A Simple Path to Biggins' Martingale Convergence for Branching Random Walk

BY RUSSELL LYONS

Abstract. We give a simple non-analytic proof of Biggins' theorem on martingale convergence for branching random walks.

Let $\mathcal{L} := \{X_i\}_{i=1}^{L}$ be a random L-tuple of real numbers, where L is also random and can take the values 0 and ∞. This can also be thought of as an ordered point process on $\mathbf{R}$. The random variable $\mathcal{L}$ is used as the basis for construction of a branching random walk in the usual way: An initial particle at the origin of $\mathbf{R}$ gives birth to L particles with displacements $X_1, X_2, \ldots$. Then each of these particles gives birth to a random number of particles with random displacements from its new position according to the same law as $\mathcal{L}$ and independently of one another and of the initial displacements. This continues in a like manner forever or until there are no more particles. For a particle σ, write $|\sigma|$ for the generation in which σ is born, $X(\sigma)$ for its displacement from its parent, and $S(\sigma)$ for its position. Denote the initial particle by 0, also known as the root of the family tree. If τ is an ancestor of σ, write $\tau < \sigma$. Thus, we have

$$S(\sigma) = \sum_{0<\tau\leq\sigma} X(\tau)\,.$$

Also, write $\mathcal{L}(\sigma)$ for the copy of $\mathcal{L}$ used to generate the children of σ. Let q be the extinction probability of the underlying Galton-Watson process.

For $\alpha \in \mathbf{R}$, define $\langle \alpha, \mathcal{L}\rangle := \sum_{i=1}^{L} e^{-\alpha X_i}$ and $m(\alpha) := \mathbf{E}[\langle \alpha, \mathcal{L}\rangle] \in (0, \infty]$. Assume that $m(0) > 1$, so that $q < 1$. If $m(\alpha) < \infty$ for some α, then the sequence

$$W_n(\alpha) := \frac{\sum_{|\sigma|=n} e^{-\alpha S(\sigma)}}{m(\alpha)^n}$$

is a martingale with a.s. limit $W(\alpha)$. Write

$$m'(\alpha) := \mathbf{E}\left[\sum_{i=1}^{L} X_i e^{-\alpha X_i}\right]$$

1991 Mathematics Subject Classification. Primary 60J80.
Key words and phrases. Galton-Watson.
Research partially supported by the Institute for Mathematics and Its Applications (Minneapolis) and NSF Grant DMS-9306954.

when this exists in $[-\infty, \infty]$ as a Lebesgue integral. Biggins (1977) has determined when $W(\alpha)$ is nontrivial:

BIGGINS' THEOREM. *Suppose that $\alpha \in \mathbf{R}$ is such that $m(\alpha) < \infty$ and $m'(\alpha)$ exists and is finite. Then the following are equivalent:*

(i) $\mathbf{P}[W(\alpha) = 0] = q$;

(ii) $\mathbf{P}[W(\alpha) = 0] < 1$;

(iii) $\mathbf{E}[W(\alpha)] = 1$;

(iv) $\mathbf{E}[\langle \alpha, \mathcal{L} \rangle \log^+ \langle \alpha, \mathcal{L} \rangle] < \infty$ *and* $\alpha m'(\alpha)/m(\alpha) < \log m(\alpha)$.

REMARK. In fact, the hypotheses here are very slightly weaker than those of Biggins (1977), Lemma 5. Moreover, the proof to follow works without the assumption that $m'(\alpha)$ be finite, except for the implication (ii) $\Rightarrow$ (iv), where it needs the assumption that $\alpha m'(\alpha) \neq -\infty$.

REMARK. The case of Biggins' Theorem where L is constant, X_i are independent and identically distributed, and $m(\alpha) = 1$ was proved also by Kahane (see Kahane and Peyrière (1976); the first condition in (iv) above follows from the assumptions that $m(\alpha) < \infty$ and $|m'(\alpha)| < \infty$ in the case that L is bounded since convexity of the function $x \mapsto x \log x$ shows that $\mathbf{E}[\langle \alpha, \mathcal{L} \rangle \log^+ \langle \alpha, \mathcal{L} \rangle] \leq |m'(\alpha)| + \| \log^+ L \|_\infty m(\alpha)$). When the conclusions of Biggins' Theorem hold in Kahane's context, the measure $\widehat{\mu}$ below is introduced by Peyrière on p. 141 of Kahane and Peyrière (1976) for another purpose. It and related constructions in other situations also occur, usually including the same direct construction as ours, in Kallenberg (1977), Hawkes (1981), Rouault (1981), Joffe and Waugh (1982), Kesten (1986), Chauvin and Rouault (1988), Chauvin, Rouault and Wakolbinger (1991), and Waymire and Williams (1996).

Evidently, (iii) implies (ii). The fact that (i) and (ii) are equivalent follows from the standard "zero-one" property of Galton-Watson processes. We shall present a simple proof of the other equivalences modelled on the proof of the Kesten-Stigum theorem in Lyons, Pemantle and Peres (1995). I am grateful to Anatole Joffe for asking me for the details of how this is done. The same method was discovered independently by Waymire and Williams (1996) for the case treated by Kahane (mentioned above). In fact, Waymire and Williams relax the condition that the X_i be i.i.d. They even relax the independence of the $\mathcal{L}(\sigma)$, which could be done here as well.

Proof. Fix α. If t is a rooted tree (with distinguishable vertices) and X is a real-valued function on the vertices of t other than its root, we call (t, X) a **labelled tree**. A **ray** in a tree is an infinite line of descent starting from the root. Given a ray ξ, the vertex on ξ in generation n is denoted ξ_n. In the space of labelled trees, let $\mathcal{F}_n$ denote the σ-field generated by the first n levels. We shall also work on the space of labelled trees with distinguished rays, (t, X, ξ); denote by $\mathcal{F}_n^*$ the σ-field generated by the first n levels there. For $\sigma \in t$, write $S(\sigma) := \sum_{0<\tau\leq\sigma} X(\tau)$ and set

$$W_n(t, X) := \frac{\sum_{|\sigma|=n} e^{-\alpha S(\sigma)}}{m(\alpha)^n} .$$

Branching random walk gives a probability measure, μ, on the set of labelled trees. We shall construct a related probability measure $\widehat{\mu}^*$ on the set of infinite labelled trees with distinguished rays. Let μ_n be the restriction of μ to $\mathcal{F}_n$. Now any $\mathcal{F}_n^*$-measurable function f can be written as

$$f(t, X, \xi) = \sum_{|\sigma|=n} f_\sigma(t, X)\mathbf{1}_{\xi_n=\sigma}$$

for some $\mathcal{F}_n$-measurable functions f_σ. Let μ_n^* be counting measure on $\{|\sigma| = n\}$ fibered over μ_n; more precisely, μ_n^* is the (non-probability) measure on $\mathcal{F}_n^*$ such that for all nonnegative $\mathcal{F}_n^*$-measurable functions f,

$$\int f(t, X, \xi)\, d\mu_n^* = \int \sum_{|\sigma|=n} f_\sigma(t, X)\, d\mu_n .$$

Then the measure $\widehat{\mu}^*$ will satisfy

$$\frac{d\widehat{\mu}_n^*}{d\mu_n^*}(t, X, \xi) = \frac{e^{-\alpha S(\xi_n)}}{m(\alpha)^n} \tag{1}$$

for all n and all labelled trees with rays (t, X, ξ), where $\widehat{\mu}_n^*$ denotes the restriction of $\widehat{\mu}^*$ to $\mathcal{F}_n^*$. The projection of $\widehat{\mu}^*$ to the space of trees, denoted $\widehat{\mu}$, then satisfies

$$\frac{d\widehat{\mu}_n}{d\mu_n}(t, X) = W_n(t, X) \tag{2}$$

for all n and all labelled trees (t, X), where $\widehat{\mu}_n$ denotes the restriction of $\widehat{\mu}$ to $\mathcal{F}_n$.

To define $\widehat{\mu}^*$, let $\widehat{\mathcal{L}}$ be a random variable whose law has Radon-Nikodym derivative $\langle \alpha, \mathcal{L}\rangle / m(\alpha)$ with respect to the law of $\mathcal{L}$. Start with one particle v_0 at the origin. Generate offspring and displacements according to a copy $\widehat{\mathcal{L}}_1$ of $\widehat{\mathcal{L}}$. Pick one of these children v_1 at random, where a child is picked with probability proportional to $e^{-\alpha X}$ when its displacement is X. The children other than v_1 give rise to ordinary independent branching random walks, while v_1 gets an independent number of offspring and displacements according to a copy $\widehat{\mathcal{L}}_2$ of $\widehat{\mathcal{L}}$. Again, pick one of the children of v_1 at random, call it v_2, with the others giving rise to ordinary independent branching random walks, and so on.

Define the measure $\widehat{\mu}^*$ as the joint distribution of the random labelled tree and the random ray $(v_0, v_1, v_2, \ldots)$. Then

$$\frac{d\widehat{\mu}^*_{n+1}}{d\mu^*_{n+1}}(t, X, \xi) = \frac{d\widehat{\mu}^*_n}{d\mu^*_n}(t, X, \xi) \cdot \frac{\langle \alpha, \mathcal{L}(\xi_n) \rangle}{m(\alpha)} \cdot \frac{e^{-\alpha X(\xi_{n+1})}}{\langle \alpha, \mathcal{L}(\xi_n) \rangle} = \frac{e^{-\alpha X(\xi_{n+1})}}{m(\alpha)} \frac{d\widehat{\mu}^*_n}{d\mu^*_n}(t, X, \xi) .$$

Thus, (1) follows by induction.

Note that for any $k \geq 0$,

$$\int X(v_k)\, d\widehat{\mu}^* = \mathbf{E}\left[\sum_{i=1}^{L} X_i \frac{e^{-\alpha X_i}}{\langle \alpha, \mathcal{L} \rangle} \frac{\langle \alpha, \mathcal{L} \rangle}{m(\alpha)} \right] = -m'(\alpha)/m(\alpha) . \tag{3}$$

Thus, by the strong law of large numbers, we have

$$S(v_n)/n \to -m'(\alpha)/m(\alpha) \quad \widehat{\mu}^*\text{-a.s.} \tag{4}$$

For any labelled tree (t, X), set $W(t, X) := \limsup W_n(t, X)$. Now by (2), we have the implications (Durrett (1991), p. 210, Exercise 3.6)

$$W(t, X) = \infty \quad \widehat{\mu}\text{-a.s.} \quad \Longrightarrow \quad W(t, X) = 0 \quad \mu\text{-a.s.} , \tag{5}$$

$$W(t, X) < \infty \quad \widehat{\mu}\text{-a.s.} \quad \Longrightarrow \quad \int W(t, X)\, d\mu = 1 . \tag{6}$$

Suppose first that (iv) fails. We have

$$W_{n+1}(t, X) \geq \frac{e^{-\alpha S(v_n)}}{m(\alpha)^{n+1}} \langle \alpha, \widehat{\mathcal{L}}_{n+1} \rangle , \tag{7}$$

with the two terms in the product being $\widehat{\mu}^*$-independent. Now, if $\alpha m'(\alpha)/m(\alpha) \geq \log m(\alpha)$, then $\limsup e^{-\alpha S(v_n)}/m(\alpha)^n = \infty$ by (4) in case $\alpha m'(\alpha)/m(\alpha) > \log m(\alpha)$ and by (3) and the Chung-Fuchs theorem in case $\alpha m'(\alpha)/m(\alpha) = \log m(\alpha)$. This implies by (7) that $W(t, X) = \infty$ $\widehat{\mu}$-a.s., whence by (5), (ii) fails. On the other hand, if $\alpha m'(\alpha)/m(\alpha) < \log m(\alpha)$, then since $\mathbf{E}[\log^+ \langle \alpha, \widehat{\mathcal{L}} \rangle] = \mathbf{E}[\langle \alpha, \mathcal{L} \rangle \log^+ \langle \alpha, \mathcal{L} \rangle]/m(\alpha) = \infty$ by assumption, we have that

$$\limsup \frac{1}{n} \log^+ \langle \alpha, \widehat{\mathcal{L}}_n \rangle = \infty \quad \widehat{\mu}^*\text{-a.s.}$$

by virtue of the Borel-Cantelli lemma. This means that the first term in the right-hand side of (7) decays exponentially while the second has superexponential explosions. Hence, again, $W(t, X) = \infty$ $\widehat{\mu}$-a.s.

Conversely, suppose that (iv) holds. Let $\mathcal{G}$ be the σ-field generated by $\{\widehat{\mathcal{L}}_k\}_{k\geq 1}$. Then

$$\begin{aligned}\mathbf{E}_{\widehat{\mu}^*}[W_n(t,X) \mid \mathcal{G}] &= \sum_{k=0}^{n-1} \frac{e^{-\alpha S(v_k)}}{m(\alpha)^{k+1}}\left(\langle \alpha, \widehat{\mathcal{L}}_{k+1}\rangle - e^{-\alpha X(v_{k+1})}\right) + \frac{e^{-\alpha S(v_n)}}{m(\alpha)^n} \\ &= \sum_{k=0}^{n-1} \frac{e^{-\alpha S(v_k)}}{m(\alpha)^{k+1}} \langle \alpha, \widehat{\mathcal{L}}_{k+1}\rangle - \sum_{k=1}^{n-1} \frac{e^{-\alpha S(v_k)}}{m(\alpha)^k}\,.\end{aligned}$$

By hypothesis and (4), the terms $e^{-\alpha S(v_k)}/m(\alpha)^k$ decay exponentially while the terms $\langle \alpha, \widehat{\mathcal{L}}_{k+1}\rangle$ grow (at most) subexponentially by the Borel-Cantelli lemma again. Therefore both series converge $\widehat{\mu}^*$-a.s., whence $\liminf W_n(t,X) < \infty$ $\widehat{\mu}$-a.s. by Fatou's lemma. In light of (2), $\{1/W_n(t,X)\}$ is a $\widehat{\mu}$-supermartingale, so that $\{W_n(t,X)\}$ converges $\widehat{\mu}$-a.s. Thus, we have $W(t,X) < \infty$ $\widehat{\mu}$-a.s. and (iii) is a consequence of (6). ∎

REFERENCES

Biggins, J. D. (1977). Martingale convergence in the branching random walk. *J. Appl. Prob.* **14**, 25–37.

Chauvin, B. and Rouault, A. (1988). KPP equation and supercritical branching Brownian motion in the subcritical speed area. Application to spatial trees, *Probab. Theory Relat. Fields* **80**, 299–314.

Chauvin, B., Rouault, A., and Wakolbinger, A. (1991). Growing conditioned trees, *Stochastic Process. Appl.* **39**, 117–130.

Durrett, R. (1991). Probability: Theory and Examples. Wadsworth, Pacific Grove, California.

Hawkes, J. (1981). Trees generated by a simple branching process, *J. London Math. Soc.* **24**, 373–384.

Joffe, A. and Waugh, W. A. O'N. (1982). Exact distributions of kin numbers in a Galton-Watson process, *J. Appl. Prob.* **19**, 767–775.

Kahane, J.-P. and Peyrière, J. (1976). Sur certaines martingales de Benoit Mandelbrot, *Adv. in Math.* **22**, 131–145.

Kallenberg, O. (1977). Stability of critical cluster fields, *Math. Nachr.* **77**, 7–43.

Kesten, H. (1986). Subdiffusive behavior of random walk on a random cluster, *Ann. Inst. H. Poincaré Probab. Statist.* **22**, 425–487.

Lyons, R., Pemantle, R. and Peres, Y. (1995). Conceptual proofs of $L \log L$ criteria for mean behavior of branching processes, *Ann. Probab.*, **23** (1995), 1125–1138.

Rouault, A. (1981). Lois empiriques dans les processus de branchement spatiaux homogènes supercritiques, *C. R. Acad. Sci. Paris Sér. I. Math.* **292**, 933–936.

Waymire, E. C. and Williams, S. C. (1996). A cascade decomposition theory with applications to Markov and exchangeable cascades, *Trans. Amer. Math. Soc.* **348**, 585–632.

Department of Mathematics, Indiana University, Bloomington, IN 47405-5701