Shapes of Binary Trees

Steven Finch

June 24, 2004

This is a sequel to our treatment of various attributes of trees [1], expressed in the language of probability. Let $\{Y_t : 0 \le t \le 1\}$ be standard Brownian excursion. Define the L_p-norm

$$
\|Y\|_p = \begin{cases} \left(\int\limits_0^1 |Y_t|^p \, dt \right)^{1/p} & \text{if } 0 < p < \infty, \\ \max\limits_{0 \le t \le 1} |Y_t| & \text{if } p = \infty \end{cases}
$$

and a (new) seminorm

$$
\langle Y \rangle_p = \begin{cases} \left(\int\limits_0^1 \int\limits_0^v \left| Y_u + Y_v - 2 \min\limits_{u \le t \le v} Y_t \right|^p du \, dv \right)^{1/p} & \text{if } 0 < p < \infty, \\ \max\limits_{0 \le u < v \le 1} \left| Y_u + Y_v - 2 \min\limits_{u \le t \le v} Y_t \right| & \text{if } p = \infty. \end{cases}
$$

We examined $\|Y\|_p$ earlier [2]; $\langle Y \rangle_p$ is a less familiar random variable but nevertheless important in the study of trees. Note that $\langle Y \rangle_p$ is not a norm since, for any constant c, $\langle c \rangle_p = 0$ even if $c \ne 0$.

Let T be an ordered (strongly) binary tree with $N = 2n+1$ vertices. The **distance** between two vertices of T is the number of edges in the shortest path connecting them. The **height** of a vertex is the number of edges in the shortest path connecting the vertex and the root.

The **Wiener index** $d_1(T)$ is the sum of all $\binom{N}{2}$ distances between pairs of distinct vertices of T, and the **diameter** $d_\infty(T)$ is the maximum such distance. If $\delta(v, w)$ denotes the distance between vertices v and w, then

$$
d_\lambda(T) = \left(\frac{1}{2} \sum_{v,w} \delta(v, w)^\lambda \right)^{1/\lambda}, \qquad \lambda > 0,
$$

includes both the Wiener index and diameter as special cases.

The **internal path length** $h_1(T)$ of a tree is the sum of all N heights of vertices of T, and the **height** $h_\infty(T)$ is the maximum such height. Let o denote the root of T. The generalization

$$h_\lambda(T) = \left(\sum_v \delta(v,o)^\lambda \right)^{1/\lambda}, \qquad \lambda > 0,$$

includes both the internal path length and height as special cases. If we restrict attention to only those $n+1$ vertices $\hat{v}_k$ that are leaves (terminal nodes) of T, listed from left to right, then a sequence $\delta(\hat{v}_1, o)$, $\delta(\hat{v}_2, o)$, ..., $\delta(\hat{v}_{n+1}, o)$ emerges. This is called the **contour** of T.

The **width** $w_\infty(T)$ of a tree is the maximum of $\zeta_l(T)$ over all $l \geq 0$, where $\zeta_l(T)$ is the number of vertices of height l in T. Note that

$$w_\lambda(T) = \left(\sum_{l=0}^{h_\infty(T)} \zeta_l(T)^\lambda \right)^{1/\lambda}, \qquad \lambda > 0,$$

includes the trivial case $w_1(T) = N$. The sequence $\zeta_0(T)$, $\zeta_1(T)$, ..., $\zeta_{h_\infty}(T)$ is known as the **profile** of T.

0.1. Uniform Combinatorial Model. In this model, the $\binom{2n}{n}/(n+1)$ ordered binary trees are weighted with equal probability, where $N = 2n + 1$ is fixed.

Janson [3] determined the joint distribution of internal path length and Wiener index:

$$\left(\frac{h_1(T)}{2N^{3/2}}, \frac{d_1(T)}{2N^{5/2}} \right) \to (\|Y\|_1, \langle Y \rangle_1)$$

as $N \to \infty$. The marginal distribution of $h_1(T)$ was obtained earlier by Takács [4, 5, 6]; the result for $d_1(T)$ is apparently new. No explicit formula for $\mathrm{P}(\langle Y \rangle_1 \leq x)$ is known; see [2] for the corresponding result for $\|Y\|_1$. We have expected values

$$\mathrm{E}(\|Y\|_1) = \frac{1}{2}\sqrt{\frac{\pi}{2}}, \qquad \mathrm{E}(\langle Y \rangle_1) = \frac{1}{4}\sqrt{\frac{\pi}{2}}$$

and correlation coefficient

$$\frac{\mathrm{Cov}(\|Y\|_1, \langle Y \rangle_1)}{\sqrt{\mathrm{Var}(\|Y\|_1)}\sqrt{\mathrm{Var}(\langle Y \rangle_1)}} = \sqrt{\frac{48 - 15\pi}{50 - 15\pi}} = 0.5519206030...$$

As an aside, we mention that $\|Y\|_1 - \langle Y \rangle_1 \geq 0$ always. Underlying the joint moment [3]

$$\mathrm{E}\left(\|Y\|_1^k (\|Y\|_1 - \langle Y \rangle_1)^l \right) = \frac{k! l! \sqrt{\pi}}{2^{(7k+9l-4)/2}\Gamma((3k+5l-1)/2)} a_{k,l}$$

is the following interesting quadratic recursion [7, 8, 9, 10, 11, 12]:

$$a_{k,l} = 2(3k + 5l - 4)a_{k-1,l} + 2(3k + 5l - 6)(3k + 5l - 4)a_{k,l-1} + \sum_{0 < i+j < k+l} a_{i,j}a_{k-i,l-j}$$

with $a_{0,0} = -1/2$, $a_{1,0} = 1 = a_{0,1}$ and $a_{k,l} = 0$ when $k < 0$ or $l < 0$. All $a_{k,l}$ but $a_{0,0}$ are positive integers when $k \geq 0$ and $l \geq 0$. Applications include the enumeration of connected graphs with n vertices and $n + m$ edges. We have asymptotics [3, 13]

$$a_{k,0} \sim \frac{1}{2\pi}6^k(k-1)!, \qquad a_{0,l} \sim C \cdot 50^l\left((l-1)!\right)^2,$$

where the precise identity of the constant

$$C = \frac{1}{50}(0.981038...) = 0.01962... = \frac{1}{50.9664...}$$

remains an unsolved problem.

Chassaing, Marckert & Yor [14] determined the joint distribution of height and width:

$$\left(\frac{h_\infty(T)}{N^{1/2}}, \frac{w_\infty(T)}{N^{1/2}}\right) \rightarrow \left(\int_0^1 \frac{dt}{Y_t}, \|Y\|_\infty\right)$$

as $N \rightarrow \infty$. The marginal distribution of height was obtained earlier by Rényi & Szekeres and Stepanov [15, 16, 17, 18, 19, 20, 21, 22, 23, 24]; earlier works on width include [25, 26, 27, 28, 29, 30]. It turns out that the marginal distributions are identical (up to a factor of 2) and that this is the first of several theta distributions [31] we will see here:

$$P\left(\frac{1}{2}\int_0^1 \frac{dt}{Y_t} \leq x\right) = P\left(\|Y\|_\infty \leq x\right) = \frac{\sqrt{2}\pi^{5/2}}{x^3}\sum_{k=1}^\infty k^2 e^{-\pi^2 k^2/(2x^2)}.$$

The expected values thus coincide:

$$E\left(\frac{1}{2}\int_0^1 \frac{dt}{Y_t}\right) = E\left(\|Y\|_\infty\right) = \sqrt{\frac{\pi}{2}}.$$

Rényi & Szekeres also computed the location of the maximum of the probability density [15]:

$$\mathrm{mode}\left(\|Y\|_\infty\right) = \frac{1}{2}(2.3151543618...) = \frac{1}{2}\sqrt{\frac{2}{0.3731385248...}}.$$

Returning to the joint distribution formula, it is clear that $h_\infty(T)$ and $w_\infty(T)$ are negatively correlated. A numerical estimate (let alone an exact expression) for the correlation coefficient evidently remains open [14, 32].

For the generalized height and diameter parameters, we have marginal distributions [3, 14, 33, 34, 35]:

$$\frac{h_\lambda(T)}{2N^{(\lambda+2)/(2\lambda)}} \to \|Y\|_\lambda, \qquad \frac{d_\lambda(T)}{2N^{(\lambda+4)/(2\lambda)}} \to \langle Y \rangle_\lambda$$

as $N \to \infty$. The latter includes the special cases of Wiener index ($\lambda = 1$, as mentioned before) and diameter ($\lambda = \infty$):

$$\mathrm{P}\left(\langle Y \rangle_\infty \le x\right) = \frac{1024\sqrt{2}\pi^{5/2}}{3x^9} \sum_{k=1}^\infty k^2 \left[\left(3 + \pi^2 k^2\right) x^4 - 36\pi^2 k^2 x^2 + 64\pi^4 k^4\right] e^{-8\pi^2 k^2/x^2},$$

which possesses expected value

$$\mathrm{E}\left(\langle Y \rangle_\infty\right) = \frac{4}{3}\sqrt{2\pi}$$

and maximum location [33]

$$\mathrm{mode}\left(\langle Y \rangle_\infty\right) = 3.2015131492... = \sqrt{\frac{8}{0.7805116813...}}.$$

Nothing is known for other values of λ (even $\lambda = 2$ seems to have been neglected). It would also be good to learn the value of the correlation coefficient of $d_\infty(T)$ and $h_\infty(T)$, or of $d_\infty(T)$ and $w_\infty(T)$.

Consider finally the minimum height $\eta(T)$ of a leaf, that is,

$$\eta(T) = \min_{1 \le k \le n+1} \delta(\hat{v}_k, o),$$

and the height $\delta(\hat{v}_{\lceil n/2 \rceil}, o)$ of the central leaf. It is known that

$$\mathrm{E}(\eta) \to \sum_{k=1}^\infty 2^{k+1-2^k} = 1.5629882961...$$

as $N \to \infty$ [36, 37]. It is also known that [38, 39, 40, 41]

$$\sqrt{n}\,\mathrm{P}\left(\frac{\delta(\hat{v}_{\lceil n/2 \rceil}, o)}{\sqrt{n}} \le x\right) \to \frac{1}{2\sqrt{\pi}}\int_0^x t^2 e^{-t^2/4}dt = \mathrm{P}\left(\sqrt{X_1^2 + X_2^2 + X_3^2} \le \frac{x}{\sqrt{2}}\right),$$

the Maxwell distribution from thermodynamics, where X_1, X_2, X_3 are independent standard normal variables. (This can also be written in terms of the chi square distribution with 3 degrees of freedom.) Can these results be related to Brownian excursion in some way? We will report more on the properties of leaves of T later.

0.2. Critical Galton-Watson Model. In this model, the size $N = 2n + 1$ is free to vary: All ordered binary trees are included but with weighting 2^{-N}. (We omit subcritical and supercritical cases for reasons of space.)

Let T be a random tree. The probability that T has precisely N vertices is clearly [42]

$$\frac{1}{n+1}\binom{2n}{n}2^{-N} \sim \sqrt{\frac{2}{\pi}}N^{-3/2};$$

hence the expected number of vertices of T is infinite. We examine this result in another way. If

$$\nu_l = \sum_{k=0}^{l} \zeta_k$$

where ζ_k is the number of vertices of height k in T, then $\mathrm{E}(\nu_l) = l + 1$ and $\mathrm{Var}(\nu_l) = (2l + 1)(l + 1)l/6$, both which $\to \infty$ as $l \to \infty$. More complicated conditional distributions are due to Pakes [43, 44]:

$$\lim_{l\to\infty} \mathrm{P}\left(\frac{\nu_l}{l^2} \le x \,|\, \zeta_l > 0\right) = \int_0^x f(t)\, dt,$$

$$\lim_{l\to\infty} \mathrm{P}\left(\frac{\nu_l}{l^2} \le x \,|\, \zeta_m > 0 \text{ for all positive integers } m\right) = \int_0^x g(t)\, dt,$$

where the first density function is given by

$$f(t) = \frac{2}{\sqrt{2\pi}t^{3/2}} \sum_{k=0}^{\infty} \left(\frac{(2k+1)^2}{t} - 1\right) \exp\left(-\frac{(2k+1)^2}{2t}\right)$$

with mean $1/3$, variance $2/45$, and Laplace transform

$$\int_0^\infty e^{-st} f(t)\, dt = \sqrt{2s}\, \mathrm{csch}\left(\sqrt{2s}\right).$$

The second density function is not explicitly known, but has mean $1/2$, variance $1/12$ and satisfies

$$\int_0^\infty e^{-st} g(t)\, dt = \mathrm{sech}^2\left(\sqrt{\frac{s}{2}}\right).$$

Consequently $g(t)$ is the convolution of $\tilde{g}(t)$ with itself, where

$$\tilde{g}(t) = \frac{1}{\sqrt{2\pi}t^{3/2}} \sum_{k=0}^{\infty} (-1)^k (2k+1) \exp\left(-\frac{(2k+1)^2}{8t}\right),$$

but this appears to be as far as we can go.

Define T_l to be the subtree of T consisting of all ν_l vertices up to and including height l. We have the parameters $d_\lambda(T_l)$, $h_\lambda(T_l)$ and $w_\lambda(T_l)$ available for study, but little seems to be known. Of course, $w_1(T_l) = \nu_l$. Athreya [45], building on [46, 47, 48], proved that $\mathrm{E}\left(w_\infty(T_l)\right) \sim \ln(l)$ as $l \to \infty$, which contrasts nicely with the fact that $\mathrm{P}\left(\zeta_k = 0\right) \to 1$ as $k \to \infty$. See also [49, 50, 51, 52, 53, 54, 55]. Kesten, Ney & Spitzer [56, 57, 58] demonstrated that $\mathrm{P}\left(h_\infty(T_l) = j\right) \sim 2/j^2$ as $j \to \infty$; further references include [59, 60, 61]. Can exact distributional results be found? What about other values of λ? Is anything known about diameter for Galton-Watson trees?

Just as the limit behavior for the uniform model is related to Brownian excursion, the limit behavior for the critical GW model is related to what is known as the two-sided three-dimensional Bessel process $\{B_t : -\infty < t < \infty\}$. That is, $\{B_t : t \geq 0\}$ and $\{B_{-t} : t \geq 0\}$ are independent copies of standard 3D radial Brownian motion $\sqrt{W_{1,t}^2 + W_{2,t}^2 + W_{3,t}^2}$, each starting from zero [35, 62]. It would be good to learn more about the concrete distributional results arising from this correspondence.

0.3. Leaves of Maximum Height. Our closing remarks are concerned not with binary trees, but instead with labeled rooted trees. Choose such a tree T with N vertices uniformly out of the N^{N-2} possibilities (we agree that the root is labeled 1). Out of all possible parameters (suitably generalized), we mention only the minimum height $\eta(T)$ of a leaf. Meir & Moon [37] computed that

$$\mathrm{E}(\eta) \to 9 \sum_{k=1}^{\infty} \frac{1}{4^k(1 + 2 \cdot 4^{-k})^2} = 1.6229713847...$$

as $N \to \infty$. A more difficult problem involves counting the leaves $\hat{v}_k$ at prescribed distance from the root. Kesten & Pittel [63] proved, for leaves of maximum height, that there exists a probability distribution q_l such that

$$\lim_{N \to \infty} \mathrm{P}\left(\zeta_{h_\infty}(T) = l\right) = q_l, \qquad l \geq 1.$$

Further, q_l is the unique nonnegative solution of the system of equations

$$l! e^l q_l = \sum_{k=1}^{\infty} k^l q_k, \qquad \sum_{k=1}^{\infty} q_k = 1$$

and thus $q_1 = 0.602...$, $q_2 = 0.248...$, $q_3 = 0.094...$, $q_4 = 0.035...$ with mean 1.636... and standard deviation 0.995.... No exact expressions for these quantities are known. What is the corresponding distribution for the uniform ordered binary tree case?

0.4.   Addendum.   Janson [64] computed that

$$
\mathrm{E}\left(\int_0^1 (1/Y_t)\, dt \cdot \|Y\|_\infty\right) = 1 + \sum_{m=1}^{\infty} \frac{\ln\left[m(m+1)\right]}{m(m+1)}
$$
$$
= 1 + 2.0462774528...
$$
$$
= \pi - 0.0953152007...,
$$

$$
\frac{\mathrm{Cov}\left(\int_0^1 (1/Y_t)\, dt, \|Y\|_\infty\right)}{\sqrt{\mathrm{Var}\left(\int_0^1 (1/Y_t)\, dt\right)}\sqrt{\mathrm{Var}\left(\|Y\|_\infty\right)}} = \frac{3(3.0462774528... - \pi)}{\pi(\pi - 3)}
$$
$$
= -0.6428251027...
$$

and the infinite series [65, 66] is a Lüroth analog of Lévy's constant $\pi^2/(6\ln(2))$. Why is the joint distribution of height and width of trees related to the ergodic theory of numbers? (Coincidences as such do not happen without a reason.)   Another recent result is

$$
C = \frac{\sqrt{15}}{20\pi^2} = 0.0196207628...
$$

in connection with the asymptotics of the sequence $a_{0,l}$, due to Kotesovec [12].

REFERENCES

[1] S. R. Finch, Otter's tree enumeration constants, *Mathematical Constants*, Cambridge Univ. Press, 2003, pp. 295–316.

[2] S. R. Finch, Variants of Brownian motion, unpublished note (2004).

[3] S. Janson, The Wiener index of simply generated random trees, *Random Structures Algorithms* 22 (2003) 337–358; http://www2.math.uu.se/~svante/papers/; MR1980963 (2004b:05186).

[4] L. Takács, On the total heights of random rooted trees, *J. Appl. Probab.* 29 (1992) 543–556; MR1174430 (93g:05127).

[5] L. Takács, The asymptotic distribution of the total heights of random rooted trees, *Acta Sci. Math. (Szeged)* 57 (1993) 613–625; MR1243312 (94i:05086).

[6] L. Takács, On the total heights of random rooted binary trees, *J. Combin. Theory Ser. B* 61 (1994) 155–166; MR1280604 (95e:05035).

[7] E. M. Wright, The number of connected sparsely edged graphs, *J. Graph Theory* 1 (1977) 317–330; MR0463026 (57 #2990).

[8] V. A. Voblyǐ, Wright and Stepanov-Wright coefficients (in Russian), *Mat. Zametki* 42 (1987) 854–862, 911; Engl. transl. in *Math. Notes* 42 (1987) 969–974; MR0934817 (89d:05103).

[9] J. Spencer, Enumerating graphs and Brownian motion, *Comm. Pure Appl. Math.* 50 (1997) 291–294; MR1431811 (97k:05105).

[10] D. Aldous, Brownian excursions, critical random graphs and the multiplicative coalescent, *Annals of Probab.* 25 (1997) 812–854; http://www.stat.berkeley.edu/~aldous/Papers/bibliog.html; MR1434128 (98d:60019).

[11] P. Flajolet and G. Louchard, Analytic variations on the Airy distribution, *Algorithmica* 31 (2001) 361–377; http://algo.inria.fr/flajolet/Publications/publist.html; MR1855255 (2002j:68115).

[12] N. J. A. Sloane, On-Line Encyclopedia of Integer Sequences, A062980 and A094199.

[13] S. Janson and P. Chassaing, The center of mass of the ISE and the Wiener index of trees, *Elec. Comm. Probab.* 9 (2004) 178–187; arXiv:math.PR/0309284; MR2108865 (2005k:60303).

[14] P. Chassaing, J. F. Marckert and M. Yor, The height and width of simple trees, *Mathematics and Computer Science. Algorithms, Trees, Combinatorics and Probabilities*, Proc. 2000 Versailles colloq., ed. D. Gardy and A. Mokkadem, Birkhäuser, 2000, pp. 17–30; http://www.labri.fr/perso/marckert/; MR1798284 (2001j:68095).

[15] A. Rényi and G. Szekeres, On the height of trees, *J. Austral. Math. Soc.* 7 (1967) 497–507; MR0219440 (36 #2522).

[16] V. E. Stepanov, The distribution of the number of vertices in strata of a random tree (in Russian), *Teor. Verojatnost. i Primenen.* 14 (1969) 64–77; Engl. transl. in *Theory Probab. Appl.* 14 (1969) 65–78; MR0246344 (39 #7648).

[17] N. G. de Bruijn, D. E. Knuth and S. O. Rice, The average height of planted plane trees, *Graph Theory and Computing*, ed. R. C. Read, Academic Press, 1972, pp. 15–22; also in *Selected Papers on Analysis of Algorithms*, CSLI, 2000, pp. 215–223; MR0505710 (58 #21737).

[18] V. F. Kolchin, The extinction time of a branching process and the height of a random tree (in Russian), *Mat. Zametki* 24 (1978) 859–870, 894; Engl. transl. in *Math. Notes* 24 (1978) 954–961; MR0522419 (80i:60118).

[19] V. F. Kolchin, *Random Mappings*, Optimization Software Inc., 1986, pp. 135–146; MR0865130 (88a:60022).

[20] P. Flajolet and A. Odlyzko, The average height of binary trees and other simple trees, *J. Comput. System Sci.* 25 (1982) 171–213; MR0680517 (84a:68056).

[21] G. G. Brown and B. O. Shubert, On random binary trees, *Math. Oper. Res.* 9 (1984) 43–65; MR0736638 (86c:05099).

[22] R. Durrett, H. Kesten and E. Waymire, On weighted heights of random trees, *J. Theoret. Probab.* 4 (1991) 223–237; MR1088403 (91m:60021).

[23] P. Flajolet, Z.-C. Gao, A. Odlyzko and B. Richmond, The distribution of heights of binary trees and other simple trees, *Combin. Probab. Comput.* 2 (1993) 145–156; http://algo.inria.fr/flajolet/Publications/publist.html; MR1249127 (94k:05061).

[24] T. Luczak, A greedy algorithm estimating the height of random trees, *SIAM J. Discrete Math.* 11 (1998) 318–329; MR1617161 (99d:68117).

[25] A. M. Odlyzko and H. S. Wilf, Bandwidths and profiles of trees, *J. Combin. Theory Ser. B* 42 (1987) 348–370; MR0888687 (89b:05077).

[26] L. Takács, Limit distributions for queues and random rooted trees, *J. Appl. Math. Stochastic Anal.* 6 (1993) 189–216; MR1238599 (94m:60194).

[27] M. Drmota and B. Gittenberger, On the profile of random trees, *Random Structures Algorithms* 10 (1997) 421–451; MR1608230 (99c:05176).

[28] G. Kersting, On the height profile of a conditioned Galton-Watson tree, unpublished manuscript (1998), http://ismi.math.uni-frankfurt.de/kersting/research/papers.de.shtml.

[29] P. Chassaing and J.-F. Marckert, Parking functions, empirical processes, and the width of rooted labeled trees, *Elec. J. Combin.* 8 (2001) R14; MR1814521 (2002b:05037).

[30] M. Drmota and B. Gittenberger, The width of Galton-Watson trees conditioned by the size, *Discrete Math. Theor. Comput. Sci.* 6 (2004) 387–400; http://www.dmg.tuwien.ac.at/drmota/litalt.html; MR2081482 (2005f:60180).

[31] L. Devroye, Simulating theta random variates, *Statist. Probab. Lett.* 31 (1997) 275–279; MR1439946 (97j:65009).

[32] C. Donati-Martin, Some remarks about the identity in law for the Bessel bridge $\int_0^1 \frac{ds}{r(s)} \overset{(law)}{=} 2\sup_{s\le 1} r(s)$, *Studia Sci. Math. Hungar.* 37 (2001) 131–144; MR1834327 (2002j:60141).

[33] G. Szekeres, Distribution of labelled trees by diameter, *Combinatorial Mathematics X*, Proc. 1982 Adelaide conf., ed. L. R. A. Casse, Lect. Notes in Math. 1036, Springer-Verlag, 1983, pp. 392–397; MR0731595 (85b:05069).

[34] T. Luczak, The number of trees with large diameter, *J. Austral. Math. Soc. Ser. A* 58 (1995) 298–311; MR1329866 (96c:05086).

[35] D. Aldous, The continuum random tree. II. An overview, *Stochastic Analysis*, Proc. 1990 Durham conf., ed. M.T. Barlow and N.H. Bingham, Cambridge Univ. Press, 1991, pp. 23–70; MR1166406 (93f:60010).

[36] G. R. Grimmett, Random labelled trees and their branching networks, *J. Austral. Math. Soc. Ser. A* 30 (1980/81) 229–237; MR0607933 (82g:05042).

[37] A. Meir and J. W. Moon, Terminal path numbers for certain families of trees, *J. Austral. Math. Soc. Ser. A* 31 (1981) 429–438; MR0638271 (82m:05039).

[38] W. Gutjahr and G. C. Pflug, The asymptotic distribution of leaf heights in binary trees, *Graphs Combin.* 8 (1992) 243–251; MR1185403 (93i:05049).

[39] W. Gutjahr and G. C. Pflug, The limiting common distribution of two leaf heights in a random binary tree, *RAIRO Inform. Théor. Appl.* 26 (1992) 1–18; MR1155342 (93b:68032).

[40] W. Gutjahr and G. C. Pflug, The asymptotic contour process of a binary tree is a Brownian excursion, *Stochastic Process. Appl.* 41 (1992) 69–89; MR1162719 (93i:60017).

[41] M. Drmota, Distribution of the height of leaves of rooted trees (in Russian), *Diskret. Mat.* 6 (1994) 67–82; Engl. transl. in *Discrete Math. Appl.* 4 (1994) 45–58; MR1273234 (95c:05045).

[42] R. Otter, The multiplicative process, *Annals of Math. Statist.* 20 (1949) 206–224; MR0030716 (11,41a).

[43] A. G. Pakes, Some limit theorems for the total progeny of a branching process, *Adv. Appl. Probab.* 3 (1971) 176–192; MR0283892 (44 #1122).

[44] A. G. Pakes, The laws of some random series of independent summands, *Advances in the Theory and Practice of Statistics. A Volume in Honor of Samuel Kotz*, ed. N. L. Johnson and N. Balakrishnan, Wiley, 1997, pp. 499–516; MR1481190 (99g:60027).

[45] K. B. Athreya, On the maximum sequence in a critical branching process, *Annals of Probab.* 16 (1988) 502–507; MR0929060 (89f:60098).

[46] H. Weiner, Moments of the maximum in a critical branching process, *J. Appl. Probab.* 21 (1984) 920–923; MR0766828 (86a:60112).

[47] K. Kämmerle and H.-J. Schuh, The maximum in critical Galton-Watson and birth and death processes, *J. Appl. Probab.* 3 (1986) 601–613; MR0855368 (87m:60188).

[48] A. G. Pakes, Remarks on the maxima of a martingale sequence with application to the simple critical branching process, *J. Appl. Probab.* 24 (1987) 768–772; MR0900316 (88g:60113).

[49] J. Bishir, Maximum population size in a branching process, *Biometrics* 18 (1962) 394–403; MR0145637 (26 #3166).

[50] S. R. Adke, The maximum population size in the first N generations of a branching process, *Biometrics* 20 (1964) 649–651; MR0181013 (31 #5243).

[51] T. Lindvall, On the maximum of a branching process, *Scand. J. Statist.* 3 (1976) 209–214; MR0426188 (54 #14134).

[52] O. Nerman, On the maximal generation size of a non-critical Galton-Watson process, *Scand. J. Statist.* 4 (1977) 131–135; MR0474531 (57 #14169).

[53] A. Spataru, Equations concerning maxima of critical Galton-Watson processes, *Rev. Roumaine Math. Pures Appl.* 37 (1992) 935–937; MR1195094 (94f:60110).

[54] V. A. Vatutin and V. A. Topchii, The maximum of critical Galton-Watson processes, and left-continuous random walks (in Russian), *Teor. Veroyatnost. i Primenen.* 42 (1997) 21–34; Engl. transl. in *Theory Probab. Appl.* 42 (1997) 17–27; MR1453327 (98g:60154).

[55] E. M. Bondarenko and V. A. Topchii, Estimates for the expectation of the maximum of a critical Galton-Watson process on a finite interval (in Russian), *Sibirsk. Mat. Zh.* 42 (2001) 249–257; Engl. transl. in *Siberian Math. J.* 42 (2001) 209–216; MR1833155 (2002b:60145).

[56] H. Kesten, P. Ney and F. Spitzer, The Galton-Watson process with mean one and finite variance, *Teor. Verojatnost. i Primenen.* 11 (1966) 579–611; Engl. transl. in *Theory Probab. Appl.* 11 (1966) 513–540; MR0207052 (34 #6868).

[57] E. Seneta, The Galton-Watson process with mean one, *J. Appl. Probab.* 4 (1967) 489–495; MR0228075 (37 #3659).

[58] K. B. Athreya and P. Ney, *Branching Processes*, Springer-Verlag, 1972, pp. 19–23; MR0373040 (51 #9242).

[59] T. E. Harris, Branching processes, *Annals of Math. Statist.* 19 (1948) 474–494; MR0027465 (10,311b).

[60] L. Castoldi, Sulla distribuzione dei tempi di estinzione nelle discendenze biologiche, *Boll. Un. Mat. Ital.* 11 (1956) 158–167; MR0079380 (18,77d).

[61] A. T. Bharucha-Reid, *Elements of the Theory of Markov Processes and their Applications*, McGraw-Hill, 1960, pp. 26–27; MR0112177 (22 #3032).

[62] B. Gittenberger, Convergence of branching processes to the local time of a Bessel process, *Random Structures Algorithms* 13 (1998) 423–438; MR1662793 (99k:60213).

[63] H. Kesten and B. Pittel, A local limit theorem for the number of nodes, the height, and the number of final leaves in a critical branching process tree, *Random Structures Algorithms* 8 (1996) 243–299; MR1603252 (98m:60137).

[64] S. Janson, On the asymptotic joint distribution of height and width in random trees, *Studia Sci. Math. Hungar.* 45 (2008) 451–467; http://www2.math.uu.se/~svante/papers/; MR2641443 (2011d:60239).

[65] S. R. Finch, Khintchine-Lévy constants, *Mathematical Constants*, Cambridge Univ. Press, 2003, pp. 59–65.

[66] W. Bosma, K. Dajani and C. Kraaikamp, Entropy and counting correct digits, Dept. of Math. Report 9925 (1999), Univ. of Nijmegen; http://www.math.ru.nl/onderzoek/reports/reports99.html.