On Complete N-ary Subtrees on Branching Family Trees

George P. Yanev

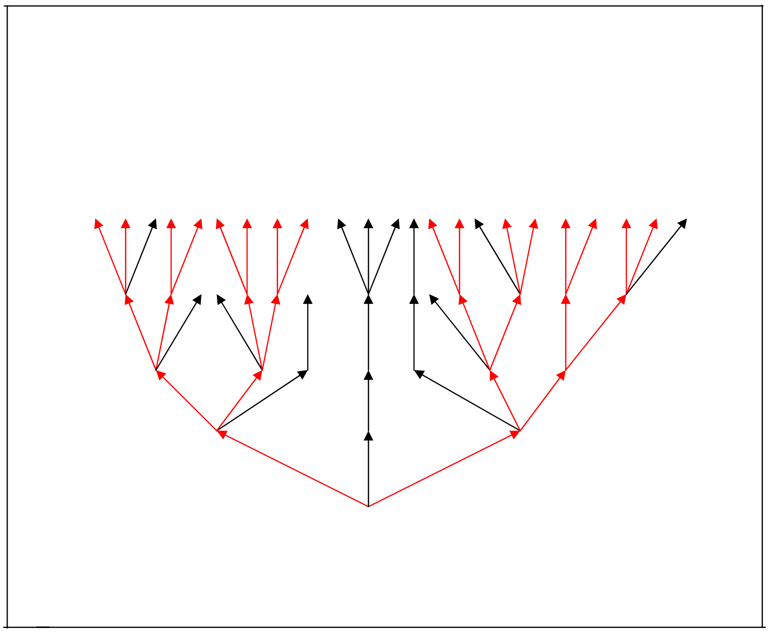
Department of Mathematics,
University of Texas - Pan American,
Edinburg, USA

"Facts are the enemy of truth."
- from the play "Man of La Mancha"
Miguel de Cervantes

Workshop on Branching Processes and Their Applications,
April 20-23, 2009, Badajoz, Spain

Outline

1. Galton-Watson processes that grow faster than binary splitting.
2. Critical phenomenon.
3. Number of subtrees with infinite height.
4. Limit theorems for the maximum height of a subtree.
5. Geometric offspring.
6. Poisson offspring.
7. One unexpected result.
8. Links to other research.
9. References.



Define

$T_2 - 1 =$ maximum height of a binary subtree rooted at the ancestor.

Note that

- $T_2 = 0$ if $Z_1 < 2$.
- $T_1 - 1 =$ maximum height of a unary subtree rooted at the ancestor, i.e., T_1 is the extinction time of $\{Z_n\}$.

Consider

- $\gamma_2 = P(T_2 < \infty)$ - there is no infinite binary subtree, i.e., the growth is slower than binary splitting;
- $\gamma_1 = P(T_1 < \infty)$ - there is no infinite unary subtree, i.e., extinction probability.

Theorem 1 [Dekking (1991)] The probability γ_2 is the smallest root in $[0, 1]$ of

$$x = f(x) + (1 - x)f'(x).$$

Main recurrent argument in the proof. If $Z_1 = k$, then the family tree does not contain a binary subtree of height $n + 1$ iff

- $k = 0$ or $k = 1$, or
- all of the k subtrees rooted at Z_1 do not have a subtree of height n; or
- all but one of the k subtrees rooted at Z_1 do not have a subtree of height n;

Therefore,

$$\gamma_2(n+1) = p_0 + p_1 + \sum_{k=2}^{\infty} [\gamma_2^k(n) + k\gamma_2^{k-1}(n)(1 - \gamma_2(n))]p_k$$

$$\gamma_2(n + 1) = f(\gamma_2(n)) + (1 - \gamma_2(n))f'(\gamma_2(n)).$$

Theorem 2 [Dekking (1991)] Probability γ_2 cannot be a continuous function of the offspring moments, nor of any other parameter that depends continuously on p_k, $k = 0, 1, \ldots$

There is a critical $m_2^c > 1$ such that

$$\begin{cases} \gamma_2 = 1, & m < m_2^c; \\ \gamma_2 < 1, & m \geq m_2^c. \end{cases}$$

This is qualitatively different to the behavior of the extinction probability, when $m_1^c = 1$, and

$$\begin{cases} \gamma_1 = 1, & m \leq 1; \\ \gamma_2 < 1, & m > 1. \end{cases}$$

This difference occurs because

$$G_2(s) = f(s) + (1 - s)f'(s),$$

is increasing but not convex; $G_2'(1) = 0$.

Number of binary subtrees with infinite height

Let V_2 be the number of distinct complete binary subtrees with infinite height and rooted at the ancestor.

Theorem 3 [Y. and Mutafchiev (2006)] For $j = 0, 1, 2, \ldots$

$$P(V_2 = j) = \sum_{k=2j}^{2j+1} \frac{(1-\gamma_2)^k}{k!} f^{(k)}(\gamma_2).$$

Consider the Taylor expansion of $f(1)$ about the point γ_2. Then $P(V_2 = j)$ is the $(j+1)$st segment of length 2 in this expansion, i.e.,

$$P(V_2 = 0) = f(\gamma_2) + (1-\gamma_2) f'(\gamma_2)$$

$$P(V_2 = 1) = \frac{(1-\gamma_2)^2}{2!} f''(\gamma_2) + \frac{(1-\gamma_2)^3}{3!} f'''(\gamma_2)$$

$$\ldots$$

$$P(V_2 = j) = \frac{(1-\gamma_2)^{2j}}{(2j)!} f^{(2j)}(\gamma_2) + \frac{(1-\gamma_2)^{2j+1}}{(2j+1)!} f^{(2j+1)}(\gamma_2)$$

Denote for $N \geq 1$

- V_N = number of distinct complete N-ary subtrees with infinite height rooted at the ancestor.
- $\gamma_N = P(V_N = 0) = P(T_N < \infty)$ - there is no infinite N-ary subtree rooted at the ancestor.

Theorem 4 [Pakes and Dekking (1991)] γ_N is the smallest root in $[0, 1]$ of

$$x = \sum_{i=0}^{N-1} \frac{(1-x)^i}{i!} f^{(i)}(x).$$

Theorem 5 [Y. and Mutafchiev (2006)] For $j = 0, 1, 2, \ldots$ and $N \geq 1$

$$P(V_N = j) = \sum_{k=jN}^{jN+N+1} \frac{(1-\gamma_2)^k}{k!} f^{(k)}(\gamma_2).$$

$P(V_N = j)$ is the $(j+1)$st segment of length N

in the Taylor expansion of $f(1)$ about γ_N.

Denote for $N \geq 2$

$$G_N(s) = \sum_{j=0}^{N-1} \frac{(1-s)^j}{j!} f^{(j)}(s),$$

$$a_N = G'_N(\gamma_n) \qquad \text{and} \qquad 2b_N = G''_N(\gamma_N).$$

Theorem 6 [L. Mutafchiev (2008)]

Assume $\gamma_N \in (0,1)$ and $N \geq 2$. Then $a_N \leq 1$.

(i) If $a_N < 1$, then as $n \to \infty$

$$P(T_N > n \mid T_N < \infty) = c_N a_N^n + O(a_N^{2n}),$$

where $c_N > 0$ is certain constant.

(ii) If $a_N = 1$ and $b_N < \infty$, then as $n \to \infty$

$$P(T_N > n \mid T_N < \infty) \sim \frac{1}{\gamma_N b_N n}.$$

Theorem 6 extends the results for

$P(Z_n > 0) = P(T_1 > n)$.

Theorem 7 [Pakes and Dekking (1991)]
Suppose $\gamma_N = 1$ and $f^{(N)}(1-) < \infty$. Then

$$P(T_N > n) \sim \exp(-k_N N^n) \qquad (n \to \infty),$$

where $k_N > 0$ is certain constant.

The critical mean for geometric offspring is

$$m_N^c = (N-1)\left(1 - \frac{1}{N}\right)^{-N}.$$

Corollary 1 Consider geometric offspring.
(i) If $m < m_N^c$, then as $n \to \infty$

$$P(T_N > n \mid T_N < \infty) = c_N a_N^n + O(a_N^{2n}),$$

where $c_N > 0$ is certain constant.

(ii) If $m = m_N^c$, then as $n \to \infty$

$$P(T_N > n \mid T_N < \infty) \sim \frac{2N}{m_N^c - N + 1}\,\frac{1}{n}.$$

(iii) If $m > m_N^c$, then as $n \to \infty$

$$P(T_N > n) \sim \exp(-k_N N^n),$$

where $k_N > 0$ is certain constant.

Next results for T_N are obtained by letting the initial population size i become large.

Theorem 8 [M. Mota and Y.]

(i) If $\gamma_N < 1$ and $a_N < 1$, then

$$\lim_{i\to\infty} P_i\left(a_N^{-T_N} \leq ix \mid T_N < \infty\right) = \exp\left(-\frac{c_N}{x}\right),$$

where $c_N > 0$ is certain constant.

(ii) If $\gamma_N < 1$, $a_N = 1$, and $f^{(N+1)}(1-) < \infty$, then

$$\lim_{i\to\infty} P_i\left(T_N \leq ix \mid T_N < \infty\right) = \exp\left(-\frac{1}{\gamma_N b_N x}\right)$$

(iii) If $\gamma_N = 1$ and $f^{(N)}(1-) < \infty$, then

$$\lim_{i\to\infty} P_i\left(\exp\left(N^{T_N}\right) \leq ix \mid T_N < \infty\right) = \exp\left(-\frac{k_N}{x}\right),$$

where $k_N > 0$ is certain constant.

Let $p_k = (1-p)p^k, \quad k \geq 0.$

The number of N-ary subtrees V_N is geometric

$$P(V_N = j) = \gamma_N (1-\gamma_N)^j \quad (j \geq 1),$$

where γ_N is the smallest solution in $[0,1]$ of

$$(1 - x + 1/m)^N = (1-x)^{N-1}.$$

The critical value for m is

$$m_N^c = (N-1)\left(1 - \frac{1}{N}\right)^{-1} \quad \text{and}$$

the probability of not having a N-ary subtree

$$\gamma_N^c = 1 - \left(1 - \frac{1}{N}\right)^N \to 1 - e^{-1} \quad (N \to \infty)$$

The mean number of N-ary subtrees is

$$EV_N^c = \frac{1-\gamma_N^c}{\gamma_N^c} \to \frac{1}{e-1} \quad (N \to \infty).$$

N	2	3	4	6	10	20	100
m_N^c	4	6.75	9.481	14.93	25.812	53.001	270.468
γ_N^c	0.75	0.704	0.684	0.665	0.651	0.641	0.634

Table 1: Values of m_N^c and γ_N^c for geometric offspring.

Poisson offspring: $f(s) = e^{m(s-1)}, \quad m > 0.$

The distribution of V_N is $(j = 0, 1, \ldots)$

$$P(V_N = j) = P(jN \leq Y_N \leq jN + N - 1),$$

where Y_N is Poisson with mean $m(1 - \gamma_N)$

and γ_N is the smallest solution in $[0, 1]$ of

$$xe^{m(1-x)} = \sum_{i=0}^{N-1} \frac{m^i}{i!}(1-x)^i$$

The critical offspring mean values are:

$m_2^c = 3.35$, $m_3^c = 5.15$, $m_4^c = 6.80$, $m_5^c = 8.37$.

$V_N =$	0	1	2	3	4	5	6	7	8	9	≥10	$E(V_N)$
$N = 2$	0	0	0.01	0.04	0.11	0.19	0.22	0.19	0.13	0.07	0.04	6.25
$N = 3$	0	0.01	0.09	0.25	0.32	0.22	0.08	0.02	0	0	0.01	4.00
$N = 4$	0	0.05	0.30	0.41	0.19	0.04	0	0	0	0	0.01	2.87
$N = 5$	0	0.17	0.51	0.28	0.04	0	0	0	0	0	0	2.19

Table 2: Probability distribution of V_N assuming Poisson offspring with mean $m = 13$.

Theorem 9 If $\gamma_N < 1$, then

$$\lim_{n\to\infty} \frac{E(Z_n \mid T_N > n)}{E(Z_n \mid T_1 > n)} = \alpha_N \, \frac{1-\gamma_1}{1-\gamma_N}$$

where

$$\alpha_N = \frac{m - \sum_{j=0}^{N-1} \frac{1}{j!}(1-\gamma_N)^j f^{(j+1)}(\gamma_N)}{m - \frac{1}{(N-1)!}(1-\gamma_N)^{N-1} f^{(N)}(\gamma_N)},$$

Corollary 2 Consider geometric offspring.

Let the offspring mean equals the critical value, i.e., $m = m_N^c$. Then for any $N = 2, 3, \ldots$

$$\lim_{n\to\infty} \frac{E(Z_n^c \mid T_N^c > n)}{E(Z_n^c \mid T_1^c > n)} = 2$$

- J. Chayes, L. Chayes, and R. Durrett (1988), studying Mendelbrot's percolation process find a condition for $\gamma_8 < 1$, when $f(s) = (1 - p + ps)^9$.

- In his study of reinforced random walks, Pemantle (1988) introduces a concept of N-infinite branching process. This notion implies the existence of a N-ary subtree.

- There is a relationship between the N-ary subtrees and the existence of a k-core in a random graph, a concept introduced by Bollabàs (1984) and studied in relation to Galton-Watson trees by Pittel et al. (1996) and Riordan (2008).

F.M. Dekking, Branching processes that grow faster than binary splitting, *Amer. Math. Monthly* 98, 728-731, 1991.

A.G. Pakes and F.M. Dekking, On family trees and subtrees of simple branching processes, *J. Theoret. Probab.* 4, 353-369, 1991.

G.P. Yanev and L. Mutafchiev, Number of complete N-ary subtrees on Galton-Watson family tree, *Methodol. Comput. Appl. Probab.* 8, 223-233, 2006.

L. Mutafchiev, Survival probabilities for N-ary subtrees on a Galton-Watson family tree, *Statist. & Probab. Letters* 78, 2165-2170, 2008.