ASYMPTOTIC DEGREE DISTRIBUTION IN RANDOM RECURSIVE TREES

SVANTE JANSON

Abstract. The distributions of vertex degrees in random recursive trees and random plane recursive trees are shown to be asymptotically normal. Formulas are given for the asymptotic variances and covariances of the number of vertices with given outdegrees. We also give functional limit theorems for the evolution as vertices are added.

The proofs use some old and new results about generalized Pólya urn models. We consider generalized Pólya urns with infinitely many types, but reduce them to the finite type case.

1. Introduction

A *random recursive tree* with n vertices is a random rooted tree obtained by starting with a single root and then adding $n-1$ vertices one by one, each time joining the new vertex to a randomly chosen old vertex; the random choices are uniform and independent of each other. If the vertices are labelled $1, 2, \dots,$ we thus obtain a tree where the labels increase along each branch as we travel from the root; the random recursive tree can also be defined as a (uniform) randomly chosen such labelled tree. (The distribution of a random recursive tree differs from the distribution of a uniform random labelled tree.) See also the survey [23].

Mahmoud and Smythe [16] studied the distribution of outdegrees in a random recursive trees and obtained a joint asymptotic normal distribution of the number of nodes of outdegrees 0, 1 and 2. They also indicated how the results in principle extend to higher degrees. We do this extension to arbitrary degrees, using some new results on generalized Pólya urn processes [11].

A variant of the random recursive tree is the *random plane recursive tree* studied by Mahmoud, Smythe and Szymański [17]. This is a random rooted plane (or ordered) tree, again obtained by starting with a single root and then adding $n-1$ vertices one by one. This time, however, the descendants of each vertex are ordered (from left to right, say), and a new vertex may be inserted at any place. At a vertex with outdegree d, there are thus $d+1$ possible places to add a new vertex, and in total a plane tree with n vertices has $2n-1$ such places. We choose one of these places at random (uniformly)

Date: February 21, 2003; revised April 13, 2004.

This is a preprint of an article accepted for publication in Random Structure & Algoritms

for vertex $n+1$. It is often useful to regard the possible places for a new vertex as external vertices; a tree with n (internal) vertices thus has $2n-1$ external vertices, and the random plane recursive tree evolves by converting a randomly chosen external vertex v to an internal vertex and adding three new external vertices: one daughter below v and one sister on each side of v.

Mahmoud, Smythe and Szymański [17] obtained, among other results, a joint asymptotic normal distribution of the number of nodes of outdegrees 0, 1 and 2 in a random plane recursive tree. We extend this result too to arbitrary degrees.

We can state the results as follows. First, let X_{ni} be the number of vertices of outdegree $i \geq 0$ in a random recursive tree with n vertices.

Theorem 1.1. *As $n \to \infty$, $n^{-1}X_{ni} \to 2^{-i-1}$ a.s., and*

$$n^{-1/2}(X_{ni} - 2^{-i-1}n) \xrightarrow{\mathrm{d}} V_i,$$

jointly for all $i \geq 0$, where the V_i are jointly Gaussian variables with means $\mathbb{E}\, V_i = 0$ and covariances $\sigma_{ij} = \operatorname{Cov}(V_i, V_j)$ given by the generating function

$$\begin{aligned}\sum_{i,j=0}^{\infty} \sigma_{ij} z^i w^j &= (1-z)(1-w)\Big(\frac{1}{2-zw} - \frac{1}{(2-z)(2-w)}\Big)\frac{1}{3-z-w} \\ &= \frac{2(1-z)^2(1-w)^2}{(2-z)(2-w)(2-zw)(3-z-w)}.\end{aligned} \tag{1.1}$$

The weaker result $X_{ni}/n \xrightarrow{\mathrm{p}} 2^{-i-1}$ follows from Meir and Moon [18], who studied asymptotics of the means and variances. Earlier, Na and Rapoport [19] had shown that $\mathbb{E}\, X_{ni}/n \to 2^{-i-1}$.

The values of σ_{jk} may be found by expanding the generating function. For example, extending a result in [16], the covariance matrix of the first four components (V_0, V_1, V_2, V_3) is (using calculations done by `Maple`)

$$\begin{pmatrix} \frac{1}{12} & -\frac{7}{72} & -\frac{5}{432} & \frac{17}{2592} \\ -\frac{7}{72} & \frac{71}{432} & -\frac{37}{864} & -\frac{269}{15552} \\ -\frac{5}{432} & -\frac{37}{864} & \frac{473}{5184} & -\frac{1831}{93312} \\ \frac{17}{2592} & -\frac{269}{15552} & -\frac{1831}{93312} & \frac{26939}{559872} \end{pmatrix}$$

It is also straightforward to write down a general formula with finite sums only. However, this formula is a bit complicated, so we omit it and leave its formulation to the interested reader.

Remark 1.2. The joint convergence of infinitely many random variables in Theorem 1.1 is, by definition, the same as joint convergence of any finite subset. This is also the same as convergence of the infinite vector $\big(n^{-1/2}(X_{ni} - 2^{-i-1}n)\big)_{i=0}^{\infty}$ in $\mathbb{R}^{\infty}$ (see [7, p. 19]). We conjecture that this can be strengthened to convergence in ℓ^1, or a suitably weighted version of ℓ^1, which would imply the convergence of more continuous functionals.

We similarly let Y_{ni} be the number of vertices of outdegree $i \geq 0$ in a random plane recursive tree with n vertices.

Theorem 1.3. *As $n \to \infty$, $n^{-1}Y_{ni} \to 4/(i+1)(i+2)(i+3)$ a.s., and*

$$n^{-1/2}\Big(Y_{ni} - \frac{4}{(i+1)(i+2)(i+3)}n\Big) \xrightarrow{\mathrm{d}} W_i,$$

jointly for all $i \geq 0$, where the W_i are jointly Gaussian variables with means $\mathbb{E}\, W_i = 0$ and covariances $\tilde{\sigma}_{ij} = \mathrm{Cov}(W_i, W_j)$ given by

$$\tilde{\sigma}_{ij} = 2\sum_{k=0}^{i}\sum_{l=0}^{j} \frac{(-1)^{k+l}}{k+l+4}\binom{i}{k}\binom{j}{l}\left(\frac{2(k+l+4)!}{(k+3)!(l+3)!} - 1 - \frac{(k+1)(l+1)}{(k+3)(l+3)}\right).$$

For example, extending a result in [17], the covariance matrix of the first four components (W_0, W_1, W_2, W_3) is (again by `Maple`)

$$\begin{pmatrix} \frac{1}{9} & -\frac{4}{45} & -\frac{1}{45} & -\frac{2}{315} \\ -\frac{4}{45} & \frac{23}{180} & -\frac{11}{630} & -\frac{11}{1260} \\ -\frac{1}{45} & -\frac{11}{630} & \frac{179}{3150} & -\frac{1}{175} \\ -\frac{2}{315} & -\frac{11}{1260} & -\frac{1}{175} & \frac{187}{6300} \end{pmatrix}$$

We can extend the theorems to functional limit theorems, describing the evoultion as new vertices are added.

Theorem 1.4. *For the random recursive tree*

$$n^{-1/2}\big(X_{\lfloor xn\rfloor,i} - 2^{-i-1}xn\big) \xrightarrow{\mathrm{d}} V_i(x) \tag{1.2}$$

in $D[0,\infty)$, jointly for all $i \geq 0$, where the $V_i(x)$ are continuous Gaussian processes with $V_i(0) = 0$, $\mathbb{E}\, V_i(x) = 0$ and covariance functions

$$\mathbb{E}\, V_j(x)V_k(y) = \frac{x^2}{y}\sum_{i=0}^{k} \frac{\ln^i(y/x)}{i!}\sigma_{k-i,j}, \qquad 0 < x \leq y. \tag{1.3}$$

Theorem 1.5. *For the random planar recursive tree*

$$n^{-1/2}\big(Y_{\lfloor xn\rfloor,i} - \tfrac{4}{(i+1)(i+2)(i+3)}xn\big) \xrightarrow{\mathrm{d}} W_i(x)$$

in $D[0,\infty)$, jointly for all $i \geq 0$, where the $W_i(x)$ are continuous Gaussian processes with $W_i(0) = 0$, $\mathbb{E}\, W_i(x) = 0$ and covariance functions, for $0 < x \leq y$,

$$\mathbb{E}\, W_i(x)W_j(y) = 2\sum_{k=0}^{i}\sum_{l=0}^{j} \frac{(-1)^{k+l}}{k+l+4}\binom{i}{k}\binom{j}{l} \cdot \left(\frac{2(k+l+4)!}{(k+3)!(l+3)!} - 1 - \frac{(k+1)(l+1)}{(k+3)(l+3)}\right)x^{(l+3)/2}y^{-(l+1)/2}.$$

For example, for $0 < x \le y$,

$$\operatorname{Cov}(V_2(x), V_2(y)) = \frac{x^2}{y}\Big(\frac{473}{5184} - \frac{37}{864}\ln\frac{y}{x} - \frac{5}{864}\ln^2\frac{y}{x}\Big),$$
$$\operatorname{Cov}(W_2(x), W_2(y)) = -\tfrac{1}{45}x^{3/2}y^{-1/2} + \tfrac{1}{105}x^2y^{-1} + \tfrac{73}{1050}x^{5/2}y^{-3/2}.$$

We prove the theorems above by the method of [16] and [17], viz. by reformulating them as results for certain urn processes. We discuss such urn processes in Section 2 and quote some results that we will use. Section 4 contains the details of the application of the general theorems to random recursive trees and Section 5 contains the arguments for the random plane recursive trees. In both cases, we consider urns with infinitely many types, but reduce them to the finite type case. It may be observed that the arguments for the two cases differ in some parts because of different eigenvalue structures. In Section 6, we briefly discuss analogous results for random recursive d-ary trees.

We finally note that problems of this type have been studied by other methods too. For example, Najock and Heyde [20] proved the asymptotic normality of X_{n0} (the number of leaves in a random recursive tree). Bergeron, Flajolet and Salvy [6] used generating functions and obtained general results for a class of random trees including the ones studied here; in particular, they show asymptotic normality of X_{n0} and Y_{n0}, and their methods apply also to higher degrees.

2. Generalized Pólya urns

Consider an urn containing a finite number of balls of different types (colours); say that the possible types are $1, \dots, q$. Then the content of the urn at time n is described by the vector $(X_{n1}, \dots, X_{nq})$, where $X_{ni} \ge 0$ is the number of balls of type i in the urn.

The urn starts with a given vector X_0 (or perhaps X_1), random or not. We are further given, for each type i, an activity $a_i \ge 0$ and a q-dimensional vector $\xi_i = (\xi_{i1}, \dots, \xi_{iq})$ with integer coordinates. In general ξ_i may be random, but in this paper and many other applications, we only consider deterministic ξ_i.

The urn evolves according to a Markov process. At each time $n \ge 1$, one of the balls in the urn is drawn at random such that the probability of drawing a particular ball of type i is proportional to the activity a_i, i.e. the probability of drawing a ball of type i is $a_i X_{n-1,i} / \sum_j a_j X_{n-1,j}$. (In particular, if every $a_i = 1$, a ball is drawn uniformly at random.) The drawn ball is returned to the urn together with, if it is of type i, ξ_{ij} balls of type j, for each $j = 1, \dots, q$.

The integers ξ_{ij} may be negative, which means removal of balls from the urn. In order to guarantee that we are not required to remove balls that do not exist, we assume that

$$\xi_{ij} \ge 0, \qquad i \ne j, \tag{2.1}$$

(thus balls of other types than the drawn are never removed) and

$$\xi_{ii} \geq -1 \tag{2.2}$$

or, more generally, that for each i there is an integer $d_i \geq 1$ such that $d_i | X_{0i}$ and $d_i | \xi_{ji}$ for every j and

$$\xi_{ii} \geq 0 \quad \text{or} \quad \xi_{ii} = -d_i. \tag{2.3}$$

(Hence $d_i | X_{ni}$ for all $n \geq 0$.) Such urns are called tenable by Bagchi and Pal [3].

We further assume that the urn is irreducible, in the sense that for any two distinct types i and j, if we start with only balls of type i, there is a later time when there is a positive probability of having a ball of type j in the urn. In particular, each $a_i > 0$.

For simplicity, cf. Remark 3.4 below, we also assume that there exists $m > 0$ such that, for every i,

$$\sum_j a_j \xi_{ij} = m. \tag{2.4}$$

(When every $a_i = 1$, this says that exactly m balls are added each time.) It follows that $a \cdot X_n = nm + a \cdot X_0 > 0$; in particular the urn never becomes empty.

Urn models of this type have been studied by many authors, see the references in [11].

3. Preliminaries

We let A denote the $q \times q$ matrix

$$A := (a_j \xi_{ji})_{i,j=1}^{q}. \tag{3.1}$$

Thus the j:th *column* of A is $a_j \xi_j$. (E.g. [16, 17] use the transpose matrix.)

By (2.1), $A + \alpha I$ is a non-negative matrix if α is large enough, so by the standard Perron–Frobenius theory, A has a largest real eigenvalue λ_1 such that every other eigenvalue λ satisfies $\operatorname{Re}\lambda < \lambda_1$ (see e.g. [21, Chapter 1 and Theorem 2.6] or [13, Appendix 2]). We order the eigenvalues with decreasing real parts: $\lambda_1 > \operatorname{Re}\lambda_2 \geq \operatorname{Re}\lambda_3 \geq \dots$.

We let a denote the (column) vector $(a_1, \dots, a_q)$ of activities, and let u_1 and v_1 denote left and right eigenvectors of A corresponding to the largest eigenvalue λ_1, i.e. vectors satisfying

$$u_1' A = \lambda_1 u_1', \qquad A v_1 = \lambda_1 v_1.$$

By the Perron–Frobenius theory [13], [21], (applied to $A + \alpha I$ for suitable α), u_1 and v_1 are unique up to scalar factors; they may be chosen with positive components and they are the only positive eigenvectors.

By (2.4) and (3.1),

$$(a'A)_j = \sum_i a_i a_j \xi_{ji} = a_j a \cdot \xi_j = m a_j.$$

So a is a positive left eigenvector and thus a multiple of u_1, and $\lambda_1 = m$. We normalize u_1 and v_1 such that $u_1 = a$ and

$$u_1 \cdot v_1 = a \cdot v_1 = 1. \tag{3.2}$$

We define P_{λ_1} as the matrix

$$P_{\lambda_1} = v_1 u_1', \tag{3.3}$$

which has rank 1 and is a projection onto the one-dimensional eigenspace $\{v : Av = \lambda_1 v\}$. We will in this paper only consider the case $\operatorname{Re}\lambda_2 < \frac{1}{2}\lambda_1$. We then further define $P_I := I - P_{\lambda_1}$, the complementary projection, and the following $q \times q$ matrices, regarding ξ_i as a column matrix,

$$B := \sum_{i=1}^{q} v_{1i} a_i \xi_i \xi_i', \tag{3.4}$$

$$\Sigma_I := \int_0^\infty P_I e^{sA} B e^{sA'} P_I' e^{-\lambda_1 s}\, ds. \tag{3.5}$$

The general result that we use is the following, which summarizes and simplifies some of the results in [11]. The a.s. convergence and the asymptotic normality (without explicit covariance matrix) are due to [1], see also [2, Section V.9]. (At least when (2.2) holds, which easily implies the general case, see [11, Remark 4.2].) See also similar results in [22], [4], [5].

Theorem 3.1. *Assume that the urn is irreducible and tenable, and that* (2.4) *holds with* $m > 0$. *Then* $n^{-1}X_n \xrightarrow{\text{a.s.}} \lambda_1 v_1$ *as* $n \to \infty$. *If further* $\operatorname{Re}\lambda_2 < \frac{1}{2}\lambda_1$, *then also the following hold.*

(i) $n^{-1/2}\bigl(X_n - n\lambda_1 v_1\bigr) \xrightarrow{\text{d}} N(0, m\Sigma_I)$.

(ii) *More generally, in* $D[0,\infty)$,

$$n^{-1/2}\bigl(X_{\lfloor xn \rfloor} - xn\lambda_1 v_1\bigr) \xrightarrow{\text{d}} V(x),$$

where $V(x)$ *is a continuous Gaussian vector-valued process with* $V(0) = 0$, *mean* $\mathbb{E}\, V(x) = 0$, *and*

$$\mathbb{E}\, V(x)V(y)' = \begin{cases} mx\Sigma_I \bigl(y/x\bigr)^{m^{-1}A'}, & 0 < x \le y, \\ my\bigl(x/y\bigr)^{m^{-1}A}\Sigma_I, & 0 < y \le x. \end{cases} \tag{3.6}$$

Proof. If (2.2) holds, then all assumptions (A1)–(A6) in [11] are satisfied: (A1) and (A2) hold by assumption, (A3) follows from $\lambda_1 = m$, and (A4)–(A6) hold because the urn is irreducible. The general tenable case is easily reduced to this case, see [11, Remark 4.2].

Hence we can apply the results of [11]. Theorem 3.21 there (or [1]) yields $n^{-1}X_n \xrightarrow{\text{a.s.}} \lambda_1 v_1$ and Theorem 3.22 (or [1]) yields asymptotic normality, with the asymptotic covariance matrix equal to $m\Sigma_I$ by Lemma 5.4 in [11]. Similarly, the functional convergence follows from Theorem 3.31(i) in [11], and (3.6) for $x \le y$ follows by Remark 5.7 in [11]; the case $y \le x$ follows by taking the transpose and relabelling. $\square$

Remark 3.2. Several formulas for easy evaluation of Σ_I and the covariances in (3.6) are given in [11, Section 5]. We will use some of them below.

Remark 3.3. The results extend to random ξ_{ij} with only minor changes. Assume that (2.1) and (2.2) or (2.3) (with d_i non-random) hold with probability 1, and that $\xi_{ij}^2 < \infty$ for all i and j. Assume further that (2.4) holds on the average, i.e. $a \cdot \mathbb{E}\,\xi_i = m$ for some $m > 0$. Define A and B by taking the expectations of the right-hand sides of (3.1) and (3.4). Then Theorem 3.1 still holds, see [11].

Remark 3.4. The assumption (2.4) is not necessary, and can be replaced by the assumptions that $\lambda_1 > 0$ and that extinction is impossible; however, the variance $m\Sigma_I$ and the covariance formula (3.6) in Theorem 3.1 have to be replaced by more complicated expressions, see [11, Theorems 3.22 and 3.31].

Remark 3.5. When $\operatorname{Re}\lambda_2 = \frac{1}{2}\lambda_1$, there are similar results but the right normalization factor is $(n \log^d n)^{-1/2}$ for some $d \geq 1$. The case $\operatorname{Re}\lambda_2 > \frac{1}{2}\lambda_1$ is quite different and asymptotic normality does not hold (at least in general). See e.g. [1], [2], [11].

4. Random recursive trees

We apply Theorem 3.1 to random recursive trees as follows.

Proof of Theorem 1.1. Following Mahmoud and Smythe [16], we observe that the distribution of outdegrees is the same as the distribution of types in a generalized Pólya urn with *infinitely* many types $\{0, 1, 2, \dots\}$, all activities $a_i = 1$, and the rule that if a ball with type i is drawn, it is removed and replaced by a ball of type $i+1$ and a ball of type 0. (We start at time 1 with a single ball of type 0.) In our notation, $\xi_i = -\delta_i + \delta_{i+1} + \delta_0$, where δ_i is the unit vector defined by $(\delta_i)_j = \delta_{ij}$.

Theorem 3.1 assumes that the number of types is finite, but luckily we can in this application truncate and lump all high degrees together. Thus, let $M \geq 1$ be an integer and use the types $\{0, 1, \dots, M\}$ only (thus $q = M+1$), where now type M represents all outdegrees $\geq M$. The replacement vectors ξ_i are as in the infinite model when $i < M$, while now $\xi_M = \delta_0$. For example, for $M = 3$ (the case treated in [16]) we have

$$A = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}. \tag{4.1}$$

Exactly one ball is added each time, so (2.4) holds with $m = 1$, and thus $\lambda_1 = 1$ and $u_1 = a = (1, 1, \dots, 1)$. It is easily verified that $v_1 = (1/2, 1/4, \dots, 2^{-M}, 2^{-M})$, i.e. $v_{1i} = 2^{-i-1}$ for $0 \leq i < M$ and $v_{1M} = 2^{-M}$. In particular, Theorem 3.1 shows that $X_{ni}/n \overset{\text{a.s.}}{\to} 2^{-i-1}$ for every $i \geq 0$ (by taking $M > i$).

To find the other eigenvalues of A, we regard $\mathbb{C}^q = \mathbb{C}^{M+1}$ as a subspace of $\ell^1 = \ell^1(\mathbb{N})$ and let $\pi_M : \ell^1 \to \mathbb{C}^q$ be the projection $(x_0, x_1, \dots) \mapsto (x_0, \dots, x_{M-1}, \sum_M^\infty x_i)$. ($\mathbb{C}$ is the set of complex numbers.) Let S be the shift operator on ℓ^1, $S(x_0, x_1, \dots) = (0, x_0, x_1, \dots)$. Then $Av = -v + (u_1 \cdot v)\delta_0 + \pi_M S v$, $v \in \mathbb{C}^q$.

Let $E' := \{v \in \mathbb{C}^q : u_1 \cdot v = 0\}$. Then A maps E' into itself (because u_1 is a left eigenvector) and, on E', $A = -I + \pi_M S$. Thus, for $v \in E'$, $(A+I)v = \pi_M S v$ and, by induction, since $\pi_M S \pi_M = \pi_M S$,

$$(A+I)^k v = \pi_M S^k v, \qquad \text{for } v \in E' \text{ and } k \geq 0. \tag{4.2}$$

In particular, $(A+I)^M v = \pi_M S^M v = 0$ for $v \in E'$, so $A + I$ is nilpotent on E' and the restriction of A to E' has the single eigenvalue -1. Since $E' + \mathbb{C}v_1 = \mathbb{C}^q$, it follows that the eigenvalues of A are 1 and -1, the latter with algebraic multiplicity M. We see further that $(A+I)^{M-1} \neq 0$ on E', so A is not diagonalizable when $M \geq 2$. (The geometric multiplicity of the eigenvalue -1 is 1.)

Since thus $\lambda_2 = -1$, Theorem 3.1 applies, and the vector $\big(n^{-1/2}(X_{ni} - 2^{-i-1}n)\big)_{i=0}^{M-1}$ converges in distribution to a Gaussian vector. Since M is arbitrary, this means joint convergence of for all i, see Remark 1.2.

It remains to find the (co)variances $\sigma_{jk} := \operatorname{Cov}(V_j, V_k)$. Thus, take any $M > j, k$. By [11, Lemma 5.1],

$$\sigma_{jk} = \int_0^\infty \sum_i v_{1i} \big(g_i(s)\big)_j \big(g_i(s)\big)_k e^{-\lambda_1 s} \lambda_1 \, ds, \tag{4.3}$$

where $g_i(s) = Ae^{sA}(I - P_{\lambda_1})\delta_i$. Since $(I - P_{\lambda_1})\delta_i \in E'$ and $P_{\lambda_1}\delta_i = v_1(u_1 \cdot \delta_i) = v_1$, it follows from (4.2) that

$$\begin{aligned} g_i(s) &= Ae^{sA}(\delta_i - v_1) = e^{-s}(A + I - I)e^{s(A+I)}(\delta_i - v_1) \\ &= e^{-s}\pi_M\big((S - I)e^{sS}(\delta_i - v_1)\big). \end{aligned}$$

Consequently, (4.3) yields

$$\sigma_{jk} = \sum_i v_{1i} \int_0^\infty \Big((S-I)e^{sS}(\delta_i - v_1)\Big)_j \Big((S-I)e^{sS}(\delta_i - v_1)\Big)_k e^{-3s} \, ds. \tag{4.4}$$

We may obtain an expression with finite sums only by expanding the terms in this integral. However, we prefer to compute the bivariate generating function of σ_{jk} instead.

In (4.4), v_1 really depends on our choice of M, and thus implicitly on j and k, but it is easily seen that we can replace v_1 by $\tilde{v}_1 := (2^{-i-1})_{i=0}^\infty$ (the limit of v_1 as $M \to \infty$), since the integrand does not depend on i for $i > \max(j, k)$. Let $\mathcal{F}$ be the map of ℓ^1 into the set of analytic functions in the unit disc given by $\mathcal{F}(v)(z) = \sum_{i=0}^\infty v_i z^i$. Note that $\mathcal{F}(Sv)(z) = z\mathcal{F}(v)(z)$. Let $\widetilde{v}(z) := \mathcal{F}(\tilde{v}_1)(z) = \sum_i 2^{-i-1} z^i = 1/(2-z)$. Then (4.4) yields, for

$|z|, |w| \le 1$, say, when absolute summability easily is checked,

$$\sum_{j,k=0}^{\infty} \sigma_{jk} z^j w^k =$$

$$\sum_{i=0}^{\infty} 2^{-i-1} \int_0^{\infty} \mathcal{F}\Big((S-I)e^{sS}(\delta_i - \tilde{v}_1)\Big)(z)\mathcal{F}\Big((S-I)e^{sS}(\delta_i - \tilde{v}_1)\Big)(w)e^{-3s}\,ds$$

$$= \sum_{i=0}^{\infty} 2^{-i-1} \int_0^{\infty} (z-1)e^{sz}\big(z^i - \widetilde{v}(z)\big)(w-1)e^{sw}\big(w^i - \widetilde{v}(w)\big)e^{-3s}\,ds$$

$$= (z-1)(w-1)\Big(\frac{1}{2-zw} - \widetilde{v}(z)\widetilde{v}(w)\Big)\int_0^{\infty} e^{-(3-z-w)s}ds,$$

which yields (1.1). □

Proof of Theorem 1.4. The functional convergence (1.2) follows immediately from Theorem 3.1(ii).

We find the covariance functions as follows. Given j and k we again take $M > j,k$ and combine all high degrees as above. Assume $0 < x \le y$. By (3.6), $\mathbb{E}\, V(y)V(x)' = x(y/x)^A\Sigma_I$. Since $u_1' X_n$ is deterministic, $u_1' V(x) = 0$ and thus $u_1'\Sigma_I = \mathbb{E}\big(u_1' V(1)V(1)'\big) = 0$. Hence we can use (4.2) and obtain, for $0 < x \le y$,

$$\mathbb{E}\, V_k(y)V_j(x) = x\Big(\Big(\frac{y}{x}\Big)^{A}\Sigma_I\Big)_{kj} = \frac{x^2}{y}\Big(\Big(\frac{y}{x}\Big)^{A+I}\Sigma_I\Big)_{kj} = \frac{x^2}{y}\Big(\Big(\frac{y}{x}\Big)^{\pi_M S}\Sigma_I\Big)_{kj}$$

$$= \frac{x^2}{y}\sum_{i=0}^{k} \frac{\ln^i(y/x)}{i!}\Big((\pi_M S)^i\Sigma_I\Big)_{kj} = \frac{x^2}{y}\sum_{i=0}^{k} \frac{\ln^i(y/x)}{i!}\sigma_{k-i,j}. \quad \square$$

5. Random plane recursive trees

Proof of Theorem 1.3. The outdegrees in the random plane recursive trees studied by Mahmoud, Smythe and Szymański [17] can be modelled using a generalized Pólya urn with infinitely many types as in Section 4; the ξ_i are the same, but now the activity $a_i = i + 1$.

It is advantageous to modify the urn so that the activities become the same. We thus replace each ball of type i by $i + 1$ balls of the same type; equivalently, we let the balls represent external vertices as in [17]. This gives a new generalized Pólya urn with infinitely many types, all activities 1, and the transitions given by

$$\xi_{ij} = -(i+1)\delta_{ij} + (i+2)\delta_{i+1,j} + \delta_{0j}.$$

The number of balls of type i in this urn is thus $(i+1)Y_{ni}$.

We truncate as in Section 4 and use the $M + 1$ types $0, \dots, M$ only, with ξ_{Mj} changed to $\delta_{0j} + \delta_{Mj}$. (Note that such truncation does not work in the original urn model representing internal nodes, since the activities vary.)

For example, for $M=3$ (the case treated in [17]) we have

$$A=\left(\xi_{ji}\right)_{i,j=0}^{M}=\begin{pmatrix}0&1&1&1\\2&-2&0&0\\0&3&-3&0\\0&0&4&1\end{pmatrix}. \tag{5.1}$$

Two balls are added each time, so (2.4) holds with $m=2$, and thus $\lambda_1=2$.

It is not difficult to compute the characteristic polynomial of A (by induction) and thus find the other eigenvalues. We will, however, instead find both the eigenvalues and eigenvectors directly.

For the left eigenvectors, $u^{(\lambda)}$ say, we have the equations

$$u_0^{(\lambda)}-(i+1)u_i^{(\lambda)}+(i+2)u_{i+1}^{(\lambda)}=\lambda u_i^{(\lambda)},\qquad 0\le i<M, \tag{5.2a}$$

$$u_0^{(\lambda)}+u_M^{(\lambda)}=\lambda u_M^{(\lambda)}. \tag{5.2b}$$

Let $\Delta u_i^{(\lambda)}:=u_{i+1}^{(\lambda)}-u_i^{(\lambda)}$, $0\le i<M$, and $\Delta u_M^{(\lambda)}:=0$. Then (5.2) can be written

$$(i+2)\Delta u_i^{(\lambda)}=(\lambda-1)u_i^{(\lambda)}-u_0^{(\lambda)},\qquad 0\le i\le M.$$

Hence, for $i<M$,

$$(i+3)\Delta u_{i+1}^{(\lambda)}-(i+2)\Delta u_i^{(\lambda)}=(\lambda-1)\big(u_{i+1}^{(\lambda)}-u_i^{(\lambda)}\big)=(\lambda-1)\Delta u_i^{(\lambda)}$$

or

$$(i+3)\Delta u_{i+1}^{(\lambda)}=(i+1+\lambda)\Delta u_i^{(\lambda)},\qquad 0\le i<M,$$

with the solution (where $c=2\Delta u_0^{(\lambda)}\in\mathbb{C}$)

$$\Delta u_i^{(\lambda)}=\frac{\prod_{j=1}^{i}(j+\lambda)}{(i+2)!}c,\qquad 0\le i\le M. \tag{5.3}$$

If all $\Delta u_i^{(\lambda)}=0$, $u_i^{(\lambda)}$ is a multiple of $a=(1,\dots,1)$; we already know that these are the left eigenvectors with eigenvalue $\lambda_1=2$. For $\lambda\ne 2$, we thus have $c\ne 0$; since $\Delta u_M^{(\lambda)}=0$, (5.3) implies that $j+\lambda=0$, i.e. $\lambda=-j$, for some $j\in\{1,\dots,M\}$. With $\lambda=-k$, (5.3) can be written

$$\Delta u_i^{(-k)}=\frac{(1-k)\cdots(i-k)}{(i+2)!}c=(-1)^i\binom{k+1}{i+2}\frac{c}{k(k+1)}=(-1)^i\binom{k+1}{i+2}c_1,$$

say. This is solved by

$$u_i^{(-k)}=(-1)^{i-1}\binom{k}{i+1}c_1+c_2. \tag{5.4}$$

In particular, $u_0^{(-k)}=-kc_1+c_2$ and $u_M^{(-k)}=c_2$. Hence (5.2b) yields $-kc_1+2c_2=-kc_2$ and thus $c_2=k(k+2)^{-1}c_1$. Conversely, for any such c_1 and c_2, (5.4) solves (5.2) with $\lambda=-k$.

Hence A has the eigenvalues 2, -1, -2, $\dots$, $-M$. In particular, $\lambda_2=-1<\frac{1}{2}\lambda_1$, so Theorem 3.1 applies. We also see that A has $q=M+1$ distinct eigenvalues and thus is diagonalizable.

For the right eigenvectors, we have the equations

$$\sum_{j=1}^{M} v_j^{(\lambda)} = \lambda v_0^{(\lambda)} \tag{5.5a}$$

$$(i+1)v_{i-1}^{(\lambda)} - (i+1)v_i^{(\lambda)} = \lambda v_i^{(\lambda)}, \qquad 1 \le i < M, \tag{5.5b}$$

$$(M+1)v_{M-1}^{(\lambda)} + v_M^{(\lambda)} = \lambda v_M^{(\lambda)}. \tag{5.5c}$$

For $\lambda = \lambda_1 = 2$, $v^{(\lambda)} = v_1$ and (5.5b) yields $(i+1)v_{1,i-1} = (i+3)v_{1,i}$ and thus $v_{1i} = c/(i+2)(i+3)$, $0 \le i < M$. With our normalization $a \cdot v_1 = \sum_0^M v_{1i} = 1$, (5.5a) yields $1 = 3v_{10}$; thus $c = 2$ and, using also (5.5c),

$$\begin{cases} v_{1i} = \dfrac{2}{(i+2)(i+3)}, & 0 \le i < M, \\ v_{1M} = \dfrac{2}{M+2}. \end{cases} \tag{5.6}$$

Theorem 3.1 thus shows that $(i+1)Y_{ni}/n \overset{\text{a.s.}}{\to} 4/(i+2)(i+3)$ as $n \to \infty$.

Moreover, it is easily verified that for each $k = 1, \dots, M$,

$$v_i^{(-k)} = \binom{i+1}{k}, \qquad 0 \le i < M,$$
$$v_M^{(-k)} = -\binom{M+1}{k+1}$$

solves (5.5) with $\lambda = -k$.

Recall that $u \cdot v = 0$ whenever u is a left and v a right eigenvector with different eigenvalues. Moreover, if $u^{(-k)}$ is given by (5.4),

$$u^{(-k)} \cdot v^{(-k)} = (u^{(-k)} - c_2 a) \cdot v^{(-k)} = \sum_{i=0}^{k-1} (-1)^{i-1} \binom{k}{i+1} c_1 \binom{i+1}{k} = (-1)^k c_1,$$

since the only non-zero term is for $i = k-1$. Consequently, if we choose $c_1 = (-1)^k$ in (5.4), i.e.

$$u_i^{(-k)} = (-1)^{k+i+1} \binom{k}{i+1} + (-1)^k \frac{k}{k+2}, \tag{5.7}$$

and write $u_j = u^{(1-j)}$, $v_j = v^{(1-j)}$, $2 \le j \le M+1$, then $u_i \cdot v_j = \delta_{ij}$ and $\{u_i\}_1^{M+1}$ and $\{v_i\}_1^{M+1}$ are dual bases of eigenvectors.

It is now easy to evaluate the integral in (3.5) and obtain, see [11, Lemma 5.3(ii)], with D the diagonal matrix with $D_{ii} = v_{1i}$,

$$\Sigma_I = \sum_{i,j=2}^{M+1} \frac{\lambda_i \lambda_j u_i' D u_j}{\lambda_1 - \lambda_i - \lambda_j} v_i v_j' = \sum_{k,l=1}^{M} \frac{kl(u^{(-k)\prime} D u^{(-l)})}{2+k+l} v^{(-k)} v^{(-l)\prime}. \tag{5.8}$$

Theorem 3.1 thus shows that

$$n^{-1/2}\Big((i+1)Y_{ni} - \frac{4}{(i+2)(i+3)} n\Big) \overset{\text{d}}{\to} \widehat{W}_i,$$

jointly for all $i = 0, \dots, M-1$, where $\widehat{W}_i$ are jointly Gaussian variables with means 0 and covariances

$$
\begin{aligned}
\operatorname{Cov}(\widehat{W}_i, \widehat{W}_j) = \bigl(m\Sigma_I\bigr)_{ij} &= 2\sum_{k,l=1}^{M} \frac{kl(u^{(-k)\prime} Du^{(-l)})}{k+l+2} v_i^{(-k)} v_j^{(-l)} \\
&= 2\sum_{k=1}^{i+1}\sum_{l=1}^{j+1} \frac{kl}{k+l+2}\binom{i+1}{k}\binom{j+1}{l} u^{(-k)\prime} Du^{(-l)} \qquad (5.9)\\
&= 2\sum_{k=1}^{i+1}\sum_{l=1}^{j+1} \frac{(i+1)(j+1)}{k+l+2}\binom{i}{k-1}\binom{j}{l-1} u^{(-k)\prime} Du^{(-l)}.
\end{aligned}
$$

The result follows, with $W_i = \widehat{W}_i/(i+1)$, by the following lemma (and $k \mapsto k+1$, $l \mapsto l+1$). ☐

Lemma 5.1. *For $1 \le k, l \le M$,*

$$
u^{(-k)\prime} Du^{(-l)} = (-1)^{k+l}\Bigl(2\frac{(k+l+2)!}{(k+2)!\,(l+2)!} - 1 - \frac{kl}{(k+2)(l+2)}\Bigr).
$$

Proof. Let, see (5.7),

$$
\begin{aligned}
w_{kl} :=& \Bigl((-1)^k u^{(-k)} - \frac{k}{k+2}u_1\Bigr)' D\Bigl((-1)^l u^{(-l)} - \frac{l}{l+2}u_1\Bigr) \\
=& \sum_{i=0}^{M} v_{1i}\binom{k}{i+1}\binom{l}{i+1} = \sum_{i=0}^{\infty} \frac{2}{(i+2)(i+3)}\binom{k}{i+1}\binom{l}{i+1},
\end{aligned}
$$

because $\binom{k}{i+1} = 0$ when $i \ge M \ge k$. Hence, w_{kl} does not depend on M, and we may regard w_{kl} as defined for all $k, l \ge 1$. Moreover, since $u^{(-k)\prime} Du_1 = u^{(-k)\prime} v_1 = 0$ and similarly $u_1' Du^{(-l)} = 0$, while $u_1' Du_1 = u_1' v_1 = 1$,

$$
w_{kl} = (-1)^{k+l} u^{(-k)\prime} Du^{(-l)} + \frac{kl}{(k+2)(l+2)}, \qquad k, l \le M. \qquad (5.10)
$$

Let

$$
\begin{aligned}
f(x) &:= \sum_{i=0}^{\infty} \frac{x^i}{(i+2)(i+3)} = \sum_{i=0}^{\infty} x^i\Bigl(\frac{1}{(i+2)} - \frac{1}{(i+3)}\Bigr) \\
&= x^{-2}\bigl(-\ln(1-x) - x\bigr) - x^{-3}\bigl(-\ln(1-x) - x - \tfrac12 x^2\bigr) \\
&= x^{-3}\bigl((1-x)\ln(1-x) + x - \tfrac12 x^2\bigr).
\end{aligned}
$$

Then, for $|z|,|w| < 1/2$, or as formal power series,

$$\begin{aligned}
\sum_{k,l=1}^{\infty} w_{kl} z^k w^l &= \sum_{i=0}^{\infty} \frac{2}{(i+2)(i+3)} \sum_{k=1}^{\infty} \binom{k}{i+1} z^k \sum_{l=1}^{\infty} \binom{l}{i+1} z^l \\
&= \sum_{i=0}^{\infty} \frac{2}{(i+2)(i+3)} \frac{z^{i+1}}{(1-z)^{i+2}} \frac{w^{i+1}}{(1-w)^{i+2}} \\
&= \frac{2zw}{(1-z)^2(1-w)^2} f\Bigl(\frac{zw}{(1-z)(1-w)}\Bigr) \\
&= \frac{2(1-z)(1-w)}{z^2w^2} \biggl(\Bigl(1-\frac{zw}{(1-z)(1-w)}\Bigr)\ln\Bigl(1-\frac{zw}{(1-z)(1-w)}\Bigr) \\
&\qquad\qquad + \frac{zw}{(1-z)(1-w)} - \frac{z^2w^2}{2(1-z)^2(1-w)^2}\biggr) \\
&= \frac{2}{z^2w^2}\Bigl((1-z-w)\bigl(\ln(1-z-w)-\ln(1-z)-\ln(1-w)\bigr)+zw\Bigr) \\
&\qquad\qquad - \frac{1}{(1-z)(1-w)}.
\end{aligned}$$

Hence

$$\begin{aligned}
w_{kl} &= 2[z^{k+2}w^{l+2}]\bigl((1-z-w)\ln(1-z-w)\bigr) - 1 \\
&= 2[z^{k+2}w^{l+2}]\Bigl(\sum_{i=2}^{\infty}\frac{(z+w)^i}{i(i-1)} - z - w\Bigr) - 1 \\
&= 2\frac{1}{(k+l+4)(k+l+3)}\binom{k+l+4}{k+2} - 1
\end{aligned}$$

and the result follows by (5.10). ☐

Proof of Theorem 1.5. This follows from Theorem 3.1(ii), using the urn for external vertices with truncation (and dividing by $i+1$) as in the proof of Theorem 1.3. The covariance matrix (3.6), for $0 < x \le y$, is by (5.8), since $v^{(-l)\prime}A' = -lv^{(-l)\prime}$,

$$mx\Sigma_I\bigl(y/x\bigr)^{m^{-1}A'} = 2x\sum_{k,l=1}^{M}\frac{kl(u^{(-k)\prime}Du^{(-l)})}{k+l+2}v^{(-k)}v^{(-l)\prime}\Bigl(\frac{y}{x}\Bigr)^{-l/2}. \tag{5.11}$$

The result follows by a calculation as in (5.9) and Lemma 5.1. ☐

6. Random recursive d-ary trees

Another related random tree is the *random recursive binary tree* defined by growing a subtree of a complete infinite binary tree; we start with the root and add new nodes, each new nodes being randomly put in one of the vacant positions as a child of an existing node. Equivalently, we start with one internal node having two external nodes as children, and in each step a randomly chosen external node is converted to an internal node, with two new external nodes as children. This is the same as a *random binary search tree*. Let Z_{ni} be the number of (internal) vertices with outdegree i when the tree has n vertices.

The distribution of out-degrees in this random tree was studied by Devroye [10], both by an urn model and by another method. In the urn method approach, we observe that the distribution of out-degrees is given by a generalized Pólya urn with 3 types $0, 1, 2$; ξ_i is again as in Section 4 but now the activity is $a_i = 2 - i$.

Alternatively, we may consider external nodes; then there are 2 types 0 and 1, $a_i = 1$ and $\xi_0 = \delta_1$, $\xi_1 = 2\delta_0 - \delta_1$.

The matrix A has eigenvalues 1 and -2 in the external version, and $1, 0, -2$ in the internal version. Hence $\lambda_2 < \frac{1}{2}\lambda_1$, and Theorem 3.1 yields a.s. convergence and asymptotic normality, as shown by [10]. For example, using e.g. [11, Lemma 5.3] to compute the variance, $n^{-1/2}(Z_{n0} - n/3) \xrightarrow{\mathrm{d}} N(0, 2/45)$ [10].

This is easily generalized to a *random recursive d-ary tree*, for any fixed $d \geq 2$. This tree is defined in the same way as a subtree of a complete infinite d-ary tree. In the internal urn model, we have $d + 1$ types $0, \dots, d$, the same ξ_i as in Section 4 and activities $a_i = d - i$. In the external urn model, we have only d types $0, \dots, d-1$, $\xi_i = d\delta_0 - (d-i)\delta_i + (d-i-1)\delta_{i+1}$ and $a_i = 1$.

Arguments as in Section 5 show that the eigenvalues are $\lambda_1 = d - 1$ and $-2, -3, \dots, -d$ in the external version (and also 0 in the internal version). Again, Theorem 3.1 applies (now to either version) and yields a.s. convergence and asymptotic normality. As in Section 5, the eigenvectors are easily computed explicitly, which gives explicit formulas. In particular, for the external version, $v_{1i} = \binom{2d-i-2}{d-1} / \binom{2d-1}{d}$, $i = 0, \dots, d-1$, which leads to

$$\frac{Z_{ni}}{n} \xrightarrow{\text{a.s.}} \binom{2d-i-2}{d-2} \bigg/ \binom{2d-1}{d-1} = \frac{(d-1)(d)_i}{(2d-1)_i}, \qquad 0 \leq i \leq d.$$

The limiting distribution is the same as the distribution of the number of white balls drawn (without replacement) before the first black from an urn with d white and $d-1$ black balls; this is known as a *negative hypergeometric* distribution (shifted by 1) [12, §2.5]. The limiting distribution of the types of external vertices is similarly given by another negative hypergeometric distribution, this time taking $d - 1$ white and d black balls.

For the leaves we have, generalizing the result by [10] given above for $d = 2$,

$$n^{-1/2}\Big(Z_{n0} - \frac{d-1}{2d-1}n\Big) \xrightarrow{\mathrm{d}} N\Big(0, \frac{d(d-1)^2}{(3d-1)(2d-1)^2}\Big).$$

We leave the verification and formulas for variances and covariances for types other than 0 as an exercise.

Note that for $d \geq 3$, this random recursive d-ary tree is *not* the same as a random d-ary search tree. The latter can also be treated by an urn model

[15], but $\operatorname{Re}\lambda_2 < \frac{1}{2}\lambda_1$ only for $d \le 26$; for larger d asymptotic normality does not hold [9]; see also [14] and [8].

References

[1] K.B. Athreya & S. Karlin, Embedding of urn schemes into continuous time Markov branching processes and related limit theorems. *Ann. Math. Statist.* **39** (1968), 1801–1817.

[2] K.B. Athreya & P.E. Ney, *Branching Processes.* Springer-Verlag, Berlin, 1972.

[3] A. Bagchi & A.K. Pal, Asymptotic normality in the generalized Pólya–Eggenberger urn model, with an application to computer data structures. *SIAM J. Algebraic Discrete Methods* **6** (1985), no. 3, 394–405.

[4] Z.D. Bai & F. Hu, Asymptotic theorems for urn models with nonhomogeneous generating matrices. *Stochastic Process. Appl.* **80** (1999), no. 1, 87–101.

[5] Z.D. Bai, F. Hu & L.-X. Zhang, Gaussian approximation theorems for urn models and their applications. *Ann. Appl. Probab.* **12** (2002), no. 4, 1149–1173.

[6] F. Bergeron, P. Flajolet & B. Salvy, Varieties of increasing trees. In *CAAP '92 (Rennes, 1992), Lecture Notes in Comput. Sci.* 581, Springer, Berlin, 1992, 24–48.

[7] P. Billingsley, *Convergence of Probability Measures.* Wiley, New York, 1968.

[8] B. Chauvin & N. Pouyanne, m-ary search trees when $m \ge 27$: a strong asymptotics for the space requirements. *Random Structures Algorithms* **24** (2004), no. 2, 133–154.

[9] H.-H. Chern & H.-K. Hwang, Phase changes in random m-ary search trees and generalized quicksort. *Random Structures Algorithms* **19** (2001), no. 3-4, 316–358.

[10] L. Devroye, Limit laws for local counters in random binary search trees. *Random Structures Algorithms* **2** (1991), no. 3, 303–315.

[11] S. Janson, Functional limit theorems for multitype branching processes and generalized Pólya urns. *Stochastic Process. Appl.* **110** (2004), no. 2, 177–245.

[12] N.L. Johnson & S. Kotz, *Urn models and their application.* Wiley, New York, 1977.

[13] S. Karlin, *A First Course in Stochastic Processes.* Academic Press, New York, 1969.

[14] W. Lew & H.M. Mahmoud, The joint distribution of elastic buckets in multiway search trees. *SIAM J. Comput.* **23** (1994), no. 5, 1050–1074.

[15] H.M. Mahmoud, The size of random bucket trees via urn models. *Acta Inform.* **38** (2002), no. 11-12, 813–838.

[16] H.M. Mahmoud & R.T. Smythe, Asymptotic joint normality of outdegrees of nodes in random recursive trees. *Random Structures Algorithms* **3** (1992), no. 3, 255–266.

[17] H.M. Mahmoud, R.T. Smythe & J. Szymański, On the structure of random plane-oriented recursive trees and their branches. *Random Structures Algorithms* **4** (1993), no. 2, 151–176.

[18] A. Meir & J.W. Moon, Recursive trees with no nodes of out-degree one. *Congr. Numer.* **66** (1988), 49–62.

[19] N.S. Na & A. Rapoport, Distribution of nodes of a tree by degree. *Math. Biosci.* **6** (1970), 313–329.

[20] D. Najock & C.C. Heyde, On the number of terminal vertices in certain random trees with an application to stemma construction in philology. *J. Appl. Probab.* **19** (1982), no. 3, 675–680.

[21] E. Seneta, *Nonnegative matrices and Markov chains.* 2nd ed. Springer-Verlag, New York, 1981.

[22] R.T. Smythe, Central limit theorems for urn models. *Stochastic Process. Appl.* **65** (1996), no. 1, 115–137.

[23] R.T. Smythe and H. Mahmoud, A survey of recursive trees. *Theory Probab. Math. Statist.* **51** (1995), 1–27.

Department of Mathematics, Uppsala University, PO Box 480, S-751 06 Uppsala, Sweden

E-mail address: `svante.janson@math.uu.se`

URL: `http://www.math.uu.se/~svante/`