New Analytic Results for Speciation Times in Neutral Models

Tanja Gernhard

Department of Mathematics, Kombinatorische Geometrie (M9), TU München, Bolzmannstr. 3, 85747 Garching, Germany

Received: 8 August 2007 / Accepted: 20 November 2007 / Published online: 3 January 2008
© Society for Mathematical Biology 2008

Abstract In this paper, we investigate the standard Yule model, and a recently studied model of speciation and extinction, the “critical branching process.” We develop an analytic way—as opposed to the common simulation approach—for calculating the speciation times in a reconstructed phylogenetic tree. Simple expressions for the density and the moments of the speciation times are obtained.

Methods for dating a speciation event become valuable, if for the reconstructed phylogenetic trees, no time scale is available. A missing time scale could be due to supertree methods, morphological data, or molecular data which violates the molecular clock. Our analytic approach is, in particular, useful for the model with extinction, since simulations of birth-death processes which are conditioned on obtaining \( n \) extant species today are quite delicate. Further, simulations are very time consuming for big \( n \) under both models.

Keywords Phylogenetics · Yule model · Critical branching process · Reconstructed process

1. Introduction

In phylogenetics, a major task is to reconstruct the evolutionary history for some extant species. With the reconstructed trees, the aim is to understand evolution better. A central question is if the species evolved under a neutral model. A neutral model assumes that throughout time, whenever a speciation or extinction event occurs, each species is equally likely to be the one undergoing that event. That is, if a phylogeny evolved under a neutral model, all species behave alike. Of course, speciation is not always just random—some lineages can speciate faster than others due to differing selective pressures and environmental factors. However, for rejecting a neutral model for some data set, we need to know properties of the neutral models (which are not present in the data set). On the other hand, if we may assume a neutral model for some phylogeny, we can make further statements about the evolution of the phylogeny—inferrred from the properties of the neutral model. In this paper, we will calculate the speciation times under a neutral model in a given tree shape.

E-mail address: gernhard@ma.tum.de.
The most popular neutral model is the so-called Yule model (Aldous, 2001; Edwards, 1970; Yule, 1924) which Yule introduced in 1924. Under the Yule model, no extinction occurs and each species has an exponential (rate $\lambda$) lifetime. The Yule model is often used as a null model, even though extinction clearly occurs in nature. But being a pure birth model with exponential waiting times, the Yule model is relatively simple to analyze which makes it attractive to use. For example, common procedures for estimating the time of undated divergence times in supertrees assume the Yule model (Bininda-Emonds et al., 2007; Purvis, 1995; Vos, 2006), even though extinction clearly occurred in the considered phylogenies.

Recently, a critical branching process as a neutral model for speciation was introduced (Aldous and Popovic, 2005; Popovic, 2004). In the critical branching process, each species has an exponential (rate $\lambda$) lifetime during which it produces offspring according to a Poisson (rate $\lambda$) process. After a species lifetime, it goes extinct (without further offspring). We condition the process to have $n$ extant—present day—species; this is the conditioned critical branching process (cCBP). Conditioning on $n$ extant species is crucial for data analysis, since a tree which we obtain from data has a fixed number of species. A lineage tree is the smallest subtree of the cCBP containing the $n$ extant species; see Fig. 1. Note that in other work, a lineage tree is also called a reconstructed process (Nee et al., 1994).

Both the Yule model and the critical branching process are special cases of a birth-death process. In a birth-death process, we again start with one individual. The extant individuals act independently from each other, and each individual has a birth rate $\lambda$ and a death rate $\mu$. So the Yule model is a birth-death process with $\mu = 0$ and the critical branching process is a birth-death process with $\lambda = \mu$. Note that for all birth-death processes where $\mu < \lambda$, the number of species is increasing exponentially fast. Only for $\mu = \lambda$, we do not have an exponential increase, which is biological more reasonable (Aldous, 2001).

In Nee et al. (1994), lineage trees which are obtained from birth-death processes are discussed. The authors investigate the lineage tree after a time $t$. Our approach differs since we investigate the lineage tree with $n$ extant species. The joint probability for the shape of the lineage tree and the times of all speciation events has been considered in Rannala and Yang (1996) for the general birth-death process and in Edwards (1970) for the Yule model. In Yang and Rannala (1997), the joint probability for the times of all speciation events in lineage trees has been considered, disregarding the shape. They all did not obtain the marginal probabilities for the times of speciation events given the tree shape.

We are interested in the time of the successive speciation events. For the two special cases of birth-death models, the Yule model and the cCBP, we calculate the time for the

Fig. 1 A complete tree (left) and its lineage tree (right).