

Stephens, Matthew; Donnelly, Peter

Inference in molecular population genetics. (With discussion). (English) Zbl 0962.62107
J. R. Stat. Soc., Ser. B, Stat. Methodol. 62, No. 4, 605-655 (2000).

Summary: Full likelihood-based inference for modern population genetics data presents methodological and computational challenges. The problem is of considerable practical importance and has attracted recent attention, with the development of algorithms based on importance sampling (IS) and Markov chain Monte Carlo (MCMC) sampling. Here we introduce a new IS algorithm. The optimal proposal distribution for these problems can be characterized, and we exploit a detailed analysis of genealogical processes to develop a practicable approximation to it.

We compare the new method with existing algorithms on a variety of genetic examples. Our approach substantially outperforms existing IS algorithms, with efficiency typically improved by several orders of magnitude. The new method also compares favourably with existing MCMC methods in some problems, and less favourably in others, suggesting that both IS and MCMC methods have a continuing role to play in this area. We offer insights into the relative advantages of each approach, and we discuss diagnostics in the IS framework.

MSC:

- 62P10 Applications of statistics to biology and medical sciences; meta analysis
- 92D10 Genetics and epigenetics
- 65C40 Numerical analysis or methods applied to Markov chains

Cited in **5** Reviews
Cited in **49** Documents

Keywords:

ancestral inference; coalescent; computationally intensive inference; importance sampling; population genetics; Markov chain Monte Carlo

Full Text: [DOI](#)