Evolution: the change in the frequency of types from one generation to the next. Biological evolution is the change in allele frequencies within a gene pool.

Exponential growth (Malthusian law)

Discrete model

Consider dividing cells in discrete generations t and let x_t be the number of cells in generation t. The difference equation $x_{t+1} = 2x_t$ has the solution $x_t = x_0 2^t$

Continuous model

Consider continuous time t and let x(t) be the continuous number (fraction) of cells at time t. Assume that **the generation time** T is exponentially distributed with average $\frac{1}{r}$, i.e., $Prob(T \leq \tau) = 1 - e^{-r\tau}$. The differential equation $x' = \frac{dx}{dt} = rx$ has the solution $x(t) = x_0\, e^{rt}$, where r is the rate of cell division

Logistic growth

- Suppose that the population has a *carrying capacity*, K.
- The logistic equation x' = rx(1 – x/K) has the solution

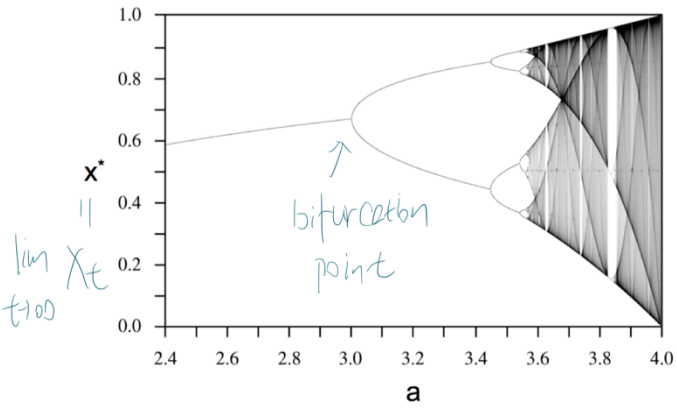
$$x(t) = \frac{Kx_0e^{rt}}{K + x_0(e^{rt} - 1)}$$

- We have $x^* = \lim_{t\to\infty} x(t) = K$

The logistic difference equation

After rescaling, assume K = 1. Then $x_{t+1} = ax_t(1 - x_t)$
The growth rate, a, corresponds to r in the logistic differential equation

Bifurcation diagram of the logistic map



For a = 4, the equation produces a deterministic chaos.

Cell death

Suppose that cells have an average life span of 1/d. The differential equation becomes $x' = (r - d)\,x$. The quantity $\boldsymbol{R_0 = r/d}$ **is called the basic reproductive ratio**. It denotes the expected number of offspring from a single individual. If $R_0 > 1$, the population expands indefinitely; If $R_0 < 1$, the population goes extinct. If $R_0 = 1$, then the population size remains constant

Selection: the result of different growth rates associated with different types of individuals.

The **fitness** of an individual is its relative growth rate.

For two independent exponentially growing types, the type with higher growth rate would be dominant but neither type go extinct.

- Consider two types of individuals,
  - type A with growth rate a and abundance x(t),
  - type B with growth rate b and abundance y(t),

  growing according to x' = ax and y' = by. $x = x_0 e^{at}$
- Neither type can go extinct, if the *fitness* values a, b > 0.
- For their ratio, $\rho(t) = x(t)/y(t)$, we have $\rho' = (x'y - xy')/y^2 = (a - b)\rho$ and thus $\rho(t) = \rho_0\, e^{(a-b)t}$.
- If a > b, then $\rho \to \infty$. Selection favors A over B.
- If a < b, then $\rho \to 0$. Selection favors B over A.
- If a = b, then $\rho(t) = \rho_0$.

relative ratio

For competing types

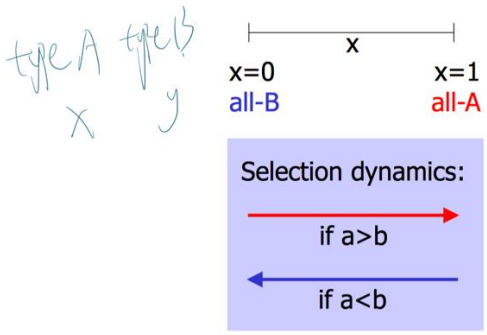
Subexponential and superexponential growth

- Consider two types with $x + y = 1$ and dynamics $x' = ax^c - \phi x$ and $y' = by^c - \phi y$, where $\phi = ax^c + by^c$.
- If c = 0, growth is linear (immigration). the growth does not depend on the size of population (different from reproduction)
- If c = 1, growth is exponential.
- If c < 1, growth is subexponential.
- If c > 1, growth is superexponential.
- The system is equivalent to $x' = x(1 - x)f(x)$, where $f(x) = ax^{c-1} - b(1 - x)^{c-1}$. (simpler form; (timer is similar with two competing type $x' = x(c$
- It has fixed points at $x = 0$, $x = 1$, and for $c \neq 1$, there is exactly one additional fixed point $x^* = 1 / (1 + (a/b)^{1/(c-1)})$ between 0 and 1. coexistence

C=1, survival of the fitter.

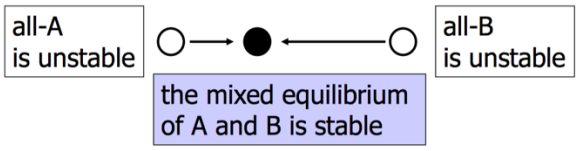
- The dynamics are described by the equations
  $x' = x(a - \phi)$,
  $y' = y(b - \phi)$,
  where $\phi = ax + by$ is the average fitness of the population. growth rate
- The terms involving ϕ ensure that $x + y = 1$.
- The system is equivalent to $x' = x(1 - x)(a - b)$.

- The equation $x' = x(1 - x)(a - b)$ has two equilibria.



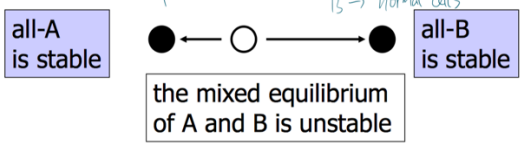
c < 1: survival of all

- If $c < 1$, then x^* is globally stable.
- Even if $a > b$, B can *invade* A (i.e., an infinitesimal small amount of type B individuals can grow in a population of almost all type A individuals).



c > 1: survival of the first

- If $c > 1$, then x^* is unstable.
  - If $x > x^*$, then A will outcompete B.
  - If $x < x^*$, then B will outcompete A.
- Even if $a > b$, a B population cannot be invaded by an A mutant.



Expansion from 2 to n types in exponential growth model:

- Consider n types with fitness values f_i, frequencies $x_i(t)$, and $x_1(t) + \ldots + x_n(t) = 1$. The type frequencies are points in the $(n - 1)$-dimensional probability simplex S_n.
- The average fitness of the population is $\phi = x_1 f_1 + \ldots + x_n f_n$.
- The selection dynamics are

$$x_i' = x_i(f_i - \phi) \qquad i = 1, \ldots, n$$

- This ODE system has a single stable equilibrium: starting from any interior point of the probability simplex, the fittest type will eventually outcompete all other types.

Mutation: during reproduction (cell division)/ without reproduction (radiation)

Basic mutation dynamics (no selection => assume two types of equal fitness a = b = 1)

Suppose that mutation rates $u_1 = Prob(A \to B)$ *and* $u_2 = Prob(B \to A)$ during reproduction.

- We have $x' = x(1 - u_1) + yu_2 - \phi x$
  $y' = xu_1 + y(1 - u_2) - \phi y$.
- Because $\phi = 1$ and $x + y = 1$, this system is equivalent to $x' = u_2 - x(u_1 + u_2)$ with stable equilibrium $x^* = u_2/(u_1 + u_2)$.
- Mutation leads to coexistence, $x^*/y^* = u_2/u_1$.
- If $u_1 \gg u_2$, we may assume $u_2 = 0$. Then $x' = -xu_1$ and A will go extinct. Thus, even without fitness differences, mutation alone can affect survival.

Expansion from 2 to n types

- Let q_{ij} = Prob(type i → type j), i,j = 1, ..., n.
- For all i, we have $q_{i1} + \ldots + q_{in} = 1$.
- Thus, $Q = (q_{ij})$ is a stochastic matrix.
- The n-dimensional mutation dynamics are

$$x_i' = \sum_{j=1}^{n} x_j q_{ji} - \phi x_i \quad i = 1, \ldots, n$$

  or $x' = xQ - \phi x$ in matrix notation.
- The equilibrium is the left hand eigenvector of Q associated with the largest eigenvalue, 1.

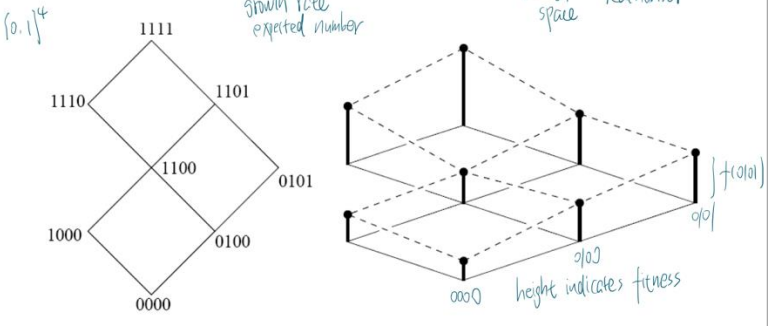
Hardy-Weinberg principle: states that allele and genotype frequencies in a population will remain constant from generation to generation in the absence of other evolutionary influences.

Sequence space

Sequence space A^L has dimension L, and consists of $|A|^L$ sequences. Distance is measured by the Hamming distance. In this metric, all sequences are close to each other: the distance between any pair of sequences is bounded by L. The genotype space is subset $G \subset A^*$.

Fitness landscapes

- A (static) fitness landscape is a mapping f : **G** → **IR**, such that $f(g)$ is the fitness of genotype g.



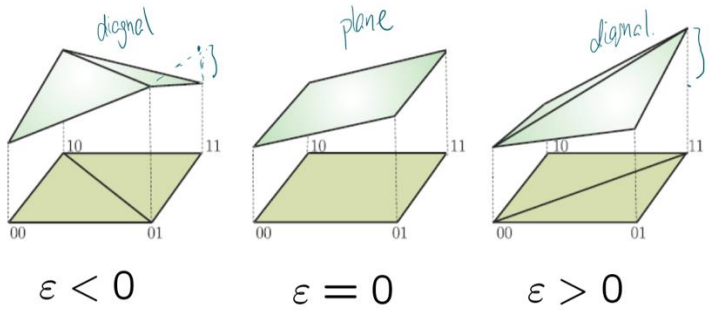
Binary sequence

Consider binary sequences of fixed length L. We assume that only point mutations occur in this population (ignoring insertion, deletions, recombination) and that they are independent and occur at the same rate u.

$q_{ij} = u^{d(i,j)}(1-u)^{L-d(i,j)}$ where $d(i,j)$ is hamming distance between genotype i and j

Epistatic interactions (Epistasis is the phenomenon wherein the effect of one gene (locus) is dependent on the presence of one or more 'modifier genes')

- Two-locus, biallelic system



$\varepsilon < 0$ $\varepsilon = 0$ $\varepsilon > 0$

- Epistasis:

$$\varepsilon = f_{00} + f_{11} - f_{01} - f_{10}$$

More precisely, f : G → P → IR because fitness depends on the phenotype of the organism. But phenotype space is more complex and less well defined. Besides, in general, fitness depends on how the individual interacts with the ecological environment and with other individuals in the population. By now we assume fitness only depends on genotype.

The quasispecies equation (species refers to chemical species)

- We consider (geno-)types i = 0, 1, 2, …, n.
- Let x(t) = (x_0(t), …, x_n(t)) denote the genotype frequencies.
- Let Q = (q_{ij}) = ($q_{i\to j}$) be a mutation matrix, f = (f_0, …, f_n) a fitness landscape, and $\phi = x \cdot f$ the average fitness of the population.
- The quasispecies equation is

$$\dot{x}_i = \sum_{j=0}^{n} x_j f_j q_{ji} - \phi x_i, \quad i = 0, \ldots, n$$

When Q = I, it means no mutation, then we recover the selection equation ("survival of the fittest").

When f = (1,…,1), it means no selection, then we recover the mutation equation.

If Q is irreducible, there is a single, globally stable equilibrium x * in the interior of the probability simplex. It means no type goes extinct

Solving the quasispecies equation

1. Suppose $\psi(t) = \int_0^t \phi(s)ds$, then $X = \sum_{j=0}^{n} X_j = e^{\psi}$ and $\dot{X} = \dot{\psi}e^{\psi} = \phi X$??
   The total population size X grows exponentially at rate ϕ
2. With the mutation-selection matrix $W = (w_{ij}) = (f_j q_{ji})$, the quasispecies equation can be written as

$$\dot{x} = xW - \phi x$$

- The equilibrium x^* is the solution of the eigenvalue problem

$$xW = \phi x$$

- The average fitness Á is the largest eigenvalue of the matrix W, and x^* is the corresponding (normalized) left eigenvector.

Adaptation of a quasispecie

Adaptation of a population is localization in sequence space at a (local) maximum of the fitness landscape.

Interpretation of equilibrium x^* of the quasispecies equation: **a mutation-selection balance of the population: selection drives individuals towards a local peak of the fitness landscape, while mutation generates types of lower fitness.**

Condition for adaptation: Adaptation is no longer possible if mutation removes the fittest types faster than they are selected (high mutton rate). There should be a necessary condition on the mutation rate and the selection strength for adaptation to occur. This condition is called the error threshold, and it exists for many, but not all, fitness landscapes.

Example

We consider binary sequences of length L. Type 0, the sequence 0...0 called *wild type* or *master sequence*, has fitness $f_0 > 1$. All other types have fitness 1.

The wild type is copied without error with probability $q = (1-u)^L$. We denote by x_0 the frequency of the wild type and by x_1 the sum of the frequencies of all other types. Ignoring back mutations to the wild type, the quasispecies equation is

$$\begin{aligned} \dot{x}_0 &= x_0 f_0 q - \phi x_0 \\ \dot{x}_1 &= x_0 f_0 (1-q) + x_1 - \phi x_1 \end{aligned}$$

or equivalently, $\dot{x}_0 = x_0[f_0 q - 1 - x_0(f_0 - 1)]$. The equilibrium is

$$x_0^* = \begin{cases} \frac{f_0 q - 1}{f_0 - 1} & \text{if } f_0 q > 1 \\ 0 & \text{otherwise} \end{cases}$$

The virus population can populate the wild type ($x_0^* > 0$) only if $f_0 q > 1$. For small mutation rates, $u \ll 1$, and moderate fitness values, $\log f_0 \approx 1$, this inequality becomes

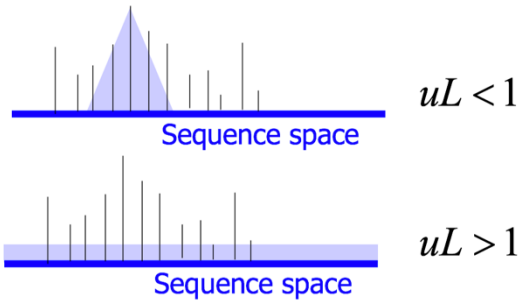
$$\log(f_0 q) = 1 + L \log(1-u) \approx 1 - Lu > \log 1 = 0$$

or

$$uL < 1$$

where uL is the genomic mutation rate, i.e., the expected total number of mutations per replication.

Therefore, the adaption only occurs below the critical mutation rate $u_c = \frac{1}{L}$, the error threshold. For longer genome, lower mutation rate is required to allow the wild (fittest) type adaptation.



$uL < 1$

$uL > 1$

- The case uL > 1, is known as *mutational meltdown*.
- Some antiretroviral drugs seem to work by increasing u.

The exponential distribution

A continuous random variable X is exponentially distributed with parameter $\lambda > 0$

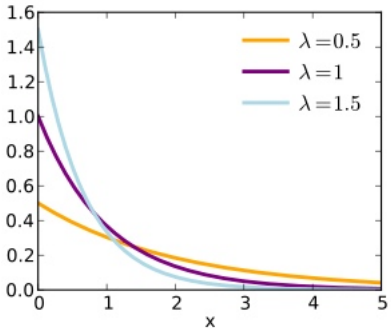
density function $f(x) = \lambda e^{-\lambda x}, \quad x \geq 0$

- The cumulative and tail probabilities are

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

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

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



Memoryless $P(X > s + t) = \exp\{-\lambda(t + s)\}$
$= \exp(-\lambda t)\exp(-\lambda s)$
$= P(X > t)\,P(X > s)$

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:

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

first failure among X and Y

and

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

Markov chains

A discrete-time *Markov chain* is a stochastic process {X(t)} with T = {0, 1, 2, …}, 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 P of a Markov chain {X(t)} is defined by

$$P_{ij}(t) = P(X(t+1) = j | X(t) = i)$$

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 $t \ge t_0$.

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

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

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

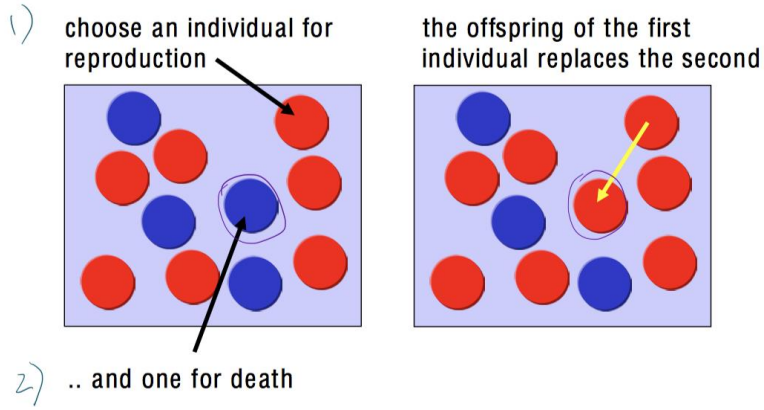
$$\Pi' P = \Pi'$$

where Π' denotes the transpose of Π.

where $\Pi' = [\pi_1, \ldots, \pi_n]$

The Moran process

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



1. It defines a discrete 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

   $$P_{i,i+1} = p(1-p)$$
   $$P_{i,i-1} = (1-p)p$$
   $$P_{i,i} = p^2 + (1-p)^2$$

   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*.

   Absorbing states (No way to escape from the current state): 1. Fixation, all A ($P_{N,N} = 1, P_{N,i} = 0 \ for \ i < N$) 2. Extinction, All B ($P_{0,0} = 1, P_{0,i} = 0 \ for \ i > 0$)

2. It is a birth-death process
   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.

Fixation probabilities

Let x_i be the probability of ending up in state N when starting from state i.

For the simple neutral Moran process this probability is $x_i = \frac{i}{N}$
Since all individuals have the same fitness, they also have the same chance of becoming the ancestor of the whole population.

The mean fixation time (how to compute ??)

- In the Moran process, for large population sizes, the mean fixation time is (fraction of type A: $\frac{i}{N}$)
$$-N^2[(1-p)\log(1-p) + p\log p]$$
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) (eventually hit the absorbing state)
decays approximately exponentially at rate 2 / N².
- This rate quantifies the amount of random genetic drift that the population is experiencing. (No difference in fitness)

When fitness of different types is different (selection)

This can be incorporated into the model if individuals with allele A have fitness f_i and individuals with allele B have fitness g_i. Suppose $\alpha_i = P_{i,i+1}$, $\beta_i = P_{i,i-1}$, $y_i = x_i - x_{i-1}$

$$\begin{aligned} P_{i,i-1} &= \frac{g_i(N-i)}{f_i \cdot i + g_i(N-i)} \cdot \frac{i}{N} \\ P_{i,i} &= 1 - P_{i,i-1} - P_{i,i+1} \\ P_{i,i+1} &= \frac{f_i \cdot i}{f_i \cdot i + g_i(N-i)} \cdot \frac{N-i}{N} \end{aligned}$$

$$x_i = \begin{cases} 0 & i = 0 \\ \beta_i x_{i-1} + (1-\alpha_i-\beta_i)x_i + \alpha_i x_{i+1} & 1 \le i \le N-1 \\ 1 & i = N \end{cases}$$

Since $x_i = \beta_i x_{i-1} + \alpha_i x_{i+1} + (1-\alpha_i-\beta_i)x_i$
This can be further simplified into: $\beta_i(x_i - x_{i-1}) = \alpha_i(x_{i+1} - x_i) \Rightarrow \beta_i y_i = \alpha_i y_{i+1}$.
Let $\gamma_i = \frac{\beta_i}{\alpha_i}$, we have $y_j = x_1 \cdot \frac{\beta_1}{\alpha_1} \cdots \frac{\beta_{j-1}}{\alpha_{j-1}} = x_1 \prod_{k=1}^{j-1} \gamma_k$
We have:

$$x_j = \sum_{i=1}^{j} y_i = x_1 + x_1 \cdot \sum_{i=1}^{j-1} \prod_{k=1}^{i} \gamma_k$$

Also by the fact that $x_N = 1$, $x_1 + x_1 \cdot \sum_{i=1}^{N-1} \prod_{k=1}^{i} \gamma_k = 1$, we can obtain that:

$$x_1 = \frac{1}{1 + \sum_{i=1}^{N-1} \prod_{k=1}^{i} \gamma_k}$$

Combining all that above, yielding:

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

When the selection is constant

- Consider exponentially distributed waiting times to the reproduction of a type A and type B individual with rates $\lambda_A = r$ and $\lambda_B = 1$, 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
  - $T_A \sim \min\{\text{Exp}(\lambda_A), \ldots, \text{Exp}(\lambda_A)\} = \text{Exp}(i\lambda_A)$
  - $T_B \sim \text{Exp}((N-i)\lambda_B)$.
- T_A and T_B are competing exponentials:

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

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

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

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

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

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

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

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) | t ≥ 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 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,

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

density

Poisson process only depends on the time interval

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, $T_1 \sim Exp(Nu)$

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 $R = Nu\rho$??

The probability that a mutant B has been fixed by time t is $1 - e^{-Rt}$

If B is neutral, then ρ= 1/N and R = u, the mutation rate.

If u is constant, then neutral mutations accumulate at a constant rate R = u, independent of population size. In this way we can determine the time scale of this process (molecular clock).

Cancer

The rate of activating oncogenes and inactivating TSGs depends on population size, mutation rates, and fitness**.**
Tissue architecture (small compartments)can affect the rate of evolution of **cancer.** The linear process delays cancer initiation.

Oncogene: increase fitness if one allele is mutated or inappropriately expressed. They are activated by 1. Point mutation 2. Amplification 3. Chromosome fusion

1. Fixation of oncogene mutations

Consider a Moran process in a compartment with effective population size N

- Mutants arise with probability u and have relative fitness r (r > 1 advantageous, r = 1 neutral, r < 1 deleterious).
- The fixation probability is $\rho = x_1 = (1 - 1/r)/(1 - 1/r^N)$. described in last lecture

The probability that a mutant has been fixed by time t is

$$P(t) = 1 - e^{-Nu\rho t}$$

fixation probability

P(t) is increasing in N if r > 1, and decreasing in N if r < 1.

Influence: Large compartments accelerate the accumulation of advantageous mutations, small compartments slow it down. Most tissues with high cell turnover are organized in many small compartments.

2. The linear process of cancer

A mutant with relative fitness difference r has fixation probability $\rho = \frac{1}{N}$, because only a mutation in the left most stem cell leads to fixation, which means only the location could impact the chance of becoming the ancestor for the population instead of fitness

Therefore, the probability that the mutant has taken over by time t is $P(t) = 1 - e^{-ut}$ and independent of r

Influence: Delays advantageous mutations. In contrast to well-mixed populations, where advantageous mutations accumulate faster, all types of mutations (advantageous, neutral, or deleterious) have the same fixation probability in the linear process

Tumor suppressor genes: somatic mutations in TSGs are recessive: inactivation of one allele is (nearly) neutral, while inactivating the second allele confers a fitness advantage. TSGs are inactivated by 1. Point mutation 2. one point mutation followed by loss of heterozygosity (LOH)

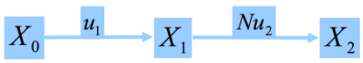
1. For neutral mutation, type 1 cells reach fixation before a type 2 cell arises in small compartments

The average fixation time of the first mutation is $1/\rho$ = N (ρ is the fixation probability) in the Moran process. The average waiting time for the second mutation (in any teach occurrence not cell) is $\frac{1}{N\mu_2}$. (??)

So, type 1 cells reach fixation before a type 2 cell arises, if $N \ll \frac{1}{N\mu_2} \Rightarrow N \ll 1/\sqrt{\mu_2}$ (small compartments)

Three-state ODE model

the dynamics can be described by the probabilities $X_i(t)$ of being in the three states i = 0, 1, 2. State 0: all cells are of type 0 ; State 1: all cells are of type 1; State 2: at least one cell is of type 2



$$\dot{X}_0 = -u_1 X_0$$
$$\dot{X}_1 = u_1 X_0 - N u_2 X_1$$
$$\dot{X}_2 = N u_2 X_1$$

Initially, X(0) = (1, 0, 0). (all cells in type 1) For $t \to \infty$, $X(t) \to (0, 0, 1)$.

Solution:

$$P(t) = X_2(t) = 1 - \frac{N u_2 e^{-u_1 t} - u_1 e^{-N u_2 t}}{N u_2 - u_1}$$

In state 0, the rate of producing type 1 cells is Nu_1. The probability that such a cell reaches fixation is $1/N$. Therefore the transition rate from state 0 to state 1 is given by the mutation rate u_1. If the population is in state 1, then the rate of producing type 2 cells is given by Nu_2.

Three time scales

1. When $t \ll \frac{1}{Nu_2}$ $P(t) \approx N u_1 u_2 t^2 / 2$

there are two rate limiting events

2. When $\frac{1}{N\mu_2} < t < \frac{1}{u_1}, P(t) \approx 1 - e^{-u_1 t}$

On this time scale, the second hit is fast and can be
Neglected, only the first hit is rate limiting.

3. For very long times (beyond human life time), t >> $\frac{1}{u_1}$,

there are no rate limiting events.

For intermediate population size

The average waiting time for a type 1 cell is $\frac{1}{Nu_1}$, which is long if N < $1/u_1$
A type 2 cell is generated before fixation of the type 1 lineage, if N > $1/\sqrt{u_2}$.

Thus, type 1 is "tunneled" in the intermediate regime (population tunnels from state 2 to state 0 without reaching state 1).

don't wait for type 1 fixation, but produce type 2 cell

$$1/\sqrt{u_2} \ll N \ll 1/u_1$$

- In this parameter region, the probability that at least one cell with two hits has arisen before time t is

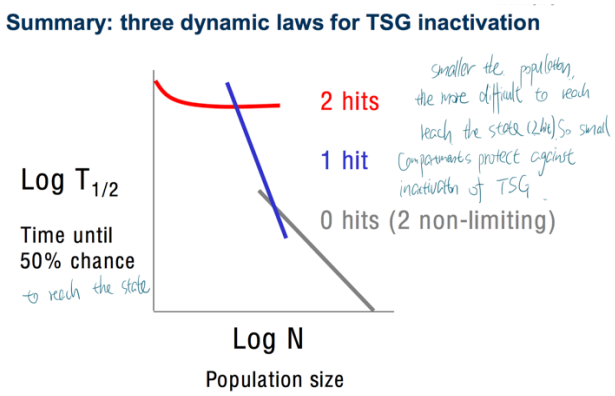
$$P(t) = 1 - e^{-Nu_1\sqrt{u_2}t}$$??

For large population size

- If $N \gg 1/u_1$ then type 1 cells are generated immediately and they grow according to $x_1(t) = N u_1 t$.
- The probability of producing a type 2 mutant during type 1 growth is

$$P(t) = 1 - \exp\left\{-u_2 \int_0^t x_1(t)dt\right\}$$
$$= 1 - \exp\left(-\frac{1}{2}Nu_1u_2t^2\right)$$

??

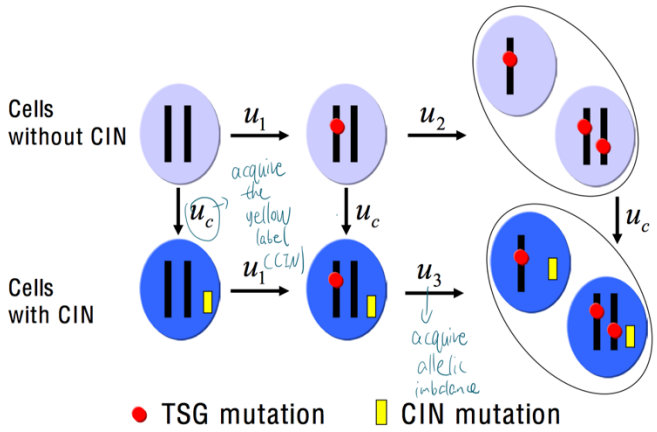


Genetic instability

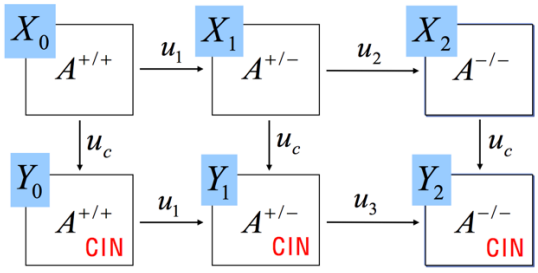
Microsatellite instability: caused by mutations in mismatch repair genes; increases the point mutation rate

Chromosomal instability: increased rate of gaining or losing whole chromosomes or large parts of it

TSG inactivation with and without CIN

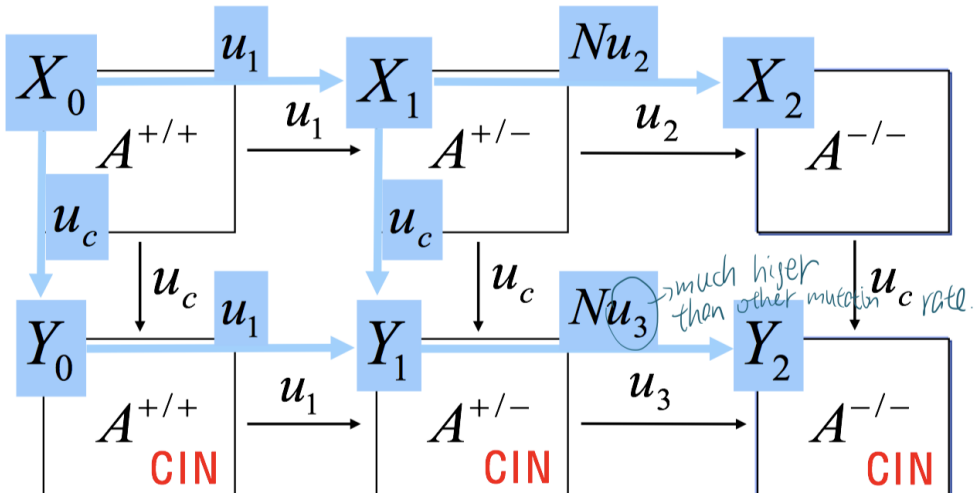


For small compartments, $N \ll 1/u_1, 1/u_2, 1/u_c$, we consider the corresponding homogeneous states



For Neutral CIN in small compartments

Suppose CIN and A +/– are neutral (ρ = 1/N), and that A –/– will be fixed immediately (ρ = 1).



- The ODE system of state probabilities (with $X_0(0) = 1$)

$$\dot{X}_0 = -(u_1 + u_c)X_0 \quad \dot{X}_1 = u_1 X_0 - (u_c + Nu_2)X_1 \quad \dot{X}_2 = Nu_2 X_1$$

$$\dot{Y}_0 = u_c X_0 - u_1 Y_0 \quad \dot{Y}_1 = u_c X_1 + u_1 Y_0 - Nu_3 Y_1 \quad \dot{Y}_2 = Nu_3 Y_1$$

has, on the relevant time scale, the approximate solution

$$X_0(t) \approx 1 \quad X_1(t) \approx u_1 t \quad X_2(t) \approx Nu_1 u_2 t^2/2$$

$$Y_0(t) \approx u_c t \quad Y_1(t) \approx u_1 u_c t^2 \approx Y_2(t) \approx u_1 u_c t^2$$

$Y_1 \to Y_2$ occurs immediately

- $Y_1 \approx Y_2$, because $u_3 \approx 10^{-2}$ and $N u_3 t \gg 1$. This step is not rate limiting: the waiting time for LOH in CIN cells is negligible compared to the other events.

For costly CIN in small compartments

- If CIN cells have fitness $r < 1$, their fixation probability in the Moran process is $\rho = (1 - 1/r)/(1 - 1/r^N)$ and the non-CIN-to-CIN transition rate is $N\rho u_c$. Nu_c: total rate ρ: depends on r
- On the relevant time scale, we find approximately

$$X_0(t) \approx 1 \quad X_1(t) \approx u_1 t \quad X_2(t) \approx Nu_1 u_2 t^2/2$$

$$Y_0(t) \approx N\rho u_c t \quad Y_1(t) \approx N\rho u_1 u_c t^2 \quad Y_2(t) \approx N\rho u_1 u_c t^2$$

Costly CIN in large compartments

- For large N and $r < 1$, the product $N\rho$ becomes vanishingly small, such that the intermediate CIN types $A^{+/+}$*CIN* and $A^{+/-}$*CIN* will not reach fixation. tunnel effect
- $A^{+/-}$*CIN* cells are produced at rate Nu_c and they remain near a mutation-selection balance with average abundance $Nu_c/(1 - r)$. They produce $A^{-/-}$*CIN* cells at rate ru_3.
- Thus, the population tunnels from X_1 to Y_2 at rate
  $R = (Nu_c ru_3)/(1 - r)$
- On the relevant time scale, we obtain approximately

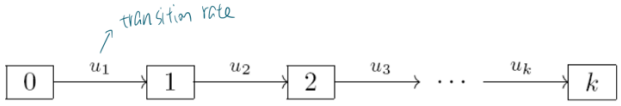
$$X_0(t) \approx 1 \quad X_1(t) \approx u_1 t \quad X_2(t) \approx Nu_1 u_2 t^2/2$$

$$Y_0(t) \approx 0 \quad Y_1(t) \approx 0 \quad Y_2(t) \approx Ru_1 t^2/2$$

Iwasa et al., Genetics 200

Cancer progression

Multistage theory (rate of incidence and time is log-log linear relationship)



normal tissue ⟶ metastases

If $u_j = u$ is small, Prob (one step by t) = $1 - e^{-ut} \approx ut$. The incidence (rate) is
$I_k(t) = (ut)^{k-1}u = u^k t^{j-1}$ => $\log I_k = k \log u + (k-1) \log t$

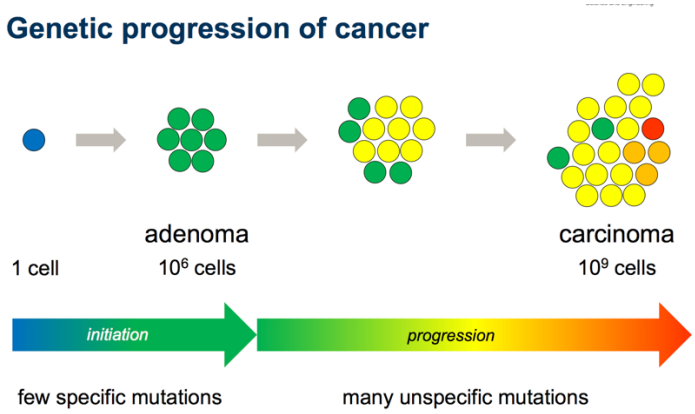
- The waiting time for each step is exponential, Exp(u_j)
- Let τ_k be the waiting time until stage k is reached. Then

$$\tau_k \sim \text{Exp}(u_1) + \cdots + \text{Exp}(u_k)$$

sum of waiting time in exponential distribution

$$\begin{aligned} E[\tau_k] &= E\left[\sum_{j=1}^{k} \text{Exp}(u_j)\right] \\ &= \sum_{j=1}^{k} E[\text{Exp}(u_j)] \\ &= \sum_{j=1}^{k} \frac{1}{u_j} \end{aligned}$$

The mutational landscape of colon cancer: few mountains (In each patient, 15 to 20 mutated genes seem to drive progression. This set of genes differs considerably among patients.), many hills



Independent mutations
Each mutation occurs independently at time $T_j \sim Exp(\lambda_j)$ (different from transition rate u)

Waiting time to any k independent mutations

- Let τ_k be the waiting time until any *k* out of *d* mutations have occurred,

$$\tau_k = \min_{\{j_1,\ldots,j_k\} \subset \{1,\ldots,d\}} \max\{T_{j_1}, \ldots, T_{j_k}\}$$

maximal individual waiting time since the mutations occur independently

- For $k = 1$, we have

$$\tau_1 = \min\{T_1, \ldots, T_d\} \sim \text{Exp}(\lambda_1 + \cdots + \lambda_d)$$

- If $\lambda_1 = \ldots = \lambda_k = \lambda$, then

$$\begin{aligned} \tau_1 &\sim \text{Exp}(d\lambda) \\ \tau_2 &\sim \tau_1 + \text{Exp}((d-1)\lambda) \\ &\vdots \\ \tau_j &\sim \tau_{j-1} + \text{Exp}((d-j+1)\lambda) \end{aligned}$$

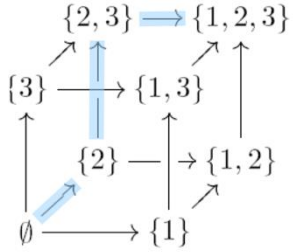
- Hence,

$$E[\tau_k] = \frac{1}{\lambda} \sum_{j=1}^{k} \frac{1}{d-j+1}$$

Each total order of mutations $j_1 < \cdots < j_k$ defines a mutational pathway in the genotype lattice. For a fixed path, let $Exit_i$ denote the set of all possible mutations in step i.

For example,

- $\text{Exit}_1 = \{1, 2, 3\}$
- $\text{Exit}_2 = \{1, 3\}$
- $\text{Exit}_3 = \{1\}$

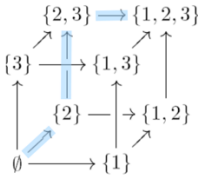


Probability and waiting time of pathways

- The probability of a pathway $P = j_1 \to \ldots \to j_k$ is

$$\text{Prob}(P) = \prod_{i=1}^{k} \frac{\lambda_{j_i}}{\sum_{j \in \text{Exit}_i} \lambda_j}$$

(competing exponentials)



- The expected waiting time of P is

$$E(\tau_P) = \sum_{i=1}^{k} \frac{1}{\sum_{j \in \text{Exit}_i} \lambda_j}$$

Waiting time to cancer

$$E(\tau_k) = \sum_{path} P(path) E(\tau_{path})$$

The Wright-Fisher process (a stochastic process that describes the evolution of an asexual)

1. It defines a Markov chain

   - Consider a haploid population of constant size N.
   - There are two different types, A and B.
   - Reproduction occurs in discrete, non-overlapping generations, i.e., individuals are synchronized.
   - Let X(t) be the number of type A individuals in generation t = 0, 1, 2, 3, …
   - X(t) has state space {0, 1, …, N}.

2. Choose parents from the previous generation randomly (Binomial sampling)

- Each generation is sampled from the previous generation according to the binomial distribution,

$$(X(t+1) \mid X(t) = i) \sim \text{Binom}(N, i/N)$$

- The transition probabilities are given by

$$P_{ij} = P(X(t+1) = j \mid X(t) = i)$$
$$= \binom{N}{j} \left(\frac{i}{N}\right)^j \left(\frac{N-i}{N}\right)^{N-j}$$

3. Properties of Neutral WF process (Proof in ex5)
   Mean: The Wright-Fisher process is unbiased: the average frequency of type A individuals does not change over time, namely $E[X(t)] = i$ for all t = 0, 1, 2, …

   Variance: The variance of X(t) is $Ni(1 - p)[1 - (1 - p)^t]$.

   Fixation probability: $x_i = \frac{i}{N}$

   Mean fixation times:
   - No simple way of calculating k_i, even approximately, is known. We will later use diffusion theory to show that

$$k_1 \approx 2\,[\log(N-1) + \gamma]$$

   where γ = 0.5772… is Euler's constant.

Wright-Fisher process for accumulating mutations

Assume that each locus (independently) undergoes mutation from 0 to 1 at rate u. We ignore back mutations from 1 to 0. A constant fitness advantage, s, per mutation.

Let $X_j(t)$ be the number of cells with j mutations (called j-cells) in generation t. The fitness of a j-cell is proportional to $(1 + s)^j$. Set $x_j(t) = \frac{X_j(t)}{N}$. Initially, $X_0(0) = N, X_1(0) = \cdots = X_d(0) = 0$

What is the probability of sampling a j-cell?

$$\theta_j(t) = \sum_{i=0}^{j} P(\text{i-cell} \to \text{j-cell})$$
$$= \sum_{i=0}^{j} P(\text{i-to-j mutation}) P(\text{i-cell parent})$$
$$= \sum_{i=0}^{j} \binom{d-i}{j-i} u^{j-i}(1-u)^{d-j} \frac{(1+s)^i x_i(t)}{\sum_\ell (1+s)^\ell x_\ell(t)}$$

Multinomial sampling

- The transition probabilities

$$P_{m,n} = P[X(t+1) = n \mid X(t) = m]$$

of the Markov chain X(t) = (X_0(t), …, X_d(t)) are given by the multinomial distribution, d+1 types

$$\frac{(n_0 + \cdots + n_d)!}{n_0! \ldots n_d!} \prod_{j=1}^{d} \theta_j(t)^{n_j}$$

where n = (n_0, …, n_d) and n_0 + … + n_d = N.

Approximating the average waiting time

Assumption: 1. each mutant wave has a Gaussian shape; 2. waves travel at constant speed; 3. s << 1, so $(1 + s)^j \approx 1 + sj$

Ansatz: number of mutations; speed for travel

$$x_j(t) = Ae^{-\frac{(j-vt)^2}{2\sigma^2}}$$

A: constants

- $A \approx 1/\sqrt{2\pi}\sigma$ by the case of continuous j.
- The two unknowns, v and σ, are determined by decoupling clonal expansion (driven by selection) and generation of new types (by mutation).

growth of C this type

- Clonal expansion is governed by the replicator equation

$$\dot{x}_j = sx_j \left[j - \sum_{i=0}^{d} i\, x_i(t) \right]$$

selection advantage

- Substituting the expression for x_j(t) yields the relation

$$v = s\sigma^2$$

Consider an additional mutation when the number of current mutations is j

- Let τ be the average time it takes to produce a new mutant.
- The velocity v = 1/τ is found by solving

$$\frac{1}{N} = x_{j+1}(\tau) = ud \int_0^\tau x_j(t)dt$$

Based on the fact that initially x_j grows exponentially. And clonal expansion follows the replicator equation.

- For the velocity, v, we eventually find

$$v \approx \frac{2s \log N}{\log\,[s/(ud)]^2}$$??

This is (an approximation of) the speed of adaptation in an asexual population evolving according to the Wright-Fisher process.

The average time it takes until the first cell with k mutations appears is approximately

$$\tau_k \approx \frac{k \log [s/(ud)]^2}{2s \log N}$$

=> **waiting time to cancer**

- Thus, $\tau_k \propto k / s$. The waiting time is linear in k.
- The selective advantage, s, has a much larger impact on the waiting time than the mutation rate, u, or the population size, N.

Diffusion theory

- Directional processes, M(p)
  - Non-zero expected change in allele frequency
  - E.g., selection, mutation, migration, recombination
  - M(p) is measured as the expected change per generation
- Nondirectional processes, V(p) stochastic
  - Zero expected change
  - Increase in the spread of the allele distribution
  - E.g., random variation in survivorship, random gamete success (drift)
  - V(p) is measured by the expected variance in the next generation

For example in Moran process

- With p = p(t) = X(t)/N,

allele frequency difference between t and t+1 which depends on the current state

$$\begin{aligned} M(p) &= E[p(t+1) - p(t) \mid p(t)] \\ &= p - p = 0 \end{aligned}$$

$$\begin{aligned} V(p) &= E\{\mathrm{Var}[p(t+1)] \mid p(t)\} \\ &= \frac{1}{N^2} E\{\mathrm{Var}[X(t+1)] \mid X(t)\} \\ &= \frac{2p(1-p)}{N^2} \end{aligned}$$

↓ constant value

Diffusion equation (Forward equation)

Assume that

- $\psi = \psi(p, t)$
  = Probability density of allele frequency p at time t
- $g = g(p, \epsilon; dt)$
  = Probability that the allele frequency changes from p to $p + \epsilon$ in time interval dt.

Any allele frequency p at time t + dt must have evolved from some distribution p – ϵ at time t,

$$\psi(p, t+dt) = \int \psi(p-\epsilon, t) g(p-\epsilon, \epsilon; dt) d\epsilon$$

- Taylor series expansion of $\psi \cdot g$ around p:

$$\psi(p, t+dt) = \int \left[\psi(p,t) g(p, \epsilon; dt) - \epsilon \frac{\partial(\psi g)}{\partial p} + \frac{\epsilon^2}{2} \frac{\partial^2(\psi g)}{\partial p^2} - \frac{\epsilon^3}{6} \frac{\partial^3(\psi g)}{\partial p^3} + \dots \right] d\epsilon$$

the order is related to the precision. If only keep the ϵ^1, it keeps the mean. stop here

- The *diffusion approximation* is derived by assuming $\epsilon^2 \gg \epsilon^3$. This assumption means that the population does not change too much in any short time interval, i.e., the effects of selection and mutation are relatively weak.

- Also, p and ϵ are independent. Thus, approximately,

$$\psi(p, t+dt) = \psi(p,t) \int g(p, \epsilon; dt)\, d\epsilon - \frac{\partial}{\partial p}\psi \int g\epsilon\, d\epsilon + \frac{1}{2}\frac{\partial^2}{\partial p^2}\psi \int g\epsilon^2\, d\epsilon$$

- Because $\int g(p, \epsilon; dt)\, d\epsilon = 1$, we have

$$\psi(p, t+dt) = \psi - \frac{\partial}{\partial p}\psi \int g\epsilon\, d\epsilon + \frac{1}{2}\frac{\partial^2}{\partial p^2}\psi \int g\epsilon^2\, d\epsilon$$

The meaning of the integrals

- $\int g\epsilon\, d\epsilon$ = E(ϵ) is the expected change over time dt. With M(p) the rate of directional change, we have

$$\int g(p, \epsilon; dt)\epsilon\, d\epsilon = M(p)dt$$

ϵ is small; rate

- $\int g\epsilon^2\, d\epsilon$ = E(ϵ^2) = Var(ϵ) + E(ϵ)^2 $\approx$ Var(ϵ). With V(p) the variance in allele frequency due to nondirectional effects,

$$\int g(p, \epsilon; dt)\epsilon^2\, d\epsilon = V(p)dt$$

Therefore, the diffusion equation (Kolmogorov forward equation, Fokker-Planck equation) is

current shape; M and P are evolutionary forces

$$\frac{\partial \psi(p,t)}{\partial t} = -\frac{\partial}{\partial p}[\psi(p,t)M(p)] + \frac{1}{2}\frac{\partial^2}{\partial p^2}[\psi(p,t)V(p)]$$

Equilibrium

- At equilibrium,

$$-\frac{\partial}{\partial p}[\psi^*(p,t)M(p)] + \frac{1}{2}\frac{\partial^2}{\partial p^2}[\psi^*(p,t)V(p)] = 0$$

- Integrating over p yields

$$\frac{1}{2}\frac{\partial}{\partial p}[\psi^*(p,t)V(p)] - \psi^*(p,t)M(p) = 0$$

- This first order homogeneous ODE can be solved.

- The equilibrium distribution of the allele frequency is

$$\psi^*(p) = \frac{C}{V(p)} \exp\left[\int_0^p \frac{2M(q)}{V(q)} dq\right]$$

where C is a constant of integration.

Computing M(p)

Selection

- Assume two alleles A_1 and A_2 with frequencies p and 1 – p, and fitness w_1 and w_2, respectively.
- The average fitness of the population is

$$\bar{w} = pw_1 + (1-p)w_2 \quad \text{and} \quad \frac{d\bar{w}}{dp} = w_1 - w_2$$

- In the next generation, the allele frequency is

$$p' = (pw_1)/\bar{w}$$

$$\begin{aligned}
\Delta p_{\text{sel}} = p' - p &= \frac{p(w_1 - \bar{w})}{\bar{w}} \\
&= \frac{p(1-p)(w_1 - w_2)}{\bar{w}} \\
&= \frac{p(1-p)}{\bar{w}}\frac{d\bar{w}}{dp} \\
&= p(1-p)\frac{d\log\bar{w}}{dp}
\end{aligned}$$

difference of allele frequence due to selection

Mutation

- Let u_1 be the A_1-to-A_2 mutation rate, and u_2 the A_2-to-A_1 mutation rate.
- The per-generation change due to mutation is $\Delta p_{mut} = -p\,u_1 + (1-p)\,u_2$.

- Thus, the directional processes of selection and mutation add up to

$$M(p) = \frac{p(1-p)}{\bar{w}}\frac{d\bar{w}}{dp} - pu_1 + (1-p)u_2$$

Combination of processes we consider $\Delta p_{sel.} + \Delta p_{mute}$

Computing V(p) (assume WF process)

$$\begin{aligned} V(p) &= E\{\mathrm{Var}[p(t+1)] \mid p(t)\} \\ &= E\{\mathrm{Var}[X(t+1)/N] \mid X(t)/N\} \\ &= \frac{1}{N^2} Np(1-p) \\ &= \frac{p(1-p)}{N} \end{aligned}$$

" $E(\mathrm{Var}(X(t+1) \mid X(t))$

$$\psi^*(p) = \frac{C}{V(p)} \exp\left[\int_0^p \frac{2M(q)}{V(q)} dq\right]$$

$$\frac{M}{V} = N\left[\frac{d\log\bar{w}}{dp} - \frac{u_1}{1-p} + \frac{u_2}{p}\right]$$

$$\Rightarrow \frac{\frac{p(1-p)}{\bar{w}}\frac{d\bar{w}}{dp} - pu_1 + (1-p)u_2}{\frac{p(1-p)}{N}} = \frac{M}{V}$$

$$\int_0^p \frac{M(q)}{V(q)} dq = N\left[\log\bar{w} + u_1\log(1-p) + u_2\log(p)\right]$$

Computing equilibrium

$$\begin{aligned} \psi^*(p) &= \frac{C}{V(p)} \exp\left[\int_0^p \frac{2M(q)}{V(q)} dq\right] \\ &= C\,\bar{w}^{2N}\,(1-p)^{2Nu_1-1}\,p^{2Nu_2-1} \end{aligned}$$

where C is given by the requirement that

$$\int_0^1 \psi^*(p)dp = 1$$

For neutral variation

For $\bar{w} = 1$ and $u = u_1 = u_2$,

No selection $w_1 = w_2 = 1$

$$\psi^*(p) \propto [p(1-p)]^{2Nu-1}$$

- We set $\theta = 2\,N\,u$, because this scaled mutation parameter determines the equilibrium distribution.

For selection

Selection

- Assume that A_1 has fitness $w_1 = 1 + s$ and A_2 has fitness $w_2 = 1$. Then the average population fitness is

$$\begin{aligned} \bar{w} &= p(1+s) + (1-p) \\ &= 1 + sp \approx e^{sp} \end{aligned}$$

$$\Rightarrow \bar{w}^{2N} = e^{2Nsp}$$

$\theta = 2Nu$

- We set $\sigma = 2\,N\,s$, because this scaled selection parameter determines the equilibrium dynamics.
- Ignoring mutation, we have

$$\psi^*(p) \propto e^{\sigma p}/p(1-p)$$

Summary

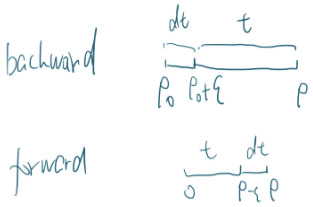
$$\psi^*(p) \propto (1-p)^{\theta-1}\, p^{\theta-1}\, e^{\sigma p}$$

with scaled mutation parameter $\theta = 2Nu$ and scaled selection parameter $\sigma = 2Ns$.

Backward equation

When the allele frequency is p0 at time t0

$$\psi(p, t+dt \mid p_0) = \int \psi(p, t \mid p_0 + \epsilon) g(p_0, \epsilon; dt) d\epsilon$$



- Taylor series expansion of $\psi(p, t \mid p_0 + \epsilon)$ around p_0:

$$\psi(p,t \mid p_0+\epsilon) = \psi(p,t \mid p_0) + \\ + \frac{\partial \psi}{\partial p_0} + \frac{\epsilon^2}{2}\frac{\partial^2 \psi}{\partial p_0^2} + \frac{\epsilon^3}{6}\frac{\partial^3 \psi}{\partial p_0^3} + \dots$$

- Write $\psi = \psi(p, t, \mid p_0)$, $g = g(p_0, \epsilon, dt)$ and approximate to

$$\psi(p,t+dt \mid p_0) = \psi \int g d\epsilon + \\ + \frac{\partial \psi}{\partial p_0}\int g\epsilon d\epsilon + \frac{1}{2}\frac{\partial^2 \psi}{\partial p_0^2}\int g\epsilon^2 d\epsilon$$

The Kolmogorov backward equation

- The integrals have the same meaning as before.
- Subtract $\psi(p, t \mid p_0)$, divide by dt, and let $dt \to 0$ to obtain

$$\frac{\partial \psi(p,t \mid p_0)}{\partial t} = \\ = M(p_0)\frac{\partial \psi(p,t \mid p_0)}{\partial p_0} + \frac{1}{2}V(p_0)\frac{\partial^2 \psi(p,t \mid p_0)}{\partial p_0^2}$$

Equilibrium

- We find

$$\frac{\partial \psi^*}{\partial p_0} = C \exp\left(-\int_0^p \frac{2M(q)}{V(q)}dq\right)$$

- Integration w.r.t. p_0 does not yield a unique solution.

- We are interested in the probability of fixation of the first allele, A_1, (p = 1) given its initial allele frequency p_0, i.e., in

$$\rho(p_0) = \psi(1, \infty \mid p_0)$$

- Clearly,

$$\rho(1) = \psi(1, \infty \mid 1) = 1$$
$$\rho(0) = \psi(1, \infty \mid 0) = 0$$

- With these boundary conditions, we can solve for ψ^*.

to get unique solution for ψ^*

$$\rho(p_0) = \frac{\int_0^{p_0} \exp\left(-\int_0^p \frac{2M(q)}{V(q)} dq\right) dp}{\int_0^1 \exp\left(-\int_0^p \frac{2M(q)}{V(q)} dq\right) dp}$$

General

Computing M(p) and V(p) (assume WF process)

Selection

- Recall that

$$\bar{w} = 1 + ps, \quad \frac{d\bar{w}}{dp} = s,$$

$$M(p) = \frac{p(1-p)}{\bar{w}} \frac{d\bar{w}}{dp} = \frac{p(1-p)s}{1+ps}$$

- For $p \ll 1$,

$V = \frac{p(1-p)}{N}$ in WF process

$$\frac{2M(p)}{V(p)} = \frac{2Ns}{1+ps} \approx 2Ns$$

because s is also small.

Computing fixation probability

- Now,

$$\int_0^p \underbrace{\frac{2M(q)}{V(q)}}_{\text{constant}} dq = 2Nsp$$

and hence

assumption: WF process $p \ll 1$

$$\rho(p_0) = \frac{1 - e^{-2Nsp_0}}{1 - e^{-2Ns}}$$

Neutral case $s=0$

- For s = 0, $\rho(p_0)$ is undefined. However,

$$\lim_{s \to 0} \rho(p_0) = \frac{-2Nsp_0}{-2Ns} = p_0$$

Apply L'Hospital's Rule $\lim_{s\to 0} \rho(p_0) = \lim_{s\to 0} \frac{1-e^{-2Nsp_0}}{1-e^{-2Ns}} = \lim_{s\to 0} \frac{\frac{\partial}{\partial s}(\,)}{\frac{\partial}{\partial s}(\,)}$

A new mutant in a large population

$$\rho(1/N) = \frac{1 - e^{-2s}}{1 - e^{-2Ns}}$$

- If N is large, $2Ns \gg 1$, then the denominator is one.
- If s is small, $s \ll 1$, then $\exp(-2s) \approx 1 - 2s$
- In this situation, we recover a classical result due to Haldane (1927): Assumption: WF process and $p \ll 1$

$$\rho(\text{new mutant}) \approx 2s$$

Mean fixation time

- Let $\tau(p_0)$ be the expected waiting time until A_1 becomes fixed, given its fixation and the initial allele frequency p_0.
- Define

$$A(p) = \int_0^p \frac{2M(q)}{V(q)} dq$$

$$S(p_0) = \int_0^{p_0} \exp\left(-A(p)\right) dp$$

$$\rho(p_0) = \frac{\int_0^{p_0} \exp\left(-\int_0^p \frac{2M(q)}{V(q)} dq\right) dp}{\int_0^1 \exp\left(-\int_0^p \frac{2M(q)}{V(q)} dq\right) dp}$$

such that $S(0) = 0$ and $\rho(p_0) = S(p_0)/S(1)$

- Using the Kolmogorov backward equation, one can derive an analytical expression for $\tau(p_0)$.

$$\tau(p_0) = 2S(1)\left[\int_{p_0}^1 \frac{\rho(y)(1-\rho(y))}{\exp(-A(y))V(y)} dy + \frac{1-\rho(p_0)}{\rho(p_0)} \int_0^{p_0} \frac{\rho(y)^2}{\exp(-A(y))V(y)} dy\right]$$

General solution available for different processes

??

For the neutral Wright-Fisher model

$$\tau(p_0) = -2N \frac{1-p_0}{p_0} \log(1-p_0)$$

For $p_0 = 1 / N$, we have approximately $\tau(1/N) = 2N$

Evolutionary game theory

Assume frequency-dependent selection

- Consider two types A and B, with frequencies x_A and x_B, respectively.
- The vector $x(t) = (x_A(t), x_B(t))^T$ describes the population.
- Denote by $f_A(x(t))$ the fitness of A and by $f_B(x(t))$ the fitness of B. The average fitness is $\phi(x) = x_A f_A(x) + x_B f_B(x)$.
- The deterministic selection dynamics are given by

$$\begin{aligned} \dot{x}_A &= x_A [f_A(x) - \phi(x)] \\ \dot{x}_B &= x_B [f_B(x) - \phi(x)] \end{aligned}$$

Equilibria

- With $x := x_A$ and $1 - x = x_B$, the system is equivalent to

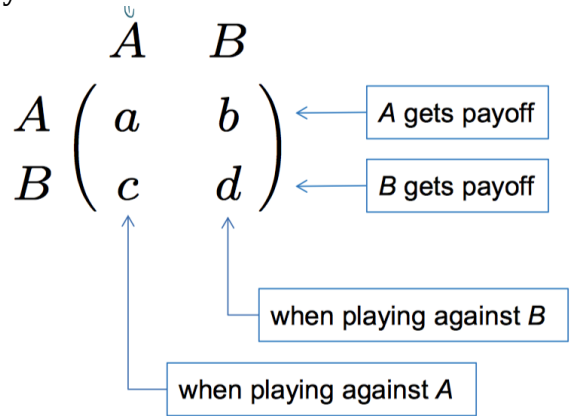
$$\dot{x} = x(1-x)\,[f_A(x) - f_B(x)]$$

- Equilibria: $x = 0$, $x = 1$, and $\{x \in (0, 1) \mid f_A(x) = f_B(x)\}$
- $x^* = 0$ is stable if $f_A(0) < f_B(0)$
- $x^* = 1$ is stable if $f_A(1) > f_B(1)$
- An interior equilibrium x^* is stable if

$$\frac{\partial f_A}{\partial x}(x^*) < \frac{\partial f_B}{\partial x}(x^*)$$

Evolutionary games with two players

Payoff matrix



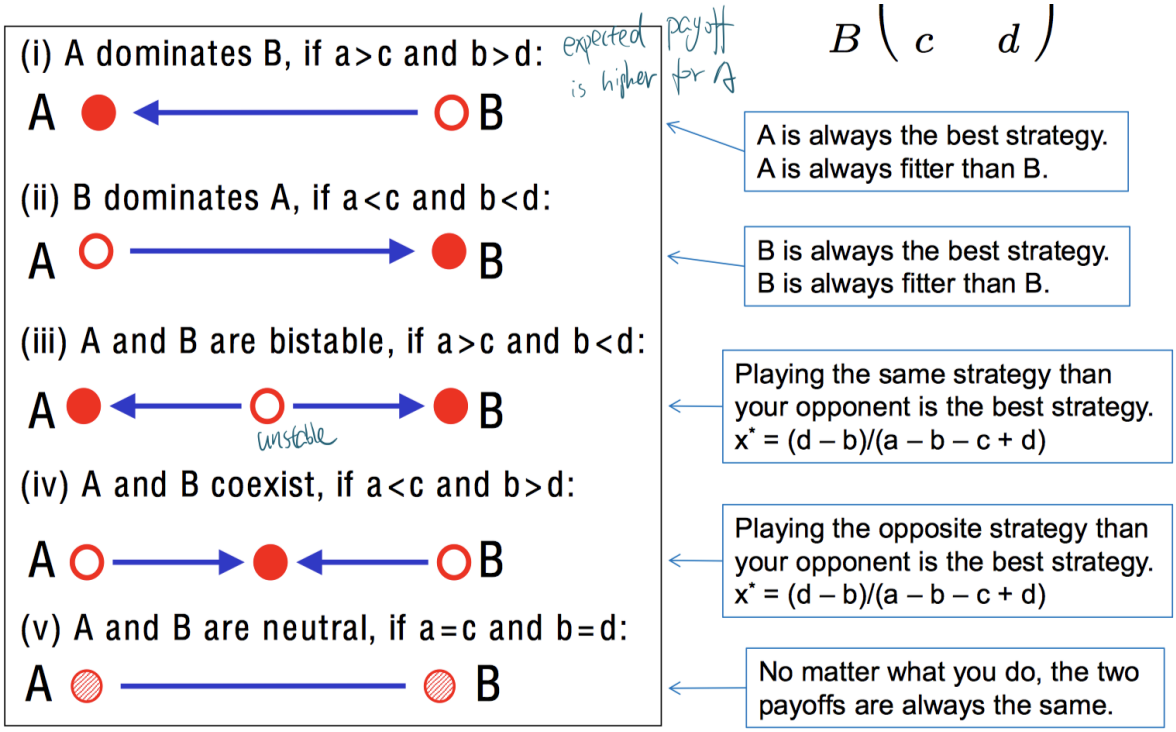
Fitness = expected payoff

$$f_A(x_A, x_B) = ax_A + bx_B$$
$$f_B(x_A, x_B) = cx_A + dx_B$$

- Setting $x := x_A$ we have $1 - x = x_B$ and we obtain the selection dynamics

$$\dot{x} = x(1-x)\left[(a-b-c+d)x + b - d\right]$$

$$B \begin{pmatrix} c & d \end{pmatrix}$$



Nash equilibrium

- **Definition**: If two players play the same strategy and neither player can increase its payoff by changing strategy, then the strategy is at *Nash equilibrium*. Mathematically,
  - A is a strict Nash equilibrium, if a > c.
  - A is a Nash equilibrium, if a ≥ c.
  - B is a strict Nash equilibrium, if d > b.
  - B is a Nash equilibrium, if d ≥ b.

$$\begin{array}{cc} & \begin{array}{cc} A & B \end{array} \\ \begin{array}{c} A \\ B \end{array} & \begin{pmatrix} a & b \\ c & d \end{pmatrix} \end{array}$$

- A Nash equilibrium is the best reply to itself.

Evolutionary stable strategy (ESS)

- Suppose an infinitesimally small quantity ϵ of B invades an all-A population. Then selection will oppose invasion, if

$$f_A(1-\epsilon) > f_B(\epsilon)$$
$$a(1-\epsilon)+b\epsilon > c(1-\epsilon)+d\epsilon$$

i.e., if in the limit as $\epsilon \to 0$

$$(a > c) \text{ or } (a = c \text{ and } b > d)$$

- **Definition**: A is ESS if either (i) $a > c$, or (ii) $a = c$ and $b > d$

$(x^* = 1\ is\ stable\ equilibrium)$

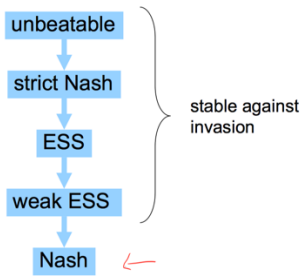
More than two strategies

$$\begin{array}{c} \\ S_1 \\ S_2 \\ \vdots \\ S_n \end{array} \begin{array}{c} \begin{array}{cccc} S_1 & S_2 & \dots & S_n \end{array} \\ \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & \ddots & & \\ \vdots & & & \vdots \\ a_{n1} & & \dots & a_{nn} \end{pmatrix} \end{array}$$

- The payoff for strategy S_i versus S_j is $E(S_i, S_j) = a_{ij}$.

- S_k is a *strict Nash* equilibrium, if $a_{kk} > a_{ik}$ for all $i \neq k$.
- S_k is a *Nash* equilibrium, if $a_{kk} \geq a_{ik}$ for all $i \neq k$.
- S_k is *ESS* if for all $i \neq k$, either
  - $a_{kk} > a_{ik}$, or
  - $a_{kk} = a_{ik}$ and $a_{ki} > a_{ii}$
- S_k is *weak ESS* if for all $i \neq k$, either
  - $a_{kk} > a_{ik}$, or
  - $a_{kk} = a_{ik}$ and $a_{ki} \geq a_{ii}$
- S_k is *unbeatable* if for all $i \neq k$, both
  - $a_{kk} > a_{ik}$, **and**
  - $a_{ki} > a_{ii}$

Hierarchy of concepts



Fitness among n players

- Fitness = expected payoff:

$$f_i(x) = f_{S_i}(x) = \sum_{j=1}^{n} x_j a_{ij}$$

- Average population fitness:

$$\phi(x) = \sum_{i=1}^{n} x_i f_i(x)$$

The replicator equation

$$\dot{x}_i = x_i \left[f_i(x) - \phi(x) \right] \qquad i = 1, \ldots, n$$

$$x_1 + \cdots + x_n = 1$$

Rock-paper-scissors (RPC)

- The general RPC game is defined by the payoff matrix

$$A = \begin{pmatrix} 0 & -a_2 & b_3 \\ b_1 & 0 & -a_3 \\ -a_1 & b_2 & 0 \end{pmatrix}$$

- Case 1: det(A) > 0
  - Unique interior equilibrium which is globally stable
  - Trajectories converge to this point as damped oscillations
- Case 2: det(A) < 0
  - Unique interior equilibrium which is unstable
  - Trajectories converge to the boundary of the simplex in oscillations with increasing amplitude

Hawks and doves

from the perspective of individual selection (rather than group selection).

There are two basic strategies: Hawks (H) escalate fights, whereas doves (D) retreat.

The benefit of winning a fight is b.

The cost of injury is c.

50% chance to win

$$\begin{array}{c} \\ H \\ D \end{array} \begin{array}{c} \begin{array}{cc} H & \quad D \end{array} \\ \begin{pmatrix} (b-c)/2 & b \\ 0 & b/2 \end{pmatrix} \end{array}$$

Hawk wins

Often, the cost of injury will be larger than the benefit of escalation (b < c). If b < c, then neither strategy is a Nash equilibrium. Hawks and doves can coexist. The equilibrium hawk

frequency is $x_H^* = \frac{b}{c}$

Mixed strategies

- Consider a strategy that plays "hawk" with probability p and "dove" with probability 1 – p.
- The space of strategies is the interval [0, 1], rather than the discrete space {H, D} as before.
- The payoff for strategy p_1 versus strategy p_2 is

$$\begin{aligned} E(p_1, p_2) \;=\; & p_1 p_2\, E(H,H) \\ & + p_1(1-p_2)\, E(H,D) \\ & + (1-p_1) p_2\, E(D,H) \\ & + (1-p_1)(1-p_2)\, E(DD) \end{aligned}$$

- We find

$$E(p_1, p_2) = \frac{b}{2}\left(1 + p_1 - p_2 - \frac{c}{b} p_1 p_2\right)$$

- The strategy p* = b/c is evolutionary stable:

$$\begin{aligned} E(p^*, p^*) &= (b/2)[1 - (b/c)] = E(p, p^*) \\ E(p^*, p) &= (b/2)[1 + (b/c) - 2p] \\ E(p, p) &= (b/2)[1 - (c/b)p^2] \end{aligned}$$

Thus p* is Nash, not strict Nash, and ESS, because E(p*, p) > E(p, p) for all $p \neq p^*$.

Cooperation and defection

The general Prisoner's Dilemma game

$$\begin{array}{c} \\ C \\ D \end{array} \begin{array}{c} \begin{array}{cc} C & D \end{array} \\ \begin{pmatrix} R & S \\ T & P \end{pmatrix} \end{array}$$

$$T > R > P > S \quad \text{and} \quad R > \frac{T+P}{2}$$

- DC: Temptation to defect
- CC: Reward for mutual cooperation
- DD: Punishment for mutual defection
- CD: Sucker's payoff

1. Direct reciprocity

- The Prisoner's Dilemma game is repeated m times.
- Consider two strategies:
  - GRIM: Cooperate initially, then as long as opponent does not defect
  - ALLD: Always defect

$$\begin{array}{c} \\ \text{GRIM} \\ \text{ALLD} \end{array} \begin{array}{c} \begin{array}{cc} \text{GRIM} & \text{ALLD} \end{array} \\ \begin{pmatrix} mR & S+(m-1)P \\ T+(m-1)P & mP \end{pmatrix} \end{array}$$

mutual defection

- If $m > (T-P)/(R-P)$, then GRIM is strict Nash, ALLD cannot invade. However, ALLD is also strict Nash.
- Hence, direct reciprocity can only stabilize cooperation. but cannot initiate cooperation

2. Defecting in the last round (know the number of rounds)

$$\begin{array}{c} \\ \text{GRIM} \\ \text{GRIM*} \end{array} \begin{array}{c} \begin{array}{cc} \text{GRIM} & \text{GRIM*} \end{array} \\ \begin{pmatrix} mR & (m-1)R+S \\ (m-1)R+T & (m-1)R+P \end{pmatrix} \end{array}$$

GRIM* dominates GRIM, GRIM → GRIM*.

But then:

GRIM → GRIM* → GRIM** → … → GRIM*$^{(m)}$ = ALLD

3. **Variable number of rounds** (Defecting in the last round is no longer possible) Suppose that after each round there is a probability w that another round will be played.
   - Thus, the expected number of rounds is

$$\begin{aligned} \bar{m} &= (1-w)+2w(1-w)+3w^2(1-w)+\ldots \\ &= 1-w+2w-2w^2+3w^2-3w^3+\ldots \\ &= 1+w+w^2+w^3+\cdots=\frac{1}{1-w} \end{aligned}$$

   - The payoff for GRIM versus ALLD is

$$\begin{array}{c} \\ \text{GRIM} \\ \text{ALLD} \end{array} \begin{array}{c} \begin{array}{cc} \text{GRIM} & \text{ALLD} \end{array} \\ \begin{pmatrix} \bar{m}R & S+(\bar{m}-1)P \\ T+(\bar{m}-1)P & \bar{m}P \end{pmatrix} \end{array}$$

   - GRIM is evolutionary stable if $\bar{m} > (T-P)/(R-P)$.

4. **Tit-for-tat**: start with cooperation, then do whatever your opponent has done in the previous round, i.e., answer C for C and D for D.

$$\begin{array}{c} \\ \text{TFT} \\ \text{ALLD} \end{array} \begin{array}{c} \begin{array}{cc} \text{TFT} & \text{ALLD} \end{array} \\ \begin{pmatrix} \bar{m}R & S+(\bar{m}-1)P \\ T+(\bar{m}-1)P & \bar{m}P \end{pmatrix} \end{array}$$

- TFT can resist invasion by ALLD if $\bar{m} > (T-P)/(R-P)$.
- TFT can resume cooperation, if the opponent cooperates.

In the long run, the payoff is as low as choosing randomly between C and D because TFT is susceptible to noise. $E(TFT, TFT) = \frac{T+R+P+S}{4} < R$

5. Reactive strategies

- The strategy S(p, q) cooperates with
  - probability p if the opponent has cooperated in the previous move,
  - probability q if the opponent has defected in the previous move.

- The repeated Prisoner's Dilemma between two reactive strategies $S_1(p_1, q_1)$ and $S_2(p_2, q_2)$ is a Markov chain with state space {CC, CD, DC, DD}, where, e.g., "CD" means "I play C and you play D".
- The transition probability matrix M is

$$\begin{array}{c} \\ CC \\ CD \\ DC \\ DD \end{array} \begin{array}{c} \begin{array}{cccc} CC & CD & DC & DD \end{array} \\ \begin{pmatrix} p_1p_2 & p_1(1-p_2) & (1-p_1)p_2 & (1-p_1)(1-p_2) \\ q_1p_2 & q_1(1-p_2) & (1-q_1)p_2 & (1-q_1)(1-p_2) \\ p_1q_2 & p_1(1-q_2) & (1-p_1)q_2 & (1-p_1)(1-q_2) \\ q_1q_2 & q_1(1-q_2) & (1-q_1)q_2 & (1-q_1)(1-q_2) \end{pmatrix} \end{array}$$

Stationary distribution of the Markov chain

- Let $x(t) = (x_{CC}(t), x_{CD}(t), x_{DC}(t), x_{DD}(t))$ be the probability distribution of the game after t rounds.
- We have $x(t + 1) = x(t) \cdot M$.
- The payoff at the stationary distribution is

$$E(S_1, S_2) = Rs_1s_2 + Ss_1(1-s_2) + \\ + T(1-s_1)s_2 + P(1-s_1)(1-s_2)$$

where $s_1(p_1, p_2, q_1, q_2)$ and $s_2(p_1, p_2, q_1, q_2)$ are the probabilities that players 1 and 2, respectively, cooperate in the stationary distribution. ??

6. **Generous Tit-for-tat (GTFT)** i.e.GTFT = S(1, 1/3)

E(GTFT, GTFT) is close to R, because GTFT can correct mistakes

GTFT for the general Prisoner's Dilemma

- GTFT = S(p, q) with

$$p = 1 \quad \text{and} \quad q = \min\left\{1 - \frac{T-R}{R-S}, \frac{R-P}{T-P}\right\}$$??

- This is the highest level of forgiveness, q, that is still resistant against invasion by ALLD.
- Among all reactive strategies that can resist ALLD, GTFT leads to the highest payoff for the population adopting it.

7. Memory-one strategies

- We now consider strategies that decide between C or D based on both the opponent's and one's own last move.
- The conditional probabilities to cooperate given that the last round was CC, CD, DC, DD are p_1, p_2, p_3, p_4.
- The Markov chain for a game between S(p_1, p_2, p_3, p_4) and S'(p'_1, p'_2, p'_3, p'_4) is defined by

$$\begin{array}{c} \\ CC \\ CD \\ DC \\ DD \end{array} \begin{array}{c} \begin{array}{cccc} CC & CD & DC & DD \end{array} \\ \begin{pmatrix} p_1p'_1 & p_1(1-p'_1) & (1-p_1)p'_1 & (1-p_1)(1-p'_1) \\ p_2p'_3 & p_2(1-p'_3) & (1-p_2)p'_3 & (1-p_2)(1-p'_3) \\ p_3p'_2 & p_3(1-p'_2) & (1-p_3)p'_2 & (1-p_3)(1-p'_2) \\ p_4p'_4 & p_4(1-p'_4) & (1-p_4)p'_4 & (1-p_4)(1-p'_4) \end{pmatrix} \end{array}$$

- ALLD = S(0, 0, 0, 0)
- ALLC = S(1, 1, 1, 1)
- TFT = S(1, 0, 1, 0)
- GTFT = (1, 1/3, 1, 1/3)
- Reactive strategies = {S(p_1, p_2, p_3, p_4) | $p_1 = p_3$, $p_2 = p_4$}

own move in the previous round doesn't matter

8. Win-stay, lose-shift (WSLS)

- WSLS = S(1, 0, 0, 1), i.e., cooperate after CC or DD, defect after CD or DC. (stay / shift / shift / stay)
- WSLS stays with high payoffs T or R, and shifts with low payoffs P or S.

- WSLS is stable against invasion by ALLD if R > (T + P)/2,

WSLS is a deterministic corrector, whereas GTFT is a stochastic corrector. WSLS dominates ALLC. GTFT does not dominate ALLC.

Evolutionary games in finite populations

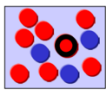
Expected payoff

- Let i denote the number of A individuals.
- The expected payoff for A and B, respectively, is

$$\begin{array}{c} \\ A \\ B \end{array} \begin{array}{c} \begin{array}{cc} A & B \end{array} \\ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \end{array}$$

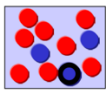
Fitness of type A individual; (i-1) times to meet A

$$F_i = \frac{(i-1)a + (N-i)b}{N-1}$$



Fitness of type B individual

$$G_i = \frac{ic + (N-i-1)d}{N-1}$$



Selection opposing invasion

- We say that *selection opposes A invading B*, if

$$\begin{aligned} & F_1 < G_1 \\ \iff \quad & b(N-1) < c + d(N-2) \end{aligned}$$

The condition is independent of a. For N = 2, we have b < c

Intensity of selection

final fitness ← $f_i = 1 - w + wF_i$ → game

$$g_i = 1 - w + wG_i$$

- If w = 0, then the game does not contribute to fitness.
- If w = 1, then fitness is entirely determined by the payoff.
- The limit $w \to 0$ is referred to as *weak selection*.
- The parameter w cancels out in the deterministic replicator equation, but plays an important role in the stochastic process describing finite populations.

Fixation probability in Moran process

- Because $P_{i,i-1} / P_{i,i+1} = g_i / f_i$, we find

A reach fixation

$$\rho_A = \frac{1}{1 + \sum_{k=1}^{N-1} \prod_{i=1}^{k} (g_i / f_i)}$$

When applying weak selection limit

- Taylor expansion gives, for $w \to 0$,

$$\rho_A \approx \frac{1}{N} \frac{1}{1-(\alpha N-\beta)w/6}$$

where α = a + 2b – c – 2d and β = 2a + b + c – 4d.

- If ρ_A > 1/N, we say that *selection favors the fixation of A*. relative to neutral process
- This condition is equivalent to $\alpha N > \beta$, or equivalently

$$a(N-2)+b(2N-1) > c(N+1)+d(2N-4)$$

- For N = 2, this becomes E(A, B) = b > c = E(B, A).

- For large N, the inequality becomes

$$a+2b > c+2d$$

- Suppose a > c and b < d, then both A and B are best replies to themselves.
  - If the frequency of A is high, then A has higher fitness
  - If the frequency of B is high, then B has higher fitness
  - The equilibrium ($F_i = G_i$) relative A-allele frequency is

} since best replies to themselves

$$x^* = \frac{d-b}{a-b-c+d}$$

The 1/3 law (weak selection, large population)

- Combining x* and the inequality, we obtain for $w \ll 1$,

$$\rho_A > 1/N \iff x^* < 1/3$$

If the unstable equilibrium occurs at a frequency smaller than 1/3, then in a large finite population, in the limit of weak selection, selection favors the fixation of A in B

Comparison with infinite population

Science and Engineering

Evolutionary stability in finite populations

- Recall that B is ESS if
  (i) d > b, or (ii) d = b and a < c.

$$\begin{array}{c} \\ A \\ B \end{array} \begin{array}{c} \begin{array}{cc} A & B \end{array} \\ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \end{array}$$

- **Definition:**
  B is ESS_N if the following two conditions hold:
  1. Selection protects against **invasion**, i.e., a single A mutant has lower fitness in a B population, $F_1 < G_1$, or equivalently

$$b(N-1) < c+d(N-2)$$

  2. Selection protects against **replacement**, i.e., ρ_A < 1/N for all w > 0. For $w \ll 1$, this condition is equivalent to

weak selection limit

$$a(N-2)+b(2N-1) < c(N+1)+d(2N-4)$$

Small versus large populations

- For N = 2, B is ESS_N if
  1. b < c
  2. b < c

  Thus traditional ESS is neither necessary nor sufficient for ESS_N.

- For large N, B is ESS_N if
  1. b < d
  2. $x^* > 1/3$

  Thus traditional ESS is necessary but not sufficient for ESS_N

Risk dominance

- A is *risk dominant* over B, if $\rho_A > \rho_B$.
- For weak selection, we have

$$\frac{\rho_A}{\rho_B} = \prod_{i=1}^{N-1} \frac{f_i}{g_i} \approx 1 + w\,[N(a+b-c-d)/2 + d - a]$$

and $\rho_A > \rho_B$ is equivalent to

$$(N-2)(a-d) > N(c-b)$$

- For large N, we obtain a – d > c – b, or equivalently x* < 1/2.
- If both A and B are strict Nash (a > c and b < d) then the risk dominant strategy has a higher fixation probability.

??

When A = TFT B = ALLD

Department of Biosystems Science and Engineering

TFT versus ALLD in finite populations

- Selection favors the replacement of ALLD by TFT if $\rho_{TFT} > 1/N$, or equivalently

$$m > \frac{T(N+1) + P(N-2) - S(2N-1)}{(R-P)(N-2)}$$

- For large N, this inequality becomes

against replacement

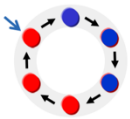
$$m > \frac{T + P - 2S}{R - P}$$

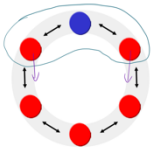
Evolutionary graph theory

Markov chain on the directed cycle

- Suppose A has fitness 1 and B has fitness r.
- We keep track of the number of B individuals, m:
  - To reduce m by one, the A individual immediately preceding the B cluster must be chosen.

$$P_{m,m-1} = \frac{1}{N - m + rm}$$



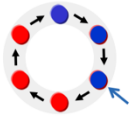


  - To increase m by one, the B individual at the end of the cluster has to be chosen.

$$P_{m,m+1} = \frac{r}{N - m + rm}$$

r → selected individual B

N − m + rm = number of type A



Fixation probability on the directed and bidirected cycle is same as Moran process since $\gamma_m = \frac{P_{m,m-1}}{P_{m,m+1}} = \frac{1}{r}$

Amplifiers and suppressors

- We write ρ_G < x (> x, = x) if the (in)equality holds for any configuration of the population, i.e., for any position of the invading mutant on the graph G.
- A graph G is an *amplifier of selection* if, for an advantageous mutant (r > 1), the fixation probability is greater than in the Moran process, $\rho_G > \rho_{Moran}$.
- A graph G is a *suppressor of selection* if, for an advantageous mutant (r > 1), the fixation probability is smaller than in the Moran process, $\rho_G < \rho_{Moran}$.
- Similar definitions apply to disadvantageous mutants (r < 1).
- The strongest suppressors of selection have ρ_G = 1/N, i.e., the fixation probability is completely independent of r.

The line and The burst

The isothermal theorem

- Define the temperature of a vertex j as $T_j = w_{1j} + \ldots + w_{nj}$.
- Hot vertices change more often than cold vertices. easily be replaced
- If all vertices have the same temperature, the graph is *isothermal*.

- **Theorem**: $\rho_G = \rho_{Moran}$ if and only if G is isothermal.

- We describe the configuration of the population on the graph by $v = (v_1, \ldots, v_N)$, where $v_i = 0$ indicates occupation of vertex i by A and $v_i = 1$ occupation by B.
- The number of B individuals is $m = v_1 + \ldots + v_N$.

only $i=B$ $j=A$ exists

$$P_{m,m+1} = \frac{r \sum_{i,j} w_{ij} v_i (1 - v_j)}{rm + N - m}$$

$i \to j$
B A

$$P_{m,m-1} = \frac{\sum_{i,j} w_{ij} (1 - v_i) v_j}{rm + N - m}$$

$i \to j$
A B

- Now, $\rho_G = \rho_{Moran}$ if and only if

$$\frac{P_{m,m-1}}{P_{m,m+1}} = \frac{1}{r}$$

for all configurations v.

  - This is the case if and only if, for all v,

$\frac{P_{m,m-1}}{P_{m,m+1}} = \frac{1}{r}$

$$\sum_{i,j} w_{ij} (1 - v_i) v_j = \sum_{i,j} w_{ij} v_i (1 - v_j)$$

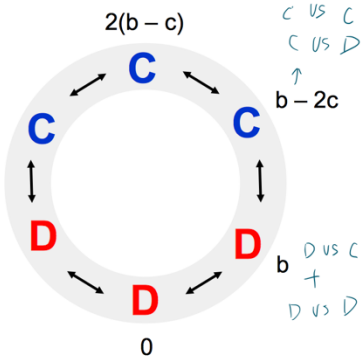
Condition for isothermal graphs

- Directed cycle
- Cycle
- All symmetric graphs, $w_{ij} = w_{ji}$

Payoff on graph

$$\begin{array}{c} \\ C \\ D \end{array} \begin{array}{c} \begin{array}{cc} C & D \end{array} \\ \begin{pmatrix} b-c & -c \\ b & 0 \end{pmatrix} \end{array}$$

- Cooperators help all neighbors, defectors do not.
- Cooperators pay a cost c for each neighbor; each neighbor of a cooperator receives a benefit b.
- Payoff from all interactions are added up.



Spatial models of the evolution of solid tumors

Cellular automata

Eden growth model

- Two states: unoccupied (S_0) and occupied (S_1)
- Neighbourhood: usually adjacent sites (von Neumann)
- Update rule: with each iteration, a site in the neighbourhood of an S_1 site switches from S_0 to S_1
- Cells are thus added to the surface of a cluster

Eden growth model variants (different way to choose sites to be updated)

- **A**vailable site-focussed: randomly choose an S_0 site that adjoins at least one S_1 site, and switch it from S_0 to S_1
- **B**ond-focussed: randomly choose an S_1 site with probability proportional to number of adjoining S_0 sites, then randomly choose an S_0 neighbour and switch it to S_1
- **C**ell-focussed: randomly choose an S_1 site that adjoins at least one S_0 site, then randomly choose an S_0 neighbour and switch it to S_1

With mutation

- Multiple occupied states $\{S_1, S_2, \ldots\}$
- Mutation: probabilities of $S_i \longrightarrow S_j$ for $i, j > 0$
- Models usually apply mutation only at the time of division
- States may confer different division rates

- Example: (cells are accumulating mutations)
  - all mutation probabilities are zero except in the case $S_i \longrightarrow S_{i+1}$
  - a site in state S_i that adjoins at least one site in state S_0 divides with probability proportional to $(1+s)^i$ (selection coeff)

Deme-based models

Each site (deme) contains multiple cells (e.g. demes correspond to glands or microenvironmental niches). Assume cells within demes are well-mixed and obey a local subpopulation model such as the Moran process or Wright-Fisher process. Cells can migrate between demes.

Assume replacement probability is weighted by cell fitness ($\frac{(1+s)n_i}{N}$) The replacement has a local parent with probability (1 − m) or a parent from a neighboring deme with probability m

The aim of the present article is to study the dynamics of the into c

- Consider a mutant invading an infinite row of demes
- Let $\mathbf{n} = \{\ldots, n_{i-1}, n_i, n_{i+2}, \ldots\}$ be the vector of mutant population sizes along the row of demes
- Transition probability densities:

prob. density that n_i increases by one: $W_i^+(\mathbf{n}) = \frac{\mu(1+s)}{N}(N - n_i)\left[n_i + \frac{m}{2}n_i''\right]$, (wild type)

prob. density that n_i decreases by one: $W_i^-(\mathbf{n}) = \frac{\mu}{N}n_i\left[(N - n_i) - \frac{m}{2}n_i''\right]$, (1)

where μ is the death rate, s is the difference in fitness, N is the deme population size, and

$$n_i'' = (n_{i-1} + n_{i+1} - 2n_i).$$

A variant of Fisher equation

Two approximations

- As in the non-spatial Moran process, we can take a
① diffusion approximation of (1), in which case we obtain

$$\frac{\partial u}{\partial t} = D[1 + s(1-u)]\frac{\partial^2 u}{\partial x^2} + \mu s u(1-u),$$

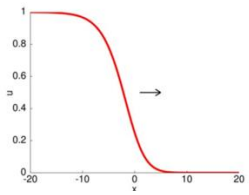
(expected frequency of mutants) (② transfer the discrete variable distance to the continuous variable x)

where $u = \langle n_i \rangle / N$, x is distance along the row of demes, and D is a diffusion coefficient

Fisher's equation

$$\frac{\partial u}{\partial t} = D\frac{\partial^2 u}{\partial x^2} + ru(1-u).$$

- Fisher's equation is difficult to solve in general
- If, at $t = 0$, u decreases monotonically and continuously from 1 to 0 over a finite distance (as below) then the equation has a *travelling wave* solution of the form $u(x, t) = U(x - ct)$, with $c = 2\sqrt{(rD)}$



Gillespie stochastic simulation algorithm

1. Initialise the system
2. Calculate event rates (birth, death, migration, etc.)
3. Randomly determine next event such that $P(\text{event} = E) = \text{rate}(E) \,/\, \sum \text{rates}$
4. Update the system
5. Advance the timer by $\delta t \sim \text{Exp}(1 \,/\, \sum \text{rates})$ (time between events)
6. Return to step 2 (until stop condition)

Branching processes in biology

(finite population but size may not constant)

Galton-Watson process

- A single ancestor lives for one unit of time after which it produces a random number of offspring, Z, according to a fixed probability distribution.
- Each offspring behaves independently and identical to the ancestor.
- Let Z_n be the number of individuals in generation n. $Z_0 = 1$, $Z_1 = Z$.
- The Galton-Watson process is the Markov chain

$$\{Z_n \mid n = 0, 1, 2, \ldots\}$$

defined on the non-negative integers.

Transition probability

- Set p_k = Prob(Z = k). Probability to produce k offsprings
- Let P(i, j) = Prob(Z_{n+1} = j | Z_n = i) be the transition probabilities of the time-homogeneous Markov chain.
- Note that P(1, k) = p_k. same transition prob over time
- P(2, j) = $p_0\, p_j + p_1\, p_{j-1} + p_2\, p_{j-2} + p_3\, p_{j-3} + \ldots + p_j\, p_0$
- In general, P(0, j) = δ_{0j}, and for i ≥ 1, $= \begin{cases} 1 & j=0 \\ 0 & j\neq 0 \end{cases}$

$$P(i,j) = p_j^{*i} = \sum_{k_1+\cdots+k_i=j} p_{k_1}\cdots p_{k_i} \quad \forall\, i \neq 0$$

- $\{p_k^{*i}\}_{k\geq 0}$ is the *i-fold convolution* of $\{p_k\}_{k\geq 0}$.

Probability generating function (pgf)

the main mathematical tool to study branching processes

- For the discrete random variable Z ~ $\{p_k\}_{k\geq 0}$, we define the *probability generating function* (pgf)

$$f(s) = \mathrm{E}\left[s^Z\right] = \sum_{k=0}^{\infty} p_k s^k \qquad s \in [0,1]$$

constant; all possible values for z

- The pgf generates the distribution *p*:

$$\frac{d^k f}{ds^k}(0) = k!\, p_k \qquad k \geq 0$$

kth derivative w.r.t s

Properties of the pgf

- Moments of *Z* :

$$\begin{aligned} \mathrm{E}[Z] &= f'(1) \\ \mathrm{Var}[Z] &= f'(1) + f''(1) - f'(1)^2 \end{aligned}$$

- Powers of *f* :

$$f(s) = \sum_j P(1,j)s^j \qquad f(1) = 1$$

$$[f(s)]^k = \sum_j P(k,j)s^j \qquad k \geq 1$$

kth fold convolution

$$
\begin{aligned}
f_{n+1}(s) &= \sum_j P_{n+1}(1,j)s^j \\
&= \sum_j \sum_k P_n(1,k)P(k,j)s^j \\
&= \sum_k P_n(1,k) \sum_j P(k,j)s^j \\
&= \sum_k P_n(1,k) f(s)^k \\
&= f_n(f(s)) = \cdots = f^{(n+1)}(s)
\end{aligned}
$$

Moments of Z_n

- We assume throughout that $p_0 + p_1 < 1$ and $p_j \neq 1$ for all j.
- If they exist, the moments of Z_n can be expressed in terms of the derivatives of f at s = 1.
- Set $m = E[Z] = E[Z_1] = f'(1)$ and $\sigma^2 = Var[Z] = f'(1) + f''(1) - f'(1)^2$. Then:

$$
\begin{aligned}
E[Z_n] &= m^n \\
Var[Z_n] &= \begin{cases} \frac{\sigma^2 m^{n-1}(m^n-1)}{m-1} & \text{if } m \neq 1 \\ n\sigma^2 & \text{if } m = 1 \end{cases}
\end{aligned}
$$

Extinction

- $Z_n = 0$ is an absorbing state.

$$
\begin{aligned}
\rho &= \text{Prob}(Z_i = 0 \text{ for some } i \geq 0) \\
&= \lim_{n\to\infty} \text{Prob}(Z_i = 0 \text{ for some } 1 \leq i \leq n) \\
&= \lim_{n\to\infty} \text{Prob}(Z_n = 0) \\
&= \lim_{n\to\infty} f_n(0)
\end{aligned}
$$

- Thus, we have to study the limit behavior of the pgf.

Extinction probability

- **Theorem**:
  The extinction probability of the Galton-Watson process $\{Z_n\}$ is the smallest non-negative root q of the equation f(s) = s.
  If $m \leq 1$ then $q = 1$. If $m > 1$ then $q < 1$.

- Criticality:

supercritical	$m > 1$	$\mathrm{E}[Z_n] \nearrow \infty$	$q < 1$
critical	$m = 1$	$\mathrm{E}[Z_n] = 1$	$q = 1$
subcritical	$m < 1$	$\mathrm{E}[Z_n] \searrow 0$	$q = 1$

Instability

- **Theorem:**

$$\lim_{n\to\infty} \mathrm{Prob}(Z_n = k) = 0 \qquad k \geq 1$$

$$\mathrm{Prob}\left(\lim_{n\to\infty} Z_n = 0\right) = q$$

$$\mathrm{Prob}\left(\lim_{n\to\infty} Z_n = \infty\right) = 1 - q$$

The multi-type Galton Watson process

1. Assume two types: type 0 (wild type) and type 1 (mutant) with counts $Z_0(t)$ and $Z_1(t)$, respectively, in generation $t \in \{0, 1, 2, \ldots\}$.
2. Each cell at the moment of division gives birth to two daughter cells, it meas the size of population is deterministic. A type 0 cell has type 1 offspring with probability α, the mutation rate. The mutation is irreversible: type 1 cells can not produce type 0 offspring

Probability generating function

- The components of the pgf $F = (F_0, F_1)$ are

$$F_0(s_0, s_1; t) = \mathrm{E}\left[s_0^{Z_0(t)} s_1^{Z_1(t)} \mid Z_0(0) = 1,\ Z_1(0) = 0\right]$$

$$F_1(s_0, s_1; t) = \mathrm{E}\left[s_0^{Z_0(t)} s_1^{Z_1(t)} \mid Z_0(0) = 0,\ Z_1(0) = 1\right]$$

$$F_0(s;t) = [(1-\alpha)F_0(s;t-1) + \alpha F_1(s;t-1)]^2$$

$$F_1(s;t) = [F_1(s;t-1)]^2$$

- We also write $F_i(t) = F_i(s; t)$, where $s = (s_0, s_1)$.

Differentiation of the recurrence equations

- Differentiation w.r.t. s_0 yields at s = (1,1), for F_1 and F_0 resp.,

$$\mathrm{E}[Z_0(t) \mid Z_i(0) = \delta_{1i}] = 2\mathrm{E}[Z_0(t-1) \mid Z_i(0) = \delta_{1i}] = 0$$

mutant cannot produce wild type

$$\mathrm{E}[Z_0(t) \mid Z_i(0) = \delta_{0i}] = 2(1-\alpha)\mathrm{E}[Z_0(t-1) \mid Z_i(0) = \delta_{0i}]$$

$$\Rightarrow \quad \mathrm{E}[Z_0(t) \mid Z_i(0) = \delta_{0i}] = [2(1-\alpha)]^t$$

the expected total number of wild type cells at time t.

The number of cells at time t

- The expected total number of cells is

$$\begin{aligned} N(t) &= \mathrm{E}[Z_0(t) + Z_1(t) \mid Z_i(0) = \delta_{0i}] \\ &= 2^t \end{aligned}$$

- Thus, the expected number of mutant cells is

$$\begin{aligned} r(t) &= \mathrm{E}[Z_1(t) \mid Z_i(0) = \delta_{0i}] \\ &= 2^t - [2(1-\alpha)]^t \\ &= 2^t[1-(1-\alpha)^t] \end{aligned}$$

The probability of a mutant-free population

- The probability of mutant cells being absent from the population at time t is

$$\begin{aligned} P_0(t) &= F_0(1,0;t) \\ &= \mathrm{E}\left[1^{Z_0(t)}\, 0^{Z_1(t)} \mid Z_i(0) = \delta_{0i}\right] \end{aligned}$$

expected value of indicator variable — whether $Z_1(t) = 0$

where

$$0^{Z_1(t)} = \begin{cases} 1 & \text{if } Z_1(t) = 0 \\ 0 & \text{else} \end{cases}$$

- Set $P_1(t) = F_1(1, 0; t)$.
- The recurrence equations at s = (1, 0) yield

$$P_0(t) = [(1-\alpha)P_0(t-1) + \alpha P_1(t-1)]^2$$
$$P_1(t) = [P_1(t-1)]^2$$

with initial conditions $P_0(0) = 1$ and $P_1(0) = 0$.

- We find $P_1(t) = 0$ for all t = 0, 1, 2, …
- Then, $P_0(1) = (1-\alpha)^2$
  $P_0(2) = [(1-\alpha)(1-\alpha)^2]^2 = (1-\alpha)^2(1-\alpha)^4$
  $P_0(3) = (1-\alpha)^2(1-\alpha)^4(1-\alpha)^8$
  …

$$P_0(t) = (1-\alpha)^{2^{t+1}-2}$$
$$P_1(t) = 0$$

=>

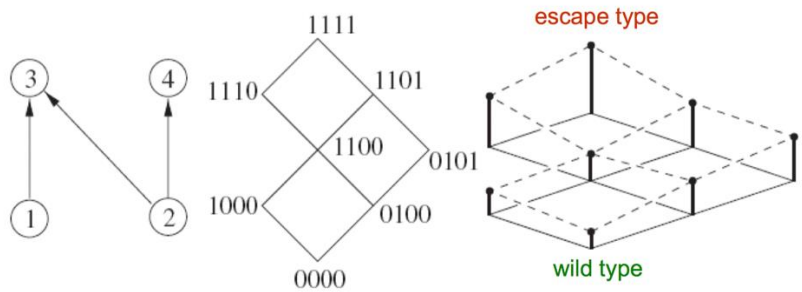
Summary of the irreversible 2-type GW process

$$N(t) = 2^t$$
$$r(t) = 2^t\left[1-(1-\alpha)^t\right]$$
$$P_0(t) = (1-\alpha)^{2(2^t-1)}$$

- For each fixed N, we can solve for P_0 to obtain

$$P_0(r) = \left(1-\frac{r}{N}\right)^{\frac{2(N-1)}{\log_2 N}}$$

Partially ordered sets ε: we don't know the order of all elements
cover relation: e 1 < e 2 is a cover relation, if there is no e' with e 1 < e' < e 2 .
order ideal: subset g in the poset ε that is closed downward. E.g. if $e_2 \in g$ *and* $e\,1 < e\,2$, *then* $e_1 \in g$
distributive lattice J(ε): set of all order ideals of ε
The genotype lattice : Let ε be a set of n = | ε | irreversible genetic events. The poset encodes constraints on the order in which mutations can accumulate. The order ideals g of J(ε) are the genotypes that can evolve subject to the order constraints. $\mathcal{G} = J(\varepsilon)$ is the genotype lattice.
Chain of length k: collection of k totally ordered subsets. $g_1 \subset g_2 \subset \cdots \subset g_k$. It actually represents the mutation pathways.
Fitness landscapes: a mapping f: $\mathcal{G} \to \mathbb{R}$



Evolutionary escape: the wildtype with lower fitness will extinct eventually. In order to escape extinction, it should be mutated fast to reach the escape state.
Mutational neighborhood: the set of genotypes h ∈ G that can be reached by mutation. $N(g) = \{h \in \mathcal{G} | g \subset h\}$
Assumption

1. Mutation: μ_e is the mutation rate of event e $\in \epsilon$. mutations are independent of each other.

The risk polynomial: $R(\mathcal{G}; f) = P_{01}(f) = \sum_{0=g_0 \subset g_1 \subset \cdots \subset g_k = 1} f_{g_1} f_{g_2} \cdots f_{g_{k-1}}$. It encodes the set of escape pathways in the genotype lattice.

The basic reproductive ratio R_g: $\frac{f_g}{1+f_g}$ if define $f_g = \frac{R_g}{1-R_g} = R_g + R_g^2 + R_g^3 + \cdots$. Assume $R_1 > 1 \; and \; R_g < 1 \; for \; g \neq 1$.

Consider a branching process on the type space G with a Poisson offspring distribution ρ_{gh}^k represents the probability that a single individual of type g produces k children of type h.

$$\rho_{gh}^k = \text{Pois}(k; u_{gh} R_g) = \frac{(u_{gh} R_g)^k}{k!} e^{-u_{gh} R_g}$$

number of event; λ; $\frac{\lambda^k}{k!} e^{-\lambda}$

ξ_g is the probability of escape (reaching 1 before extinction) starting with one individual of type g.

$$1 - \xi_g = \prod_{h \supseteq g} \sum_{k=0}^{\infty} (1 - \xi_h)^k \rho_{gh}^k$$

consider all lineages from g; $\binom{k}{k} \rho_{gh} (1-\xi_h)^k$ k offsprings extinct

Substituting the Poisson distribution, simplifying, and taking logarithms, we find

$$\log(1 - \xi_g) = -\sum_{h \supseteq g} \xi_h u_{gh} R_g$$

Recurrence equations

For $g \neq 1, \xi_g \ll 1 \; and \; (R_g)^2 \approx 0$,

$$\xi_g \approx R_g \sum_{h \supseteq g} \xi_h u_{gh}$$

probability of escape

$$\approx \frac{R_g}{1 - R_g} \sum_{h \supset g} \xi_h u_{gh} = f_g \sum_{h \supset g} \xi_h u_{gh}$$

$$\xi_0 \approx \xi_1 f_0 \prod_{e \in \mathcal{E}} \mu_e \mathcal{R}(\mathcal{G}; \mathbf{f})$$

probability to escape for totally mutants; product of all mutations

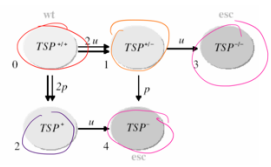
The risk of escape for N wild type pathogens is $1 - (1 - \xi_0)^N \approx 1 - e^{-\xi_0 N}$

The critical population size

- We define the critical population size
$$N^* = 1/\xi_0$$
- If N ≫ N*, then escape is almost certain.
- If N = N*, then the risk of escape is 1 − 1/e and the probability of successful intervention is 1/e. all N wild types evolve
- If N ≪ N*, then escape is almost impossible.

Example

- Remember: $\xi_g = f_g \sum_{h \supset g} \xi_h u_{gh}$



$$\xi_- = \xi_{-/-} = 1$$

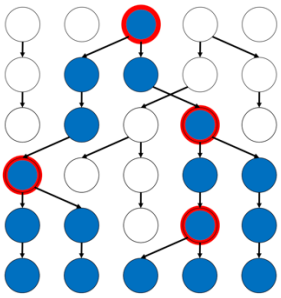
$$\xi_+ = f_+ u \xi_- = au$$

$$\xi_{+/-} = f_{+/-}(u\xi_{-/-} + p\xi_-) = a(u+p)$$

$$\xi_{+/+} = f_{+/+}(2p\xi_+ + 2u\xi_{+/-} + 2pu\xi_- + u^2\xi_{-/-}) = a(1+2a)\cdot u(u+2p)$$

The coalescent theory

Coalescent events in **WF process**



The probability that j genes have no common ancestor in the previous generation:

$\prod_{i=1}^{j-1}\left(1-\frac{i}{N}\right) = 1 - \binom{j}{2}N^{-1} + O(N^{-2})$

The coalescent time

- We measure time in units of *N* generations. ?
- Let T(*j*) be the coalescence time between *j* and *j* − 1 genes:

$$P(T(j) > t) = \left[\prod_{i=1}^{j-1}\left(1-\frac{i}{N}\right)\right]^{Nt} \longrightarrow \exp\left[-\binom{j}{2}t\right]$$

as $N \to \infty$.

The coalescence time is distributed **exponentially with parameter (j choose 2)=$\frac{j(j-1)}{2}$**

Time to the Most recent common ancestor (MRCA)

- For a sample of size *n*, the time to MRCA is

$$T_{\mathrm{MRCA}}(n) = \sum_{j=2}^{n} T(j)$$

- E[*T*(*j*)] = 1 / (*j* choose 2) = 2 / [*j*(*j* − 1)], hence:

$$\mathrm{E}[T_{\mathrm{MRCA}}(n)] = \sum_{j=2}^{n} \mathrm{E}[T(j)] = \sum_{j=2}^{n} \frac{2}{j(j-1)}$$

$$= 2\sum_{j=2}^{n}\left(\frac{1}{j-1} - \frac{1}{j}\right) = 2\left(1 - \frac{1}{n}\right)$$

the last coalesce event

- Note that $\mathrm{E}[T(2)] = 1$ and $\lim_{n\to\infty} \mathrm{E}[T_{\mathrm{MRCA}}(n)] = 2$.

- T(*j*) are independent and var[*T*(*j*)] = 1 / (*j* choose 2)^2

$$\begin{aligned}\mathrm{var}[T_{\mathrm{MRCA}}(n)] &= \sum_{j=2}^{n} \mathrm{var}[T(j)] = \sum_{j=2}^{n} \left(\frac{2}{j(j-1)}\right)^2 \\ &= 4\sum_{j=2}^{n}\left(\frac{1}{j-1}-\frac{1}{j}\right)^2 \\ &= 8\sum_{j=1}^{n}\frac{1}{j^2}+\frac{4}{n^2}-8\left(1-\frac{1}{n}\right)-4\end{aligned}$$

- $\mathrm{var}[T(2)] = 1,\ \lim_{n\to\infty} \mathrm{var}[T_{\mathrm{MRCA}}(n)] = \frac{8\pi^2}{6} - 12 \approx 1.16$

Therefore $\boldsymbol{T_{MRCA}(n)}$ **is dominated by** $\boldsymbol{T(2)}$

Estimation of mutation rate

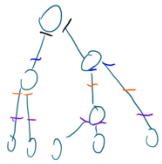
Assumption

1. We assume **a Poisson process** that puts down mutations independently on all branches at rate $\theta/2$, where θ = 2 N u is the scaled mutation rate.
2. We assume an infinite number of sites (loci) and each mutation to affect a different nucleotide site. **The infinite sites model** is appropriate for long DNA sequences under **neutral evolution**.

Number of segregating sites, S: number of sites in the genome where lineages are segregated

- Under the infinite sites model, *S* is equal to the total number of mutations of the genealogy.
- The total branch length is

$$T_{\mathrm{tot}}(n) = \sum_{j=2}^{n} jT(j)$$



- Hence,

$$\mathrm{E}[S] = \frac{\theta}{2}\mathrm{E}[T_{\mathrm{tot}}(n)] = \frac{\theta}{2}\sum_{j=2}^{n} j\frac{1}{\binom{j}{2}} = \theta\sum_{j=2}^{n}\frac{1}{j-1} = \theta\, c_n$$

constant that depend on n

For average pairwise nucleotide distance, K, $E[K] = \frac{\theta}{2}2E[T(2)] = \theta$

Under the neutral infinite sites model, we have two different estimates of the mutation rate:

$$\mathrm{E}[K] = \theta = c_n^{-1}\mathrm{E}[S]$$

Selection changes the allele frequencies in the population and affects these two estimates in different ways:

- *S* ignores allele frequency changes, but is sensitive to low-frequency deleterious alleles. allele frequency does not matter to S
- *K* is strongly affected by allele frequencies, but largely insensitive to low-frequency deleterious alleles.

detecting selection (deviation from neutrality)

Tajima's D

$$D = \frac{\hat{K} - c_n^{-1}\hat{S}}{\sqrt{\hat{V}}} \quad \text{(Tajima's D)}$$

difference between estimates

where $\hat{K}$, $\hat{S}$, $\hat{V}$ are estimates of *K*, *S*, and the variance of $\hat{K} - c_n^{-1}\hat{S}$, respectively.

- The distribution of *D* under the null hypothesis of no selection is approximated by simulations of the coalescent.

Then compare the empirical estimate of D to the distribution under coalescent simulation