STATS 700-002 Class 7. Complex Population Dynamics and the Coalescent Under Neutrality

Aaron King and Edward Ionides

October 9, 2025

Outline

Volz, E. M. (2012) Complex population dynamics and the coalescent under neutrality. Genetics **190**: 187–201. doi:10.1534/genetics.111.134627

- The first approach to phylodynamic likelihood for a compartment model with a structured population
- Supposes that the model dynamics are determined by a system of differential equations

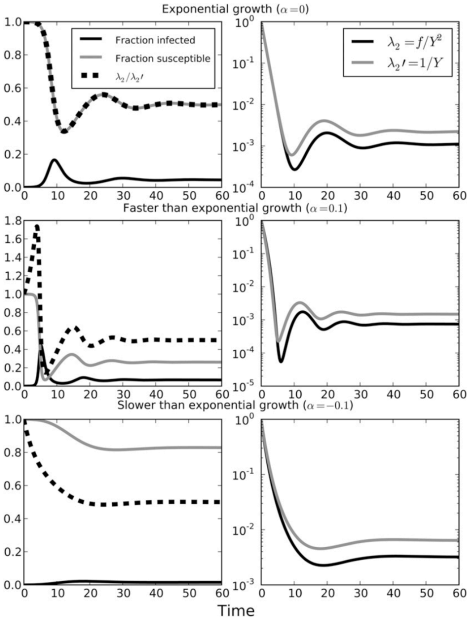


Figure 1 (Left) The fraction of the population susceptible and infected is shown over time for model (16). (Right) The rates of coalescence $\lambda_2 = f/Y^2$ and $\lambda_2' = 1/Y$. In all solutions to Equation 16, $N = 10^4$, $\beta = 2$, $\gamma = 1$, $\eta = \frac{1}{10}$. The incidence scaling factor α is varied for each row: $\alpha = 0$ (top), $\alpha = \frac{1}{10}$ (middle), and $\alpha = -\frac{1}{10}$ (bottom).

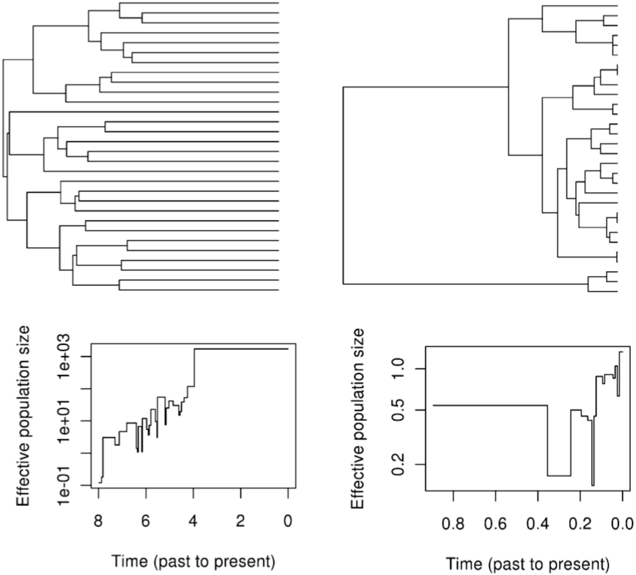


Figure 2 Simulated genealogies (top) and corresponding skyline estimates of N_e (bottom) for exponential growth (left) and FTE growth (right). Simulations were of a pure-birth process with monotonically increasing population sizes. Samples of 30 taxa were taken during a period of growth (either exponential or FTE) at the point when a population size of $Y = 2 \times 10^4$ was reached. In the exponential case, the skyline is unbiased for the harmonic mean of $Y/2\beta$ within each interval. In the FTE case, the skyline underestimates population size.

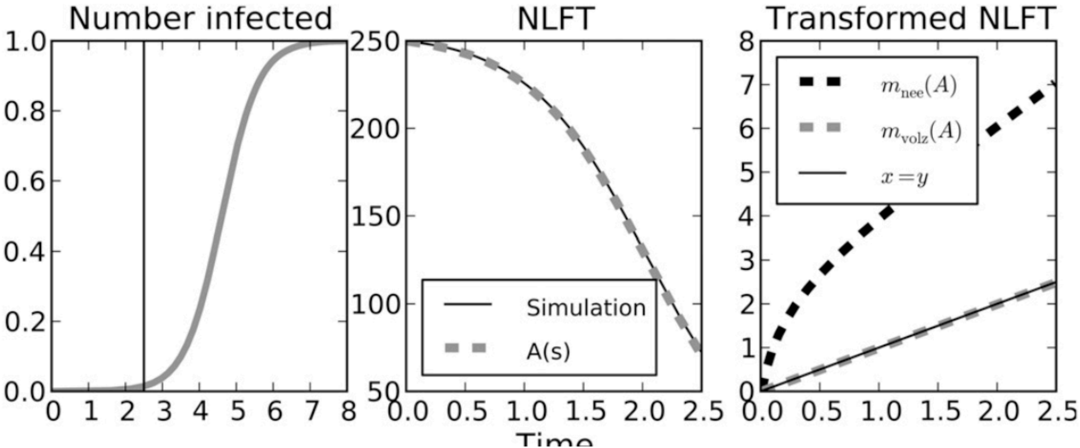


Fig 3. $n = 250$ samples at $t = 2.5$.

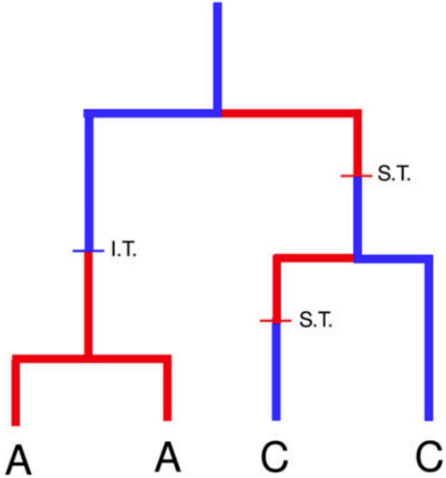
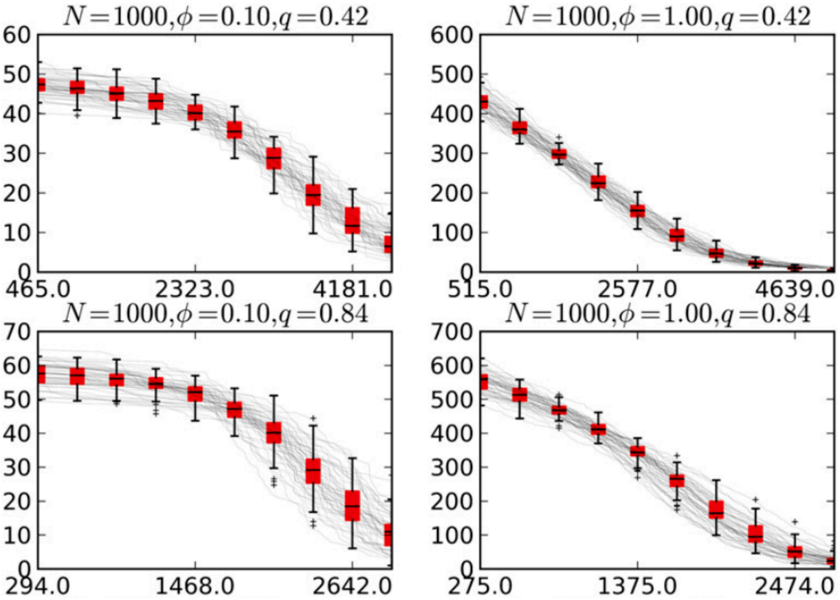
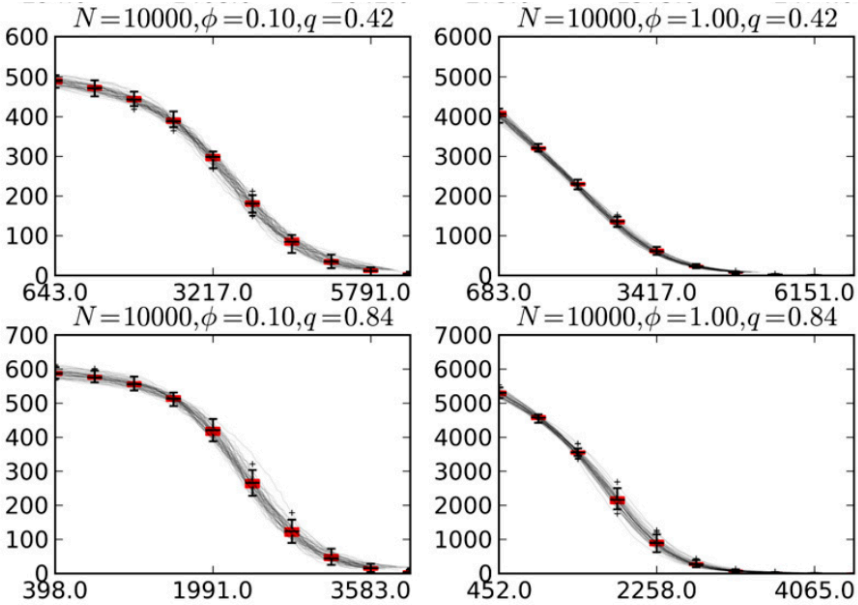


Figure 4 An example gene genealogy that could be generated by the HIV model (Equation 29). Red branches correspond to stage-1 infected hosts. Blue branches correspond to stage 2.



N=1000,ϕ=0.10,q=0.42
N=1000,ϕ=1.00,q=0.42
N=1000,ϕ=0.10,q=0.84
N=1000,ϕ=1.00,q=0.84

NLFT (Fig 5)



$N=10000, \phi=0.10, q=0.42$
600
500
400
300
200
100
0
643.0
3217.0
5791.0
$N=10000, \phi=1.00, q=0.42$
6000
5000
4000
3000
2000
1000
0
683.0
3417.0
6151.0
$N=10000, \phi=0.10, q=0.84$
700
600
500
400
300
200
100
0
398.0
1991.0
3583.0
$N=10000, \phi=1.00, q=0.84$
7000
6000
5000
4000
3000
2000
1000
0
452.0
2258.0
4065.0

NLFT (Fig 5)

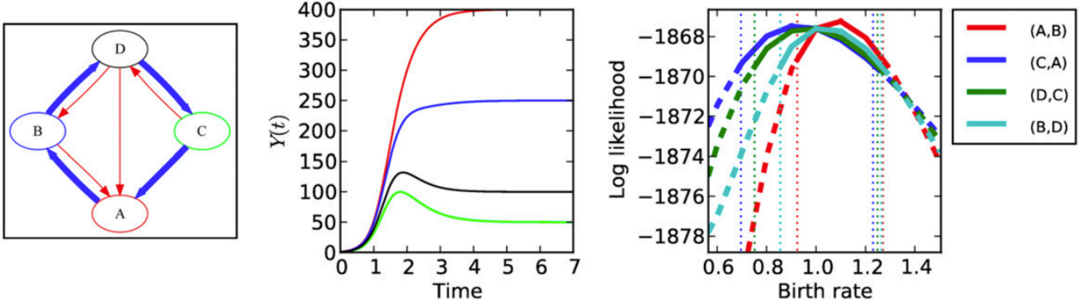


Fig 6. (Left) Model with m = 5 states, four birth terms, and seven migration terms. Blue arrows are logistic birth terms. Red arrows are migration. (Center) The population size Y_k over time for each of 5 states. (Right) Likelihood profile of four (relative) birth rates and 95% CIs.

Branching process approximations

- Why does Volz describe his method as a branching process approximation?
- How is the branching process approximation related to the assumption of a large population with a low sampling fraction?
- How would you assess the inaccuracy incurred by the branching process approximation in a particular application?

Deterministic population dynamics

- What are the benefits and weaknesses for data analysis of making an assumption of deterministic population dynamics?
  - This is a question about the population model, not its relationship to phylodynamic data.
- Is a branching process approximation to the phylodynamic model more suitable in a deterministic or stochastic population model, or are those decisions separate?

The Riccati equation

- How do you solve Eq. (20)?
- Is the proposed solution in Eq. (22) correct?