

Statistical population genetics

Lecture 3: The Moran model

Xavier Didelot

Dept of Statistics, Univ of Oxford

didelot@stats.ox.ac.uk

The Moran model

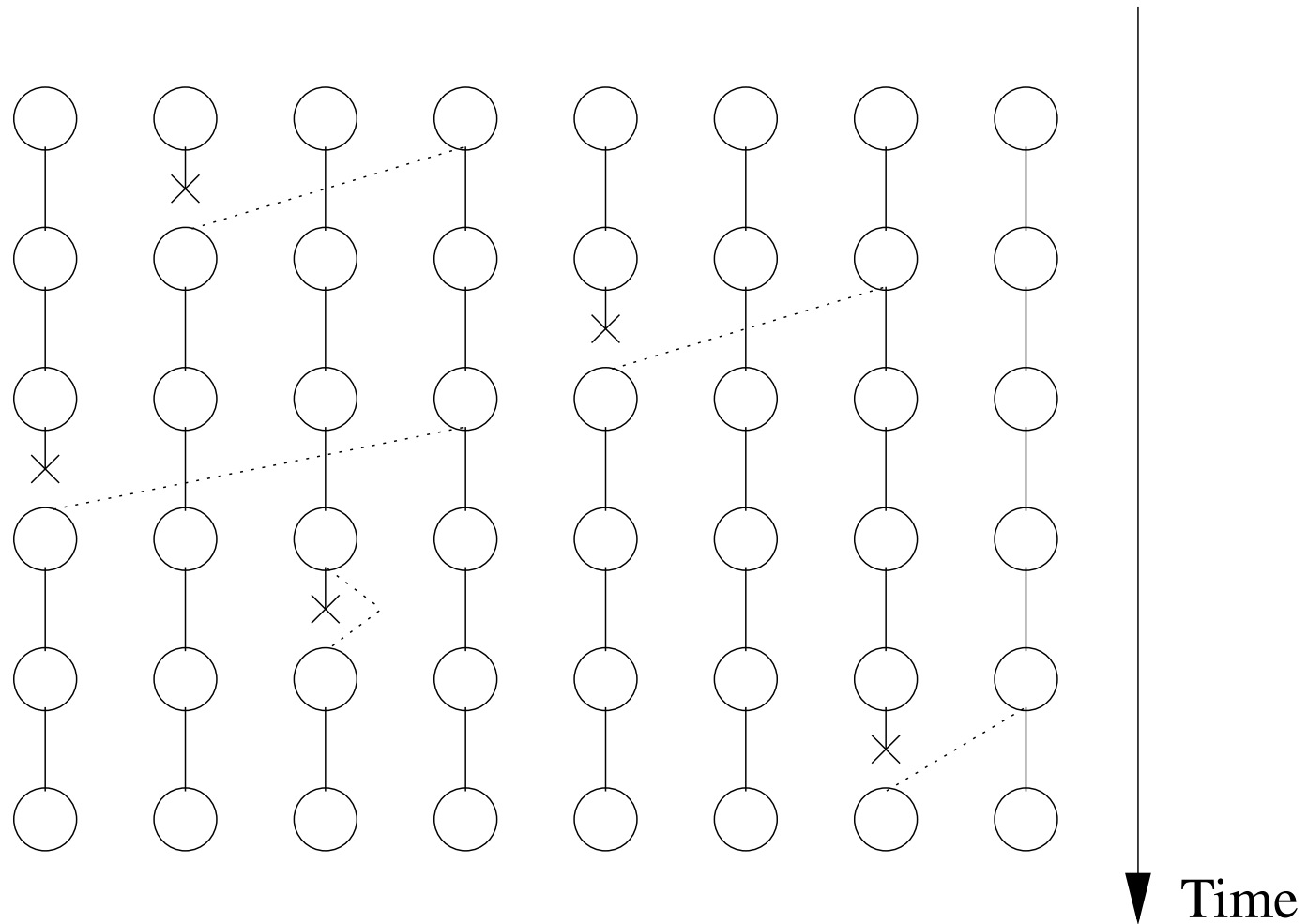
This model due to Moran (1958) is less popular than the WF model amongst biologists, but represents a mathematically attractive alternative.

Definition (The Moran model).

A population is said to evolve under the Moran model if:

- *The population has a **constant size** M ;*
- *Generations are **allowed to overlap**;*
- *At discrete time intervals, two members a and b of the population are chosen at random (a and b can be the same). a **reproduces** and b **dies**.*

The Moran model



Biallelic Moran model

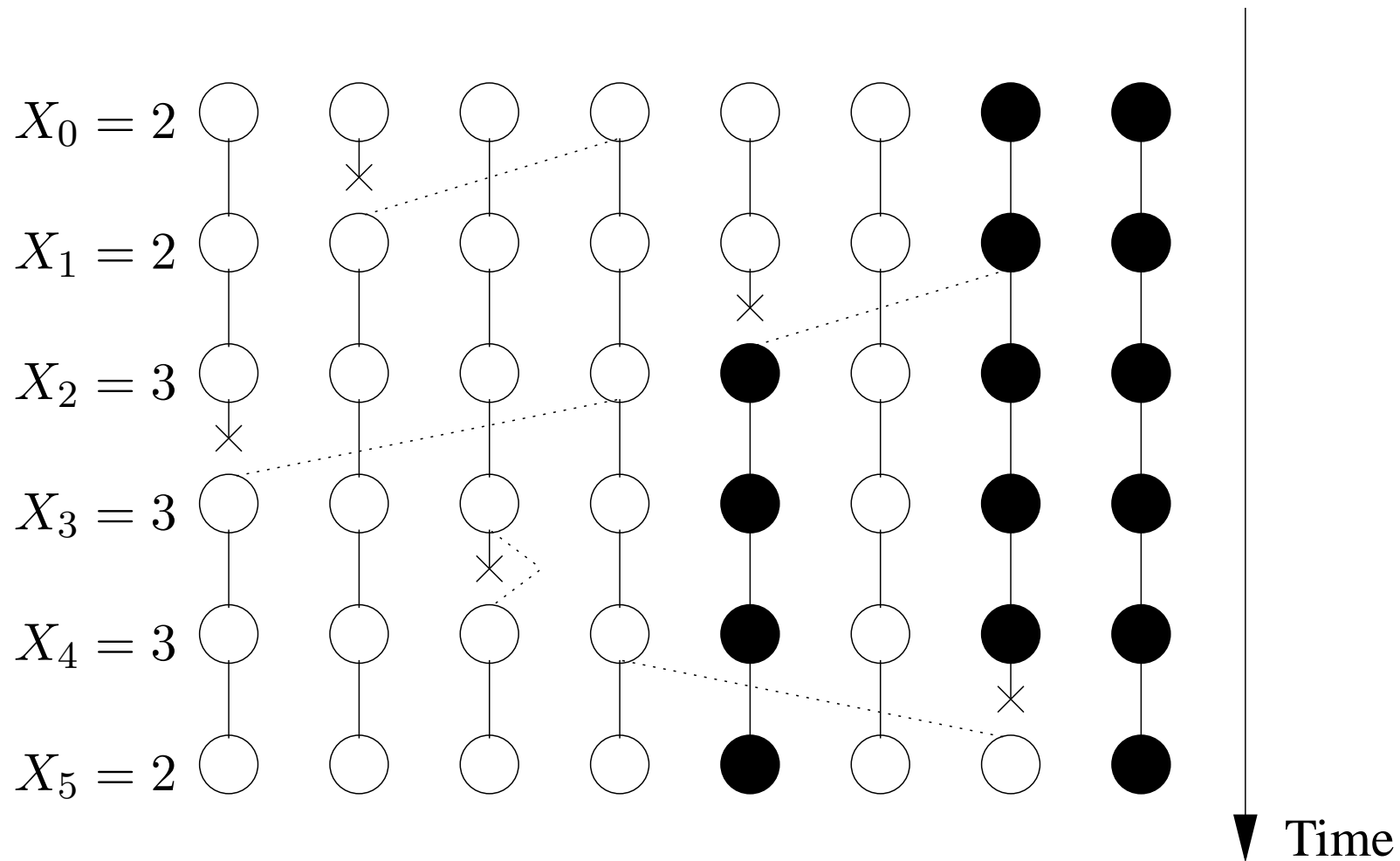
The biallelic Moran model can be defined as a **Markov chain**. If we consider (as we did in the WF model) that there are two alleles A and a in the population and that X_t denotes the number of copies of A at time t , then:

$$\mathbb{P}(X_{t+1} = j | X_t = i) = P_{i,j} = \begin{cases} \frac{i(M-i)}{M^2} & \text{if } j = i \pm 1 \\ \frac{i^2}{M^2} + \frac{(M-i)^2}{M^2} & j = i \\ 0 & \text{otherwise.} \end{cases}$$

It follows that:

$$\mathbb{E}(X_{t+1} | X_t) = X_t \text{ and } \text{var}(X_{t+1} | X_t) = 2 \frac{X_t}{M} \left(1 - \frac{X_t}{M} \right)$$

Biallelic Moran model



Heterozygosity

Theorem (Heterozygosity in the Moran model).

Under the biallelic Moran model, the expected heterozygosity decays approximately at rate $2/M^2$ when M is large.

Recall the definition of heterozygosity in a biallelic model:

$$H_t = 2 \frac{X_t}{M} \left(1 - \frac{X_t}{M} \right)$$

Heterozygosity

Proof.

$$\begin{aligned}\mathbb{E}(H_{t+1}) &= \frac{2}{M} \mathbb{E} \left(X_{t+1} \left(1 - \frac{X_{t+1}}{M} \right) \right) \\ &= \frac{2}{M} \left\{ \mathbb{E}(X_{t+1}) - \frac{\mathbb{E}(X_{t+1}^2)}{M} \right\} \\ &= \frac{2}{M} \left\{ \mathbb{E}(X_{t+1}) - \frac{\text{var}(X_{t+1}) + \mathbb{E}(X_{t+1})^2}{M} \right\} \\ &= \frac{2}{M} \left\{ X_t - 2\frac{X_t}{M} + 2\frac{X_t^2}{M^2} + \frac{X_t^2}{M} \right\} \\ &= H_t \left(1 - \frac{2}{M^2} \right)\end{aligned}$$

So by induction we find:

$$\mathbb{E}(H_t) = H_0 \left(1 - \frac{2}{M^2} \right)^t \approx H_0 \exp(-2t/M^2)$$

Fixation

As in the WF model, 0 and M are absorbing states, and we have:

Theorem (Probability of fixation).

The probability that A will reach fixation rather than a given its initial frequency is equal to its initial frequency.

Fixation

Proof. As for the WF model, this is a direct consequence of the fact that the expectation remains the same since $P_{i,i-1} = P_{i,i+1}$.

A more formal (and expendable) approach is to let q_i denote the probability of fixation of A given $X_t = i$ and notice that:

$$q_i = P_{i,i+1}q_{i+1} + P_{i,i}q_i + P_{i,i-1}q_{i-1}$$

This simplifies to:

$$2q_i = q_{i+1} + q_{i-1}$$

which can be solved with boundary conditions $q_0 = 0$ and $q_M = 1$ to find the result $q_i = i/M$. □

Time before fixation

Theorem (Time before fixation in the Moran model).

Let τ_i be the expected time before fixation given that $X_0 = i$. Then:

$$\tau_i = M \left(\sum_{a=1}^i \frac{M-i}{M-a} + \sum_{a=i+1}^{M-1} \frac{i}{a} \right)$$

Time before fixation

Proof. Let τ_i be the expected time before fixation given that $X_0 = i$. The same reasoning as in the proof for the WF model gives:

$$\tau_0 = 0, \tau_M = 0, \tau_i = 1 + \sum_{j=0}^M P_{i,j} \tau_j$$

We have:

$$\tau_i = 1 + P_{i,i-1} \tau_{i-1} + P_{i,i} \tau_i + P_{i,i+1} \tau_{i+1}$$

$$\tau_i = 1 + \frac{i(M-i)}{M^2} \tau_{i-1} + \frac{i^2 + (M-i)^2}{M^2} \tau_i + \frac{i(M-i)}{M^2} \tau_{i+1}$$

Time before fixation

$$\tau_{i-1} - 2\tau_i + \tau_{i+1} = -\frac{M^2}{i(M-i)}$$

Using the boundary conditions $\tau_0 = 0$ and $\tau_M = 0$ we can solve this to get the result:

$$\tau_i = M \left(\sum_{a=1}^i \frac{M-i}{M-a} + \sum_{a=i+1}^{M-1} \frac{i}{a} \right)$$

□

Remark. If M is large, we can approximate τ_i by the continuous function $\tau(p)$ with $p = i/M$ and we get:

$$\tau(p) \approx -M^2(p \log(p) + (1-p) \log(1-p))$$

Note that this is the same as the fixation time in the WF model if we consider that $M/2$ Moran time units correspond to a WF generation. The heterozygosity scales in the same way.

Summary

- Like the **Wright-Fisher model**, the **Moran model** is a model of **genetic drift**.
- Exact calculations under the **WF model with mutations** are often impossible.
- Exact calculations under the Moran model are often possible, but the results are usually so complex that the **effects** of the different evolutionary parameters are obscured.
- **Diffusion approximation** is a solution to study the **forward in time** properties of genetic drift and other forces.
- Here we will take another approach: look **backward in time** to explain the data currently observed.