

The sampled protracted speciation model in macro-evolution: Calculating its probability density and a comparison with two other models

Siyamak Kaffashi*

May 2012

Abstract

There have been several models suggested to model macro-evolution. Recently, it has been shown that in the birth-death process, diversified sampling is a better fit to empirical data (phylogenetic trees) than complete sampling and other sampling methods. This implies that taxon sampling (species sampling) has a large impact on estimating the diversification parameters. Besides, the protracted speciation model has been suggested to explain the observed slowdowns in the lineages through time plots. The protracted speciation model assumes that speciation events do not happen instantaneously, and it takes some time for a new born species (incipient species) to become a good species. The incipient lineages with dead good parent are also considered as good in the protracted speciation model. The likelihood functions have been derived for the diversified sampling and the random sampling models. Deriving the likelihood function for the protracted speciation model has faced some difficulties. In the present thesis, we try to calculate the probability density for the protracted speciation model by applying a method that has been suggested to calculate the probability of the Binary State Speciation Extinction (BiSSE) model. Unfortunately, we cannot provide an analytical closed form solution of the probability density function. Thus, we suggest the sampled protracted speciation model based on the sampling probability for good and incipient species, and explain the calculation of the probability density, using BiSSE approach. We compare the sampled protracted speciation model with the diversified and random sampling models, using simulated trees and empirical data. In this thesis, we show that the inference on the simulated trees is more accurate when the simulation and the inference methods are the same. We also show the effect of the transition (speciation completion) rate on the inferred parameters of the sampled protracted speciation model. The thesis also illustrates that the sampling probability of the incipient species does not affect the estimated parameters.

*Postal address: Mathematical Statistics, Stockholm University, SE-106 91, Sweden. E-mail: Siyamak.kaffashi@gmail.com. Supervisor: Sebastian Höhna.