

Revision

Review of probability theory

Especially conditional probability & Bayes' Rule

Markov chains

Definition

Chapman-Kolmogorov equations (including proof)

Stationary frequencies (and how to calculate them)

Conditions for stationarity

Reversibility (definition, conditions, proof; detailed balance)

MCMC

sampling procedure

proof that this yields the specified equilibrium frequency vector
(via detailed balance)

know how and why MCMC is applied in Bayesian statistics

Gibbs sampling

- What is it and what is it useful for

- Proof of detailed balance

Continuous-time Markov processes

- Definition

- Chapman-Kolmogorov equations (and proof)

- Generator matrix (and relationship to transition probability matrix)

- Stationary distribution of continuous-time chain (and how to calculate it)

- Reversibility

Hidden Markov models

- Definition

- Know how to apply them

- Algorithms (viterbi, forward, backward)

Molecular phylogenetics

- Uses continuous-time Markov processes

- Scaling the generator matrix to units of genetic distance

- Maximum likelihood estimate of genetic distance between a pair of aligned sequences

- Felsenstein's pruning algorithm (calculate the likelihood of a tree)

- Basic ideas of phylogenetic inference, especially Bayesian phylogenetic inference

Sequence alignment

Needleman-Wunsch dynamic programming algorithm to find highest-scoring alignment

Probabilistic alignment using pair-HMMs

pair-HMM algorithms

Analysis of high throughput genotype data

Understand the nature of genotype data

The phasing problem

Clarke's algorithm

EM algorithm

method of PHASE software

Inferring population structure

method of STRUCTURE software