Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

D-BSSE
Department of Biosystems
Science and Engineering

Stochastic models of finite populations

Niko Beerenwinkel

Let's gamble!

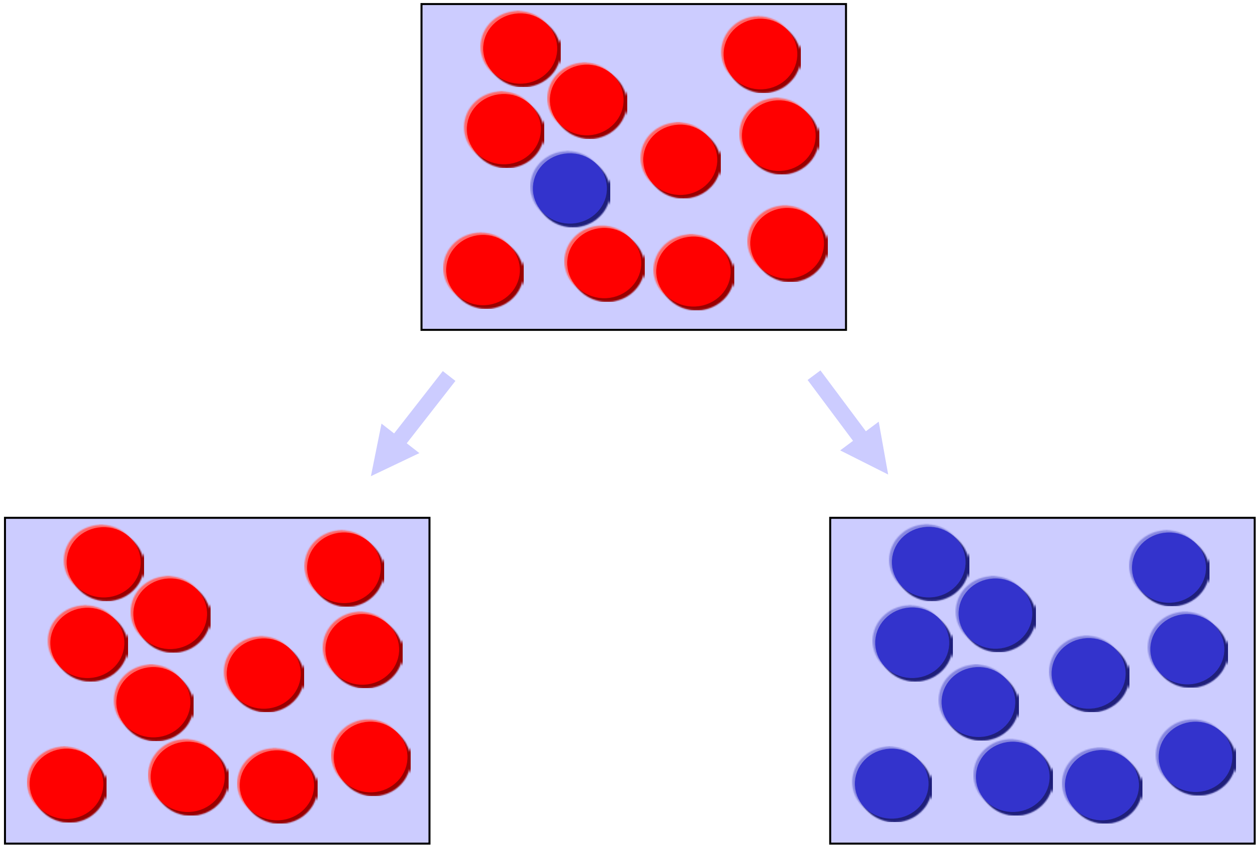
- You play against the bank.
- In each round, a **fair** coin is flipped and the loser pays 1 SFr to the winner.
- The game ends when one party has nothing left.
- Are you willing to play?

Let’s gamble!

- You play against the bank.
- In each round, a **fair** coin is flipped and the loser pays 1 SFr to the winner.
- The game ends when one party has nothing left.
- Are you willing to play?

- Your fate is well-known in probability theory as the *gambler’s ruin*.
- The reason is that *a random walk goes somewhere*.

Finite populations



Outline

- Some basic probability
- Markov chains

- Moran process
- Birth-death process
- Fixation probability
- Mean fixation time
- Moran process with selection

Conditional probabilities

- Let X and Y be (discrete) random variables with probability distributions P(X) and P(Y).
- The joint probability of X and Y is denoted P(X,Y).
- The *conditional probability* of X given Y is

$$P(X \mid Y) = \frac{P(X, Y)}{P(Y)}$$

Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

D-BSSE
Department of Biosystems
Science and Engineering

Bayes' theorem

- X and Y are independent, if $P(Y \mid X) = P(Y)$.
- Bayes' theorem states that

$$P(Y \mid X) = \frac{P(X \mid Y) P(Y)}{P(X)}$$

$P(Y \mid X)$ is the posterior probability,, $P(Y)$ is the prior probability

- If $y_1, \ldots, y_n$ are disjoint outcomes of Y, then for any r.v. X, we can write $P(X) = \sum_{i=1,\ldots,n} P(X \mid Y = y_i)\, P(Y = y_i)$ and hence

$$P(Y \mid X) = \frac{P(X \mid Y) P(Y)}{\sum_i P(X \mid y_i) P(y_i)}$$

The law of total probability

- "The prior probability is equal to the expected value of the posterior probability"

$$
\begin{aligned}
P(X = x) &= \sum_y P(X = x, Y = y) \\
&= \sum_y P(X = x \mid Y = y)\, P(Y = y) \\
&= E_Y[P(X = x \mid Y)]
\end{aligned}
$$

so $P(X) = E_Y[P(X \mid Y)]$ for any r.v. Y.

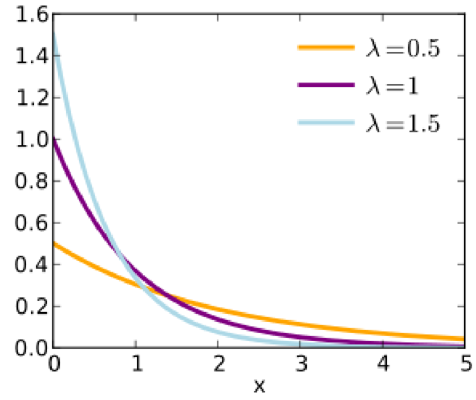
ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

D-BSSE
Department of Biosystems
Science and Engineering

The exponential distribution

- A continuous random variable X is *exponentially distributed with parameter* $\lambda > 0$ if its density function is

$$f(x) = \lambda e^{-\lambda x}, \qquad x \geq 0$$



- The cumulative and tail probabilities are

$$P(X \leq x) = \int_0^x \lambda e^{-\lambda y} dy = 1 - e^{-\lambda x}$$

$$P(X > x) = 1 - P(X \leq x) = e^{-\lambda x}$$

Moments of the exponential distribution

- Expectation

$$E(X) = \int_0^\infty x f(x)\, dx = \frac{1}{\lambda}$$

- Variance

$$V(X) = E(X^2) - E(X)^2 = \frac{1}{\lambda^2}$$

Memoryless property

- A random variable X is *memoryless* if for all s, t > 0,

$$P(X > s + t \mid X > t) = P(X > s)$$

- If X is a failure time, it means that the chance to fail in the next moment is always the same, no matter when.
- Because (X > s + t, X > t) is equivalent to just (X > s + t), the memoryless property is equivalent to

$$P(X > s + t) = P(X > t)\, P(X > s)$$

- The exponential distribution is memoryless:

$$\begin{aligned} P(X > s + t) &= \exp\{-\lambda(t + s)\} \\ &= \exp(-\lambda t)\exp(-\lambda s) \\ &= P(X > t)\, P(X > s) \end{aligned}$$

Competing exponentials

- We write X ~ Exp(λ) if X is an exponential random variable with rate λ.
- Consider X ~ Exp(λ) and Y ~ Exp(μ).
- Assume that X and Y are independent, i.e., P(X | Y) = P(X). Then:

$$\min(X, Y) \sim \text{Exp}(\lambda + \mu)$$

and

$$P(X < Y) = \frac{\lambda}{\lambda + \mu} \qquad \text{(competing exponentials)}$$

Markov chains

- A *stochastic process* is an indexed collection of random variables $\{X(t) \mid t \in T\}$ with common state space $\mathcal{S}$.
- $X(t)$ is the state of the process at time t.
- Stochastic processes can be discrete or continuous in both time and state space.

- A discrete-time *Markov chain* is a stochastic process $\{X(t)\}$ with $T = \{0, 1, 2, \ldots\}$, in which each next state only depends on the current state, that is

$$P(X(t+1) \mid X(0), \ldots, X(t)) = P(X(t+1) \mid X(t))$$

Transition matrix

- The transition matrix P of a Markov chain $\{X(t)\}$ is defined by $P_{ij}(t) = P(X(t + 1) = j \mid X(t) = i)$.
- The Markov chain is time-homogeneous if P_{ij} does not depend on t for all i and j.
- A state x* is an *absorbing* state if $X(t) = x^*$ for all $t \geq t_0$.

Ergodicity

- A Markov chain is **ergodic** if it is
  1) aperiodic (return to any state is always possible),
  2) irreducible (any state is accessible from any other), and
  3) positive recurrent (any state will eventually be reached with probability 1 and the mean recurrence time is finite).

- An ergodic Markov chain has a unique stationary distribution $\Pi = (\pi_i)_{i \in \mathcal{S}}$ such that

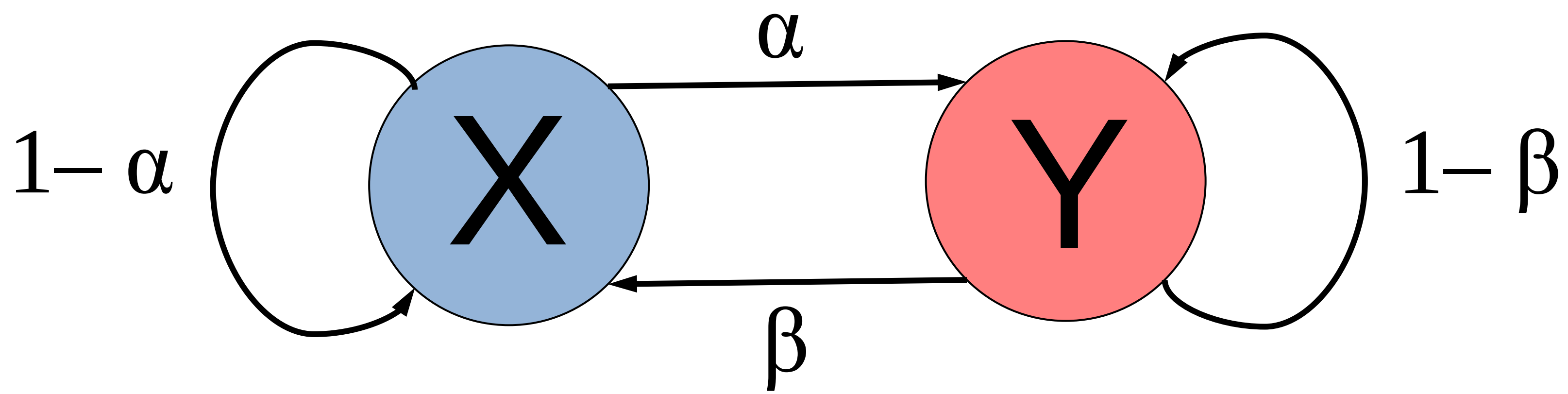
$$P_{ij}(t) \rightarrow \pi_j \quad \text{as} \quad t \rightarrow \infty$$

for all $i, j \in \mathcal{S}$, and

$$\Pi' P = \Pi'$$

where Π' denotes the transpose of Π.

Example of a two-state Markov chain

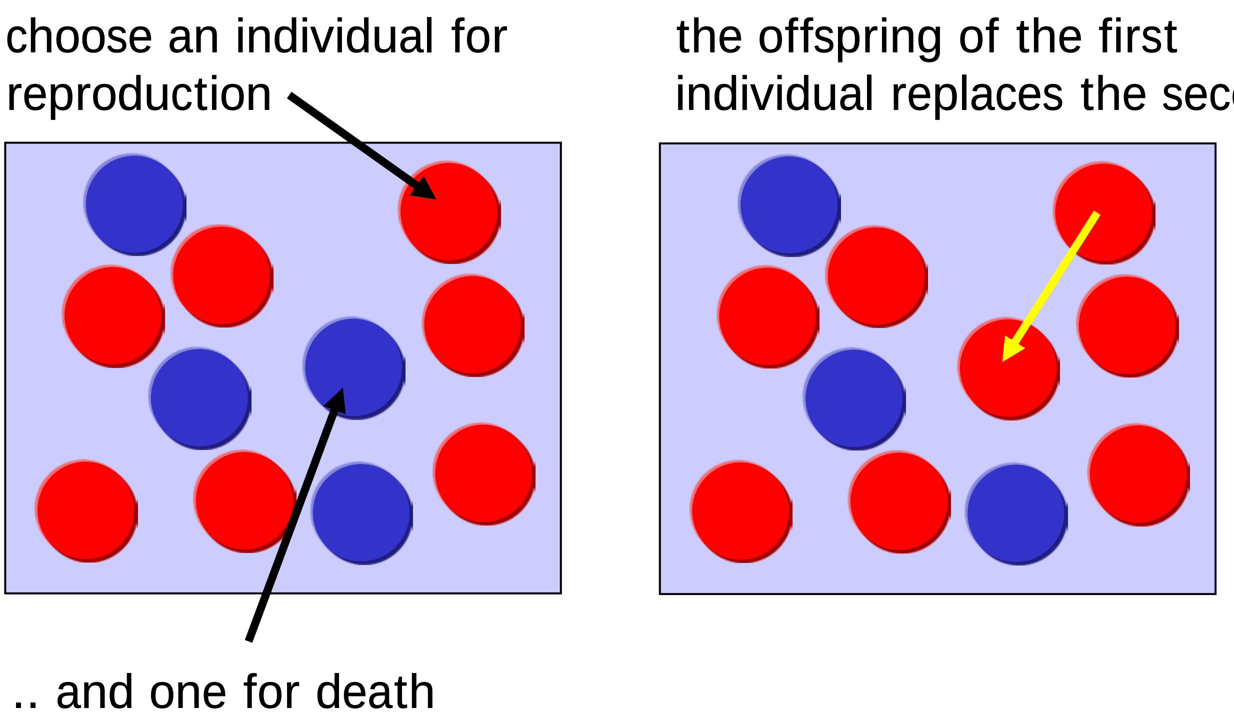


$$P = \begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix}$$

$$\Pi' = \Pi' P \iff \pi_X = \frac{\beta}{\alpha+\beta}, \quad \pi_Y = \frac{\alpha}{\alpha+\beta}$$

The Moran process

- We consider a finite population of constant size N with individuals of two types, A and B.



Patrick Alfred Pierce Moran (1917-1988)

The Moran process defines a Markov chain

- The state space is i = 0, …, N, the number of A individuals.
- Let p = i / N be the allele frequency of A.
- The transition matrix is given by

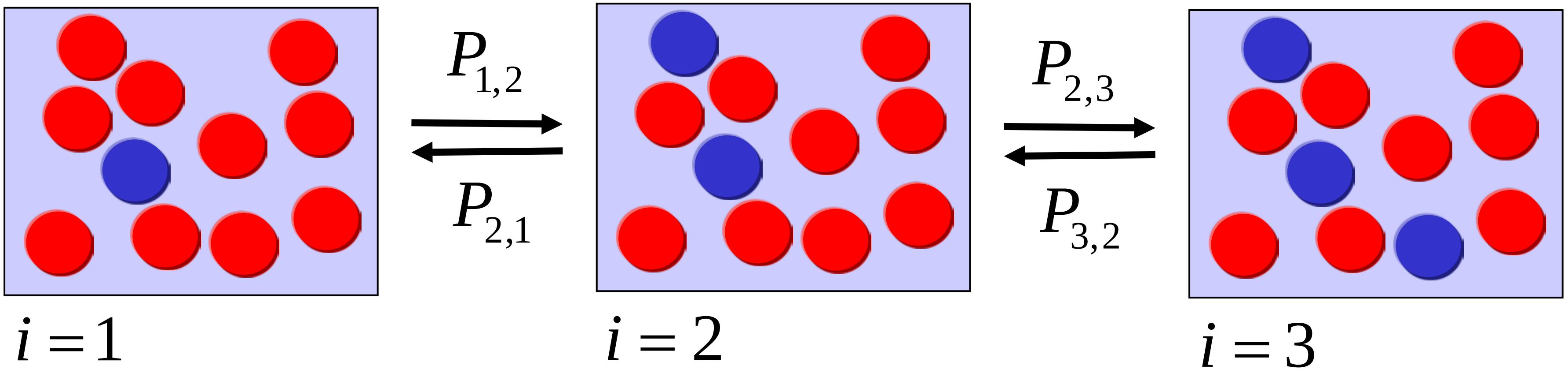
$$\begin{aligned} P_{i,i+1} &= p(1-p) \\ P_{i,i-1} &= (1-p)p \\ P_{i,i} &= p^2 + (1-p)^2 \end{aligned}$$

All other entries are zero. P is a tri-diagonal matrix.

- Both types have the same probability of reproduction and death. The changes in allele frequency are only due to random fluctuations, a phenomenon called *neutral drift*.

The Moran process is a birth-death process

- Because P is tri-diagonal, the number of A individuals can change only by one in each step. A stochastic process with this property is called a *birth-death process.*



Absorbing states

- For the Moran process, we have
  - $P_{0,0} = 1$ and $P_{0,i} = 0$ for all $i > 0$
  - $P_{N,N} = 1$ and $P_{N,i} = 0$ for all $i < N$

There are two absorbing states: all-red and all-blue

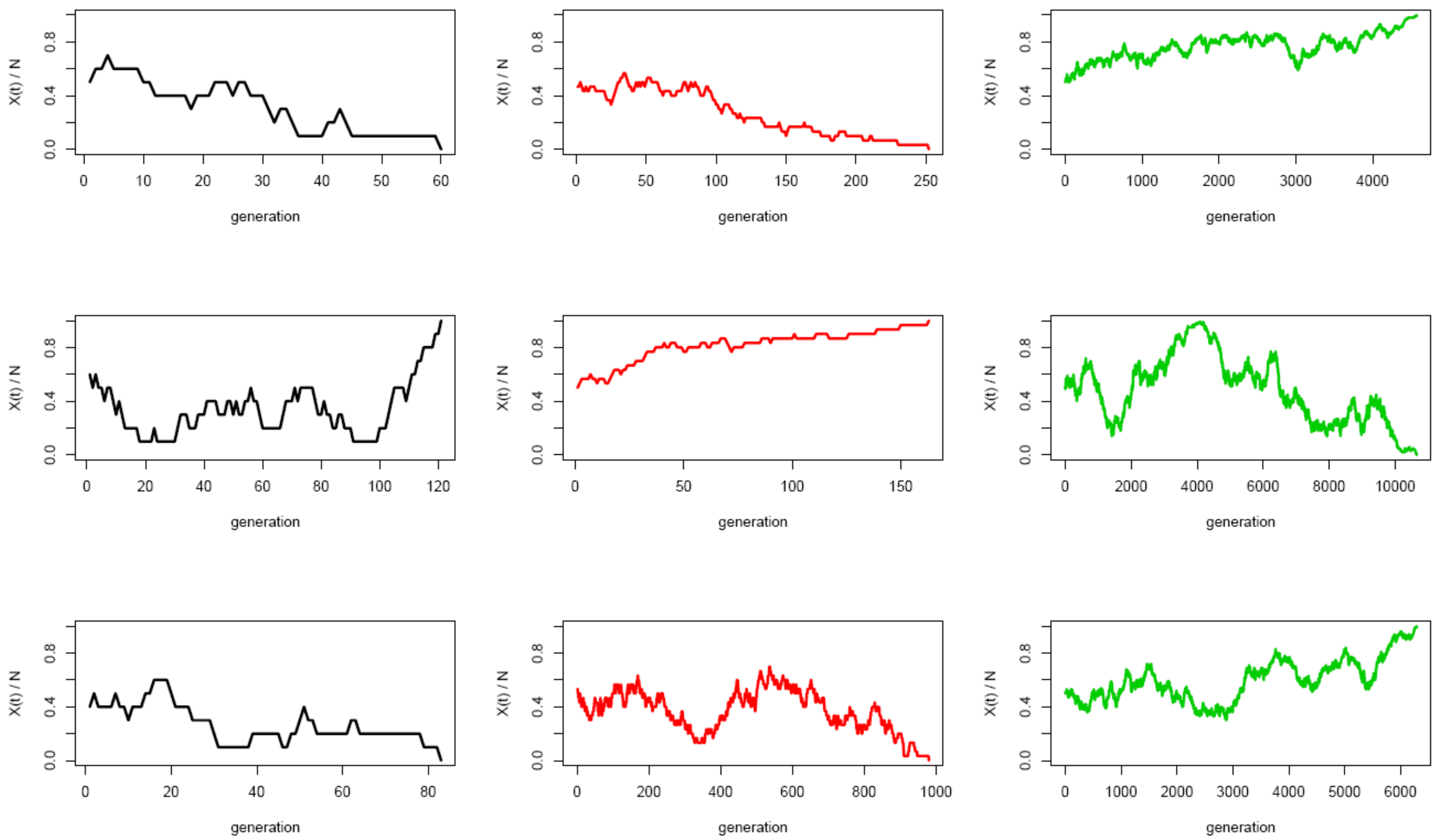
$i = 0$ $\qquad\qquad$ $i = N$

Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

D-BSSE
Department of Biosystems
Science and Engineering

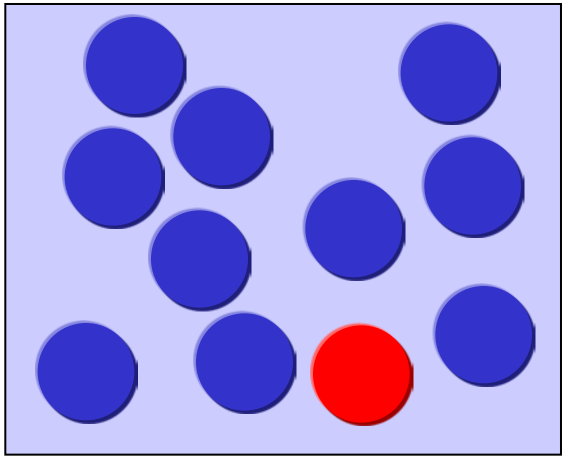
Dynamics

N = 10 N = 30 N = 100



Fixation probabilities

- Let x_i be the probability of ending up in state N when starting from state i.
- Clearly, $x_i = i / N$ for all $i = 0, \ldots, N$, because each allele has the same chance of being fixated.



Absorption probabilities in a birth-death process

- We consider a more general birth-death process with transition probabilities $P_{i,i+1} = \alpha_i$ and $P_{i,i-1} = \beta_i$.
- Assume that 0 and N are absorbing states, $\alpha_0 = \beta_N = 0$.
- Set $\gamma_i = \beta_i / \alpha_i$. Then:

$$x_i = \frac{1 + \sum_{j=1}^{i-1} \prod_{k=1}^{j} \gamma_k}{1 + \sum_{j=1}^{N-1} \prod_{k=1}^{j} \gamma_k}$$

is the probability of ending in state N (all-A) when starting in state i.

Mean fixation time

- In the Moran process, for large population sizes, the mean fixation time is

$$-N^2[(1-p)\log(1-p) + p\log p]$$

  generations (steps consisting of one reproduction and one death).
- The diversity (or *heterozygosity*) of the population

$$H(t) = 2\,(X(t)/N)\,(1 - X(t)/N)$$

  decays approximately exponentially at rate $2 / N^2$.
- This rate quantifies the amount of random genetic drift that the population is experiencing.

Moran process with constant selection

- Consider exponentially distributed waiting times to the reproduction of a type A and type B individual with rates λ_A = r and λ_B = 1, respectively.
  - If r > 1, then A has a fitness advantage over B.
  - If r = 1, we have the neutral process again.
- The waiting times to the next birth are
  - $T_A \sim \min\{\mathrm{Exp}(\lambda_A), \ldots, \mathrm{Exp}(\lambda_A)\} = \mathrm{Exp}(i\lambda_A)$
  - $T_B \sim \mathrm{Exp}((N - i)\lambda_B)$.
- T_A and T_B are competing exponentials:

$$P(T_A < T_B) = \frac{ri}{ri + (N - i)}$$

$$P(T_A > T_B) = \frac{N - i}{ri + (N - i)}$$

Transition probabilities

$$P_{i,i+1} = \frac{ri}{ri + N - i} \frac{N - i}{N}$$

$$P_{i,i-1} = \frac{N - i}{ri + N - i} \frac{i}{N}$$

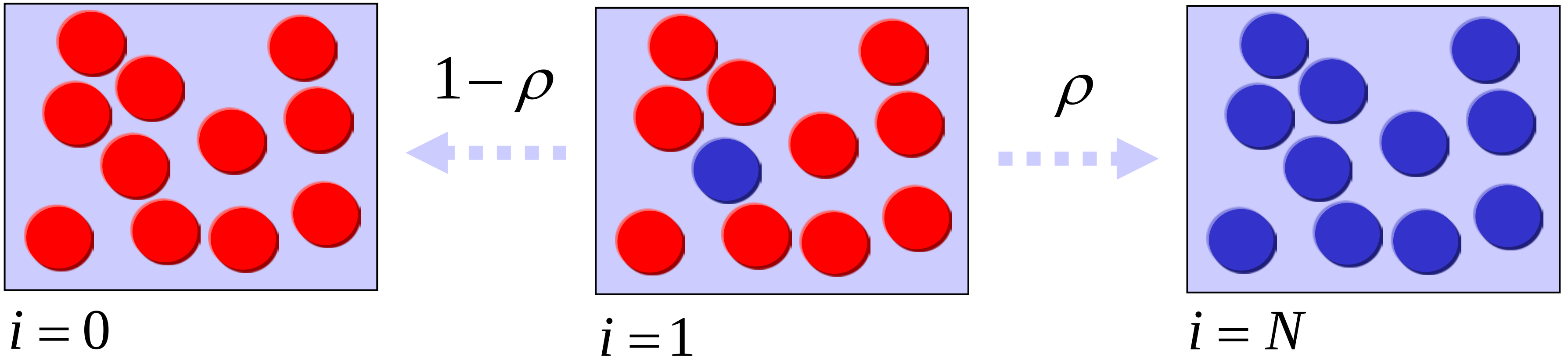
$$P_{i,i} = 1 - P_{i,i+1} - P_{i,i-1}$$

Fixation probabilities

- Because $\gamma_i = P_{i,i-1} / P_{i,i+1} = 1/r$, we find the absorption probabilities, or *fixation probabilities*

$$x_i = \frac{1 - 1/r^i}{1 - 1/r^N}$$

$\rho = x_1$

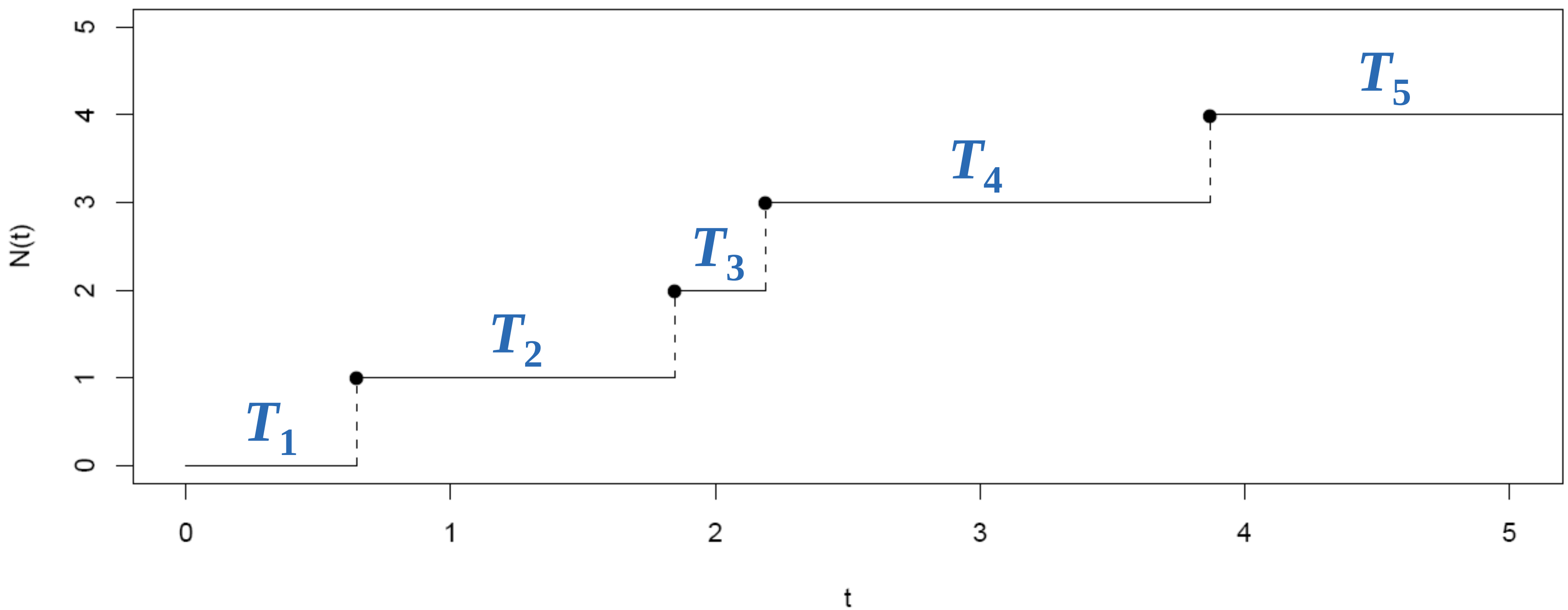


Poisson process

- A Poisson process is a stochastic counting process:
- A Poisson process is a continuous-time Markov chain with independent Poisson distributions in each interval.
- More precisely, $\{N(t) \mid t \geq 0\}$ is a Poisson process if
  - $N(0) = 0$
  - The number of events in an interval depends only on the length of the interval, and the number of events in disjoint intervals are independent.
  - The number of events in each interval of length t is Poisson distributed with mean λt,

$$P(N(t+s) - N(s) = k) \;=\; e^{-\lambda t} \frac{(\lambda t)^k}{k!}$$

Inter-arrival times



Inter-arrival times of a Poisson process are exponential

- Let $\{T_n \mid n = 1, 2, \ldots\}$ be the inter-arrival times.
- $T_1 \sim \text{Exp}(\lambda)$, because

$$P(T_1 > t) = P(N(t) = 0) = e^{-\lambda t}$$

- By the law of total probability,

$$\begin{aligned} P(T_2 > t) &= E_{T_1}[P(T_2 > t) \mid T_1)] \\ &= \int_s P[N(s+t) = N(s) \mid T_1 = s] f_{T_1}(s) ds \\ &= \int_s P(N(t) = 0) f_{T_1}(s) ds \\ &= e^{-\lambda t} \end{aligned}$$

ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

D-BSSE
Department of Biosystems
Science and Engineering

The rate of evolution

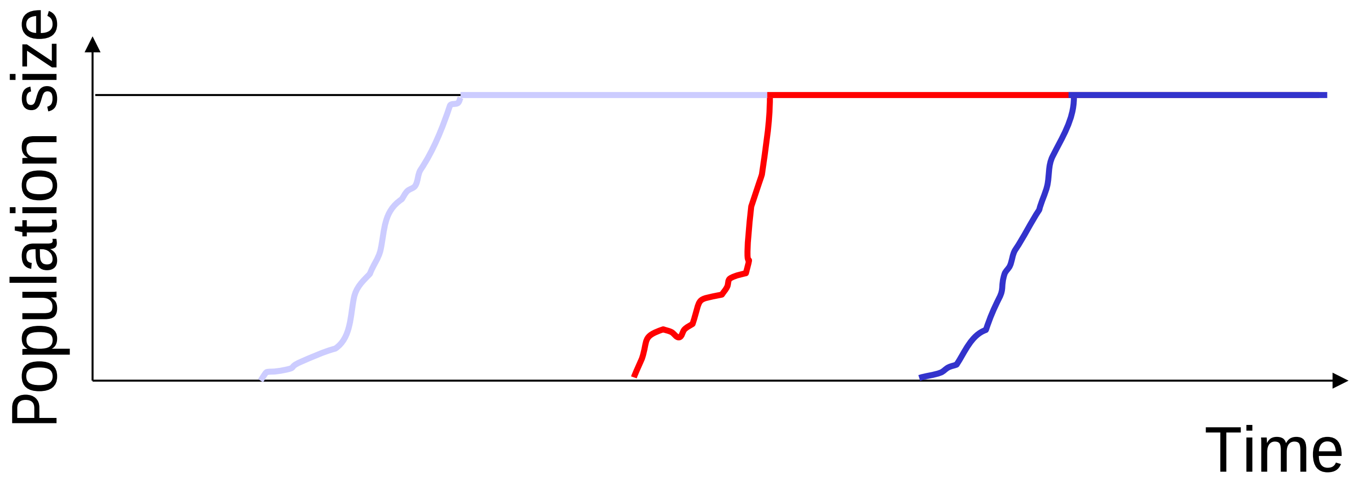
- Consider an all-A population where a B mutant occurs rarely at mutation rate u.
- The Poisson process is a good model for counting the mutations. In particular, $T_1 \sim \mathrm{Exp}(Nu)$.
- Suppose that type B has a selective advantage r. Then the fixation probability is $\rho = x_1$.
- The rate of evolution from all-A to all-B is

$$R = N\,u\,\rho$$

- If B is neutral, then $\rho = 1/N$ and $R = u$, the mutation rate.

The molecular clock of neutral evolution

- If u is constant, then neutral mutations accumulate at a constant rate R = u, independent of population size.



- *The Neutral Theory of Molecular Evolution*, Motoo Kimura, 1993.

Summary

- The Moran process is a birth-death process, an integer-valued Markov chain that changes by at most 1 in each step.
- The Moran process with two types has two absorbing states: fixation and extinction.
- In the Moran process, we can calculate analytically the fixation probability of a neutral and of a selectively advantageous mutant.
- In the neutral case, we can also determine the time scale of this process.

- Exercises: #7, #8, #9, #10