves labelled** $\{1, 2, ..., k\}$ and positive edge lengths $\{l_e;\ e \text{ edges}\}$.
- Each such tree has $2k-3$ edges. When edge lengths are ignored, there are $\prod_{i=1}^{k-2}(2i-1)$ many possible shapes $\hat{t}$ for the tree.

Lemma (Aldous (1993), [1])

There exists a family $(\mathcal{R}(k);\ k \geq 1)$ *of such random binary trees s.t.*

- $\mathcal{R}(k)$ *has density*

$$\mathbb{P}\big(\text{shape}(\mathcal{R}(k)) = \hat{t}, L_1 \in \mathrm{d}l_1, ..., L_{2k-3} \in \mathrm{d}l_{2k-3}\big)$$
$$= s \cdot \exp\big(-s^2/2\big)\mathrm{d}l_1 ... \mathrm{d}l_{2k-3},$$

where $s := \sum_{i=1}^{2k-3} l_i$, *and*

- *for each* $k \in \mathbb{N}$, *the subtree spanned by* $j \leq k$ *leaves sampled randomly from* $\{1, 2, ..., k\}$ *equals in distribution the tree* $\mathcal{R}(k)$.

$$\mathbb{P}\big(\mathrm{shape}(\mathcal{R}(k)) = \hat{t}, L_1 \in \mathrm{d}l_1, ..., L_{2k-3} \in \mathrm{d}l_{2k-3}\big)$$
$$= s \cdot \exp\big(-s^2/2\big)\mathrm{d}l_1 ... \mathrm{d}l_{2k-3}, \quad s := \sum_{i=1}^{2k-3} l_i.$$

Remarks.

1. The shape is uniform on the set of possible shapes, the edge lengths are independent of the shape and edge lengths are exchangeable.
2. If $k = 2$, then $\mathcal{R}(2)$ has 2 leaves, 1 possible shape, 1 edge, no internal node. The single edge's length is **Rayleigh distributed**, i.e.,
$$\mathbb{P}\big(L \in \mathrm{d}l\big) = l \cdot \exp\big(-l^2/2\big)\mathrm{d}l.$$

Exercise. Show that the right hand side of the above expression is indeed a probability density function.

1. Let $(C_1, C_2, C_3, \ldots)$ be the times of a **non-homogeneous Poisson point process** with **rate** $r(t) = t$, i.e., for example,

$$\mathbb{P}\{C_1 > t\} = \mathbb{P}\{\text{no point in } [0,t]\} = e^{-\int_0^t \mathrm{d}s\, r(s)} = e^{-\frac{t^2}{2}},$$

and

$$\begin{aligned}\mathbb{P}\{C_2 > t\} &= \int_0^t \mathrm{d}s\, \mathbb{P}(C_2 > t | C_1 = s)\mathbb{P}(C_1 \in \mathrm{d}s) \\ &= \int_0^t \mathrm{d}s\, \mathbb{P}\{\text{no point in } [s,t]\} \cdot s e^{-\frac{s^2}{2}} \\ &= \int_0^t \mathrm{d}s\, e^{-\int_s^t \mathrm{d}u\, r(u)} s e^{-\frac{s^2}{2}} = \frac{t^2}{2} e^{-\frac{t^2}{2}}.\end{aligned}$$

Aldous' CRT: The line breaking construction

1. Let $(C_1, C_2, C_3, ...)$ be the times of a **non-homogeneous Poisson point process** with **rate** $r(t) = t$, i.e., for example,

$$\mathbb{P}\{C_1 > t\} = \mathbb{P}\{\text{no point in } [0,t]\} = e^{-\int_0^t \mathrm{d}s\, r(s)} = e^{-\frac{t^2}{2}},$$

and

$$\begin{aligned}\mathbb{P}\{C_2 > t\} &= \int_0^t \mathrm{d}s\, \mathbb{P}(C_2 > t | C_1 = s)\mathbb{P}(C_1 \in \mathrm{d}s) \\ &= \int_0^t \mathrm{d}s\, \mathbb{P}\{\text{no point in } [s,t]\} \cdot s e^{-\frac{s^2}{2}} \\ &= \int_0^t \mathrm{d}s\, e^{-\int_s^t \mathrm{d}u\, r(u)} s e^{-\frac{s^2}{2}} = \frac{t^2}{2} e^{-\frac{t^2}{2}}.\end{aligned}$$

2. Let $\mathcal{R}(1)$ be a **line of length** C_1 from a root to leaf 1.

Aldous' CRT: The line breaking construction

1. Let $(C_1, C_2, C_3, \ldots)$ be the times of a **non-homogeneous Poisson point process** with **rate** $r(t) = t$, i.e., for example,

$$\mathbb{P}\{C_1 > t\} = \mathbb{P}\{\text{no point in } [0,t]\} = e^{-\int_0^t \mathrm{d}s\, r(s)} = e^{-\frac{t^2}{2}},$$

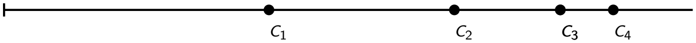
and

$$\begin{aligned}\mathbb{P}\{C_2 > t\} &= \int_0^t \mathrm{d}s\, \mathbb{P}(C_2 > t | C_1 = s)\mathbb{P}(C_1 \in \mathrm{d}s) \\ &= \int_0^t \mathrm{d}s\, \mathbb{P}\{\text{no point in } [s,t]\} \cdot se^{-\frac{s^2}{2}} \\ &= \int_0^t \mathrm{d}s\, e^{-\int_s^t \mathrm{d}u\, r(u)} se^{-\frac{s^2}{2}} = \frac{t^2}{2} e^{-\frac{t^2}{2}}.\end{aligned}$$

2. Let $\mathcal{R}(1)$ be a **line of length** C_1 from a root to leaf 1.
3. Inductively, **obtain** $\mathcal{R}(k+1)$ **from** $\mathcal{R}(k)$ **by attaching an edge of length** $C_{k+1} - C_k$ **to a uniform random point of** $\mathcal{R}(k)$ (i.e., sampled with respect to the normalized Lebesgue measure on the edges), labeling a new leaf $k+1$.

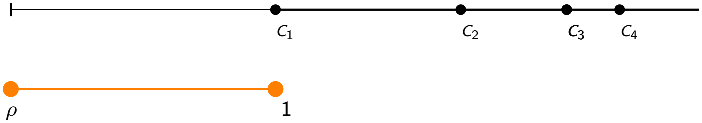
The line breaking construction: an illustration

1. Let $(C_1, C_2, C_3, \ldots)$ be the times of a non-homogeneous Poisson point process with **rate** $r(t) = t$.



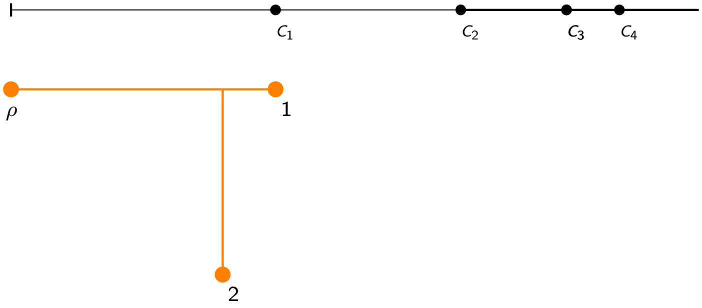
The line breaking construction: an illustration

1. Let $(C_1, C_2, C_3, \ldots)$ be the times of a non-homogeneous Poisson point process with **rate** $r(t) = t$.
2. Let $\mathcal{R}(1)$ be a line of length C_1 from a root to leaf 1.



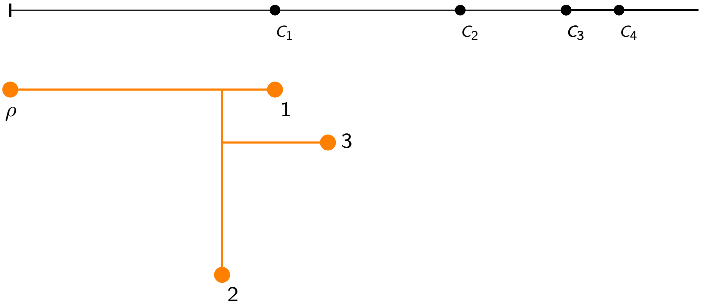
The line breaking construction: an illustration

1. Let $(C_1, C_2, C_3, \ldots)$ be the times of a non-homogeneous Poisson point process with **rate** $r(t) = t$.
2. Let $\mathcal{R}(1)$ be a line of length C_1 from a root to leaf 1.
3. Inductively, obtain $\mathcal{R}(k+1)$ from $\mathcal{R}(k)$ by attaching an edge of length $C_{k+1} - C_k$ to a uniform random point of $\mathcal{R}(k)$ labeling a new leaf $k+1$.



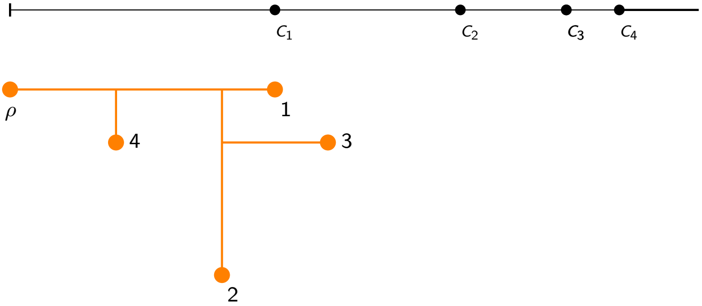
The line breaking construction: an illustration

1. Let $(C_1, C_2, C_3, \dots)$ be the times of a non-homogeneous Poisson point process with **rate** $r(t) = t$.
2. Let $\mathcal{R}(1)$ be a line of length C_1 from a root to leaf 1.
3. Inductively, obtain $\mathcal{R}(k+1)$ from $\mathcal{R}(k)$ by attaching an edge of length $C_{k+1} - C_k$ to a uniform random point of $\mathcal{R}(k)$ labeling a new leaf $k+1$.



The line breaking construction: an illustration

1. Let $(C_1, C_2, C_3, \ldots)$ be the times of a non-homogeneous Poisson point process with **rate** $r(t) = t$.
2. Let $\mathcal{R}(1)$ be a line of length C_1 from a root to leaf 1.
3. Inductively, obtain $\mathcal{R}(k+1)$ from $\mathcal{R}(k)$ by attaching an edge of length $C_{k+1} - C_k$ to a uniform random point of $\mathcal{R}(k)$ labeling a new leaf $k+1$.



- We have seen that the density of C_1 is the right Rayleigh distribution. We proceed by induction. Let $(\mathfrak{t}^*, x_1^*, ..., x_{2k+1}^*)$ be a binary tree with $k+1$ leaves, shape $\mathfrak{t}$ and $2k+1$ edge lengths x_1^*, ..., x_{2k+1}^*, and Let $(\mathfrak{t}, x_1, ..., x_{2k-1})$ be the associated binary tree spanned by the leaves $\{1, 2, ..., k\}$.

Aldous' CRT: analysing the density

- We have seen that the density of C_1 is the right Rayleigh distribution. We proceed by induction. Let $(\mathfrak{t}^*, x_1^*, ..., x_{2k+1}^*)$ be a binary tree with $k+1$ leaves, shape $\mathfrak{t}$ and $2k+1$ edge lengths x_1^*, ..., x_{2k+1}^*, and Let $(\mathfrak{t}, x_1, ..., x_{2k-1})$ be the associated binary tree spanned by the leaves $\{1, 2, ..., k\}$.
- By construction, $\mathfrak{t}^*$ is obtained from $\mathfrak{t}$ by splitting an edge x_j for some $j = 1, ..., 2k-1$ into two edges of lengths $x_{j_1}^*$ and $x_{j_2}^*$ with $x_j = x_{j_1}^* + x_{j_2}^*$, and joining leaf $k+1$ to that new internal vertex by an edge $x_{j_3}^* = s^* - s$, say.

- We have seen that the density of C_1 is the right Rayleigh distribution. We proceed by induction. Let $(\mathfrak{t}^*, x_1^*, ..., x_{2k+1}^*)$ be a binary tree with $k+1$ leaves, shape $\mathfrak{t}$ and $2k+1$ edge lengths x_1^*, ..., x_{2k+1}^*, and Let $(\mathfrak{t}, x_1, ..., x_{2k-1})$ be the associated binary tree spanned by the leaves $\{1, 2, ..., k\}$.
- By construction, $\mathfrak{t}^*$ is obtained from $\mathfrak{t}$ by splitting an edge x_j for some $j = 1, ..., 2k-1$ into two edges of lengths $x_{j_1}^*$ and $x_{j_2}^*$ with $x_j = x_{j_1}^* + x_{j_2}^*$, and joining leaf $k+1$ to that new internal vertex by an edge $x_{j_3}^* = s^* - s$, say.
- That is,

$$f\left(\mathfrak{t}^*, x_1^*, ..., x_{2k+1}^*\right) = f\left(\mathfrak{t}, x_1, ..., x_{2k-1}\right) s^* \cdot e^{-\frac{1}{2}((s^*)^2 - s^2)} \cdot s^{-1},$$

where s^{-1} is the probability density that the $(k+1)^{\text{st}}$ edge is attached at a particular place in the existing tree.

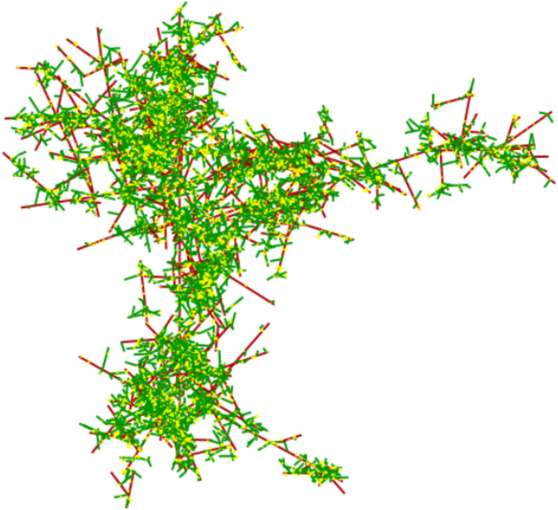
- We have seen that the density of C_1 is the right Rayleigh distribution. We proceed by induction. Let $(\mathfrak{t}^*, x_1^*, ..., x_{2k+1}^*)$ be a binary tree with $k+1$ leaves, shape $\mathfrak{t}$ and $2k+1$ edge lengths x_1^*, ..., x_{2k+1}^*, and Let $(\mathfrak{t}, x_1, ..., x_{2k-1})$ be the associated binary tree spanned by the leaves $\{1, 2, ..., k\}$.
- By construction, $\mathfrak{t}^*$ is obtained from $\mathfrak{t}$ by splitting an edge x_j for some $j = 1, ..., 2k-1$ into two edges of lengths $x_{j_1}^*$ and $x_{j_2}^*$ with $x_j = x_{j_1}^* + x_{j_2}^*$, and joining leaf $k+1$ to that new internal vertex by an edge $x_{j_3}^* = s^* - s$, say.
- That is,

$$f\left(\mathfrak{t}^*, x_1^*, ..., x_{2k+1}^*\right) = f\left(\mathfrak{t}, x_1, ..., x_{2k-1}\right) s^* \cdot e^{-\frac{1}{2}((s^*)^2 - s^2)} \cdot s^{-1},$$

  where s^{-1} is the probability density that the $(k+1)^{\text{st}}$ edge is attached at a particular place in the existing tree.
- Finally, by exchangeability of the edge lengths consistency immediately follows.

The Continuum Random Tree (CRT): an illustration

Several simulations of THE CRT can be found on the home page of Jean-François Marckert, e.g.,



Consequences of the stick breaking construction

- Let $(C_1, C_2, C_3, \ldots)$ be the times of a **non-homogeneous Poisson point process** with **rate** $r(t) = t$.
- Let $\mathcal{R}(1)$ be a **line of length** C_1 from a root to leaf 1.
- Inductively, **obtain** $\mathcal{R}(k+1)$ **from** $\mathcal{R}(k)$ **by attaching an edge of length** $C_{k+1} - C_k$ **to a uniform random point of** $\mathcal{R}(k)$ (i.e., sampled with respect to the normalized Lebesgue measure on the edges), labeling a new leaf $k+1$.

Theorem (Aldous (1991), [2])

For a realization $\mathfrak{t}(2) \subseteq \mathfrak{t}(3) \subseteq \ldots$ *of* $\mathcal{R}(2) \subseteq \mathcal{R}(3) \subseteq \ldots$, *let* T *be the completion of* $\bigcup_{\mathfrak{t}}(k)$. *The resulting random tree* $\mathcal{T}$ *satisfies:*

- $\mathcal{T}$ *is compact, almost surely.*
- *There is a* **mass measure** μ *on* $\mathcal{T}$ *with* $\mu(\mathcal{T}) = 1$ *but* $\mu(\bigcup_k \mathcal{R}(k)) = 0$, *characterized as the weak limit of the uniform distribution on the leaves* $\{1, 2, \ldots, k\} \subset \mathcal{T}$.
- *The* **total length** D_k *of the edges of* $\mathcal{R}(k)$ *has distribution*

$$\mathbb{P}\big(D_k > d\big) = \mathbb{P}\big(N(d^2/2) \le k - 1\big),$$

where $N(\nu)$ *has Poisson*(ν)*-distribution.*

Aldous' CRT is the Brownian CRT

Definition (Aldous' CRT)

Let us define the **Aldous' CRT** as the random tree $\mathcal{T}$ arising from the line-breaking construction, and additionally equipped with the mass measure.

Theorem (Aldous (1993), [1])

The Brownian CRT and Aldous' CRT are the same.

Strategy of proof.

1. Aldous introduced the following notion of convergence: a sequence of "measured $\mathbb{R}$-trees" converges to a limiting measured $\mathbb{R}$-tree if and only if

   all subtrees spanned by a finite sample converge weakly to the respective subtree in the discrete topology.

Strategy of proof.

1. Aldous introduced the following notion of convergence: a sequence of "measured $\mathbb{R}$-trees" converges to a limiting measured $\mathbb{R}$-tree if and only if

   all subtrees spanned by a finite sample converge weakly to the respective subtree in the discrete topology.

2. He then shows that GW-tree conditioned to have total progeny n and with edge lengths rescaled by $\frac{1}{\sqrt{n}}$ converges to the Aldous' tree.

Aldous' CRT is the Brownian CRT

Strategy of proof.

1. Aldous introduced the following notion of convergence: a sequence of "measured $\mathbb{R}$-trees" converges to a limiting measured $\mathbb{R}$-tree if and only if

   all subtrees spanned by a finite sample converge weakly to the respective subtree in the discrete topology.

2. He then shows that GW-tree conditioned to have total progeny n and with edge lengths rescaled by $\frac{1}{\sqrt{n}}$ converges to the Aldous' tree.
3. As we will see in Part III the latter characterizes the limiting tree uniquely.

Strategy of proof.

1. Aldous introduced the following notion of convergence: a sequence of "measured $\mathbb{R}$-trees" converges to a limiting measured $\mathbb{R}$-tree if and only if

   all subtrees spanned by a finite sample converge weakly to the respective subtree in the discrete topology.

2. He then shows that GW-tree conditioned to have total progeny n and with edge lengths rescaled by $\frac{1}{\sqrt{n}}$ converges to the Aldous' tree.
3. As we will see in Part III the latter characterizes the limiting tree uniquely.
4. We know from the converge result of contour functions that the limit must be the Brownian CRT.

Aldous' rescaling result

Theorem (Aldous (1993), [1])

Let $\mathcal{G}_n$ be the GW-tree with critical offspring distribution of finite variance $\sigma^2 > 0$ conditioned on n leaves labelled by $\{1, 2, ..., n\}$. Assign length $\frac{\sigma}{\sqrt{n}}$ to each edge of $\mathcal{G}_n$. Let $\mathcal{R}(n, k)$ be the subtree of $\mathcal{T}_n$ spanned by vertices $\{1, 2, ..., k\}$. Then for each fixed $k \geq 2$,

$$\mathcal{R}(n, k) \xrightarrow[N \to \infty]{\mathrm{w}} \mathcal{R}(k)$$

in the sense that the joint distributions of shape and edge lengths converge to the distribution of Aldous' CRT.

A useful representation of tree-lengths

Theorem (Aldous (1993), [1])

Let B^{ext} be the standard Brownian excursion, and U_1, U_2, ... independent uniform on $[0,1]$ variables, independent of B^{ext}. For each $n \geq 1$, let $\mathcal{T}_n$ be the subtree of the Brownian CRT $[0,1]|_{\sim_{2B^{\text{ext}}}}$ spanned by 0, U_1, U_2, ..., and denote the length of $\mathcal{T}_n$ by Θ_n. Then

$$(\Theta_1, \Theta_2, \Theta_3, \dots) \stackrel{d}{=} \left(\sqrt{2X_1}, \sqrt{2(X_1 + X_2)}, \sqrt{2(X_1 + X_2 + X_3)}, \dots\right),$$

where X_1, X_2, ... are independent rate 1 exponentially distributed.

Proof. We rely on the line-breaking construction for Aldous' CRT.

- For $k = 1$, notice that for all $x > 0$

$$\mathbb{P}\{\Theta_1 > x\} = \mathbb{P}\{D_1 > x\} = e^{-\frac{x^2}{2}},$$

while on the other hand

$$\mathbb{P}\{\sqrt{2X_1} > x\} = \mathbb{P}\{X_1 > \tfrac{x^2}{2}\} = e^{-\frac{x^2}{2}}.$$

- The general case I will leave for you as an **exercise**.

Yet another home work problem

Exercise. Use the latter to show that $(D_1, ..., D_k)$ has joint density

$$f_{(D_1,\dots,D_k)}(\ell_1, ..., \ell_k) = \ell_1 \cdot \ell_2 \cdot ... \cdot \ell_k e^{-\frac{\ell_k^2}{2}} \mathbf{1}\{0 < \ell_1 < \ell_2 < ... < \ell_k\}.$$

Janson's result for Galton-Watson trees

Theorem (Janson (2006), [5])

Let $\mathcal{G}_n$ be the GW-tree with offspring distribution $p(\cdot)$ conditioned to have n vertices. Assume that $p(\cdot)$ is critical, $p(1) < 1$, and $p(\cdot)$ has finite variance σ^2. Then

$$\mathbb{P}\{N(\mathcal{G}_n, \rho) \geq x\sqrt{n}\sigma\} \underset{n\to\infty}{\longrightarrow} e^{-x^2/2}.$$

Factorial moment formula

If $v_1, \ldots, v_k$ are vertices in the rooted tree (T, ρ), denote by

$$L_T(v_1, \ldots, v_k)$$

the number of edges in the subtree of T spanned by $\{\rho, v_1, \ldots, v_k\}$.

Lemma (Factorial moments)

For any rooted tree (T, r), the factorial moments of $N(T, \rho)$ are given by

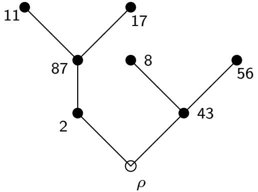
$$\mathbb{E}\big[N(T,\rho)\big(N(T,\rho)-1\big)\cdot\ldots\big(N(T,\rho)-k+1\big)\big]$$
$$= k! \sum_{v_1,\ldots,v_k}^{**} \frac{1}{L_T(v_1)\cdot L_T(v_1,v_2)\cdot L_T(v_1,\ldots,v_k)}$$

with $\sum^{**}$ denoting the sum over all $v_1, \ldots, v_k$ are distinct, $\neq \rho$, and such that v_i is not a descendent of v_j when $i < j$. In particular,

$$\mathbb{E}\big[N(T,\rho)\big] = \sum_{v \neq \rho} \frac{1}{h(v)}.$$

Equivalent formulation in terms of records

- We use the equivalence of $N(T, \rho)$ and $N_{\mathrm{vertex}}(T, \rho)$.
- $N_{\mathrm{vertex}}(T, \rho)$ appears also when we are consider **records** in a tree.
- Let each vertex v have a random value λ_e attached to it, and assume that these values are i.i.d. with a continuous distribution.
- Say that a value λ_e is a **record** if it is the largest value in the path from the root to e.
- Then the number of records equals in distribution $N_{\mathrm{vertex}}(T, \rho)$.



- To see this, generate first the values λ_e, and then cut the tree: each time choosing the vertex with the largest λ_e among the remaining ones.

Proof of factorial moment formula

Write

$$N_{\mathrm{vertex}}(T,\rho) := \sum_{v\neq\rho} \mathbf{1}_{A_v},$$

where A_v denotes the event that **"v is a record"**. Thus

$$\begin{aligned} & N_{\mathrm{vertex}}(T,\rho)\big(N_{\mathrm{vertex}}(T,\rho)-1\big)\cdot\ldots\big(N_{\mathrm{vertex}}(T,\rho)-k+1\big) \\ &= \sum_{v_1,v_2,\ldots,v_k\in V(T)\setminus\{\rho\}} \mathbf{1}_{A_{v_1}}\cdot\ldots\cdot\mathbf{1}_{A_{v_k}} \\ &= k! \sum_{v_1,v_2,\ldots,v_k\in V(T)\setminus\{\rho\}} \mathbf{1}_{\mathcal{E}(v_1,\ldots,v_k)}, \end{aligned}$$

where

$$\begin{aligned} &\mathcal{E}(v_1,\ldots,v_k) \\ &:= \big\{\lambda_{v_1} < \ldots < \lambda_{v_k} \text{ and all are records in } T'\big\} \\ &= \big\{\lambda_{v_j} \text{ is largest value in } T'(v_1,\ldots,v_j) \text{ for every } j=1,\ldots,k\big\}. \end{aligned}$$

Proof continued

$$N_{\mathrm{vertex}}(T,\rho)\big(N_{\mathrm{vertex}}(T,\rho)-1\big)\cdot\ldots\big(N_{\mathrm{vertex}}(T,\rho)-k+1\big)$$
$$= k! \sum_{v_1,v_2,\ldots,v_k\in V(T)\setminus\{\rho\}} \mathbf{1}_{\left\{\lambda_{v_j} \text{ is largest value in } T'(v_1,\ldots,v_j) \text{ for every } j=1,\ldots,k\right\}}.$$

Thus

$$\mathbb{E}\big[N(T,\rho)\big(N(T,\rho)-1\big)\cdot\ldots\big(N(T,\rho)-k+1\big)\big]$$
$$= k! \sum_{v_1,v_2,\ldots,v_k\in V(T)}^{**} \mathbb{P}\big\{\lambda_{v_j} \text{ is largest value in } T'(v_1,\ldots,v_j)\ \forall j=1,\ldots,k\big\}$$
$$= k! \sum\nolimits^{**}_{v_1,v_2,\ldots,v_k\in V(T)} \prod_{j=1}^{k} \frac{1}{L_T(v_1,\ldots,v_j)}. \qquad \square$$

Convergence to the corresponding moments of the Brownian CRT

Lemma (Janson (2006), [5])

Let $\mathcal{G}_n$ be the GW-tree with critical offspring distribution of finite variance $\sigma^2 > 0$ conditioned on total progeny n, and $(\mathcal{R}(k);\, k \in \mathbb{N})$ the leaf labelled finite trees from the line-breaking construction of Aldous' tree. Then k^{th}-factorial moments of $N(\mathcal{G}_n)$ rescaled by $\sigma^{-k} n^{-\frac{k}{2}}$ converges to

$$k!\mathbf{E}\Big[\big(D_1 \cdot D_2 \cdot D_k \cdot D_k\big)^{-1}\Big],$$

where D_k denotes the total length of $\mathcal{R}(k)$, $k \in \mathbb{N}$.

Proof. We use that $\frac{1}{\sigma\sqrt{n}}\mathcal{G}_n$ converges weakly to Aldous' CRT, and that the family of k^{th}-factorial moments of $N(\mathcal{G}_n)$ indexed by $n \in \mathbb{N}$ is uniformly integrable, as

$$\sum\nolimits^{**}_{v_1,v_2,\dots,v_k \in V(\mathcal{G})_n} \prod_{j=1}^{k} \frac{1}{L_{\mathcal{G}_n}(v_1,\dots,v_j)} \le \Big(\sum_{v \in V(\mathcal{G}_n)} L^{-1}_{\mathcal{G}_n}(v)\Big)^k. \quad \square$$

Let Y be Rayleigh distributed with density $f_Y(\mathrm{d}y) = ye^{-\frac{y^2}{2}}$.

Lemma (Janson (2006), [5])

Let $(\mathcal{R}(k);\ k \in \mathbb{N})$ the leaf labelled finite trees from the line-breaking construction of Aldous' tree, and denote by D_k the total length of $\mathcal{R}(k)$, $k \in \mathbb{N}$. Then for $k \geq 1$, $k!\mathbf{E}\big[\big(D_1 \cdot D_2 \cdot D_k \cdot D_k\big)^{-1}\big] = \mathbb{E}\big[Y^k\big]$.

Proof. Recall the joint density

$$f_{(D_1,\dots,D_k)}(\ell_1, \dots, \ell_k) = \ell_1 \cdot \ell_2 \cdot \dots \cdot \ell_k e^{-\frac{\ell_k^2}{2}} \mathbf{1}\{0 < \ell_1 < \ell_2 < \dots < \ell_k\}.$$

of the Aldous' tree lengths. Therefore the left hand side equals

$$\begin{aligned} & k!\mathbb{E}\big[\big(D_1 \cdot D_2 \cdot \dots D_k\big)^{-1}\big] \\ & = k! \int_{\{0<\ell_1<\ell_2<\dots<\ell_k\}} \mathrm{d}\ell_1 \mathrm{d}\ell_2 \dots \mathrm{d}\ell_{k+1} \big(\ell_1 \cdot \ell_2 \cdot \dots \cdot (\ell_k)\big)^{-1} \ell_1 \cdot \ell_2 \cdot \dots \cdot \ell_k e^{-\frac{\ell_k^2}{2}} \\ & = k! \int_0^\infty \mathrm{d}\ell_k \, \frac{\ell_k^{k-1}}{(k-1)!} e^{-\frac{\ell_k^2}{2}} = \int_0^\infty \mathrm{d}\ell \, \ell^{k+1} e^{-\frac{\ell^2}{2}}. \qquad \square \end{aligned}$$

Literature

- David Aldous (1993). The continuum random tree III, Annals of Probability.
- Jürgen Bennies & Götz Kersting (2000). A random walk approach to Galton-Watson trees, Journal of theoretical probability.
- Phillipe Biane & Marc Yor (1987). Valeurs principales associees aux temps locaux Browniens, Bull. Sci Math.
- Endre Csaki & S. Mohanty (1981). Excursion and meander in random walk, Canadian Journal Statistics.
- Thomas Duquesne (2003). A limit theorem for the contour process of conditioned Galton-Watson trees, Annals of Probability.
- Rick Durrett, Donald Iglehart & Douglas Miller (1977). Weak convergence of Brownain meander and Brownian excursion, Annals of Probability.
- Thomas Duquesne & Jean-François Le Gall (2002). Random trees, Lévy processes and spatial branching processes, Astérisque.
- Götz Kersting (1998). On the Height Profile of a Conditioned Galton-Watson Tree, unpublished.
- Jean-François Marckert & Abdelkader Mokkadem (2003). The depth first process of Galton-Watson trees converge to the same Brownian excursion, Annals of Probabiity.
- Alois Panholzer (2006). Cutting down very simple trees, Quaestiones Mathematicae.