



TARTU ÜLIKOOL



An evolution model that satisfies detailed balance

Jüri Lember

with C. Watkins, F. Zucca and D. Bertacchi

University of Tartu, Estonia

23.06.2021

Breeding and mutations

$\mathcal{X} = \{1, \dots, K\}$ is a finite set of possible genomes (types).

$\xi = \xi_1, \xi_2, \dots$ is an **exchangeable** \mathcal{X} -valued stochastic process – **the breeding process**.
For any **population of n genomes** $(x_1, \dots, x_n) \in \mathcal{X}^n$, we define

$$P_\xi(x_1, \dots, x_n) := P(\xi_1 = x_1, \dots, \xi_n = x_n)$$

By **DeFinetti theorem**, there exists a **prior measure** π on the set of all probability measures on \mathcal{X} (simplex), denoted by $\mathcal{P} = \{(q_1, \dots, q_K) : q_k \geq 0, \sum_k q_k = 1\}$, such that for every $(x_1, \dots, x_n) \in \mathcal{X}^n$ (with $q(k) = q_k$)

$$P_\xi(x_1, \dots, x_n) = \int_{\mathcal{P}} \prod_{i=1}^n q(x_i) \pi(dq) = \int_{\mathcal{P}} \prod_{k=1}^K q_k^{n_k} \pi(dq),$$

where $n_k(x_1, \dots, x_n)$ is the number of k -types in (x_1, \dots, x_n) , $n_1 + \dots + n_K = n$.

Breeding and mutations: an important special case

When π is $\text{Dir}(\alpha_1, \dots, \alpha_K)$, then ξ is (generalized) Polya urn model also known as Dirichlet-categorical process, where $(\alpha = (\alpha_1, \dots, \alpha_K))$

$$P_\xi(\xi_{n+1} = k \mid x_1, \dots, x_n, \alpha) := \frac{n_k + \alpha_k}{n + |\alpha|}, \quad |\alpha| = \alpha_1 + \dots + \alpha_K.$$

It follows that

$$P_\xi(x_1, \dots, x_n) = \frac{(\alpha_1)_{n_1} \cdots (\alpha_K)_{n_K}}{(|\alpha|)_n}, \quad (\alpha)_n := \alpha(\alpha + 1) \cdots (\alpha + n - 1).$$

By the formula above, it is clear that the random vector of counts has the distribution

$$P_\xi(n_1, \dots, n_K) = \frac{n!}{n_1! \cdots n_K!} \frac{(\alpha_1)_{n_1} \cdots (\alpha_K)_{n_K}}{(|\alpha|)_n}. \quad (0.1)$$

Connection with Moran model

Basic Moran model (without mutation and selection): A population (n_1, \dots, n_K) is considered as an urn consisting of n balls of K colors/types, n_k is the number of balls of color k . Two balls are drawn with replacement. First of them is considered to **breed** and another to **die**. If the first (the type that breeds) is i and another (the one that dies) is j ($i < j$), then then there is transition

$$(n_1, \dots, n_i, \dots, n_j, \dots, n_K) \rightarrow (n_1, \dots, n_i+1, \dots, n_j-1, \dots, n_K), \quad \text{with probability } \frac{n_i}{n} \frac{n_j}{n}.$$

One obtains a Markov chain (with state space being count vectors) that has K absorbing states: $(0, \dots, 0, n, 0, \dots, 0)$. Sooner or later the model is in an absorbing state – **fixation**.

Connection with Moran model

Several ways to add **mutations** to basic Moran model.

Let $u_i > 0$ $i = 1, \dots, K$ be the probability that randomly chosen ball mutates into i type, $u = u_1 + \dots + u_K$ is the overall probability of mutation and then the transition above (i ball is replaced by j ball) is

$$\left(\frac{n_i}{n} (1 - u) + u_j \right) \frac{n_j}{n}.$$

Now the Markov chain is irreducible and has stationary measure (0.1):

$$P_\xi(n_1, \dots, n_K) = \frac{n!}{n_1! \cdots n_K!} \frac{(\alpha_1)_{n_1} \cdots (\alpha_K)_{n_K}}{(|\alpha|)_n},$$

where

$$\alpha_k = \frac{nu_k}{1 - u} \Leftrightarrow u_k = \frac{\alpha_k}{|\alpha| + n}, \quad k = 1, \dots, K, \quad |\alpha| = \frac{nu}{1 - u} \Leftrightarrow u = \frac{|\alpha|}{|\alpha| + n}.$$

Connection with Moran model

This establishes the connection with Dirichlet-categorical process – the probability that i type is born is now (recall $u_i = \frac{\alpha_i}{|\alpha|+n}$)

$$\frac{n_i}{n}(1 - u) + u_i = \frac{n_i}{n} \frac{n}{|\alpha| + n} + \frac{\alpha_i}{|\alpha| + n} = \frac{\alpha_i + n_i}{|\alpha| + n} = P_\xi(\xi_{n+1} = i \mid x_1, \dots, x_n, \alpha).$$

Hence including mutations (in a certain way) into basic Moran model yields to Dirichlet-categorical breeding process. In terms of Polya urn – there are $\alpha_1, \dots, \alpha_K$ prior balls and drawing a prior ball means mutation (makes it possible to see unseen ball in population).

There are other ways of including mutation such that the stationary distribution would be (0.1). Our breeding process ξ more general because it allows another prior measures, not only Dirichlet.

Selection schemes and fitness

Given the population x_1, \dots, x_n , the x_{n+1} -th element is bred from

$$P_\xi(\cdot | x_1, \dots, x_n).$$

Now the population consists on $n + 1$ individuals and one of them is chosen to die/discard. The probability of being discard depends on the **fitness** w of type. Here

$$w : \mathcal{X} \mapsto (0, \infty),$$

is the fitness and we order the types so that $w(1) > w(2) \geq \dots \geq w(K)$. **So only type 1 is the most fit.** There are several ways – **selection schemes** – incorporate the fitness into selection.

Selection schemes and fitness

Single tournament selection:

1. Sample $x_{n+1} \sim P_{\xi}(\cdot \mid x_1, \dots, x_n)$ (uniformly)
2. Sample l randomly from $\{1, \dots, n\}$
3. With probability

$$\frac{w(x_{n+1})}{w(x_l) + w(x_{n+1})}$$

replace x_l with x_{n+1} and discard x_l , otherwise discard x_{n+1} (the tournament between x_l and x_{n+1}).

Inverse fitness selection:

1. Sample $x_{n+1} \sim P_{\xi}(\cdot \mid x_1, \dots, x_n)$
2. Sample l from $\{1, \dots, n+1\}$ so that the probability that l is sampled is proportional to $\frac{1}{w(x_l)}$
3. When $l < n+1$, replace x_l by x_{n+1} .

Selection schemes and fitness

Both schemes introduce Markov chains with state space \mathcal{X}^n (vectors (x_1, \dots, x_n)). It is easy to show that both schemes satisfy **detailed balance equations** with stationary distribution being (when state space is \mathcal{X}^n)

$$\underbrace{P_n(x_1, \dots, x_n)}_{\text{stationary distribution}} = \frac{1}{Z_n} \underbrace{P_\xi(x_1, \dots, x_n)}_{\text{breeding term}} \underbrace{w(x_1) \cdots w(x_n)}_{\text{fitness term}}.$$

Due to the exchangeability, one can define P_n on the set of counts $\{(n_1, \dots, n_K) : n_1 + \dots + n_K = n\}$

$$\underbrace{P_n(n_1, \dots, n_K)}_{\text{counts-version of } P_n} = \frac{n!}{n_1! \cdots n_K!} P_\xi(\underbrace{1, \dots, 1}_{n_1}, \dots, \underbrace{K, \dots, K}_{n_K}).$$

Connection with Moran model

Recall – ξ is Dirichlet-categorical process and the probability that i -th type is born is

$$P_{\xi}(\xi_{n+1} = i \mid x_1, \dots, x_n, \alpha) = \frac{n_i}{n}(1 - u) + u_i, \quad u_i = \frac{\alpha_i}{|\alpha| + n}.$$

Under single tournament selection the probability that type j is chosen to die is

$$\frac{n_j}{n} \cdot \frac{w(i)}{w(i) + w(j)}$$

P(the type j is chosen to fight)P(type j loses the tournament to type i).

Under inverse fitness selection the probability that type j is chosen to die is

$$\frac{n_j w^{-1}(j)}{\sum_k n_k w^{-1}(k) + w^{-1}(i)}.$$

Connection with Moran model

Now the probability of transition that type i is born and type j discarded, equivalently, the probability of transition $(\dots, n_i, n_j, \dots) \rightarrow (\dots, n_i + 1, n_j - 1, \dots)$ is

$$\left(\frac{n_i}{n}(1-u) + u_i\right) \cdot \frac{n_j}{n} \cdot \frac{w(i)}{w(i) + w(j)} \quad \text{single tournament}$$

$$\left(\frac{n_i}{n}(1-u) + u_i\right) \cdot \frac{n_j w^{-1}(j)}{\sum_k n_k w^{-1}(k) + w^{-1}(i)} \quad \text{inverse fitness}$$

In both cases the stationary distribution is the counts version of P_n :

$$P_n(n_1, \dots, n_K) = \frac{1}{Z_n} \frac{n!}{n_1! \cdots n_K!} \frac{(\alpha_1)_{n_1} \cdots (\alpha_K)_{n_K}}{(|\alpha|)_n} \cdot w(1)^{n_1} \cdots w(K)^{n_K}.$$

More ways to incorporate w into Moran model so that stationary distribution is P_n .

Continuous time version

Consider the model with n individuals x_1, \dots, x_n . Every individual has lifetime

$$T_l \sim \text{Exp}\left(\frac{1}{w(x_l)}\right), \quad ET_l = w(x_l),$$

less fit individuals have shorter lifetimes. When an individual, say x_l dies, then it is replaced x'_l that is generated from $P_\xi(\cdot | x_1, \dots, x_{l-1}, x_{l+1}, \dots, x_n)$, where

$$(x_1, \dots, x_{l-1}, x_{l+1}, \dots, x_n)$$

is the **population without x_l at the time when x_l dies**. We then have a continuous time Markov chain with state space \mathcal{X}^n , and it is possible to show that the chain has the **same stationary distribution P_n** .

Equivalently, we can consider the same chain with states being the counts (n_1, \dots, n_K) . In a special case $K = 2$ and $i := n_1$, $n_2 = n - i$, we have continuous time MC with state space $\{0, 1, \dots, n\}$ and transition rates

$$i \rightarrow i + 1 \text{ at rate } \frac{(n - i)}{w(2)} \cdot P_\xi(1|i, n - i - 1)$$

$$i \rightarrow i - 1 \text{ at rate } \frac{i}{w(1)} \cdot P_\xi(2|i - 1, n - i).$$

When ξ is Dirichlet-categorical (Beta-Bernoulli), then we have a version of continuous time Moran model.

$$i \rightarrow i + 1 \text{ at rate } \frac{(n - i)}{w(2)} \cdot \frac{i + \alpha_1}{n - 1 + \alpha_1 + \alpha_2}$$

$$i \rightarrow i - 1 \text{ at rate } \frac{i}{w(1)} \cdot \frac{n - i + \alpha_2}{n - 1 + \alpha_1 + \alpha_2}$$

To recapitulate: We introduce a general evolution model that naturally incorporates mutation, generalizes Moran model, satisfies detailed balance equation and has stationary distribution

$$\underbrace{P_n(x_1, \dots, x_n)}_{\text{stationary distribution}} = \frac{1}{Z_n} \underbrace{P_\xi(x_1, \dots, x_n)}_{\text{breeding term}} \underbrace{w(x_1) \cdots w(x_n)}_{\text{fitness term}}.$$

Limit theorems

The main object of study is the measure (stationary distribution) on \mathcal{X}^n

$$P_n(x_1, \dots, x_n) = \frac{1}{Z_n} P_\xi(x_1, \dots, x_n) w(x_1) \cdots w(x_n) = \frac{1}{Z_n} \int_{\mathcal{P}} \prod_{l=1}^n q(x_l) w(x_l) \pi(dq).$$

Using counts version $P_n(n_1, \dots, n_K)$, we define a counterpart of P_n on simplex \mathcal{P} :

$$Q_n\left(\frac{n_1}{n}, \dots, \frac{n_K}{n}\right) := P_n(n_1, \dots, n_K).$$

Observe: the domain of P_n is \mathcal{X}^n that depends on n , Q_n is always defined on \mathcal{P} .

Limit theorems

What happens with P_n (with Q_n) when $n \rightarrow \infty$?

For limit theorems, we let w depend on n as follows:

$$w_n(k) = \exp\left[-\frac{\phi(k)}{n^\lambda}\right], \quad 0 \leq \phi(1) < \phi(2) \leq \dots \leq \phi(K), \quad \lambda \geq 0.$$

Hence $\lambda = 0$ corresponds to the constant prior, the bigger λ , the quicker the differences disappear.

In what follows, P_n and Q_n are defined with w_n .

Limit theorems: the modes of convergence

Q_n is defined on \mathcal{P} for every n .

Question: is there prob measure Q^* such that $Q_n \Rightarrow Q^*$ (weak convergence)?

For P_n , we consider the existence of **limit process** as follows:

$$X_{1,1} \sim P_1$$

$$(X_{2,1}, X_{2,2}) \sim P_2$$

...

$$(X_{n,1}, X_{n,2}, \dots, X_{n,n}) \sim P_n$$

...

Limit theorems: the modes of convergence

Question: Is there a stochastic process X_1, X_2, \dots (limit population) so that for every m the following convergence holds

$$(X_{1,n}, \dots, X_{m,n}) \Rightarrow (X_1, \dots, X_m)?$$

This is equivalent to the existence of probability measure P^* on \mathcal{X}^∞ so that

$$P_n(x_1, \dots, x_m) \rightarrow P^*(x_1, \dots, x_m), \quad \forall m, \quad \forall (x_1, \dots, x_m) \in \mathcal{X}^m.$$

If this holds, we say that **the limit process exists**.

Limit theorems: First theorem

Theorem: (constant priors) Assume the support of the prior π is \mathcal{P} .

1) If $\lambda \in [0, 1)$, then $Q_n \Rightarrow \delta_{q^*}$ and the limit process exists with

$$P^*(x_1, \dots, x_m) = \prod_{i=1}^m q^*(x_i), \text{ where } q^* = (1, 0, \dots, 0).$$

2) If $\lambda = 1$, then $Q_n \Rightarrow \bar{\pi}$ and the limit process exists with

$$P^*(x_1, \dots, x_m) = \int \prod_{i=1}^m q(x_i) \bar{\pi}(dq),$$

where for every $E \subset \mathcal{P}$,

$$\bar{\pi}(E) = \frac{1}{Z} \int_E \exp[-\langle \phi, q \rangle] \pi(dq), \quad Z = \int \exp[-\langle \phi, q \rangle] \pi(dq), \quad \langle \phi, q \rangle = \int \phi dq.$$

3) If $\lambda > 1$, then $Q_n \Rightarrow \pi$ and the limit process exists with

$$P^*(x_1, \dots, x_m) = P_{\xi}(x_1, \dots, x_m).$$

Limit theorems: First theorem

This is our first **phase transition theorem**:

- 1) when $\lambda \in [0, 1)$ (w_n is constant or differences vary slowly), then the limit process X_1, X_2, \dots has only one trajectory: 1111... i.e. **only the fittest genotype survives**;
- 2) when $\lambda = 1$, then the limit process has nondegenerate prior $\bar{\pi}$ (not iid process);
- 3) when $\lambda > 1$, then the differences between fitnesses decrease so fast that the influence of selection disappears and **the limit process equals to breeding process ξ** .

So, in a sense $\lambda = 1$ is the proper scaling. In literature often $w(k) = (1 + s_k)$, with $s_k \cdot n \rightarrow \gamma_k$ (constant), and then

$$w(k) \approx \left(1 + \frac{\gamma_k}{n}\right) \approx \exp\left[\frac{\gamma_k}{n}\right].$$

Limit theorems: Dirichlet prior

The weights as previously

$$w_n(k) = \exp\left[-\frac{\phi(k)}{n^\lambda}\right], \quad 0 \leq \phi(1) < \phi(2) \leq \dots \leq \phi(K), \quad \lambda \geq 0.$$

We now consider Dirichlet priors (important special case, connection to Moran model):

$$\pi_n = \text{Dir}(n^{1-\lambda}\alpha_1, \dots, n^{1-\lambda}\alpha_K), \quad \lambda \in [0, 1]$$

so that π_n depends on n as well!

Here $\alpha := (\alpha_1, \dots, \alpha_K)$, $\alpha_k > 0$ and $|\alpha| := \sum_k \alpha_k$.

Limit theorems: Dirichlet prior

Recall that π_n has density

$$\pi_n(q) \propto \prod_{k=1}^K (q_k)^{n^{1-\lambda}\alpha_k}, \quad q = (q_1, \dots, q_K) \in \mathcal{P}.$$

Also recall that when $(Y_1, \dots, Y_K) \sim \pi_n$, then

$$E(Y_1, \dots, Y_K) = (\alpha_1/|\alpha|, \dots, \alpha_K/|\alpha|), \quad \text{Var}(Y_k) = \frac{\frac{\alpha_k}{|\alpha|} \left(1 - \frac{\alpha_k}{|\alpha|}\right)}{|\alpha| n^{1-\lambda}}.$$

Hence the bigger n or the smaller λ , the smaller variance and the more π_n is concentrated over its expectation.

Hence λ tunes the prior influence. $\lambda = 1$ is the constant prior case.

Limit theorems: Second theorem

Theorem: (Dirichlet priors)

1) If $\lambda = 0$, then $Q_n \Rightarrow \delta_{r^*}$ and the limit process exists with

$$P^*(x_1, \dots, x_m) = \prod_{i=1}^m r^*(x_i), \quad \text{where } r^*(k) \propto w(k)q^*(k), \quad q^*(k) = \frac{\alpha_k}{(1 + |\alpha|) - \frac{w(k)}{\theta}}$$

and θ is a parameter such that q^* is probability measure.

2) If $\lambda \in (0, 1)$, then $Q_n \Rightarrow \delta_{q^*}$ and the limit process exists with

$$P^*(x_1, \dots, x_m) = \prod_{i=1}^m q^*(x_i), \quad \text{where } q^*(k) = \frac{\alpha_k}{\phi(k) + |\alpha| - \theta}, \quad k = 1, \dots, K,$$

and θ is a parameter.

Limit theorems: Second theorem

3) If $\lambda = 1$, then $Q_n \Rightarrow \bar{\pi}$ and the limit process exists with

$$P^*(x_1, \dots, x_m) = \int \prod_{i=1}^m q(x_i) \bar{\pi}(dq),$$

where $\bar{\pi}$ has density

$$\bar{\pi}(q) \propto \exp[-\langle \phi, q \rangle] \prod_{k=1}^K (q_k)^{\alpha_k - 1}.$$

Limit theorems: Second theorem

This is our second **phase transition theorem**:

- 1) When $\lambda = 0$, then the limit process X_1, X_2, \dots is i.i.d process with distribution r^* .
- 2) When $\lambda \in (0, 1)$, then limit process X_1, X_2, \dots is i.i.d process with distribution q^* .
Observe that **the limit process does not depend on λ !** However, the speed of convergence does.
- 3) The last case $\lambda = 1$ corresponds to the case of previous theorem (constant prior).
The limit density

$$\bar{\pi}(q) \propto \exp[-\langle \phi, q \rangle] \prod_{k=1}^K (q_k)^{\alpha_k - 1}$$

can be found in textbooks (for both Moran and Wright-Fisher model), but typically (always?) obtained via **diffusion approximation**, our proof is much simpler.

Limit theorems: Dirichlet prior, different λ 's

A generalization:

$$w_n(k) = \exp\left[-\frac{\phi(k)}{n^{\lambda_1}}\right], \quad \pi_n = \text{Dir}(n^{1-\lambda_2}\alpha_1, \dots, n^{1-\lambda_2}\alpha_K).$$

Then

- 1) When $\lambda_1 > \lambda_2$, and $\lambda_2 < 1$ then limit process X_1, X_2, \dots is i.i.d process with distribution $(\frac{\alpha_1}{|\alpha|}, \dots, \frac{\alpha_K}{|\alpha|})$. This is again the case where the influence of fitness vanishes.
- 2) When $0 \leq \lambda_1 < \lambda_2$, then the limit process X_1, X_2, \dots has only one trajectory: 1111... i.e. only the fittest genotype survives. The influence of prior vanishes.
- 3) The case $0 \leq \lambda_1 = \lambda_2 = \lambda \leq 1$ is considered above.

Uncountable \mathcal{X}

Surprisingly hard to generalize these results to the case when \mathcal{X} is uncountable ($\mathcal{X} = \mathbb{R}$ or Polish space or $[0, 1]$).

So let $\xi = \xi_1, \xi_2, \dots$ be \mathcal{X} -valued exchangeable breeding process with prior π on \mathcal{P} . Now \mathcal{P} is the space of all probability measures on \mathcal{X} .

The special important case is when π is **Dirichlet process** $\text{Dir}(m, \bar{\alpha})$, where $\bar{\alpha}$ is **non-atomic probability measure (base measure)** on \mathcal{X} and $m > 0$ is **concentration parameter**.

Recall that when $P \sim \text{Dir}(m, \bar{\alpha})$, then for any finite measurable partition A_1, \dots, A_k of \mathcal{X} ,

$$(P(A_1), \dots, P(A_k)) \sim \text{Dir}(m\bar{\alpha}(A_1), \dots, m\bar{\alpha}(A_k)).$$

Uncountable \mathcal{X} : DP breeding

With $\pi = \text{Dir}(m, \bar{\alpha})$, for every $n \geq 1$ and measurable $A \subset \mathcal{X}$,

$$P(\xi_{n+1} \in A | \xi_1 = x_1, \dots, \xi_n = x_n) = \frac{m}{m+n} \bar{\alpha}(A) + \frac{n}{m+n} \cdot \frac{1}{n} \sum_{i=1}^n \delta_{x_i}.$$

If x_1^*, \dots, x_k^* are the distinct values of x_1, \dots, x_n with n_1, \dots, n_k being their frequencies, the conditional distribution above can be interpreted as follows:

$$\xi_{n+1} | \xi_1 = x_1, \dots, \xi_n = x_n \sim \begin{cases} \bar{\alpha}, & \text{with probability } \frac{m}{m+n}; \\ \delta_{x_j^*}, & \text{with probability } \frac{n_j}{m+n} \quad j = 1, \dots, k. \end{cases}$$

When $\bar{\alpha}$ is non-atomic, then every draw from $\bar{\alpha}$ is different from all previously seen types – a **mutation**. The probability of mutation is then $\frac{m}{m+n}$ and taking $m_n = c \cdot n$, where $c > 0$, we obtain that the probability of mutation is constant: $c/(c+1)$.

Uncountable \mathcal{X} : kernels and measure P_n

With the selection schemes – single tournament or inverse fitness – the corresponding Markov chain have uncountable state space \mathcal{X}^n and have **transition kernels** $P(\mathbf{x}, A)$ instead of transition matrices, where $\mathbf{x} = (x_1, \dots, x_n)$. These kernels (for both schemes) satisfy **detailed balance equations**

$$\int_B P(\mathbf{x}, A) P_n(d\mathbf{x}) = \int_A P(\mathbf{x}, B) P_n(d\mathbf{x}), \quad \mathbf{x} = (x_1, \dots, x_n) \in \mathcal{X}^n, \quad A, B \subset \mathcal{X}$$

with stationary measure P_n being now

$$P_n(A) = \frac{1}{Z_n} \int_A \prod_{j=1}^n w_n(x_j) P_\xi^n(d\mathbf{x}) = \frac{1}{Z_n} \int_{\mathcal{P}} \int_A \prod_{j=1}^n w_n(x_j) q(dx_j) \pi_n(dq).$$

The counterpart Q_n on \mathcal{P} easy to define.

Uncountable \mathcal{X} : the fitness

Fitness w_n is a function on \mathcal{X} :

$$w_n(x) = \exp\left[-\frac{\phi(x)}{n^\lambda}\right], \quad \phi : \mathcal{X} \rightarrow [0, \infty) \text{ continuous, bounded, unique minimum } x_o.$$

Thus x_o is the single most fit type.

Uncountable \mathcal{X} : the first theorem

Theorem: (constant priors) Assume π is independent of n .

1) Suppose the support of π contains δ_{x_0} . If $\lambda \in [0, 1)$, then (under additional mild assumption on ϕ) then the limit process exists and has a.s. the trajectory x_0, x_0, \dots . If \mathcal{X} is **compact**, then $Q_n \Rightarrow \delta_{q^*}$, where $q^* = \delta_{x_0}$

2) If $\lambda = 1$, then the limit process exists and the the corresponding prior is

$$\bar{\pi}(E) = \frac{1}{Z} \int_E \exp[-\langle \phi, q \rangle] \pi(dq), \quad Z = \int \exp[-\langle \phi, q \rangle] \pi(dq), \quad \langle \phi, q \rangle = \int \phi dq.$$

If \mathcal{X} is **compact**, then $Q_n \Rightarrow \bar{\pi}$.

3) If $\lambda > 1$, then the limit process is ξ . If \mathcal{X} is **compact**, then $Q_n \Rightarrow \pi$.

This theorem directly generalizes the previous one with finite \mathcal{X} , but the proof is way more involved and much more technical.

Uncountable \mathcal{X} : the second theorem, $\lambda = 0$

DP-prior, $\lambda = 0$, so $\pi_n = \text{DP}(cn, \bar{\alpha})$ (constant mutation probability), fitness $w_n = w$ (independent of n).

The limit process depends on the following inequality

$$\int \frac{\bar{w}}{\bar{w} - w(x)} \bar{\alpha}(dx) \geq \frac{1+c}{c}, \quad \bar{w} = w(x_0). \quad (0.2)$$

The inequality relates $\bar{\alpha}$ with w , it fails when α has very low mass around x_0 if w has sharp peak at x_0 .

Uncountable \mathcal{X} : the second theorem, $\lambda = 0$

Define a probability measure r^* on \mathcal{X} :

$$r^*(A) = \begin{cases} \frac{1}{\theta} \int_A w(x) f_\theta(x) \bar{\alpha}(dx), & \text{when (0.2) holds;} \\ \frac{1}{\theta_o} \left(\int_A w(x) f_{\theta_o}(x) \bar{\alpha}(dx) + (1 - \beta) \bar{w} \delta_{x_o}(A) \right), & \text{when (0.2) fails} \end{cases},$$

where

$$f_\theta(x) = \frac{c}{\left(1 + c - \frac{w(x)}{\theta}\right)}, \quad \theta = \int w(x) f_\theta(x) \bar{\alpha}(dx), \quad \theta_o = \frac{\bar{w}}{1 + c}, \quad \beta = \int f_{\theta_o}(x) \bar{\alpha}(dx).$$

When inequality (0.2) holds, then $\beta = 1$ and $r^* \ll \bar{\alpha}$, otherwise $\beta < 1$ and r^* contains atom (with mass $1 - \beta$) at x_o .

Theorem: (Dirichlet prior, $\lambda = 0$) Let $\lambda = 0$ and \mathcal{X} be compact. Then $Q_n \Rightarrow \delta_{r^*}$.
Then also the limit process X_1, X_2, \dots exists and it is an iid process, with $X_i \sim r^*$.

Uncountable \mathcal{X} : the second theorem, $\lambda = 0$

The density (with respect to $\bar{\alpha}$)

$$\frac{w(x)f(x)}{\theta} = \frac{w(x)c}{\theta(1+c) - w(x)}$$

is continuous counterpart of r^* for discrete \mathcal{X} (now $|\alpha| = c$)

$$\frac{w(k)\alpha_k}{\theta(1+|\alpha|) - w(k)},$$

but when (0.2) fails, an atom at x_0 appears !

Uncountable \mathcal{X} : the second theorem, $\lambda \in (0, 1)$

DP-prior, $\lambda \in (0, 1)$, so

$$\pi_n = \text{DP}(cn^{1-\lambda}, \bar{\alpha}), \quad w_n(x) = \exp\left[-\frac{\phi(x)}{n^\lambda}\right].$$

In this case, the following inequality is crucial

$$\int \frac{1}{\phi(x) - \phi(x_0)} \bar{\alpha}(dx) \geq \frac{1}{c}. \quad (0.3)$$

Define the probability measure q^* on \mathcal{X} :

$$q^*(A) = \begin{cases} \int_A f_\theta(x) \bar{\alpha}(dx), & \text{when (0.3) holds;} \\ \int_A f_{\theta_0}(x) \bar{\alpha}(dx) + (1 - \beta) \bar{w} \delta_{x_0}(A), & \text{when (0.3) fails} \end{cases},$$

$$f_\theta(x) = \frac{c}{\phi(x) + c - \theta}, \quad \theta = \int \phi(x) f_\theta(x) \bar{\alpha}(dx), \quad \theta_0 = c + \phi_0, \quad \beta = \int f_{\theta_0}(x) \bar{\alpha}(dx).$$

Uncountable \mathcal{X} : the second theorem, $\lambda \in (0, 1)$

Theorem: (Dirichlet prior, $\lambda \in (0, 1)$) Let $\lambda \in (0, 1)$ and \mathcal{X} be compact and $\bar{\alpha}(x_0) = 0$. Then $Q_n \Rightarrow \delta_{q^*}$. Then also the limit process X_1, X_2, \dots exists and it is an iid process, with $X_i \sim q^*$.

When (0.3) holds, then the density (with respect to $\bar{\alpha}$)

$$f(x) = \frac{c}{\phi(x) + c - \theta}$$

is continuous counterpart of q^* for discrete \mathcal{X} (now $|\alpha| = c$)

$$q^*(k) = \frac{\alpha_k}{\phi(k) + |\alpha| - \theta}$$

but when (0.3) fails, an atom at x_0 appears ! Again the limit is independent of λ ! The proofs of last two theorems are technical and based on large deviation principle on \mathcal{P} . Hence the need for compactness (compact \mathcal{X} ensures compact \mathcal{P}).



TARTU ÜLIKOOL

J. Lember, C. Watkins (2020). An Evolutionary Model that Satisfies Detailed Balance.
Methodol. Comput. Appl.

This is now really the very last slide!

Thank you!



unitartu



tartuylikool

