Several topics in Experimental Mathematics

Andrew Lohr

[1]Rutgers University, New Brunswick, New Jersey (ajl213@math.rutgers.edu)

April 2, 2018

Galton-Watson Random Trees: Definition

Defined in terms of some finite set $S \subset \mathbb{N}^+$
For the root, Either it is a leaf, or its number of children is in S.

Our question

Given a tree T sampled from a Galton-Watson distribution. Compute the "total height".

$$\sum_{v \in T} h(v)$$

What distribution does this have if we generate the tree with certain classes of Galton-Watson processes?

introducing generating functions

A generating function $f(x)$ for any sequence y_n is the FORMAL sum

$$f(x) = \sum_{n=0}^{\infty} y_n x^n$$

A different way of viewing sequences, so that you can apply algebraic manipulations to the functions

An application of Generating functions

Suppose we wanted to get a nice formula for x_n, with $x_0 = 0$, $x_1 = 1$ satisfying

$$x_{n+2} = 2x_{n+1} + x_n$$

Suppose $F(x)$ is a generating function for x_n

$$\begin{aligned} F(x) &= x + (2f_1 + f_0)x^2 + (2f_2 + f_1)x^3 + \cdots \\ &= x + 2x(f_1 x + f_2 x^2 + \cdots) + x^2(f_0 + f_1 x + \cdots) \\ &= x + 2xF(x) + x^2 F(x) \end{aligned}$$

Application of Generating functions (Continued)

$$F(x) = \frac{x}{1 - 2x - x^2} = \frac{A}{x - r_1} + \frac{B}{x - r_2}$$

$r_1 = 1 + \sqrt{2}$, $r_2 = 1 - \sqrt{2}$,
Geometric sum formula gets us that x_n is a sum of a power of r_1 and r_2.

our approach

We construct a bivariate generating function, $F(x, y)$, where the coefficient of $x^i y^j$ denotes the number of trees on i vertices with total height j

our approach

We construct a bivariate generating function, $F(x, y)$, where the coefficient of $x^i y^j$ denotes the number of trees on i vertices with total height j
If we let

$$P(x) = 1 + \sum_{i \in S} x^i,$$

We can show that F satisfies

$$F(x, y) = xP(F(xy, y))$$

Lets parse out the right hand side

1. the x comes from picking the root
2. F(xy,y) denotes counting each vertex in the tree as having one more depth
3. The function P indicates that we have the indicated number of subtrees (independent choice means multiplication)

computing the moments

We want to know, as the number of vertices in our tree gets larger, what happens to the distribution of total heights?

Take some number of partial derivatives in y, then plug in $y = 1$.

1. no derivatives – number of elements
2. one derivative – expectation(once you normalize)
3. two derivatives – second factorial moment (once you normalize)
4. ...

results

For the simple example of $S = \{2\}$, get dominant terms of $2\sqrt{\pi n^3}$ for expectation, $(40/3 - 4\pi)\, n^3$ for variance, $\sqrt{n^9 \pi (16\pi^2 - 50)}$ for skewness.

Approach can apply to any other S, all the steps are automated.

Definitions

Dyck Path

An (a,b,n)-Dyck Path is a sequence $\{w_1, w_2, \ldots w_{an+bn}\}$ of steps (1,0) and (0,1) so that for every i, the point $(x, y) = \sum_{j=1}^{i} w_j$ satisfies $ax >= by$.

In other words, it is paths from $(0, 0)$ to (bn, an) that stay on or below the line $y = a/bx$.

Results for 2D lattice walks

Theorem[Duchon'00]

The number of paths of length n that stay on or below the line $y = (a/b)x$ are $\Theta(\frac{1}{n}\binom{an+bn}{an})$.

Results for 2D lattice walks

Theorem[Duchon'00]

The number of paths of length n that stay on or below the line $y = (a/b)x$ are $\Theta(\frac{1}{n}\binom{an+bn}{an})$.

really:

Theorem[Duchon'00]

If d_n is the number of paths of length n that stay on or below the line $y = (a/b)x$,

$$\frac{1}{an+bn}\binom{an+bn}{bn} \leq d_n \leq \frac{1}{an+1}\binom{an+bn}{bn}$$

Proving an Upper Bound

Proposition

the number of (a,b,n) Dyck Paths is $\leq \frac{1}{an}\binom{an+bn}{bn}$

Definition

A conjugation of a word $w_1 u w_2$ with respect to letter u is the word $w_2 u w_1$

- u conjugating forms an equivalence class of size an+1
- only one word in each equivalence class can correpond to an (a,b,n)-Dyck Path
- at most $\frac{1}{an+1}\binom{an+bn}{an}$ (a,b,n)-Dick Paths

Exact Results for 2D lattice walks

Theorem

When a=1, the exact solution given by Fuss-Catalan number: $\frac{1}{bn+1}\binom{(b+1)n}{n}$.

Exact Results for 2D lattice walks

Theorem

When a=1, the exact solution given by Fuss-Catalan number: $\frac{1}{bn+1}\binom{(b+1)n}{n}$.

$\frac{2}{7n-1}\binom{7n-1}{2n}$ result for a only slightly different question.

A Nasty Formula

Theorem[Bizley'54]

The number of (a,b,n) Dyck Paths is exactly

$$\sum \frac{F_1^{k_1} F_2^{k_2} \cdots}{k_1! k_2! \cdots}$$

where the summation is taken over all $\{k_i\} \subseteq (\mathbb{Z}^+)^*$ such that $\sum ik_i = n$ and

$$F_j = \frac{1}{j(a+b)} \binom{ja+jb}{ja}$$

Getting Estimates

- Can quickly get a lot of data using a dynamic programming technique.

Getting Estimates

- Can quickly get a lot of data using a dynamic programming technique.
- Once we have a lot of data, we do a statistical fit to the model we want and get a good estimate for the coefficient.

Some Numerics!

a\b	1	2	3	4	5	6
1	1	0.500	0.333	0.250	0.200	0.167
2	0.500	1	0.241	0.500	0.161	0.333
3	0.333	0.241	1	0.160	0.137	0.500
4	0.250	0.500	0.160	1	0.12	0.241
5	0.200	0.161	0.137	0.120	1	0.096
6	0.166	0.333	0.500	0.241	0.096	1

3D Lattice walks

Definition

An (a,b,c) lattice walk is a sequence of steps in either (1,0,0),(0,1,0), or (0,0,1) so that all of the partial sums satisfy

$$ax \leq by \leq cz$$

for the case that one of $\{a, c\}$ are zero, we have a 2D lattice walk.

3D Lattice walks

If $a = b = c = 1$. It is the classical three dimensional Catalan numbers which have formula

$$\frac{2}{(n+1)(n+2)^2}\binom{3n}{n,n,n}$$

coefficients

- We saw that $(1,1,1)$ lattice paths grew like $\frac{1}{n^3}\binom{3n}{n,n,n}$.
- This exponenital part is the same for other choices of (a,b,c). However, the rational part grows with different exponents on n.

3D Numerics

Table: a=1

b\c	1	2	3	4	5	6	7
1	3.0	2.7	2.6	2.5	2.5	2.4	2.4
2	3.7	3.3	3.0	2.9	2.8	2.7	2.6
3	4.3	3.7	3.4	3.2	3.1	3.0	2.9
4	4.8	4.1	3.8	3.5	3.4	3.2	3.1
5	5.2	4.5	4.1	3.8	3.6	3.4	3.3
6	5.7	4.8	4.4	4.1	3.8	3.6	3.5
7	6.0	5.2	4.6	4.3	4.0	3.8	3.7

Counting Time Above the Line

Theorem[Andersen]

Of all paths of length n, the number of paths with k edges above the line $y = x$ is given exactly by:

$$\binom{2n-2k}{n-k}\binom{2k}{k}$$

for even k.

- As n goes to infinity, this looks like arcsine
- Originally in a continuous context, but the lattice path version is also usually attributed to him

Counting Time Above the Line

Theorem[Chung,Feller]

Of all paths ending at (n,n), the number of paths that have an even number k edges above $y = x$ is independent of k.

Counting Time Above the Line

Theorem[Chung,Feller]

Of all paths ending at (n,n), the number of paths that have an even number k edges above $y = x$ is independent of k.

- Distribution is definitly not uniform for slanted line

Hypergeometric summations: Gentle introduction

It's a basic fact due to the binomial theorem that, for example

$$\sum_k \binom{n}{k} = 2^n$$

Hypergeometric summations: Gentle introduction

It's a basic fact due to the binomial theorem that, for example

$$\sum_k \binom{n}{k} = 2^n$$

rewrite as

$$\sum_k \binom{n}{k} 1^k 1^{n-k} = (1+1)^n$$

It's a basic fact due to the binomial theorem that, for example

$$\sum_k \binom{n}{k} = 2^n$$

rewrite as

$$\sum_k \binom{n}{k} 1^k 1^{n-k} = (1+1)^n$$

$$\sum_k \binom{n}{k} 2^k = 3^n$$

and many more

Definition of Hypergeometric

We want to generalize. What is it about exponentials and binomial coefficients that are so nice?

Definition of Hypergeometric

We want to generalize. What is it about exponentials and binomial coefficients that are so nice?
$H(n, k)$ is called hypergeometric if it satisfies:

$$\frac{H(n+1, k)}{H(n, k)} \in \mathbb{Q}(n, k)$$

and

$$\frac{H(n, k+1)}{H(n, k)} \in \mathbb{Q}(n, k)$$

past some initial value, For example, you may give $H(0, 0) = 1$.

Sister Celine's Technique

Given $H(n, k)$ hypergeometric, try to find a good description of

$$y_n = \sum_k H(n, k)$$

Sister Celine's Technique

Given $H(n, k)$ hypergeometric, try to find a good description of

$$y_n = \sum_k H(n, k)$$

A good description means a hypergeometric description.
That is, what rational function is equal to:

$$\frac{y_{n+1}}{y_n}$$

Moving from Hypergeometric to Linear Recurrent

Not all seqences are hypergeometric.
By settling for a slightly less friendly description, we can describe many more sums
A sequence y_n is linear recurrent of order d if there are q_i each rational functions of n so that

$$y_{n+d} = q_0 y_n + q_1 y_{n+1} + \cdots + q_{d-1} y_{n+d-1}$$

If we have such a description, we call y_n Q-finite or holonomic.

Linear Recurrent Sums

Using very similar techniques to Sister celine's method, can take any linear recurrent x_k and hypergeometric $H(n, k)$, and find a recurrence for

$$y_n = \sum_k x_k^d H(n, k)$$

application to king walks

A chess king walking in a d dimensional chessboard.
Can step from $v \in \mathbb{Z}^d$ to $u \in \mathbb{Z}^d$ so long as $||v - u||_\infty = 1$.

Analyzing the problem

Generating functions again!
Let f be a generating function (on variables $z_1, \ldots, z_n$ so that the coefficient of

$$\prod_{j=1}^{d} z_j^{n_j}$$

represents number of paths from 0 to $(n_1, n_2, \ldots n_j)$

building the generating function, a simple case

What if had one dimension?

building the generating function, a simple case

What if had one dimension?

Then, at each step, we can pick either left or right, so, multiplying by either z_1^{-1} or z_1. so, we would have $f_n(z_1) = (z_1^{-1} + z_1)^n$.

building the generating function, a simple case

What if had one dimension?

Then, at each step, we can pick either left or right, so, multiplying by either z_1^{-1} or z_1. so, we would have $f_n(z_1) = (z_1^{-1} + z_1)^n$.

That was kind of boring. We'd just get binomial coefficients, as we'd expect.

building the generating function

How about higher dimensions?

building the generating function

How about higher dimensions?

If we allow standing still, All dimensions are independent! We just undo that afterwards, getting a single step is:

$$\left(\prod_{j=1}^{d}\left(z_i^{-1}+1+z_i\right)\right)-1$$

building the generating function

How about higher dimensions?

If we allow standing still, All dimensions are independent! We just undo that afterwards, getting a single step is:

$$\left(\prod_{j=1}^{d}\left(z_i^{-1}+1+z_i\right)\right)-1$$

each step is independent, so,

$$f_n(z_1,\ldots,z_d)=\left(\left(\prod_{j=1}^{d}\left(z_i^{-1}+1+z_i\right)\right)-1\right)^n$$

Rephrasing the question

The binomial theorem lets us write:

$$\begin{aligned} f_n(z_1, \dots, z_d) &= \sum_{k=0}^{n} \left(\prod_{j=1}^{d} \left(z_i^{-1} + 1 + z_i \right) \right)^k (-1)^{n-k} \binom{n}{k} \\ &= \sum_{k=0}^{n} \left(\prod_{j=1}^{d} \left(z_i^{-1} + 1 + z_i \right)^k \right) (-1)^{n-k} \binom{n}{k} \end{aligned}$$

Rephrasing the question

The binomial theorem lets us write:

$$f_n(z_1, \ldots, z_d) = \sum_{k=0}^{n} \left(\prod_{j=1}^{d} \left(z_i^{-1} + 1 + z_i\right) \right)^k (-1)^{n-k} \binom{n}{k}$$
$$= \sum_{k=0}^{n} \left(\prod_{j=1}^{d} \left(z_i^{-1} + 1 + z_i\right)^k \right) (-1)^{n-k} \binom{n}{k}$$

Take constant term to get back to where we started.

$$\sum_{k=0}^{n} \text{Constant Term}(\left(z^{-1} + 1 + z\right)^k)^d (-1)^{n-k} \binom{n}{k}$$

results

Recurrences too long to list.

For the two dimensional king, the number of paths of length n is:

$$c_2 \frac{8^n}{n} \left(1 - \frac{4}{9n} + \frac{1}{18n^2} + O\left(\frac{1}{n^3}\right)\right),$$

For three dimensions the number is:

$$c_3 \frac{26^n}{n^{\frac{3}{2}}} \left(1 - \frac{11}{18n} + \frac{683}{5832n^2} + O\left(\frac{1}{n^3}\right)\right).$$

and for four dimensions the number is:

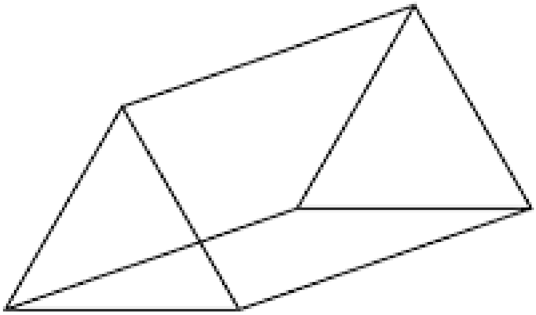
$$c_4 \frac{80^n}{n^2} \left(1 - \frac{25}{9n} + \frac{36439}{6561n^2} + O\left(\frac{1}{n^3}\right)\right).$$

Other applications

1. computing binomial transforms of sequences
2. evaluating multiple summations

Bunkbed Graphs

Starting with any finite graph G, take an isomorphic copy of it, G'. Put an edge between each vertex in G and the vertex that it corresponds to in G'
For example, the triangle, C_3 would become



Percolation

For each edge, remove it independently with some probability.

vertices stay the same.

Our proof works for slightly more general random model.

Bunkbed Conjecture

Fix any two vertices s and f. The probability that s is connected to f is at least the probability s is connected to f'. Intuitive, but hard to prove. Originally asked in 1985 by Kasteleyn, in it's current version, it was stated by Häggström in 1998.

Partial results so far

The conjecture has been shown in the following cases:

1. If only one vertical edges exists using FKG inequality.
2. If G is outerplanar
3. If G is complete and $p \geq \frac{1}{2}$

Partial results so far

The conjecture has been shown in the following cases:

1. If only one vertical edges exists using FKG inequality.
2. If G is outerplanar
3. If G is complete and $p \geq \frac{1}{2}$

We can show it for the case of exactly two vertical edges.

Thank You