

Mathematical Statistics Stockholm University Master Thesis **2015:3** http://www.math.su.se

## A phase-type distribution approach to coalescent theory

Tuure Antonangeli\*

## May 2015

## Abstract

A phase-type distribution is defined as the distribution of the time until a continuous time Markov chain with a finite state space reaches an absorbing state. These distributions are fully determined by the infinitesimal generator matrix of the Markov chain, together with a probability vector declaring the probability for the Markov chain to start in a specific state. In this paper, we propose a strategy for analyzing population genetics using phase-type distributions. In particular, we consider a continuous time coalescent process tracking the ancestry of a population sampled at present time, and apply the phase-type approach to determine the properties for such coalescent processes. These properties include the tree height and the total tree length, as well as mutation probabilities. We begin by applying this strategy to the Kingman's coalescent, and we thereafter extend the method to a population complication known as the symmetric island model. The advantage of this approach is that it leads to clear and intuitive derivations for the properties, and we show that the proposed strategy results in a very compact matrix analytic description of the coalescent.

<sup>\*</sup>Postal address: Mathematical Statistics, Stockholm University, SE-106 91, Sweden. E-mail: tuure.antonangeli@hotmail.com. Supervisor: Ola Hössjer.