The Subtree Size Profile of Plane-oriented Recursive Trees

Michael Fuchs

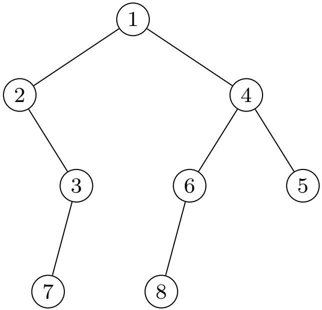
Department of Applied Mathematics
National Chiao Tung University

Hsinchu, Taiwan

ANALCO11, January 22nd, 2011

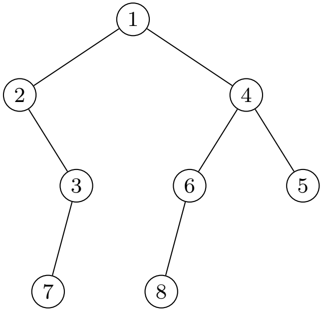
Profiles of Trees

Rooted tree of size n.



Profiles of Trees

Rooted tree of size n.

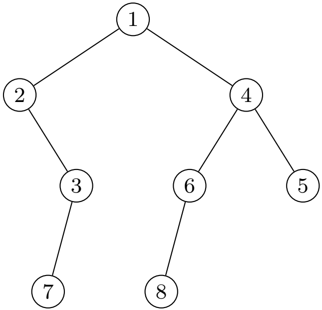


Node profile:

# of nodes at level k.

Profiles of Trees

Rooted tree of size n.



Node profile:

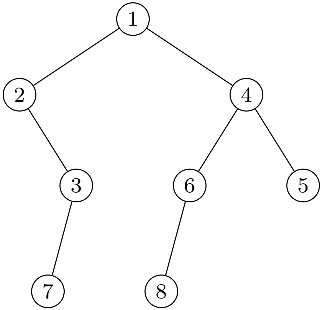
$\#$ of nodes at level k.

Subtree size profile:

$\#$ of subtrees of size k.

Profiles of Trees

Rooted tree of size n.



Node profile:

$\#$ of nodes at level k.

Subtree size profile:

$\#$ of subtrees of size k.

Both are a double-indexed sequence $X_{n,k}$.

Recent Studies of Profile

Node Profile: extensively studied for many classes of trees.

Drmota and Gittenberger (1997); Chauvin, Drmota, Jabbour-Hattab (2001); Chauvin, Klein, Marckert, Rouault (2005); Drmota and Hwang (2005); Fuchs, Hwang, Neininger (2006); Hwang (2007); Drmota, Janson, Neininger (2008); Park, Hwang, Nicodeme, Szpankowski (2009); Drmota and Szpankowski (2010); etc.

Recent Studies of Profile

Node Profile: extensively studied for many classes of trees.

Drmota and Gittenberger (1997); Chauvin, Drmota, Jabbour-Hattab (2001); Chauvin, Klein, Marckert, Rouault (2005); Drmota and Hwang (2005); Fuchs, Hwang, Neininger (2006); Hwang (2007); Drmota, Janson, Neininger (2008); Park, Hwang, Nicodeme, Szpankowski (2009); Drmota and Szpankowski (2010); etc.

Subtree Size Profile: mainly studied for binary trees and increasing trees.

Feng, Miao, Su (2006); Feng, Mahmoud, Su (2007); Feng, Mahmoud, Panholzer (2008); Fuchs (2008); Chang and Fuchs (2010); Dennert and Grübel (2010); etc.

Why studying the Subtree Size Profile?

- Fine shape characteristic

Why studying the Subtree Size Profile?

- Fine shape characteristic, e.g.,

$$T_n := \text{total path length};$$
$$W_n := \text{Wiener index},$$

then,

$$T_n = \sum_{k\geq 0} kX_{n,k} \qquad \text{and} \qquad W_n = \sum_{k\geq 0}(n-k)kX_{n,k}.$$

Why studying the Subtree Size Profile?

- Fine shape characteristic, e.g.,

$$T_n := \text{total path length};$$
$$W_n := \text{Wiener index},$$

then,

$$T_n = \sum_{k\geq 0} kX_{n,k} \qquad \text{and} \qquad W_n = \sum_{k\geq 0}(n-k)kX_{n,k}.$$

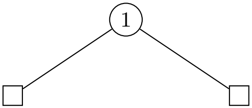
- Contains information about occurrences of patterns.

  Important in many fields such as Computer Science (compressing, etc.), Mathematical Biology (phylogenetics), etc.

Random Binary Trees

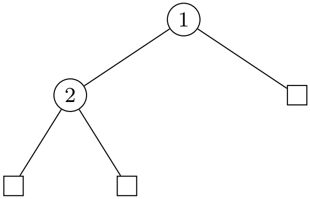
Randomly pick an external node and replace it by a cherry.

Random Binary Trees



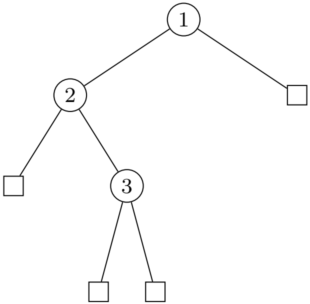
Randomly pick an external node and replace it by a cherry.

Random Binary Trees



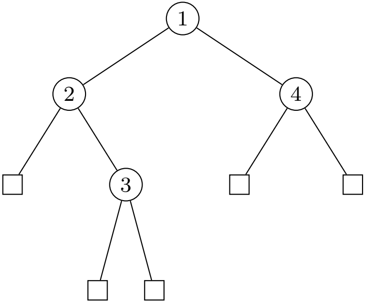
Randomly pick an external node and replace it by a cherry.

Random Binary Trees



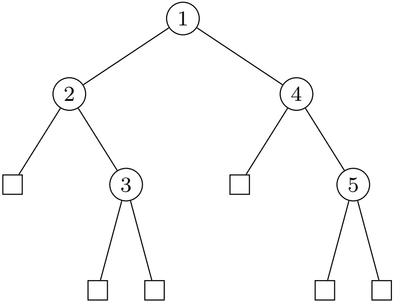
Randomly pick an external node and replace it by a cherry.

Random Binary Trees



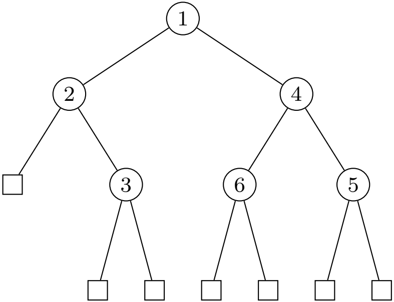
Randomly pick an external node and replace it by a cherry.

Random Binary Trees



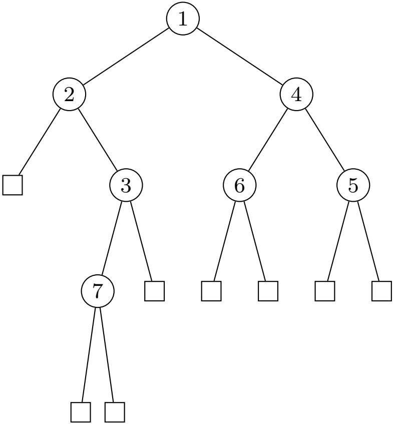
Randomly pick an external node and replace it by a cherry.

Random Binary Trees



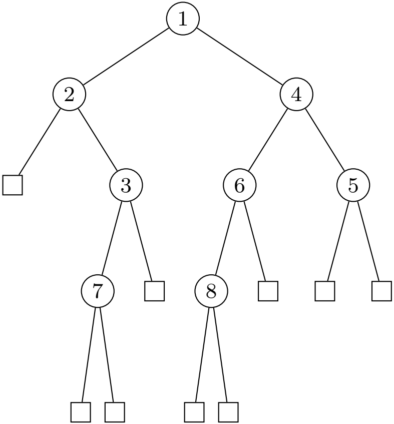
Randomly pick an external node and replace it by a cherry.

Random Binary Trees



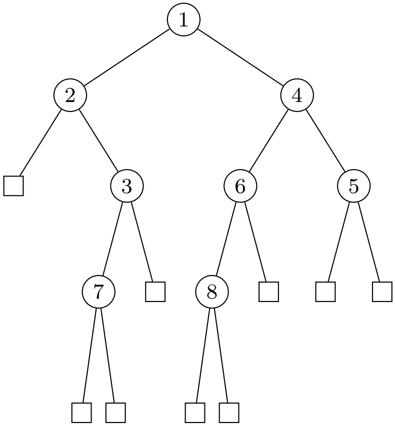
Randomly pick an external node and replace it by a cherry.

Random Binary Trees



Randomly pick an external node and replace it by a cherry.

Random Binary Trees



Randomly pick an external node and replace it by a cherry.

Equivalent to random binary search tree.

Random Recursive Trees

Randomly pick an external node and replace it by an internal node.

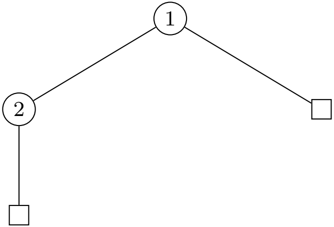
Add external nodes.

Random Recursive Trees

Randomly pick an external node and replace it by an internal node.

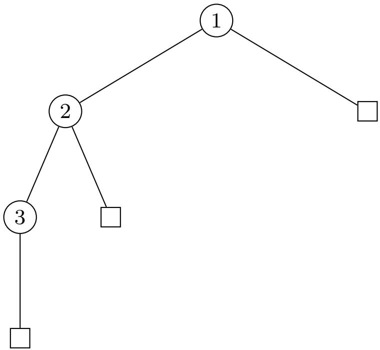
Add external nodes.

Random Recursive Trees



Randomly pick an external node and replace it by an internal node.

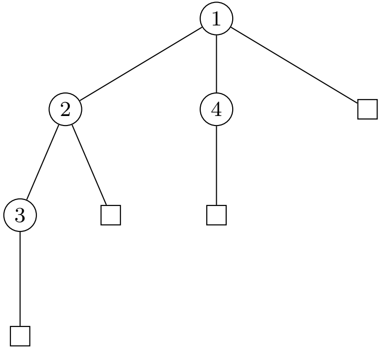
Add external nodes.



Randomly pick an external node and replace it by an internal node.

Add external nodes.

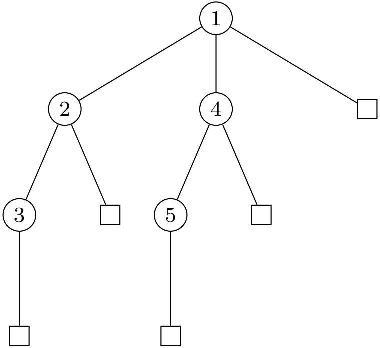
Random Recursive Trees



Randomly pick an external node and replace it by an internal node.

Add external nodes.

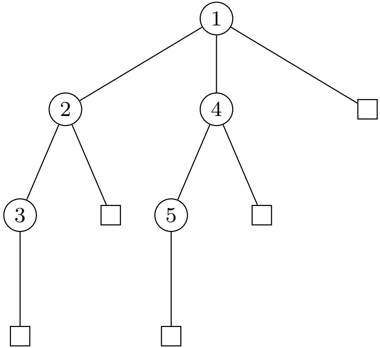
Random Recursive Trees



Randomly pick an external node and replace it by an internal node.

Add external nodes.

Random Recursive Trees



Randomly pick an external node and replace it by an internal node.

Add external nodes.

Same as uniform model for non-plane trees with increasing labels.

Random Plane-oriented Recursive Trees (PORTs)

Randomly pick an external node and replace it by an internal node.

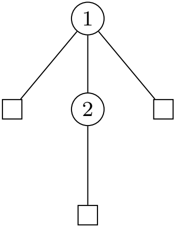
Add external nodes.

Random Plane-oriented Recursive Trees (PORTs)

Randomly pick an external node and replace it by an internal node.

Add external nodes.

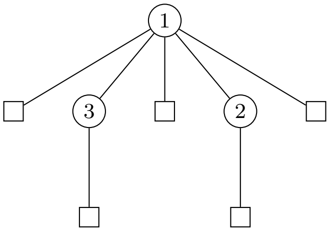
Random Plane-oriented Recursive Trees (PORTs)



Randomly pick an external node and replace it by an internal node.

Add external nodes.

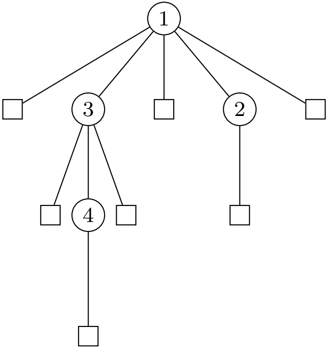
Random Plane-oriented Recursive Trees (PORTs)



Randomly pick an external node and replace it by an internal node.

Add external nodes.

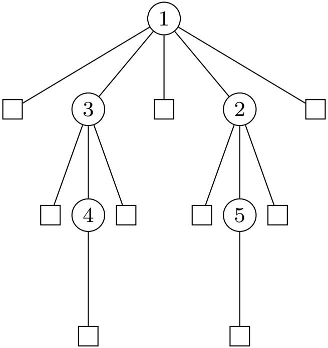
Random Plane-oriented Recursive Trees (PORTs)



Randomly pick an external node and replace it by an internal node.

Add external nodes.

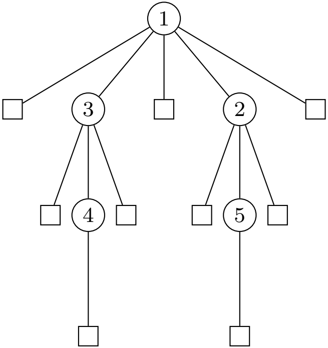
Random Plane-oriented Recursive Trees (PORTs)



Randomly pick an external node and replace it by an internal node.

Add external nodes.

Random Plane-oriented Recursive Trees (PORTs)



Randomly pick an external node and replace it by an internal node.

Add external nodes.

Same as uniform model for plane trees with increasing labels.

Importance of the Random Models

- **Random Binary Trees**

  Binary search tree, Quicksort, Yule-Harding Model in Phylogenetics, Coalescent Model, etc.

- **Random Recursive Trees**

  Simple model for spread of epidemics, for pyramid schemes, for stemma construction of philology, etc. Also, used in computational geometry and in Hopf algebras.

- **Random PORTs**

  One of the simplest network models (for instance for WWW).

Limit Laws for Random Binary Trees

Theorem (Feng, Mahmoud, Panholzer; F.)

(i) *(Normal range) Let* $k = o\left(\sqrt{n}\right)$*. Then,*

$$\frac{X_{n,k} - \mu_{n,k}}{\sigma_{n,k}} \xrightarrow{d} \mathcal{N}(0,1).$$

(ii) *(Poisson range) Let* $k \sim c\sqrt{n}$*. Then,*

$$X_{n,k} \xrightarrow{d} \mathrm{Poisson}(2c^{-2}).$$

(iii) *(Degenerate range) Let* $k < n$ *and* $\sqrt{n} = o(k)$*. Then,*

$$X_{n,k} \xrightarrow{L_1} 0.$$

Limit Laws for Random Binary Trees

Theorem (Feng, Mahmoud, Panholzer; F.)

(i) *(Normal range) Let* $k = o(\sqrt{n})$. *Then,*

$$\frac{X_{n,k} - \mu_{n,k}}{\sigma_{n,k}} \xrightarrow{d} \mathcal{N}(0,1).$$

(ii) *(Poisson range) Let* $k \sim c\sqrt{n}$. *Then,*

$$X_{n,k} \xrightarrow{d} \mathrm{Poisson}(2c^{-2}).$$

(iii) *(Degenerate range) Let* $k < n$ *and* $\sqrt{n} = o(k)$. *Then,*

$$X_{n,k} \xrightarrow{L_1} 0.$$

Similar result holds for random recursive trees as well.

Consequences for Occurrences of Pattern Sizes

Pattern=Subtree on the fringe of the tree.

# of patterns of size k:

$$C_k = \frac{1}{k+1}\binom{2k}{k} \sim \frac{4^k}{\sqrt{\pi}k^{3/2}}.$$

Consequences for Occurrences of Pattern Sizes

Pattern=Subtree on the fringe of the tree.

$\#$ of patterns of size k:

$$C_k = \frac{1}{k+1}\binom{2k}{k} \sim \frac{4^k}{\sqrt{\pi}k^{3/2}}.$$

Hence, all patterns can just occur up to a size $\mathcal{O}(\log n)$.

Consequences for Occurrences of Pattern Sizes

Pattern=Subtree on the fringe of the tree.

# of patterns of size k:

$$C_k = \frac{1}{k+1}\binom{2k}{k} \sim \frac{4^k}{\sqrt{\pi}k^{3/2}}.$$

Hence, all patterns can just occur up to a size $\mathcal{O}(\log n)$.

On the other hand, our result shows:

- Pattern sizes occur until $o(\sqrt{n})$.
- Pattern sizes sporadically exist around $\sqrt{n}$.
- Patterns with sizes beyond $\sqrt{n}$ are unlikely.

Why are Pattern Sizes beyond $\sqrt{n}$ unlikely?

Why are Pattern Sizes beyond $\sqrt{n}$ unlikely?

T_n =total path length.

It is well-known that T_n is of order $n \log n$ with high probability.

Why are Pattern Sizes beyond $\sqrt{n}$ unlikely?

T_n =total path length.

It is well-known that T_n is of order $n \log n$ with high probability.

Recall that

$$T_n = \sum_{k=0}^{n-1} k X_{n,k}$$

Why are Pattern Sizes beyond $\sqrt{n}$ unlikely?

T_n =total path length.

It is well-known that T_n is of order $n \log n$ with high probability.

Recall that

$$T_n = \sum_{k=0}^{n-1} k X_{n,k}$$

Hence, if all pattern sizes up to k_0 exist, then

$$\Theta(k_0^2) = \sum_{k \le k_0} k \le T_n = \Theta(n \log n).$$

Thus, pattern sizes beyond $\sqrt{n \log n}$ are very unlikely.

Method of Moments

Theorem

Z_n, Z random variables. If

$$\mathbb{E}(Z_n^m) \longrightarrow \mathbb{E}(Z^m)$$

for all $m \geq 1$ and Z is uniquely determined by its moments, then

$$Z_n \xrightarrow{d} Z.$$

Method of Moments

Theorem

Z_n, Z random variables. If

$$\mathbb{E}(Z_n^m) \longrightarrow \mathbb{E}(Z^m)$$

for all $m \geq 1$ and Z is uniquely determined by its moments, then

$$Z_n \xrightarrow{d} Z.$$

If Z is standard normal, then

$$\mathbb{E}(Z_n^m) = \begin{cases} 0, & \text{if } m \text{ is odd;} \\ m!/(2^{m/2}(m/2)!), & \text{if } m \text{ is even.} \end{cases}$$

Proof of the Limit Laws

Feng, Mahmoud, Panholzer:

- Derived a complicated explicit expression for centered moments.
- Used the exact expression to obtain first order asymptotics.

  Many cancellations!

Proof of the Limit Laws

Feng, Mahmoud, Panholzer:

- Derived a complicated explicit expression for centered moments.
- Used the exact expression to obtain first order asymptotics.

  Many cancellations!

Our approach:

- Derived a recurrence for centered moments.
- Derived first order asymptotics via induction.

Proof of the Limit Laws

Feng, Mahmoud, Panholzer:

- Derived a complicated explicit expression for centered moments.
- Used the exact expression to obtain first order asymptotics.

  Many cancellations!

Our approach:

- Derived a recurrence for centered moments.
- Derived first order asymptotics via induction.

Our approach is easier and can be applied to other random trees.

Theorem (F.)

(i) *(Normal range) Let* $k = o(\sqrt{n})$. *Then,*

$$\frac{X_{n,k} - \mu n}{\sigma\sqrt{n}} \xrightarrow{d} \mathcal{N}(0,1),$$

where $\mu = 1/(2k^2)$ *and*

$$\sigma^2 = \frac{8k^2 - 4k - 8}{(4k^2 - 1)^2} - \frac{(2k-3)!!^2}{(k-1)!^2 4^{k-1} k(2k+1)}.$$

(ii) *(Poisson range) Let* $k \sim c\sqrt{n}$. *Then,*

$$X_{n,k} \xrightarrow{d} \mathrm{Poisson}(2c^{-2}).$$

(iii) *(Degenerate range) Let* $k < n$ *and* $\sqrt{n} = o(k)$. *Then,*

$$X_{n,k} \xrightarrow{L_1} 0.$$

Proof for Fixed k

Set

$$\bar{A}_k^{[m]}(z) = \sum_{n\geq 1} \tau_n \mathbb{E}(X_{n,k} - \mu n)\frac{z^n}{n!}$$

Proof for Fixed k

Set

$$\bar{A}_k^{[m]}(z) = \sum_{n\geq 1} \tau_n \mathbb{E}(X_{n,k} - \mu n)\frac{z^n}{n!}$$

Then,

$$\frac{\mathrm{d}}{\mathrm{d}z}\bar{A}_k^{[m]}(z) = \frac{\bar{A}_k^{[m]}(z)}{1-2z} + \bar{B}_k^{[m]}(z),$$

where $\bar{B}_k^{[m]}(z)$ is a function of $\bar{A}_k^{[i]}(z)$ with $i < m$.

Proof for Fixed k

Set

$$\bar{A}_k^{[m]}(z) = \sum_{n\geq 1} \tau_n \mathbb{E}(X_{n,k} - \mu n)\frac{z^n}{n!}$$

Then,

$$\frac{\mathrm{d}}{\mathrm{d}z}\bar{A}_k^{[m]}(z) = \frac{\bar{A}_k^{[m]}(z)}{1-2z} + \bar{B}_k^{[m]}(z),$$

where $\bar{B}_k^{[m]}(z)$ is a function of $\bar{A}_k^{[i]}(z)$ with $i < m$.

Above ODE has solution

$$\bar{A}_k^{[m]}(z) = \frac{1}{\sqrt{1-2z}}\int_0^z \bar{B}_k^{[m]}(t)\sqrt{1-2t}\mathrm{d}t.$$

Proof for Fixed k

Set

$$\bar{A}_k^{[m]}(z) = \sum_{n\geq 1} \tau_n \mathbb{E}(X_{n,k} - \mu n)\frac{z^n}{n!}$$

Then,

$$\frac{\mathrm{d}}{\mathrm{d}z}\bar{A}_k^{[m]}(z) = \frac{\bar{A}_k^{[m]}(z)}{1-2z} + \bar{B}_k^{[m]}(z),$$

where $\bar{B}_k^{[m]}(z)$ is a function of $\bar{A}_k^{[i]}(z)$ with $i < m$.

Above ODE has solution

$$\bar{A}_k^{[m]}(z) = \frac{1}{\sqrt{1-2z}}\int_0^z \bar{B}_k^{[m]}(t)\sqrt{1-2t}\mathrm{d}t.$$

Asymptotics of centered moments via induction ("moment-pumping").

Singularity Analysis

Consider

$$\Delta = \{z \ : \ |z| < r, \ z \neq 1/2, \ |\arg(z-1/2)| > \varphi\},$$

where $r > 1$ and $0 < \varphi < \pi/2$.

Theorem (Flajolet and Odlyzko)

(i) *For* $\alpha \in \mathbb{C} \setminus \{0, -1, -2, \cdots\}$

$$[z^n](1-2z)^{-\alpha} \sim \frac{2^n n^{\alpha-1}}{\Gamma(\alpha)} \left(1 + \frac{\alpha(\alpha-1)}{2n} + \cdots\right).$$

(ii) *Let* $f(z)$ *be analytic in* Δ. *Then,*

$$f(z) = \mathcal{O}\left((1-2z)^{-\alpha}\right) \quad \Rightarrow \quad [z^n]f(z) = \mathcal{O}\left(2^n n^{\alpha-1}\right).$$

Closure Properties

Theorem

Let $f(z)$ be analytic in Δ with

$$f(z)=\sum_{j=0}^{J}c_j(1-2z)^{\alpha_j}+\mathcal{O}\left((1-2z)^{A}\right),$$

where $\alpha_j, A\neq 1$. Then, $\int_0^z f(t)\mathrm{d}t$ is analytic in Δ and

(i) *if $A>-1$, then for some explicit c*

$$\int_0^z f(t)\mathrm{d}t=-\frac{1}{2}\sum_{j=0}^{J}\frac{c_j}{\alpha_j+1}(1-2z)^{\alpha_j+1}+c+\mathcal{O}\left((1-2z)^{A+1}\right);$$

(ii) *if $A<-1$, then as above but without c.*

Asymptotic Expansions

Proposition

$\bar{A}_k^{[m]}(z)$ *is analytic in* Δ.

Moreover,

$$\bar{A}_k^{[2m-1]}(z) = \mathcal{O}\left((1-2z)^{3/2-m}\right)$$

and

$$\bar{A}_k^{[2m]}(z) = \frac{(2m)!(2m-3)!!\sigma^{2m}}{4^m m!}(1-2z)^{1/2-m} + \mathcal{O}\left((1-2z)^{1-m}\right).$$

Asymptotic Expansions

Proposition

$\bar{A}_k^{[m]}(z)$ is analytic in Δ.

Moreover,

$$\bar{A}_k^{[2m-1]}(z) = \mathcal{O}\left((1-2z)^{3/2-m}\right)$$

and

$$\bar{A}_k^{[2m]}(z) = \frac{(2m)!(2m-3)!!\sigma^{2m}}{4^m m!}(1-2z)^{1/2-m} + \mathcal{O}\left((1-2z)^{1-m}\right).$$

From this, by singularity analysis,

$$\mathbb{E}(X_{n,k}-\mu n)^m = \begin{cases} 0, & \text{if } m \text{ is odd;} \\ m!/(2^{m/2}(m/2)!), & \text{if } m \text{ is even.} \end{cases}$$

Proof for Varying k

Set $\bar{A}_{n,k}^{[m]} = \mathbb{E}(X_{n,k} - \mu n)^m$.

Proof for Varying k

Set $\bar{A}^{[m]}_{n,k} = \mathbb{E}(X_{n,k} - \mu n)^m$. Then,

$$\bar{A}^{[m]}_{n,k} = \sum_{1 \le j < n} \pi_{n,j}(\bar{A}^{[m]}_{j,k} + \bar{A}^{[m]}_{n-j,k}) + \bar{B}^{[m]}_{n,k},$$

where $\bar{B}^{[m]}_{n,k}$ is a function of $\bar{A}^{[i]}_{n,k}$ with $i < m$ and

$$\pi_{n,j} = \frac{2(n-j)C_j C_{n-j}}{nC_n}.$$

Proof for Varying k

Set $\bar{A}_{n,k}^{[m]} = \mathbb{E}(X_{n,k} - \mu n)^m$. Then,

$$\bar{A}_{n,k}^{[m]} = \sum_{1 \le j < n} \pi_{n,j} (\bar{A}_{j,k}^{[m]} + \bar{A}_{n-j,k}^{[m]}) + \bar{B}_{n,k}^{[m]},$$

where $\bar{B}_{n,k}^{[m]}$ is a function of $\bar{A}_{n,k}^{[i]}$ with $i < m$ and

$$\pi_{n,j} = \frac{2(n-j)C_j C_{n-j}}{nC_n}.$$

We have

$$\bar{A}_{n,k}^{[m]} = \sum_{k+1 \le j \le n} \frac{(n+1-j)C_j}{C_n} \bar{B}_{j,k}^{[m]}.$$

Asymptotic Expansions

Proposition

We have,

$$\bar{A}_{n,k}^{[2m-1]} = o\left(\left(\frac{n}{k^2}\right)^{m-1/2}\right), \qquad \bar{A}_{n,k}^{[2m]} \sim g_m \left(\frac{n}{2k^2}\right)^m,$$

where

$$g_m = (2m)!/(2^m m!).$$

Asymptotic Expansions

Proposition

Uniformly in n, k, m

$$\bar{A}_{n,k}^{[m]} = \mathcal{O}\left(\max\left\{\frac{n}{k^2}, \left(\frac{n}{k^2}\right)^{m/2}\right\}\right).$$

Proposition

We have,

$$\bar{A}_{n,k}^{[2m-1]} = o\left(\left(\frac{n}{k^2}\right)^{m-1/2}\right), \qquad \bar{A}_{n,k}^{[2m]} \sim g_m \left(\frac{n}{2k^2}\right)^{m},$$

where

$$g_m = (2m)!/(2^m m!).$$

Simple Classes of Increasing Trees (i)

Consider rooted, plane trees with increasing labels.

Let ϕ_r be a *weight sequence* with $\phi_0 > 0$ and $\phi_r > 0$ for some $r \geq 2$. Denote by $\phi(\omega)$ the OGF of ϕ_r.

Simple Classes of Increasing Trees (i)

Consider rooted, plane trees with increasing labels.

Let ϕ_r be a *weight sequence* with $\phi_0 > 0$ and $\phi_r > 0$ for some $r \geq 2$. Denote by $\phi(\omega)$ the OGF of ϕ_r.

Define the weight of a tree T as

$$\omega(T) = \prod_{v \in T} \phi_{d(v)},$$

where $d(v)$ is the out-degree of v.

Simple Classes of Increasing Trees (i)

Consider rooted, plane trees with increasing labels.

Let ϕ_r be a *weight sequence* with $\phi_0 > 0$ and $\phi_r > 0$ for some $r \geq 2$. Denote by $\phi(\omega)$ the OGF of ϕ_r.

Define the weight of a tree T as

$$\omega(T) = \prod_{v \in T} \phi_{d(v)},$$

where $d(v)$ is the out-degree of v.

Set

$$\tau_n = \sum_{\#T=n} \omega(T)$$

which is the cumulative weight of all trees of size n.

Simple Classes of Increasing Trees (ii)

Define the probability of a tree of size n as

$$P(T) = \frac{\omega(T)}{\tau_n}.$$

Simple Classes of Increasing Trees (ii)

Define the probability of a tree of size n as

$$P(T) = \frac{\omega(T)}{\tau_n}.$$

Previous Models

- Random binary trees: $\phi_0 = 1, \phi_1 = 2, \phi_2 = 1$ and $\phi_r = 0$ for $r \geq 3$;
- Random increasing trees: $\phi_r = 1/r!$.
- Random PORTs: $\phi_r = 1$.

Simple Classes of Increasing Trees (ii)

Define the probability of a tree of size n as

$$P(T) = \frac{\omega(T)}{\tau_n}.$$

Previous Models

- Random binary trees: $\phi_0 = 1, \phi_1 = 2, \phi_2 = 1$ and $\phi_r = 0$ for $r \geq 3$;
- Random increasing trees: $\phi_r = 1/r!$.
- Random PORTs: $\phi_r = 1$.

All these models can be obtained from a tree evolution process.

Grown Simple Classes of Increasing Trees

Theorem (Panholzer and Prodinger)

All simple classes of increasing tree which can be obtained via a tree evolution process are

- *Random d-ary trees:* $\phi(\omega) = \phi_0(1 + ct/\phi_0)^d$ *with* $c > 0, d \in \{2, 3, \dots\}$*;*
- *Random increasing trees:* $\phi(\omega) = \phi_0 e^{ct/\phi_0}$ *with* $c > 0$*;*
- *Generalized random PORTs:* $\phi(\omega) = \phi_0(1 - ct/\phi_0)^{-r-1}$ *with* $c > 0, r > 1$*.*

Grown Simple Classes of Increasing Trees

Theorem (Panholzer and Prodinger)

All simple classes of increasing tree which can be obtained via a tree evolution process are

- *Random d-ary trees:* $\phi(\omega) = \phi_0(1 + ct/\phi_0)^d$ *with* $c > 0, d \in \{2, 3, \ldots\}$;
- *Random increasing trees:* $\phi(\omega) = \phi_0 e^{ct/\phi_0}$ *with* $c > 0$;
- *Generalized random PORTs:* $\phi(\omega) = \phi_0(1 - ct/\phi_0)^{-r-1}$ *with* $c > 0, r > 1$.

For stochastic properties, it is sufficient to consider the cases with $\phi_0 = c = 1$.

Mean for Grown Simple Classes of Increasing Trees

Proposition

For random d-ary trees,

$$\mu_{n,k} := \mathbb{E}(X_{n,k}) = \frac{d((d-1)n+1)}{((d-1)k+d)((d-1)k+1)}.$$

Mean for Grown Simple Classes of Increasing Trees

Proposition

For random d-ary trees,

$$\mu_{n,k} := \mathbb{E}(X_{n,k}) = \frac{d((d-1)n+1)}{((d-1)k+d)((d-1)k+1)}.$$

Proposition

For generalized random PORTs,

$$\mu_{n,k} := \mathbb{E}(X_{n,k}) = \frac{(r-1)(rn-1)}{(rk+r-1)(rk-1)}.$$

Limit Laws for d-ary Trees

Theorem (F.)

(i) *(Normal range) Let* $k = o(\sqrt{n})$ *and* $k \to \infty$. *Then,*

$$\frac{X_{n,k} - \mu_{n,k}}{\sqrt{\mu_{n,k}}} \xrightarrow{d} \mathcal{N}(0,1).$$

(ii) *(Poisson range) Let* $k \sim c\sqrt{n}$. *Then,*

$$X_{n,k} \xrightarrow{d} \mathrm{Poisson}(2c^{-2}).$$

(iii) *(Degenerate range) Let* $k < n$ *and* $\sqrt{n} = o(k)$. *Then,*

$$X_{n,k} \xrightarrow{L_1} 0.$$

Limit Laws for Generalized Random PORTs

Theorem (F.)

(i) *(Normal range) Let* $k = o\left(\sqrt{n}\right)$ *and* $k \to \infty$. *Then,*

$$\frac{X_{n,k} - \mu_{n,k}}{\sqrt{\mu_{n,k}}} \xrightarrow{d} \mathcal{N}(0,1).$$

(ii) *(Poisson range) Let* $k \sim c\sqrt{n}$. *Then,*

$$X_{n,k} \xrightarrow{d} \mathrm{Poisson}(2c^{-2}).$$

(iii) *(Degenerate range) Let* $k < n$ *and* $\sqrt{n} = o(k)$. *Then,*

$$X_{n,k} \xrightarrow{L_1} 0.$$

Summary

- Inductive approach for deriving limit laws of subtree size profile.

Summary

- Inductive approach for deriving limit laws of subtree size profile.
- Approach can be applied to many classes of random trees and shows universality of the phenomena observed by Feng, Mahmoud and Panholzer.

Summary

- Inductive approach for deriving limit laws of subtree size profile.

- Approach can be applied to many classes of random trees and shows universality of the phenomena observed by Feng, Mahmoud and Panholzer.

  Simple classes of trees (without labels) can also be treated. For instance, random Catalan trees were considered in

  Chang and Fuchs (2010). Limit Laws for Patterns in Phylogenetic Trees, *J. Math. Biol.*, **60**, 481-512.

Summary

- Inductive approach for deriving limit laws of subtree size profile.

- Approach can be applied to many classes of random trees and shows universality of the phenomena observed by Feng, Mahmoud and Panholzer.

  Simple classes of trees (without labels) can also be treated. For instance, random Catalan trees were considered in

  Chang and Fuchs (2010). Limit Laws for Patterns in Phylogenetic Trees, *J. Math. Biol.*, **60**, 481-512.

- Approach can be refined to obtain Berry-Esseen bounds, LLT and Poisson approximation results.