On a Tail Bound for Root-Finding in Randomly Growing Trees

Sam Justice* and N. D. Shyamalkumar*

May 18, 2019

Abstract

We re-examine a lower-tail upper bound for the random variable

$$X = \prod_{i=1}^{\infty} \min\left\{ \sum_{k=1}^{i} E_k, 1 \right\},$$

where $E_1, E_2, \ldots \overset{iid}{\sim} \text{Exp}(1)$. This bound has found use in root-finding and seed-finding algorithms for randomly growing trees, and was initially proved as a lemma in the context of the uniform attachment tree model. We first show that X has a useful representation as a compound product of uniform random variables that allows us to determine its moments and refine the existing nonasymptotic bound. Next we demonstrate that the lower-tail probability for X can equivalently be written as a probability involving two independent Poisson random variables, an equivalence that yields a novel general result regarding indpendent Poissons and that also enables us to obtain tight asymptotic bounds on the tail probability of interest.

1 Introduction

Concentration inequalities, which provide bounds on the tail probabilities of random variables, represent a vital and highly popular subject of mathematical research. These inequalities find applications in a variety of fields, including geometry, statistics, empirical process theory, machine learning, and randomized algorithms, to name a few. In this brief note, we reconsider an inequality that has been used in a particular area of the theory of randomly growing trees. A randomly growing tree consists of a sequence of inductively defined trees $\{T_n\}_{n \in \mathbb{N}}$, where T_n is formed from T_{n-1} by introducing a new vertex and attaching it to an existing vertex of T_{n-1} according to some probabilistic rule. The inequality was first proved and utilized in Bubeck et al. (2017) in the context of root-finding algorithms for the uniform attachment tree model, and it was subsequently used again in Devroye and Reddad (2019) to prove a result regarding seed-finding in this same model. A variation on the inequality

---

*Department of Statistics and Actuarial Science, University of Iowa; contact email: samuel-justice@uiowa.edu

was also proved in Khim and Loh (2017) to aid in the analysis of root-finding for diffusions over regular trees. Note that the term root-finding algorithm refers to any method that identifies (in the limit) the first vertex ("root") of a randomly growing tree with some specified probability. Similarly, a seed-finding algorithm seeks to partially or fully recover the first several vertices ("seed") of a randomly growing tree with high probability. The tail bound appears as Lemma 2 in Bubeck et al. (2017) and is stated as follows:

Lemma 2 from Bubeck et al. (2017): Let $E_1, E_2, \ldots \overset{iid}{\sim} \text{Exp}(1)$ and let

$$X = \prod_{i=1}^{\infty} \min \left\{ \sum_{k=1}^{i} E_k, 1 \right\}. \tag{1}$$

Then for any $t > 0$, we have that

$$\mathbb{P}(X \leq t) \leq 6t^{1/4}. \tag{2}$$

We note that (2) is intended to be used for small t, as it gives a nontrivial bound only for $t < (1/6)^4 \approx 0.000772$. Lemma 12 of Khim and Loh (2017) proved that the same tail bound holds for

$$X = \prod_{i=1}^{\infty} \min \left\{ -\sum_{k=1}^{i} \log(B_k), 1 \right\}, \tag{3}$$

where $B_1, B_2, \ldots \overset{iid}{\sim} \text{Beta}(\beta, 1)$ for $\beta \in (0, 1]$. Since $B_1 \overset{\text{d}}{=} U^{1/\beta}$ for $U \sim \text{U}(0, 1)$, we have using the uniform-exponential relationship that $-\log(B_1) \sim \text{Exp}(\beta)$. Thus the case $\beta = 1$ corresponds to the definition of X in (1).

The random variable X arises in the context of a particular method for finding the root/seed of a randomly growing tree. Specifically, letting $V(T)$ denote the vertex set of a tree T and $(T, v)_{u\downarrow}$ denote the subtree starting at u in the rooted tree (T, v), consider the function $\varphi_T : V(T) \to \mathbb{N}$ defined by

$$\varphi_T(v) = \prod_{u \in V(T) \setminus \{v\}} |(T, v)_{u\downarrow}|. \tag{4}$$

Thus $\varphi_T(v)$, which is referred to in Khim and Loh (2017) as the subtree product estimator, is the product of all the sizes of the subtrees of the rooted tree (T, v). It can be thought of as a measure of vertex centrality in T, with smaller values of φ_T corresponding to more central vertices. Given the product form of φ_T, it is perhaps unsurprising that X comes into play in proofs involving the estimator. Interestingly, the vertex minimizing the function φ_T turns out to be the maximum likelihood estimate for the root of a diffusion over a regular tree, and φ_T can also be seen as a relaxation of the likelihood functions for the root and seed in the uniform attachment model. It has been proven for these models (see Bubeck et al. (2017), Khim and Loh (2017), and Devroye and Reddad (2019)) that picking the K vertices with the smallest values of φ_T produces a successful root/seed-finding algorithm for K large enough. The variable X appears in an intermediate step of these proofs that involves upper-bounding the probability that a vertex $v \in V(T)$ becomes more central than the root vertex r in the limit of a sequence of randomly growing trees $\{T_n\}_{n \in \mathbb{N}}$:

$$\limsup_{n \to \infty} \mathbb{P}(\varphi_{T_n}(v) \leq \varphi_{T_n}(r)).$$

In this paper, we consider a general setting that encompasses both of the preceding definitions of X. In particular, we define X as in (1), but now we let $E_1, E_2, \ldots \overset{iid}{\sim} \mathrm{Exp}(\lambda)$ for an arbitrary $\lambda > 0$. We carry out a careful examination of X en route to obtaining both nonasymptotic and asymptotic bounds for $\mathbb{P}(X \leq t)$. To begin, we demonstrate that X has a useful representation as a compound (Poisson) product of iid standard uniform random variables. This representation allows us to easily calculate the moments of X, and it also leads to a quick refinement of the $t^{1/4}$ rate shown in (2). After this, we prove that $\mathbb{P}(X \leq t)$ can be equivalently expressed as a probability involving two independent Poisson random variables. This equivalence immediately yields a general result regarding independent Poissons, and we also use it to obtain tight asymptotic bounds on $\mathbb{P}(X \leq t)$.

2   A Useful Representation and Some Consequences

We begin by proving that X can be written as a Poisson product of iid standard uniform random variables. Our proof uses some fundamental results from the theory of Poisson processes. In particular, we recall that for a Poisson process $\{N(t) : t \in [0, \infty)\}$ with rate λ, the number of arrivals in any time interval of length t has a $\mathrm{Pois}(\lambda t)$ distribution. Also, conditional on the number of arrivals by time t, the arrival times of $N(t)$ have the same joint distribution as the order statistics from a $\mathrm{U}(0, t)$ distribution. See, for example, Ross (1996) for these and other basic details on Poisson processes.

Lemma 1. *For $\lambda > 0$, let $E_1, E_2, \ldots \overset{iid}{\sim} Exp(\lambda)$ and let*

$$X = \prod_{i=1}^{\infty} \min \left\{ \sum_{k=1}^{i} E_k, 1 \right\}. \tag{5}$$

Then we have that

$$X \overset{\mathrm{d}}{=} \prod_{i=1}^{N} U_i, \tag{6}$$

where $N \sim Pois(\lambda)$, $\{U_i\}_{i \in \mathbb{N}}$ is an iid sequence of $U(0, 1)$ random variables independent of N, and the product is understood to equal 1 if $N = 0$.

Proof. Consider a Poisson process with rate λ and arrival times given by the sequence

$$\left\{ \sum_{k=1}^{i} E_k \right\}_{i \in \mathbb{N}}.$$

Moreover, let N denote the number of arrivals by time 1 for this Poisson process, so that $N \sim \mathrm{Pois}(\lambda)$. We clearly have that

$$X = \begin{cases} 1 & \text{if } N = 0, \\ \prod_{i=1}^{N} \sum_{k=1}^{i} E_k & \text{if } N > 0. \end{cases}$$

3

Given that $N = n$, we also see that

$$\left(E_1, E_1 + E_2, \ldots, \sum_{k=1}^{n} E_k \right) \stackrel{\mathrm{d}}{=} \left(U_{(1)}, \ldots, U_{(n)} \right),$$

where $U_{(1)}, \ldots, U_{(n)}$ are the order statistics corresponding to a sample of size n from a U$(0, 1)$ distribution. It follows that

$$X \stackrel{\mathrm{d}}{=} X^* = \begin{cases} 1 & \text{if } N = 0, \\ \prod_{i=1}^{N} U_{(i)} = \prod_{i=1}^{N} U_i & \text{if } N > 0, \end{cases}$$

where $\{U_i\}_{i \in \mathbb{N}}$ is an iid sequence of U$(0, 1)$ random variables independent of N. $\qquad \square$

This compound product representation easily yields the moments of X.

Corollary 1. *For any $\beta > -1$, we have that $\mathbb{E}\left[X^\beta \right] = e^{-\frac{\beta\lambda}{1+\beta}}$. In particular, this implies that $\mathbb{E}\left[X \right] = e^{-\lambda/2}$ and $\mathrm{Var}(X) = e^{-2\lambda/3} - e^{-\lambda}$.*

The next result is a simple consequence of Corollary 1. It shows that we can obtain a rate of t^α for any $\alpha \in (0, 1)$ for the lower tail of X, and it furthermore gives the optimal rate corresponding to such α. Note that this is a refinement of the original result shown in (2), which gives the rate $t^{1/4}$.

Theorem 1. *For any $\alpha \in (0, 1)$, we have for all $t > 0$ that*

$$\mathbb{P}(X \leq t) \leq e^{\frac{\alpha\lambda}{1-\alpha}} t^\alpha. \tag{7}$$

In particular, by optimizing the bound with respect to α, it follows that

$$\mathbb{P}(X \leq t) \leq \exp\left\{ -\left(\sqrt{-\log t} - \sqrt{\lambda} \right)_+^2 \right\}, \tag{8}$$

where $x_+ = \max\{x, 0\}$.

Proof. Using Corollary 1, we have for $\alpha \in (0, 1)$ and $t > 0$ that

$$\begin{aligned} \mathbb{P}(X \leq t) &= \mathbb{P}(X^{-\alpha} \geq t^{-\alpha}) \\ &\leq \mathbb{E}\left[X^{-\alpha} \right] t^\alpha \qquad \text{(Markov's Inequality)} \\ &= e^{\frac{\alpha\lambda}{1-\alpha}} t^\alpha. \end{aligned}$$

Optimizing the preceding expression with respect to α, we see that it is minimized at $(1 - \sqrt{\lambda/(-\log t)})_+$, which yields (8). $\qquad \square$

Remark 1. We note that Theorem 1 provides a better constant than that given in (2). Also, we observe that since $1/X$ does not have a finite moment generating function in a neighborhood of zero, we resort to finding the best moment bound. In this connection it is worth pointing out that the best moment bound is in general always tighter than the Chernoff bound, see Nelson and Philips (1995). Another interesting aside is that $-\log X$ is a sub-exponential random variable.

4

3 Tight Asymptotic Bounds

The following result further exploits the characterization of X from Lemma 1 as a compound product of uniforms to show that $\mathbb{P}(X \leq t)$ has a novel representation as a probability involving two independent Poisson random variables. While the result is intriguing in its own right, we will also see later that it provides the key to obtaining tight asymptotic bounds for the lower tail of X that in particular demonstrate that the bound of Theorem 1 is at least asymptotically tight.

Theorem 2. *For any $t \in (0,1)$, we have that*

$$\mathbb{P}(X \leq t) = \mathbb{P}(N > N^*), \tag{9}$$

where $N \sim Pois(\lambda)$ and $N^* \sim Pois(-\log t)$ are independent.

Proof. Using Lemma 1, we have for $t \in (0,1)$ that

$$\mathbb{P}(X \leq t) = \mathbb{P}\left(\prod_{i=1}^{N} U_i \leq t, \ N > 0\right)$$

$$= \mathbb{P}\left(-\sum_{i=1}^{N} \log U_i \geq -\log t, \ N > 0\right).$$

Now since $\{-\log U_i\}_{i\in\mathbb{N}}$ is an iid sequence of $\text{Exp}(1)$ random variables, we can associate with it a rate 1 Poisson process that is independent of N. Letting N^* denote the number of arrivals by time $-\log t$ for this Poisson process, so that $N^* \sim \text{Pois}(-\log t)$, we see that

$$\left\{-\sum_{i=1}^{n} \log U_i \geq -\log t\right\} = \{n > N^*\}.$$

We can therefore conclude that

$$\mathbb{P}(X \leq t) = \mathbb{P}(N > N^*),$$

where N and N^* are independent. $\qquad\square$

Remark 2. In view of Theorem 2, it is worth mentioning that the use of the union bound and concentration inequalities for the Poisson distribution fails to recover Theorem 1. Nevertheless, Theorem 2 suggests that the bound of Theorem 1 can be generalized to one involving two independent Poisson random variables; this is done in the following corollary.

Corollary 2. *Let M_μ and M_ν be two independent Poisson random variables with means μ and ν, respectively. Then we have that*

$$\mathbb{P}(M_\mu \geq M_\nu) \leq \exp\left\{-\left(\sqrt{\nu} - \sqrt{\mu}\right)_+^2\right\}. \tag{10}$$

Proof. Let $\beta > 0$. Using Markov's Inequality, we observe that

$$\begin{aligned}
\mathbb{P}(M_\mu \geq M_\nu) &= \mathbb{P}(e^{\beta M_\mu} \geq e^{\beta M_\nu}) \\
&\leq \mathbb{E}\left[e^{\beta M_\mu}\right]\mathbb{E}\left[e^{-\beta M_\nu}\right] \\
&= \exp\left\{\mu(e^\beta - 1) + \nu(e^{-\beta} - 1)\right\}.
\end{aligned}$$

Minimizing the above bound with respect to β yields (10). $\square$

We are now ready to state and prove our asymptotic result. The proof makes use of a couple simple sets of inequalities which are given in the appendix.

Theorem 3. *We have that*

$$\log \mathbb{P}(X \leq t) = -\left(\sqrt{-\log(t)} - \sqrt{\lambda}\right)^2 + O\left(\log(-\log t)\right). \tag{11}$$

Proof. By Theorem 2, it suffices to consider $\mathbb{P}(N > N^*)$, where $N \sim \text{Pois}(\lambda)$ and $N^* \sim \text{Pois}(-\log t)$ are independent. For a lower bound, we have that

$$\begin{aligned}
\mathbb{P}(N > N^*) &= \sum_{k=0}^\infty \mathbb{P}(N^* = k)\mathbb{P}(N > N^* | N^* = k) \\
&= \sum_{k=0}^\infty \frac{e^{-(-\log t)}(-\log t)^k}{k!}\mathbb{P}(N \geq k+1) \\
&\geq t\sum_{k=0}^\infty \frac{(-\log t)^k}{k!}\frac{e^{-\lambda}\lambda^{k+1}}{(k+1)!} && \text{(Lemma 2)} \\
&= te^{-\lambda}\left(\sum_{k=1}^\infty \frac{(-\log t)^k\lambda^{k+1}}{k!(k+1)!} + \lambda\right) \\
&\geq te^{-\lambda}\left(\sum_{k=1}^\infty(-\log t)^k\lambda^{k+1}\frac{2^{2k}}{\sqrt{2\pi k}(2k+1)!} + \lambda\right) && \text{(Lemma 3)} \\
&\geq te^{-\lambda}\left(\frac{1}{\sqrt{2\pi}}\sum_{k=1}^\infty(-\log t)^k\lambda^{k+1}\frac{2^{2k}}{(2k+2)(2k+1)!} + \lambda\right) \\
&= te^{-\lambda}\left(-\frac{1}{4\sqrt{2\pi}\log t}\sum_{k=1}^\infty\frac{(2\sqrt{-\lambda\log t})^{2k+2}}{(2k+2)!} + \lambda\right) \\
&= te^{-\lambda}\left(-\frac{1}{4\sqrt{2\pi}\log t}\left(\frac{e^{2\sqrt{-\lambda\log t}} + e^{-2\sqrt{-\lambda\log t}}}{2} - 1 - \frac{(2\sqrt{-\lambda\log t})^2}{2}\right) + \lambda\right) \\
&\geq te^{-\lambda}\left(-\frac{1}{4\sqrt{2\pi}\log t}\left(\frac{e^{2\sqrt{-\lambda\log t}}}{2} - 1 + 2\lambda\log t\right) + \lambda\right).
\end{aligned}$$

We prove an analogous upper bound for $\lambda < 2$ (see Remark 3 following this proof for

6

comments on how to extend it to the case $\lambda \geq 2$). To this end, we observe that

$$
\begin{aligned}
\mathbb{P}(N > N^*) &= \sum_{k=0}^{\infty} \frac{e^{-(-\log t)}(-\log t)^k}{k!} \mathbb{P}(N \geq k+1) \\
&\leq t \sum_{k=0}^{\infty} \frac{(-\log t)^k}{k!} \frac{e^{-\lambda}\lambda^{k+1}}{(k+1)!} \frac{1}{1-\lambda/(k+2)} \qquad \text{(Lemma 2)} \\
&\leq \frac{te^{-\lambda}}{1-\lambda/2} \left(\sum_{k=1}^{\infty} \frac{(-\log t)^k \lambda^{k+1}}{k!(k+1)!} + \lambda \right) \\
&\leq \frac{te^{-\lambda}}{1-\lambda/2} \left(\sum_{k=1}^{\infty} (-\log t)^k \lambda^{k+1} \frac{(2\sqrt{2})2^{2k}}{\sqrt{2\pi k}(2k+1)!} + \lambda \right) \qquad \text{(Lemma 3)} \\
&\leq \frac{te^{-\lambda}}{1-\lambda/2} \left(\frac{1}{\sqrt{\pi}} \sum_{k=1}^{\infty} (-\log t)^k \lambda^{k+1} \frac{2^{2k+1}}{(2k+1)!} + \lambda \right) \\
&= \frac{te^{-\lambda}}{1-\lambda/2} \left(\sqrt{\frac{\lambda}{-\pi \log t}} \sum_{k=1}^{\infty} \frac{(2\sqrt{-\lambda \log t})^{2k+1}}{(2k+1)!} + \lambda \right) \\
&\leq \frac{te^{-\lambda}}{1-\lambda/2} \left(\sqrt{\frac{\lambda}{-\pi \log t}} e^{2\sqrt{-\lambda \log t}} + \lambda \right).
\end{aligned}
$$

$\square$

Remark 3. Note that our use of the upper bound provided by Lemma 2 in the proof is only valid for $\lambda < 2$, and it is slightly messier (though straightforward) to handle the case $\lambda \geq 2$. For this case, we simply observe that

$$
\begin{aligned}
\mathbb{P}(N > N^*) &= \sum_{k=0}^{\infty} \frac{e^{-(-\log t)}(-\log t)^k}{k!} \mathbb{P}(N \geq k+1) \\
&= t \sum_{k=0}^{\lfloor \lambda \rfloor} \frac{(-\log t)^k}{k!} \mathbb{P}(N \geq k+1) + t \sum_{k=\lfloor \lambda \rfloor+1}^{\infty} \frac{(-\log t)^k}{k!} \mathbb{P}(N \geq k+1) \\
&\leq t \sum_{k=0}^{\lfloor \lambda \rfloor} \frac{(-\log t)^k}{k!} + t \sum_{k=\lfloor \lambda \rfloor+1}^{\infty} \frac{(-\log t)^k}{k!} \mathbb{P}(N \geq k+1).
\end{aligned}
$$

The first term above is $O(t(-\log t)^{\lfloor \lambda \rfloor})$, while the second term can be handled exactly as in the proof of Theorem 3 (i.e., by applying the upper bound in Lemma 2) since $k+2 > \lambda$ for all $k \geq \lfloor \lambda \rfloor + 1$. The result of the theorem remains unchanged.

References

Bubeck, S., L. Devroye, and G. Lugosi (2017). Finding adam in random growing trees. *Random Structures & Algorithms 50*(2), 158–172.

Devroye, L. and T. Reddad (2019). On the discovery of the seed in uniform attachment trees. *Internet Mathematics*.

Feller, W. (1968). *An introduction to probability theory and its applications* (3 ed.), Volume 1. Wiley New York.

Khim, J. and P.-L. Loh (2017). Confidence sets for the source of a diffusion in regular trees. *IEEE Transactions on Network Science and Engineering 4*(1), 27–40.

Nelson, R. and T. K. Philips (1995). The moment bound is tighter than chernoff's bound for positive tail probabilities. *The American Statistician 49*(2), 175–178.

Ross, S. M. (1996). *Stochastic Processes* (2 ed.). Wiley New York.

4 Appendix

We borrow from folklore the following simple set of inequalities for the upper tail of a Poisson random variable.

Lemma 2. *Let $Y \sim Pois(\mu)$. Then for all $\mu < n + 1$, we have that*

$$\frac{e^{-\mu}\mu^n}{n!} \leq \mathbb{P}(Y \geq n) \leq \frac{e^{-\mu}\mu^n}{n!} \frac{1}{1 - \mu/(n+1)}. \tag{12}$$

Proof. The lower bound is trivial. For the upper bound, we observe for $\mu < n + 1$ that

$$\mathbb{P}(Y \geq n) = \sum_{k=n}^{\infty} \frac{e^{-\mu}\mu^k}{k!} \leq \frac{e^{-\mu}\mu^n}{n!} \sum_{k=0}^{\infty} \frac{\mu^k}{(n+1)^k} = \frac{e^{-\mu}\mu^n}{n!} \frac{1}{1 - \mu/(n+1)}.$$

$\square$

The next lemma is an easy consequence of Stirling's approximation.

Lemma 3. *For all $n \in \mathbb{N}$, we have that*

$$\frac{2^{2n}}{\sqrt{2\pi n}(2n+1)!} \leq \frac{1}{n!(n+1)!} \leq 2\sqrt{2} \frac{2^{2n}}{\sqrt{2\pi n}(2n+1)!}. \tag{13}$$

Proof. Let $n \in \mathbb{N}$. We begin by observing that the inequalities

$$\frac{(n!)^2}{2(2n)!} \leq \frac{n!(n+1)!}{(2n+1)!} \leq \frac{(n!)^2}{(2n)!}$$

hold, which combined with Stirling's bounds (see, e.g., Feller (1968)),

$$e^{\frac{1}{12n+1}}\sqrt{2\pi}n^{n+1/2}e^{-n} \leq n! \leq e^{\frac{1}{12n}}\sqrt{2\pi}n^{n+1/2}e^{-n},$$

yield that

$$\frac{e^{\frac{2}{12n+1}-\frac{1}{24n}}\sqrt{2\pi n}}{2^{2n+3/2}} \leq \frac{n!(n+1)!}{(2n+1)!} \leq \frac{e^{\frac{1}{6n}-\frac{1}{24n+1}}\sqrt{2\pi n}}{2^{2n+1/2}}.$$

The preceding inequalities simplify to the inequalities

$$\frac{1}{2\sqrt{2}}\frac{\sqrt{2\pi n}}{2^{2n}} \leq \frac{n!(n+1)!}{(2n+1)!} \leq \frac{\sqrt{2\pi n}}{2^{2n}},$$

which are an easy rearrangement of those in (13). $\qquad\qquad$ □