The Annals of Applied Probability
1994, Vol. 4, No. 3, 923–932

A STRONG LAW FOR THE HEIGHT OF RANDOM BINARY PYRAMIDS[1]

BY HOSAM M. MAHMOUD

George Washington University

By embedding in a suitable continuous-time process, we find a strong law for h_n, the height of a random binary pyramid of order n. We show that $h_n/\ln n$ converges almost surely to a constant limit and we determine that limit.

1. Introduction. Chain letters are a business practice of a pyramidal hierarchy in which a promoter sells an initial letter carrying a certain list containing a number of repetitions of his name. A purchaser crosses off the name at the top of the list and adds his own name at the bottom, then tries to sell copies of this modified letter. The process repeats recursively and at each stage all participants in the chain letter venture compete, under certain rules, to sell copies of their letters. Each time someone purchases a letter, he or she must pay a certain amount of money to the seller and an equal amount of money to the promoter. The promoter's claim is that the scheme is quite lucrative to all participants, a claim refuted in Gastwirth and Bhattacharya (1984).

In some pyramid schemes, a participant cannot sell more than a certain quota (m letters), unless the participant buys another copy of the letter. This model can be represented as a growing m-ary tree with a nonuniform probability model as follows. The promoter is represented by a *root* node labeled 1. The nodes of the tree at any stage have a bound m on the number of children they may have (the number of children a node has as its *outdegree*). At the $(n+1)$st stage, a new node labeled $n+1$ is adjoined to the tree; a node with outdegree m is *saturated*, but a node of outdegree $k < m$ has an equal chance, like any other unsaturated node, of attracting the new node. Thus the probability that an unsaturated node (participant) succeeds in attracting the new node (selling a copy of the letter to a new participant) is $1/U_n$, where U_n is the total number of unsaturated nodes after n participants join the venture.

It is helpful to speak of the *extended pyramid*. This is a pyramid or tree obtained from the original pyramid by adding to every unsaturated node one *external node*. These external nodes represent the insertion positions of the next participant in the hierarchy and they are the equally likely objects in the growth rule. The number of external nodes added is U_n. Figure 1

Received September 1993; revised December 1993.

[1] Research supported in part by NSA contract MDA904-92-H3086.

AMS 1991 *subject classifications*. Primary 60J80; secondary 60F15.

Key words and phrases. Random trees, stochastic processes, strong laws.

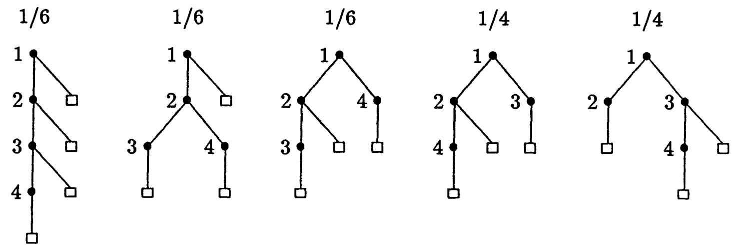

FIG. 1. *Extended binary pyramids of order* 4 *and their probabilities.*

illustrates all possible binary pyramids ($m = 2$) of order 4; the top row of numbers in the figure denotes the probabilities of these trees.

Although pyramids existed in the literature for over 25 years, very few of their properties are known, except for the case of unbounded outdegree. Trees corresponding to that latter case are known as recursive trees. There exists quite a substantial literature on recursive trees as surveyed in Smythe and Mahmoud (1994). By contrast, only two papers were written on pyramids with bounded node outdegree [Bhattacharya and Gastwirth (1983) and Gastwirth and Bhattacharya (1984)]. In these two papers, Markov chains were used as a model to study the proportion of nodes of outdegree i in a random pyramid of size n. Nodes of outdegree 0 represent the "shut-outs" or the frustrated participants who join the venture but are unable to sell any copies of their letters. In the language of trees, the shut-outs are the *leaves* of the tree.

One reason for the difficulty in analyzing parameters in pyramids with bounded node outdegrees is that the number of insertion positions, U_n, is a random variable, unlike its deterministic counterpart in many other well-studied classes of random trees, such as the binary search tree [see Mahmoud (1992), for example] and, noticeably, such as recursive trees, where simply $U_n = n$. This difficulty is finessed in this paper, as it suffices for our purpose to work with probabilistic limits of U_n.

In this paper we analyze the *height* of a random binary pyramid, or the length of the longest root-to-leaf path in the pyramid. This parameter may have the following significance in the legal aspect of the business venture. Suppose a frustrated participant who is not able to reach his quota decides to take legal action against the promoter. Lawyers involved may then want to trace the case back to the promoter, talking first with the parent node of the frustrated participant, and if no compromise is reached talking with the parent of that parent and so forth, until the root is reached. The height then represents the worst case amount of time and effort in preparing the case (and hence of legal fees) across the whole pyramid.

The method of analysis applies in principle to all m, with the proviso U_n is explicitly characterized. The probabilistic behavior of this quantity will be derived in the binary case from the results of Gastwirth and Bhattacharya (1984).

2. An associated birth-and-death process. Consider the following birth-and-death process. An ancestor to a species is born at time $t = 0$. The ancestor produces its first child after time τ_1, which is distributed like EXP(1) [the notation EXP(s) stands for a random variable following the exponential distribution with parameter s]. A second child is born to the ancestor after time $\tau_1 + \tau_2$, where τ_2, the interbirth time, is also distributed like EXP(1) and is independent of τ_1. Immediately after the birth of the second child, the ancestor dies. The lifetime of the ancestor $L = \tau_1 + \tau_2$ is thus a random variable. Therefore, $X(t)$, the number of children born to the ancestor at time t, is determined by randomly stopping a Poisson process with rate 1; that is, if $Z(t)$ represents a Poisson process with rate 1, then

$$X(t) = \begin{cases} Z(t), & \text{if } t \le L, \\ 2, & \text{if } t > L. \end{cases}$$

The instant a child is born, it mimics its ancestor, thus behaving according to a (time-shifted) Poisson process with rate 1 and with the same random stopping rule applied to its own lifetime. All members of the species are assumed to behave independently of their parents, of other members, and of their own birth time. The children of the ancestor are called the first generation, their children are the second generation and so on.

The process may be represented by a family tree that grows with time. Suppose the nth member of the species is born at time t_n (with probability 1, the numbers t_n are distinct). Define a sequence of nested family trees $\{T_n\}_{n=1}^{\infty}$, where T_n is the family tree within the time interval

$$(1) \qquad t_n \le t < t_{n+1}.$$

At time t, some members of the species may be dead, but each living member is in the process of producing a child. This may be thought of as if each living member has an embryo, which may be represented by a partly grown edge emanating out of the parent in the family tree. The remaining time until the embryo matures into a child is, of course, still distributed like EXP(1) by the memoryless property of the exponential distribution. Assume that at time t belonging to the interval specified in (1), the number of living members (subsequently, the number of growing embryos) is r_n. According to the independence assumption and the memoryless property, any embryo is equally likely to mature into a child, and the next birth [of the $(n+1)$st member] will take place after a period of time distributed (conditional on r_n) like

$$a_{n+1} =_{\text{def}} \min\{e_1, \dots, e_{r_n}\} =_{\mathscr{L}} \text{EXP}(r_n),$$

where each e_i is an independent EXP(1) random variable, for $1 \le i \le r_n$. Thus the sequence of family trees is recursively built according to a rule which, at any stage, chooses with equal probability one of the candidate embryos for the next birth (just like the equal likelihood of unsaturated nodes in a binary pyramid to succeed next in selling their letter). The boundary conditions are the same, too; that is, the start of this recursive process (at

$t = 0$) coincides with the basis of the inductive process of building a random binary pyramid. Hence the sequence of family trees has the same distribution as random binary pyramids, and r_n must have the same distribution as U_n, and we have the conditional relation

$$a_{n+1} | U_n =_{\mathscr{L}} \mathrm{EXP}(U_n). \tag{2}$$

Hence it suffices to study properties of T_n to develop any parameter concerning binary pyramids. [Extension to the m-ary case is obviously done by considering m children, instead of just two, with interbirth times that are EXP(1) and a stopping rule in which a parent dies immediately after giving birth to the mth child.] Our birth-and-death process is an instance of a general stochastic process known as the Crump–Mode process, named after Crump and Mode (1968).

Kingman (1975) developed an important result concerning B_k, the time of the first birth in the kth generation in a Crump–Mode process. Conditioned on the event of eternal survival of the species (a condition automatically met in our case), the time of the first birth in generation k satisfies the strong law:

$$\frac{B_k}{k} \to \gamma \quad \text{a.s.},$$

where

$$\gamma = \sup\{z \,|\, \mu(z) < 1\},$$
$$\mu(z) = \inf\{e^{z\theta}\phi(\theta) \,|\, \theta > 0\}$$

and $\phi(\theta)$ is a characteristic function of the process given by

$$\mathbf{E}\left[\sum_{j=1}^{X(\infty)} e^{-b_j\theta}\right],$$

where b_j is the birth time of the jth child of the root. However, with probability 1, at $t = \infty$, the number of children born to the ancestor is 2. Thus, in our case,

$$\phi(\theta) = \mathbf{E}[e^{-\tau_1\theta}] + \mathbf{E}[e^{-(\tau_1+\tau_2)\theta}] = \frac{2+\theta}{(1+\theta)^2}.$$

The function

$$\xi_z(\theta) =_{\text{def}} e^{z\theta}\phi(\theta) = \frac{e^{z\theta}(2+\theta)}{(1+\theta)^2}$$

achieves its minimum at a value $\theta_0 = \theta_0(z)$, which may be obtained by setting the derivative to 0. One sees that θ_0 is a root of the quadratic equation

$$z\theta^2 + (3z-1)\theta + (2z-3) = 0.$$

One root is always less than -2, and is of no consequence to our purpose. The other root is less than 0, if $z > 3/2$; thus, the infimum of $\xi_z(\theta)$ is reached at $\theta = 0$, as $\xi_z(\theta)$ is increasing for $\theta > 0$. Hence for $z > 3/2$, $\inf\{\xi_z(\theta)|\theta > 0\} = \xi_z(0) = 2$. On the other hand, if $0 < z \le 3/2$, the infimum of $\xi_z(\theta)$ is reached at

$$\theta_0 = \theta_0(z) = \tfrac{1}{2}\Big(-(3z-1) + \sqrt{(3z-1)^2 - 4z(2z-3)}\,\Big).$$

The value of the infimum is thus $\xi_z(\theta_0(z))$, an increasing function for $0 < z \le 3/2$. If z is negative, the function $\xi_z(\theta)$ decreases monotonically as $\theta > 0$ increases, and its infimum is obviously 0 (reached at ∞). The function $\mu(z)$ is plotted in Figure 2.

Thus γ is indeed well defined by the unique value of z that solves the equation $\mu(z) = 1$. This yields the value $\gamma = 0.405634\dots$.

3. A strong law for the height of a pyramid. We first establish a connection between the height of the tree and the time of birth of the nth member of the species.

LEMMA 1.

$$\frac{t_n}{h_n} \to \gamma = 0.405634\dots \quad a.s.$$

PROOF. When the growing family tree has n nodes at time t_n, its height is h_n (possibly reached by a node indexed with less than n). So, we have

$$B_{h_n} \le t_n < B_{h_n+1}.$$

However, $t_n = 0 + \mathrm{EXP}(U_1) + \mathrm{EXP}(U_2) + \cdots + \mathrm{EXP}(U_{n-1})$, which is stochastically larger than $0 + \mathrm{EXP}(1) + \mathrm{EXP}(2) + \cdots + \mathrm{EXP}(n-1)$, as $U_n \le n$. Thus, as $n \to \infty$, $t_n \to \infty$ almost surely and so must h_n. Dividing throughout by h_n and applying Kingman's theorem, we have the result. □

We wish to determine the asymptotic behavior of t_n. For this we first need to develop a lower bound for U_n.

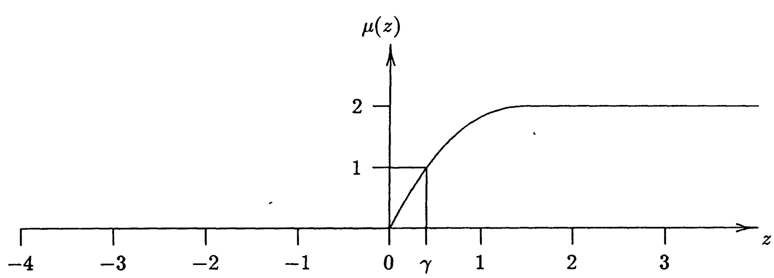

FIG. 2. *The function* $\mu(z)$.

LEMMA 2.

$$U_n \geq \frac{n+1}{2}.$$

PROOF. A pyramid is *complete* if its nodes on all levels (except possibly the last two levels) are saturated. A pyramid is *leftmost complete* if it is complete and all the nodes on its last level are as far left as possible. For example, in Figure 1 the last three pyramids (counting from the left) are complete, but only the third and fourth pyramids are leftmost complete.

We demonstrate next that the leftmost complete pyramid of order n has the minimum number $U_n^{\min}$ of unsaturated nodes among all the binary pyramids of order n. To see this, assume we have a pyramid T_n, of order n, with $U_n^{\min}$ unsaturated (or equivalently external) nodes. If the pyramid is leftmost complete, we are done. Suppose the pyramid T_n is complete, but not leftmost complete. Let v be the rightmost node on the last level and let u be the parent of the leftmost external node on the last level. There are four cases according to whether v is the only child of its parent or not and whether u is a leaf or not. One case cannot arise, and that is the case that v is the only child of its parent and u has one child, for in this case if we transfer v to become the second child of u, we obtain a pyramid with a fewer number of unsaturated nodes, contrary to our assumption that T_n has $U_n^{\min}$ unsaturated nodes. If v's parent is saturated and u is a leaf, transfer both v and its sibling to become children of u, resulting in a pyramid with the same minimal number of unsaturated nodes. In the remaining two cases, just transfer v to become a child of u, a transformation that does not change the number of external nodes.

In any of the three feasible cases, if the resulting pyramid is leftmost complete, we are done; if not, continue transferring nodes on the last level to the left as before, filling the gaps systematically until the process becomes no longer possible. This series of transformations produces a leftmost complete pyramid with the fewest possible number of unsaturated nodes.

If T_n is not complete to begin with, there must be an unsaturated node u on some level before the last two. Let v be the rightmost node on the last level. Again, there are four cases according to whether u is a leaf (or has only one child) and whether v's parent is saturated or not. One of these cases cannot arise in a pyramid with $U_n^{\min}$ unsaturated nodes, and that is the case that u has one child and v is the only child of its parent. For in this case, by transferring v to be the second child of u, the new pyramid has fewer than $U_n^{\min}$, a contradiction. A second case is when u is a leaf and v's parent is saturated. In this case transfer both v and its sibling to become children of u, producing a pyramid with the same number of saturated nodes as T_n. For the remaining two cases, just transferring v to become a child of u will produce a pyramid with the same number of unsaturated nodes as T_n. Thus in all the feasible cases we have a transformation of T_n into a new pyramid and the transformation leaves the number of external nodes unchanged.

As long as there are unsaturated nodes on levels before the last two, continue with the transformation, resulting in pyramids with the same number of external nodes. When it is no longer possible to proceed with the transformation, the final pyramid is complete and must have $U_n^{\min}$ unsaturated nodes. Once a complete binary pyramid with $U_n^{\min}$ unsaturated nodes is obtained, we can proceed as above to transform it into a leftmost complete binary pyramid with $U_n^{\min}$ unsaturated nodes. Observe also that the last pyramid in the chain of transformations has the minimum height $\lfloor \log_2 n \rfloor$ among all possible binary pyramids.

We next compute $U_n^{\min}$. In the leftmost complete pyramid of order n, let y_n be the number of leaves on the last level and let v_n be the number of unsaturated nodes on the level before last. Clearly,

$$y_n = n - (2^0 + 2^1 + \cdots + 2^{\lfloor \log_2 n \rfloor - 1}) = n + 1 - 2^{\lfloor \log_2 n \rfloor}. \tag{3}$$

The total number of saturated nodes on the level before last is $\lfloor y_n/2 \rfloor$, and $v_n = 2^{\lfloor \log_n \rfloor - 1} - \lfloor y_n/2 \rfloor$. Thus,

$$\begin{aligned} U_n^{\min} &= v_n + y_n \\ &= \frac{2^{\lfloor \log_n \rfloor}}{2} + \left\lceil \frac{y_n}{2} \right\rceil. \end{aligned}$$

It follows from (3) and the observation that $2^{\lfloor \log_2 n \rfloor - 1}$ is an integer, for all $n \geq 2$, that

$$U_n^{\min} = \left\lceil \frac{n+1}{2} \right\rceil. \qquad \square$$

LEMMA 3.

$$\frac{t_n}{\ln n} \to \frac{2}{\sqrt{5} - 1} \quad a.s.$$

PROOF. For the random variable t_n, we have recurrence

$$t_n = t_{n-1} + a_n,$$

and a_n, as discussed in Section 2, is distributed like the minimum of U_n independent exponential random variables with parameter 1. Unwinding the recurrence,

$$t_n = \sum_{j=1}^{n} a_j.$$

Thus,

$$\mathbf{E}[t_n] = \sum_{j=1}^{n} \mathbf{E}[a_j]. \tag{4}$$

As the interbirth times $\{a_j\}_{j=1}^{\infty}$ are independent, we have a similar expression for the variance:

$$\text{Var}[t_n] = \sum_{j=1}^{n} \text{Var}[a_j]. \tag{5}$$

By the conditional relation (2), we have

$$\mathbf{E}\left[a_{j+1} | U_j\right] = \frac{1}{U_j}, \qquad \text{Var}\left[a_{j+1} | U_j\right] = \frac{1}{U_j^2}. \tag{6}$$

Thus, the unconditional average is

$$\mathbf{E}[a_{j+1}] = \mathbf{E}\left[1/U_j\right]. \tag{7}$$

Gastwirth and Bhattacharya (1984) have shown that, as $n \to \infty$,

$$\mathbf{E}[U_n] \sim \frac{\sqrt{5} - 1}{2} n, \qquad \text{Var}[U_n] \sim \frac{6\sqrt{5} - 13}{11} n.$$

Thus by an application of Chebychev's inequality, we have

$$\frac{U_n}{n} \to_P \frac{\sqrt{5} - 1}{2}.$$

According to Lemma 2, $n/U_n \le 2$. In view of this uniform bound, the last convergence implies convergence of all the moments of n/U_n. Thus,

$$\mathbf{E}[1/U_n] \sim \frac{2}{(\sqrt{5} - 1)n}, \qquad \mathbf{E}\left[1/U_n^2\right] \sim \frac{4}{(6 - 2\sqrt{5})n^2}. \tag{8}$$

Plugging (6) into the conditional variance formula, we get

$$\begin{aligned} \text{Var}[a_{j+1}] &= \mathbf{E}\left[\text{Var}\left[a_{j+1} | U_j\right]\right] + \text{Var}\left[\mathbf{E}\left[a_{j+1} | U_j\right]\right] \\ &= \mathbf{E}\left[1/U_j^2\right] + \text{Var}\left[1/U_j\right]. \end{aligned}$$

We conclude from (8) that, as $n \to \infty$,

$$\text{Var}[a_n] \sim \frac{4}{(6 - 2\sqrt{5})n^2} + o\left(\frac{1}{n^2}\right). \tag{9}$$

Therefore, from (4), (7) and (8) we have

$$\mathbf{E}[t_n] \sim O(1) + \sum_{j=1}^{n} \frac{2}{(\sqrt{5} - 1)j} \sim \frac{2}{\sqrt{5} - 1} \ln n,$$

and similarly from (5) and (9) we obtain

$$\text{Var}[t_n] = O(1).$$

A standard application of Kolmogorov's theorem on the convergence of the random series $t_n - \mathbf{E}[t_n] = \sum_{j=1}^{n}(a_j - \mathbf{E}[a_j])$ completes the proof. □

Combining Lemmas 1 and 3, we arrive at the main result.

THEOREM 1. *The height of a random binary pyramid satisfies*

$$\frac{h_n}{\ln n} \to \frac{2}{(\sqrt{5}-1)\gamma} = 3.989120\dots \quad a.s.$$

4. Discussion. We have determined the probabilistic behavior of the height of a binary pyramid to be about $3.98912 \ln n$. The method used employs a birth-and-death process with a random stopping rule, an instance of the Crump–Mode process.

The method would also work for m-ary pyramids with higher branching ($m > 2$) if we have an explicit characterization of the probabilistic behavior of the total number of unsaturated nodes, which would be a generalization of the estimate of Gastwirth and Bhattacharya (1984) in the binary case. As m increases, one would expect the limiting height to come down because nodes with outdegree between 2 and $m - 1$ continue to attract newcomers, whereas their analogue in the binary case is saturated. In other words, newcomers attracted by nodes having outdegree between 2 and m would have been directed to deeper levels in the binary pyramid, increasing the chance of a larger height in the binary case. When m increases without bound, the pyramid scheme is equivalent to the uniform recursive tree, whose height has been determined by a stochastic process similar to ours [Pittel (1993)], the one difference being that no nodes are ever saturated; that is, each node continues to produce children indefinitely, or no member of the species ever dies. The process there is thus a pure birth process with no stopping. It is shown in Pittel (1993) that h_n^{rec}, the height of a recursive tree of size n, satisfies the strong law

$$\frac{h_n^{\mathrm{rec}}}{\ln n} \to e = 2.71828\dots \quad \text{a.s.}$$

This latter result is implied in a subtle way in Devroye (1987), who proves a similar probabilistic law for the height of UNION-FIND trees; Pittel (1994) makes the connection explicit by a construction that links the height of recursive trees to the height of UNION-FIND trees.

Recursive trees provide a natural lower bound on the height of pyramids with any bound on the outdegree. One expects the following tree to provide an upper bound. The *binary tree* is a tree with n nodes and each node has no children, one left child, one right child or two distinct children (one left and one right). Binary trees are extended by adding the appropriate number of external nodes to make all node outdegrees uniformly equal to 2. While a uniform probability model on binary trees of order n produces the large asymptotic expected height $2\sqrt{\pi n}$ [Flajolet and Odlyzko, (1982)], a nonuniform model corresponding to a natural growth rule produces heights of logarithmic order. The uniform distribution is suitable for use when the binary trees are used as syntax trees as in the theory of compiling [Kemp (1984)]. On the other hand, when binary trees are used as data structures for efficient data retrieval, the appropriate model of randomness is the random permutation model [see Mahmoud (1992)]. This latter probability model corresponds to choosing any external node with equal probability as the position for the next insertion, and binary trees grown under this model are called *binary search trees*. These trees and their growth rule are similar to binary pyramids, but an important difference is that leaves in binary search trees have two external nodes. Thus the affinity of a leaf is twice as much as a node of outdegree 1 in binary search trees. In particular, the collection of leaves at the deepest level in the tree will have a higher probability of attracting the next node than their counterpart in the binary pyramid. That is, the few leaves at the deepest level in a binary search tree attract nodes that would have been attracted at higher levels in the pyramid, and one expects the height h_n^{bst} of a binary search tree of order n to be a bit larger than the height of a pyramid of the same order. Indeed

$$\frac{h_n^{\mathrm{bst}}}{\ln n} \to 4.31107\ldots \quad \text{a.s.},$$

as proved in Pittel (1985, 1994) and Devroye (1986, 1987). Thus, the limiting normalized height (normalized by $\ln n$) of pyramids of different branching is sandwiched between the limiting normalized height of recursive trees and the limiting normalized height of binary search trees.

REFERENCES

BHATTACHARYA, P. K. and GASTWIRTH, J. L. (1983). A non-homogeneous Markov model of a chain-letter scheme. In *Recent Advances in Statistics: Papers in Honor of Herman Chernoff* (M. H. Rizvi, J. S. Rustagi and D. Siegmund, eds.). Academic, New York.

CRUMP, K. and MODE, C. (1968). A general age-dependent branching process I. *J. Math. Anal. Appl.* **24** 494–508.

DEVROYE, L. (1986). A note on the height of binary search trees. *J. Assoc. Comput. Mach.* **33** 489–498.

DEVROYE, L. (1987). Branching processes in the analysis of the height of trees. *Acta Inf.* **24** 277–298.

FLAJOLET, P. and ODLYZKO, A. (1982). The average height of binary trees and other simple trees. *J. Comput. Syst. Sci.* **25** 171–213.

GASTWIRTH, J. L. and BHATTACHARYA, P. K. (1984). Two probability models of pyramid or chain letter schemes demonstrating that their promotional claims are unreliable. *Oper. Res.* **32** 527–536.

KEMP, R. (1984). *Fundamentals of the Average-Case Analysis of Particular Algorithms*. Wiley, New York.

KINGMAN, J. (1975). The first birth problem for an age-dependent branching process. *Ann. Probab.* **3** 790–801.

MAHMOUD, H. M. (1992). *Evolution of Random Search Trees*. Wiley, New York.

PITTEL, B. (1985). On growing random binary trees. *J. Math. Anal. Appl.* **103** 461–480.

PITTEL, B. (1994). Note on the heights of random recursive trees and random m-ary search trees. *Random Structures and Algorithms* **5** 337–347.

SMYTHE, R. T. and MAHMOUD, H. M. (1994). A survey of recursive trees. Unpublished report.

DEPARTMENT OF STATISTICS
STATISTICAL COMPUTING
GEORGE WASHINGTON UNIVERSITY
WASHINGTON, DC 20052