

Inference of natural selection from allele frequency time series data using exact simulation techniques

Jaromir Sant.

University of Warwick.

12-th February 2021, 1pm UK time

Abstract

A standard problem in population genetics is to infer evolutionary and biological parameters such as the effective population size, mutation rates, and strength of natural selection from DNA samples extracted from a contemporary population. That all samples come only from the present-day has long been known to limit statistical inference; there is potentially more information available if one also has access to ancient DNA so that inference is based on a time-series of historical changes in allele frequencies. In this talk I will introduce a Markov Chain Monte Carlo method for Bayesian inference from allele frequency time-series data based on an underlying Wright-Fisher diffusion model of evolution. The chief novelty is that we show this method to be exact in the sense that it is possible to enable mixing by augmenting the state space with the unobserved diffusion trajectory, despite the fact that the transition function of the diffusion is intractable. We develop an efficient method in which trajectory updates and accept/reject probabilities can be calculated without error, and illustrate its performance on simulated data.

This is joint work with Paul Jenkins, Jere Koskela, and Dario Spano.