The Equivalence of Two Constructions of Galton-Watson Processes

Kenji KURATA *   Nariyuki MINAMI †‡

Abstract

In most textbooks on branching processes, the Galton-Watson process is defined as an integer valued Markov chain with a special kind of transition probability, somtimes supplemented with an intuitive description of random family trees. A precise construction of the Galton-Watson process on the space of trees was first given by R. Otter in 1949, and later independently by J. Neveu in 1986. In this note, we show that these two constructions are actually equivalent.

Keywords : branching process; Galton-Watson process; random trees; construction of probability space

1 Introduction

In his elegant paper [2], J. Neveu gave a convenient definition of a σ-field $\mathcal{F}$ and a probability measure P on the space Ω of trees, obtaining a probability space which describes every detail of a Galton-Watson branching process. Neveu was motivated by the lack of fundamental theory of this kind in the existing works on branching processes, but apparently he did not know that already in 1949, R. Otter [3] had given a similar definition which turns out to be equivalent to Neveu's definition. The purpose of this note is to show this equivalence.

Otter's construction looks more honest to our intuition of branching processes, if not as concise as Neveu's, and is closely related to the topological structure of the space Ω of trees. Also, it is remarkable that Otter already emphasizes in [3] the conceptual advantage of realizing Galton-Watson processes on the space of trees, instead of treating them merely as Markov chains with a special kind of transition probabilities.

2 The construction by Neveu ([2])

Let U be the totality of finite sequences of strictly positive integers, of which the empty sequence is denoted by ϕ. We can write $U = \sum_{n\geq 0}(\mathbf{N}^*)^n$, where $\mathbf{N}^* = \{1, 2, \ldots\}$, $(\mathbf{N}^*)^0 =$

*Institute of Mathematics, University of Tsukuba, Tsukuba, 305-8571, Japan
†Institute of Mathematics, University of Tsukuba, Tsukuba, 305-8571, Japan
‡e-mail : minami@math.tsukuba.ac.jp

$\{\phi\}$, and $(\mathbf{N}^*)^n$ is the totality of sequences $u = j_1 j_2 \cdots j_n$ of length n. By defintion, a tree is a subset $\omega \subset U$ which satisfies the following three conditions: a) $\phi \in \omega$; b) $uj \in \omega$ implies $u \in \omega$ whenever $u \in U$ and $j \in \mathbf{N}^*$, where we write $uj = j_1 j_2 \cdots j_n j$ if $u = j_1 j_2 \cdots j_n$; c) for each $u \in \omega$, there is an integer $\nu_u(\omega) \geq 0$ such that for any $j \in \mathbf{N}^*$, $uj \in \omega$ if and only if $1 \leq j \leq \nu_u(\omega)$. Let Ω be the totality of trees in this sense. Then $\mathcal{F}$ is defined to be the smallest σ-field containing all subsets $\Omega_u := \{\omega \in \Omega \mid \omega \ni u\}$, $u \in U$.

Given a probability distribution $\Pi = \{p_n\}_{n=0}^{\infty}$ on non-negative integers $\mathbf{N}$, let $(\Omega^*, \mathcal{F}^*, P^*)$ be the product space $(\mathbf{N}, \Pi)^{\otimes U}$, $\mathcal{F}^*$ being the σ- field generated by the coordinate maps $\nu_v^* : \omega^* = (\omega_u^*)_{u \in U} \mapsto \omega_v^*$, $v \in U$. If we define the map $\psi : \Omega^* \mapsto \Omega$ by

$$\psi(\omega^*) = \{u = j_1 \ldots j_p \in U |\ j_{k+1} \leq \nu_{j_1 \cdots j_k}(\omega^*) \ , \ \text{for } 0 \leq k < p\} \ ,$$

then ψ is $\mathcal{F}^*/\mathcal{F}$-measurable, and we can consider the induced measure $P = P^* \circ \psi^{-1}$ on $(\Omega, \mathcal{F})$. The resulting probability space $(\Omega, \mathcal{F}, P)$ then describes every detail of the Galton-Watson process with the offspring distribution Π. In particular, if $Z_n(\omega)$ is the cardinality of $\omega \cap (\mathbf{N}^*)^n$, then under this probability measure P, $\{Z_n(\omega)\}_{n=0}^{\infty}$ is an $\mathbf{N}$-valued Markov chain with transition probability $P(x, y) = (\Pi^{*x})(\{y\})$, where " $*$"deotes the convolution.

3 The construction by Otter ([3])

Since the definition of trees by Otter ([3]) is obviously seen to be equivalent to Neveu's definition, we shall work on the same space Ω introduced in the previous section.

For each tree ω, we define the set $\mathcal{I}(\omega)$ of its inner points and the set $\mathcal{E}(\omega)$ of its endpoints as follows:

$$\mathcal{I}(\omega) := \{u \in \omega \mid \nu_u(\omega) > 0\} \quad ; \quad \mathcal{E}(\omega) := \{u \in \omega \mid \nu_u(\omega) = 0\} \ .$$

For two trees $\omega, \omega' \in \Omega$, we say that ω extends ω' and write $\omega \geq \omega'$, if the following two conditions hold: i) $\omega \supset \omega'$; ii) $\nu_u(\omega) = \nu_u(\omega')$ for each $u \in \mathcal{I}(\omega')$.

Next let Ω^f be the totality of finite trees. For a $T \in \Omega^f$ and $\Lambda = (\lambda_e; e \in \mathcal{E}(T)) \in \mathbf{N}^{\mathcal{E}(T)}$, define

$$[T; \Lambda] = \{\omega \in \Omega \mid \omega \geq T \ , \ \nu_e(\omega) = \lambda_e \text{ for } e \in \mathcal{E}(T)\} \ ,$$

which Otter called a "neighborhood ".

Given a probability distribution $\Pi = \{p_n\}_{n=0}^{\infty}$ on $\mathbf{N}$, Otter constructs a σ-field $\mathcal{B}$ and a probability measure Q on $(\Omega, \mathcal{B})$ in the following way.

First he let $\mathcal{S}$ be the class of subsets $S \subset \Omega$ of the form $S = \sum_{\Lambda \in A} [T; \Lambda]$ with $T \in \Omega^f$ and $A = \prod_{e \in \mathcal{E}(T)} A_e \subset \mathbf{N}^{\mathcal{E}(T)}$. Then $\mathcal{S}$ is shown to be a semi-algebra. Next he defines a set function $\tilde{Q}$ on $\mathcal{S}$ by

$$\tilde{Q}([T; \Lambda]) = \left(\prod_{u \in \mathcal{I}(T)} p_{\nu_u(T)} \right) \left(\prod_{e \in \mathcal{E}(T)} p_{\lambda_e} \right) \quad ; \quad \tilde{Q}(S) = \sum_{\Lambda \in A} \tilde{Q}([T; \Lambda]) \ ,$$

where the value of $\tilde{Q}(S)$ turns out to be independent of the expression of S. He then shows that $\tilde{Q}$ is countably additive on $\mathcal{S}$, and hence is extended to a probability measure Q on the σ-field $\mathcal{B}$ generated by $\mathcal{S}$ (or equivalently generated by $\mathcal{N}$).

Otter's discussion is actually quite sketchy, and in order to fill its details, it is necessary to prepare several lemmas concerning the sets $[T;\Lambda]$, which are not as trivial as regarded by Otter. Among those lemmas, the following one will be useful in the rest of this note:

Lemma 1 *The class $\mathcal{N}$ defined by*

$$\mathcal{N} := \{[T;\Lambda] \mid T \in \Omega^f \ , \ \Lambda \in \mathbf{N}^{\mathcal{E}(T)}\} \cup \{\emptyset\}$$

is a π-system, namely is closed under the formation of finite intersection.

Proof. We shall show that if $[T_i;\Lambda_i] \in \mathcal{N}$, $i = 1, 2$, have non-empty intersection, then we can construct a $T_3 \in \Omega^f$ and $\Lambda_3 \in \mathbf{N}^{\mathcal{E}(T_3)}$ such that

$$[T_1;\Lambda_1] \cap [T_2;\Lambda_2] = [T_3;\Lambda_3] \ .$$

For this purpose, let $T_3 = T_1 \cup T_2$. Then obviously $T_3 \in \Omega^f$. It is also easy to see that $\mathcal{I}(T_3) = \mathcal{I}(T_1) \cup \mathcal{I}(T_2)$, and that

$$\mathcal{E}(T_3) = [\mathcal{E}(T_1) \setminus T_2] + [\mathcal{E}(T_2) \setminus T_1] + [\mathcal{E}(T_1) \cap \mathcal{E}(T_2)] \ .$$

Now suppose $\omega \in [T_1;\Lambda_1] \cap [T_2;\Lambda_2]$. If $u \in \mathcal{I}(T_1) \cap \mathcal{I}(T_2)$, then since $\omega \geq T_i$, $i = 1, 2$, we have $\nu_u(T_1) = \nu_u(\omega) = \nu_u(T_2)$. In this case, we also have $\nu_u(T_3) = \nu_u(\omega)$. If on the other hand, $u \in \mathcal{I}(T_1) \setminus \mathcal{I}(T_2)$, we see from $\omega \geq T_1$ that $\nu_u(T_1) = \nu_u(T_3) = \nu_u(\omega)$, and if $u \in \mathcal{I}(T_2) \setminus \mathcal{I}(T_1)$, we see from $\omega \geq T_2$ that $\nu_u(T_2) = \nu_u(T_3) = \nu_u(\omega)$. Thus $T_3 \geq T_i$, $i = 1, 2$, and whenever $u \in \mathcal{I}(T_3)$, one has $\nu_u(T_3) = \nu_u(\omega)$, showing $\omega \geq T_3$.

Let us write $\Lambda_i = (\lambda_e^i; e \in \mathcal{E}(T_i))$ for $i = 1, 2$. If $e \in \mathcal{E}(T_1) \setminus T_2$, then $\omega \in [T_1;\Lambda_1]$ implies $\nu_e(\omega) = \lambda_e^1$, and if $e \in \mathcal{E}(T_2) \setminus T_1$, then $\omega \in [T_2;\Lambda_2]$ implies $\nu_e(\omega) = \lambda_e^2$. Finally if $e \in \mathcal{E}(T_1) \cap \mathcal{E}(T_2)$, then we must have $\lambda_e^1 = \nu_e(\omega) = \lambda_e^2$ by $\omega \in [T_1;\Lambda_1] \cap [T_2;\Lambda_2]$. Thus if we define $\Lambda_3 = (\lambda_e^3; e \in \mathcal{E}(T_3))$ by

$$\lambda_e^3 = \begin{cases} \lambda_e^1, & \text{for } e \in \mathcal{E}(T_1) \setminus T_2 \\ \lambda_e^2, & \text{for } e \in \mathcal{E}(T_2) \setminus T_1 \\ \lambda_e^1 = \lambda_e^2, & \text{for } e \in \mathcal{E}(T_1) \cap \mathcal{E}(T_2) \ , \end{cases}$$

then we must have $\omega \in [T_3;\Lambda_3]$. Since ω was arbitrarily chosen from $[T_1;\Lambda_1] \cap [T_2;\Lambda_2]$, we conclude $[T_1;\Lambda_1] \cap [T_2;\Lambda_2] \subset [T_3;\Lambda_3]$.

To prove the converse inclusion, let $w \in [T_3;\Lambda_3]$ be arbitrary. Since $T_3 \geq T_i$, $i = 1, 2$, we have $w \geq T_i$, $i = 1, 2$. If $e \in \mathcal{E}(T_1) \setminus \mathcal{I}(T_2)$, then $e \in \mathcal{E}(T_3)$ and $\nu_e(w) = \lambda_e^3 = \lambda_e^1$. If on the other hand, $e \in \mathcal{E}(T_1) \cap \mathcal{I}(T_2)$, then $\nu_e(w) = \nu_e(T_2)$. But we are assuming $[T_1;\Lambda_1] \cap [T_2;\Lambda_2] \neq \emptyset$, so that there exists an ω such that $\omega \geq T_2$ and that $\nu_e(\omega) = \lambda_e^1$. Hence $\nu_e(T_2) = \nu_e(\omega) = \lambda_e^1$. In any case, $e \in \mathcal{E}(T_1)$ implies $\nu_e(w) = \lambda_e^1$. Thus $w \in [T_1;\Lambda_1]$. Similarly one can show $w \in [T_2;\Lambda_2]$, completing the proof of the lemma.

4 The equivalence of two constructions

We are now ready to prove our main assertion.

Proposition 1 *$\mathcal{B} = \mathcal{F}$ and $Q = P$, hence the probability space constructed by Otter and Neveu coincide.*

Proof. Since we can write, with the convention $u0 = u$,

$$[T;\Lambda] = \left(\bigcap_{u\in T} \Omega_u\right) \cap \left(\bigcap_{u\in \mathcal{I}(T)} \Omega^c_{u(\nu_u(T)+1)}\right) \cap \left(\bigcap_{u\in \mathcal{E}(T)} \Omega_{u\lambda_u} \setminus \Omega_{u(\lambda_u+1)}\right) ,$$

we see $[T;\Lambda] \in \mathcal{F}$ and hence $\mathcal{B} \subset \mathcal{F}$.

Conversely we can also write

$$\Omega_u = \bigcup \{[T;\Lambda] \mid u \in T \in \Omega^f \ , \ \Lambda \in \mathbf{N}^{\mathcal{E}(T)}\} ,$$

hence $\Omega_u \in \mathcal{B}$ for any $u \in U$, proving $\mathcal{F} \subset \mathcal{B}$.

On the other hand, since

$$\psi^{-1}([T;\Lambda]) = \left(\bigcap_{u\in \mathcal{I}(T)} \{\omega^* \mid \nu^*_u(\omega^*) = \nu_u(T)\}\right) \cap \left(\bigcap_{u\in \mathcal{E}(T)} \{\omega^* \mid \nu^*_u(\omega^*) = \lambda_u\}\right) ,$$

we have

$$P([T;\Lambda]) = P^*(\psi^{-1}([T;\Lambda])) = \left(\prod_{u\in \mathcal{I}(T)} p_{\nu_u(T)}\right)\left(\prod_{e\in \mathcal{E}(T)} p_{\lambda_e}\right) = Q([T;\Lambda]) .$$

Thus P and Q coincide on the π-system $\mathcal{N}$, hence on $\mathcal{F} = \mathcal{B}$, which is generated by $\mathcal{N}$.

5 Ω as a metric space

As was mentioned by Otter ([3]), and as is obvious from Lemma 1, we can make Ω into a topological space by calling "open"those subsets $G \subset \Omega$ which are written as unions of $[T;\Lambda]$'s. In particular, our σ-field $\mathcal{F}$ is the Borel σ-field corresponding to this topology. Let us briefly show that this topology is generated by a metric on Ω.

For $n \geq 0$, let $z_n(\omega) = \omega \cap (\mathbf{N}^*)^n$. Given $\omega \ , \ \omega' \in \Omega$, let us define

$$\mu(\omega,\omega') = \sup\{n \geq 0 \mid z_n(\omega) = z_n(\omega') \} ,$$

where we let $\mu(\omega,\omega') = \infty$ if $z_n(\omega) = z_n(\omega')$ for all $n \geq 1$, namely if $\omega = \omega'$. Since we have

$$\mu(\omega,\omega'') \geq \min\{\mu(\omega,\omega') \ , \ \mu(\omega',\omega'') \} ,$$

we can define a metric d on Ω by letting $d(\omega,\omega') := \exp\{-\mu(\omega,\omega')\}$.

Now let $G \subset \Omega$ be open in the sense already defined. Then for each $\omega \in G$, we can choose a $[T;\Lambda]$ such that $\omega \in [T;\Lambda] \subset G$. Let $h(T) = \max\{n \geq 0 | T \cap (\mathbf{N}^*)^n \neq \emptyset\}$ be the "height "of the finite tree T. Then $\omega' \in [T;\Lambda]$ as soon as $z_n(\omega') = z_n(\omega)$ for $0 \leq n \leq h(T)+1$, or equivalently as soon as $d(\omega',\omega) \leq \exp\{-(h(T)+1)\}$. This shows that G is open with respect to the metric d. Conversely, suppose G is open in the metric d. Then for each $\omega \in G$, there is an integer $m \geq 1$ such that $\omega' \in G$ whenever $z_n(\omega') = z_n(\omega)$, $0 \leq n \leq m$. If we define $T = \cup_{n=0}^{m-1} z_n(\omega)$ and $\Lambda = (\lambda_e \ ; \ e \in \mathcal{E}(T))$ with $\lambda_e = \nu_e(\omega)$, then the last condition is equivalent to $\omega' \in [T; \Lambda]$. Thus $\omega \in [T; \Lambda] \subset G$ and G is open in the original sense.

Since $\mathcal{N} = \{[T; \Lambda]\} \cup \{\emptyset\}$ is a countable basis of topology, our metric space (Ω, d) thus obtained is separable. It is also complete. To see this, let $\{\omega^{(k)}\}_{k=1}^{\infty}$ be a Cauchy sequence with respect to the metric d. Then we can choose an increasing sequence $N_n \nearrow \infty$ such that $d(\omega^{(k)}, \omega^{(\ell)}) \leq e^{-n}$ for all $k, \ell \geq N_n$. In other words, we have $z_j(\omega^{(k)}) = z_j(\omega^{(\ell)})$, $1 \leq j \leq n$ for all $k, \ell \geq N_n$. If we let $\omega^{(\infty)} = \{\phi\} \cup \bigcup_{n=1}^{\infty} z_n(\omega^{(N_n)})$, then $\omega^{(\infty)}$ is a tree, and it is clear that $d(\omega^{(k)}, \omega^{(\infty)}) \to 0$ as $k \to \infty$.

Noting that $\mathcal{N}$ is also a π-system, we can apply Theorem 2.2 of [1], to obtain the following criterion for the weak convergence of probability measures on Ω.

Proposition 2 *Let P and P_n, $n = 1, 2, \ldots$ be probability measures on Ω. If $P_n([T; \Lambda]) \to P([T; \Lambda])$ for each $[T; \Lambda] \in \mathcal{N}$, then one has $P_n \Rightarrow P$.*

As an immediate corollary of this proposition, we see that if P_n and P are the probability measures for Galton-Watson processes with offspring distributions Π_n and Π respectively, and if $\Pi_n \Rightarrow \Pi$ on $\mathbf{N}$, the $P_n \Rightarrow P$ on Ω.

Acknowledgment. The authors are grateful to Professor M. Kanda for his continual encouragement. The second author is also grateful to Professors K. Fukuyama and Y. Higuchi for valuable discussion and for giving him opportunity to lecture on this subject.

References

[1] Billingsley, P. (1999). *Convergence of Probability Measures.*, 2nd edn. John Wiley, New York.

[2] Neveu, J. (1986). Arbres et processus de Galton-Watson. *Ann. Inst. H. Poincaré, Probabilités et Statistique* **22,** 199–207.

[3] Otter, R. (1949). The Multiplicative Process. *Ann. Math. Stat.* **20,** 206–224.