MCMC Diagnosis for Bayesian Phylogenetic Inference

Tracy Heath

Department of Integrative Biology, University of California, Berkeley

2013 Workshop on Molecular Evolution
Český Krumlov, Czech Republic
Overview
Convergence and MCMC diagnosis
  • tools and metrics for evaluating MCMC samples
  • Tracer demo
  • AWTY demo

break

Tutorial — Conor Meehan and Instructors
Phylogenetic reconstruction in MrBayes
  • MCMC diagnostics in Tracer
  • Partitioned phylogenetic inference in MrBayes

beer
Andrey Markov

The Markov Chain

“An example of statistical investigation of the text Eugene Onegin concerning the connection of samples in chains.”

23* January 1913

*Julian calendar
Convergence

- Infinite MCMC samples will converge on the stationary/target distribution
- With a finite number of iterations, we can never be certain that our samples represent the underlying stationary distribution
The Markov chain reaches stationarity when the mean, variance, and autocorrelation structure do not change over time.
A chain has good “mixing time” if it rapidly samples the stationary distribution from an arbitrary starting state.

A chain is said to mix poorly if it does not reach stationarity and if the samples are highly correlated.
Burn-in refers to the common practice of discarding a percentage of initial states that were sampled before the chain reached stationarity.

This practice is not necessary for “adequately-long” chain lengths.

Average standard deviation of split frequencies

19929900  - (-96831.448) [-96812.848] (-96855.650) (-96850.506) * (-96814.997) ... - 0:18:41
19930000  - (-96826.655) [-96812.054] (-96853.963) (-96832.659) * [-96822.263] ... - 0:18:39

Average standard deviation of split frequencies: 0.002567

19930100  - (-96833.675) [-96806.706] (-96850.378) (-96842.572) * (-96824.515) ... - 0:18:38
19930200  - (-96827.767) [-96809.219] (-96838.541) (-96849.400) * (-96824.812) ... - 0:18:36

- Comparing the clade frequencies across multiple, independent runs
- ASDSF should approach 0.0 as the runs sample the same distribution
- MrBayes computes this statistic for multiple runs
ESS

The **Effective Sample Sizes** of the parameter values

This statistic gives you an idea of how independent your samples are and measures if you ran the chain sufficiently long.
Tracer

http://tree.bio.ed.ac.uk/software/tracer/