Uniform sampling in a structured branching population

Aline Marguet

November 3rd, 2016

A model to describe the dynamic of a cell population

We consider a structured cell population: each cell has a trait (its age, size, number of parasites...).

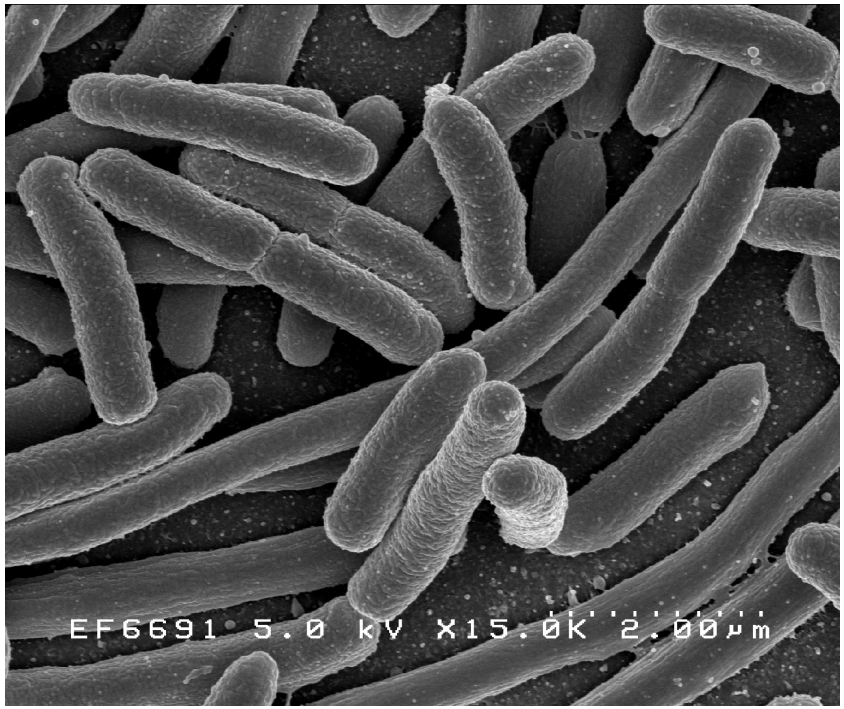


Image : NIAD

The model

- The trait $(X_t^u)_{t\geq 0}$ of each individual u follows a Markov process with infinitesimal generator $\mathcal{G}$.

The model

- The trait $(X_t^u)_{t\geq 0}$ of each individual u follows a Markov process with infinitesimal generator $\mathcal{G}$.
- An individual u dies at time t at rate $B(X_t^u)$, i.e.

$$\mathbb{P}\left(\beta(u) > t | \alpha(u),\ (X_s^u, \alpha(u) \leq s \leq t)\right) = \exp\left(-\int_{\alpha(u)}^{t} B(X_s^u) ds\right),$$

where $\beta(u)$ denotes the lifetime of u et $\alpha(u)$ its birth time.

The model

- The trait $(X_t^u)_{t\geq 0}$ of each individual u follows a Markov process with infinitesimal generator $\mathcal{G}$.
- An individual u dies at time t at rate $B(X_t^u)$, i.e.

$$\mathbb{P}\left(\beta(u) > t | \alpha(u),\ (X_s^u, \alpha(u) \leq s \leq t)\right) = \exp\left(-\int_{\alpha(u)}^{t} B(X_s^u) ds\right),$$

  where $\beta(u)$ denotes the lifetime of u et $\alpha(u)$ its birth time.
- At his death, an individual with trait x is replaced by A_u descendants. The traits at birth are chosen according to a probability law depending on x and A_u.

The model

- The trait $(X_t^u)_{t\geq 0}$ of each individual u follows a Markov process with infinitesimal generator $\mathcal{G}$.
- An individual u dies at time t at rate $B(X_t^u)$, i.e.

$$\mathbb{P}\left(\beta(u) > t | \alpha(u),\ (X_s^u, \alpha(u) \leq s \leq t)\right) = \exp\left(-\int_{\alpha(u)}^{t} B(X_s^u) ds\right),$$

  where $\beta(u)$ denotes the lifetime of u et $\alpha(u)$ its birth time.
- At his death, an individual with trait x is replaced by A_u descendants. The traits at birth are chosen according to a probability law depending on x and A_u.
- Conditionally to the trait of their parents, the dynamics of the descendants are independent.

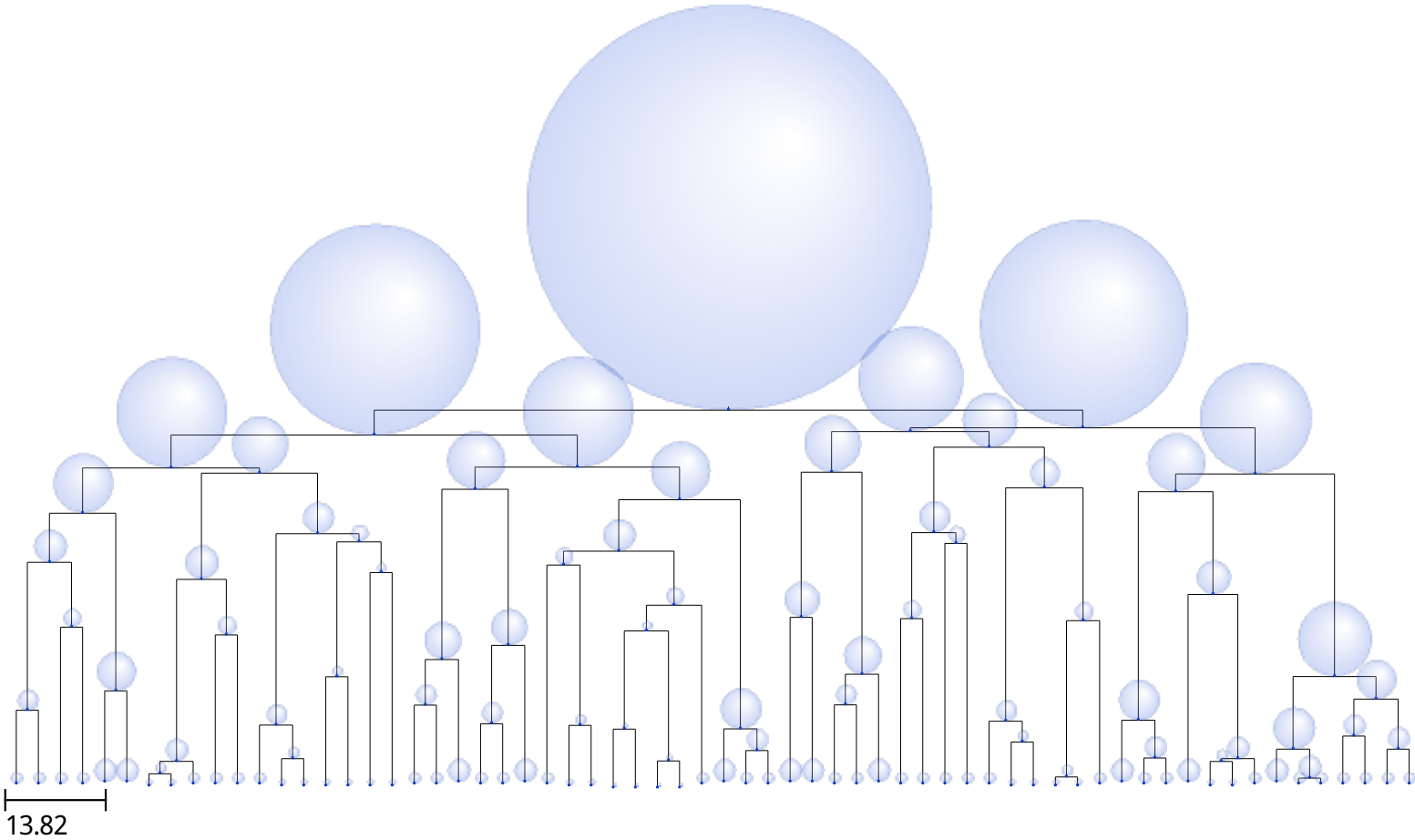


Figure: Descending genealogy from an individual with size 1 until time $T = 50$ of a size-structured population. Each cell grows exponentially at rate 0.01 and divide at rate $B(x) = x$.

Measure-valued process description

Following the work of Fournier and Méléard (2004), we set:

$$Z_t = \sum_{u \in V_t} \delta_{X_t^u},$$

- V_t : individuals alive at time t,

Measure-valued process description

Following the work of Fournier and Méléard (2004), we set:

$$Z_t = \sum_{u \in V_t} \delta_{X_t^u},$$

- V_t : individuals alive at time t,
- $N_t = \#V_t$.

Measure-valued process description

Following the work of Fournier and Méléard (2004), we set:

$$Z_t = \sum_{u \in V_t} \delta_{X_t^u},$$

- V_t : individuals alive at time t,
- $N_t = \#V_t$.

To define properly our process as the strongly unique solution of a SDE (and to avoid blow up in finite time for the population), we make two kind of assumptions:

- on the division events,
- on the dynamic of the trait.

Assumptions on the division events

Assumption A

1. There exist $b_1, b_2 \geq 0$ and $\gamma \geq 0$ such that for all $x \in \mathcal{X}$,

$$B(x) \leq b_1 |x|^\gamma + b_2.$$

Assumptions on the division events

Assumption A

1. There exist $b_1, b_2 \geq 0$ and $\gamma \geq 0$ such that for all $x \in \mathcal{X}$,

$$B(x) \leq b_1 |x|^\gamma + b_2.$$

2. For all $t \geq 0$, for all $x \in \mathcal{X}$, if $A_u = k$ and a is uniformly distributed on $[0, 1]$:

$$\sum_{i=1}^{k} F_i^{(k)}(x, a) \leq x, \text{ a.s.}$$

Assumptions on the division events

Assumption A

1. There exist $b_1, b_2 \geq 0$ and $\gamma \geq 0$ such that for all $x \in \mathcal{X}$,

$$B(x) \leq b_1 |x|^\gamma + b_2.$$

2. For all $t \geq 0$, for all $x \in \mathcal{X}$, if $A_u = k$ and a is uniformly distributed on $[0, 1]$:

$$\sum_{i=1}^{k} F_i^{(k)}(x, a) \leq x, \text{a.s.}$$

3. There exists $\overline{m} \geq 0$ such that for all $x \in \mathcal{X}$,

$$m(x) = \sum_k k p_k(x) \leq \overline{m}.$$

Assumption on the dynamic of the trait

Assumption B

There exist $c_1, c_2 \geq 0$ such that for all $x \in \mathcal{X}$:

$$\mathcal{G} h_\gamma(x) \leq c_1 h_\gamma(x) + c_2,$$

with γ previously defined and $h_\gamma(x) = |x|^\gamma = \left(\sum_{i=1}^d x_i\right)^\gamma$, for all $x \in \mathbb{R}_+^d$.

Assumption on the dynamic of the trait

Assumption B

There exist $c_1, c_2 \geq 0$ such that for all $x \in \mathcal{X}$:

$$\mathcal{G}h_\gamma(x) \leq c_1 h_\gamma(x) + c_2,$$

with γ previously defined and $h_\gamma(x) = |x|^\gamma = \left(\sum_{i=1}^d x_i\right)^\gamma$, for all $x \in \mathbb{R}_+^d$.

In particular, if we consider:

$$\mathcal{G}f(x) = r(x,t)f'(x) + \sigma^2(x,t)f''(x),$$

the previous assumption is satisfied if:

$$r(x,t) \leq a_r x + b_r \text{ et } \sigma(x,t) \leq a_\sigma x + b_\sigma.$$

The trait of a uniformly sampled individual

Let $U(t)$ denote the random variable corresponding to a uniform choice of an individual in the population at time t. We want to describe the law of:

$$\left(X_s^{U(t)}\right)_{s \leq t}$$

The trait of a uniformly sampled individual

Let $U(t)$ denote the random variable corresponding to a uniform choice of an individual in the population at time t. We want to describe the law of:

$$\left(X_s^{U(t)}\right)_{s \leq t}$$

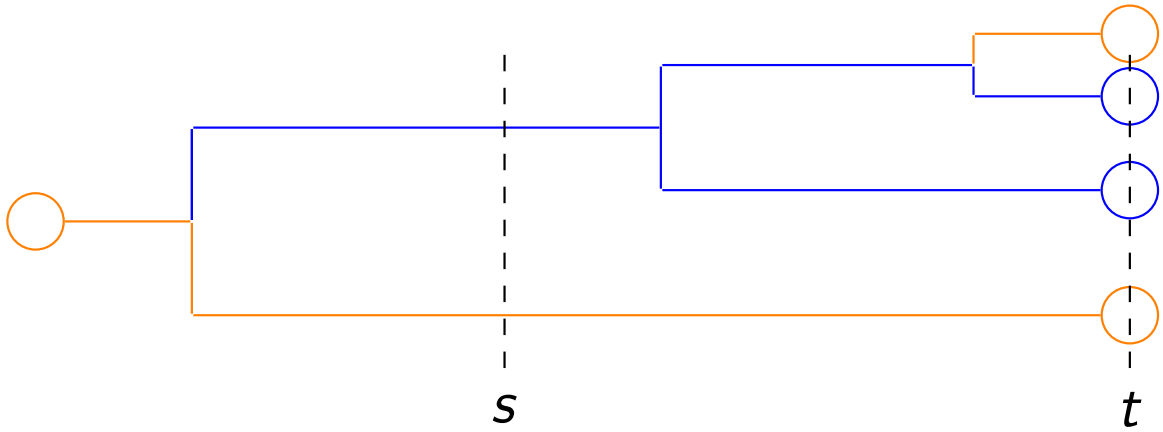
Why is it different from the law of a "tagged-cell" ?

The trait of a uniformly sampled individual

Let $U(t)$ denote the random variable corresponding to a uniform choice of an individual in the population at time t. We want to describe the law of:

$$\left(X_s^{U(t)}\right)_{s \leq t}$$

Why is it different from the law of a "tagged-cell" ?



History

- Size-biased tree and spine
  - Chauvin, Rouault (1988), Chauvin, Rouault, Wakolbinger (1991), Gorostiza, Roelly, Wakolbinger (1991): size-biased tree,
  - Lyons, Peres, Pemantle (1995): Kesten-Stigum theorem,
  - Kurtz, Lyons, Peres, Pemantle (1997), Athreya (2000): extensions to the multitype case,
  - Georgii & Baake (2003): study of the ancestral trait of a typical individual in a spectral framework,
  - Hardy & Harris (2006): new formalization of the spine

History

- Size-biased tree and spine
  - Chauvin, Rouault (1988), Chauvin, Rouault, Wakolbinger (1991), Gorostiza, Roelly, Wakolbinger (1991): size-biased tree,
  - Lyons, Peres, Pemantle (1995): Kesten-Stigum theorem,
  - Kurtz, Lyons, Peres, Pemantle (1997), Athreya (2000): extensions to the multitype case,
  - Georgii & Baake (2003): study of the ancestral trait of a typical individual in a spectral framework,
  - Hardy & Harris (2006): new formalization of the spine
- Auxiliary process approach
  - Bansaye, Delmas, Marsalle, Tran (2011): continuous model in the neutral case, law of large numbers,
  - Cloez (2011): non-neutral case of the model of Bansaye et al. in a spectral framework.

Many-to-One formula

Theorem

For all $t > 0$, $x \in \mathcal{X}$, for all positive measurable function $F : \mathbb{D}([0,t], \mathcal{X}) \to \mathbb{R}$ we have:

$$\mathbb{E}_{\delta_{x_0}}\left[\sum_{u \in V_t} F\left(X_s^u, s \leq t\right)\right] = \mathbb{E}_{\delta_{x_0}}(N_t)\mathbb{E}_{x_0}\left[F\left(Y_s^{(t)}, s \leq t\right)\right],$$

where $\left(Y_s^{(t)}\right)_{s \leq t}$ is an inhomogeneous Markov process with infinitesimal generators $\left(\mathcal{A}_s^{(t)}\right)_{s \leq t}$ given by:

$$\mathcal{A}_s^{(t)} f(x) = \widehat{\mathcal{G}}_{s,t} f(x) + \widehat{B}_{s,t}(x)\left(\int_{\mathcal{X}} \left(f(y) - f(x)\right) \widehat{P}_{s,t}(x, dy)\right).$$

The auxiliary process

$$m(x, dy) = \sum_{k\geq 0} p_k(x) \sum_{i=1}^{k} P_i^{(k)}(x, dy),\ m(x, s, t) = \mathbb{E}(N_t \big| Z_s = \delta_x),$$

The auxiliary process

$m(x, dy) = \sum_{k\geq 0} p_k(x) \sum_{i=1}^{k} P_i^{(k)}(x, dy)$, $m(x, s, t) = \mathbb{E}(N_t \big| Z_s = \delta_x)$,

- the auxiliary process jumps at an accelerated rate:

$$\widehat{B}_{s,t}(x) = B(x) \int_{\mathcal{X}} \frac{m(y, s, t)}{m(x, s, t)} m(x, dy),$$

The auxiliary process

$m(x, dy) = \sum_{k\geq 0} p_k(x) \sum_{i=1}^{k} P_i^{(k)}(x, dy)$, $m(x, s, t) = \mathbb{E}(N_t \big| Z_s = \delta_x)$,

- the auxiliary process jumps at an accelerated rate:

$$\widehat{B}_{s,t}(x) = B(x) \int_{\mathcal{X}} \frac{m(y, s, t)}{m(x, s, t)} m(x, dy),$$

- the trait of the unique descendant is chosen according to the following kernel:

$$\widehat{P}_{s,t}(x, dy) = \frac{m(y, s, t)}{m(x, s, t)} m(x, dy) \left(\int_{\mathcal{X}} \frac{m(y, s, t)}{m(x, s, t)} m(x, dy) \right)^{-1}.$$

The auxiliary process

$m(x, dy) = \sum_{k\geq 0} p_k(x) \sum_{i=1}^{k} P_i^{(k)}(x, dy)$, $m(x, s, t) = \mathbb{E}(N_t \big| Z_s = \delta_x)$,

- the auxiliary process jumps at an accelerated rate:

$$\widehat{B}_{s,t}(x) = B(x) \int_{\mathcal{X}} \frac{m(y, s, t)}{m(x, s, t)} m(x, dy),$$

- the trait of the unique descendant is chosen according to the following kernel:

$$\widehat{P}_{s,t}(x, dy) = \frac{m(y, s, t)}{m(x, s, t)} m(x, dy) \left(\int_{\mathcal{X}} \frac{m(y, s, t)}{m(x, s, t)} m(x, dy) \right)^{-1}.$$

- the dynamic of the trait between jumps is given by:

$$\widehat{\mathcal{G}}_{s,t} f(x) = \frac{\mathcal{G}\left(m(\cdot, s, t) f\right)(x) - f(x)\, \mathcal{G}\left(m(\cdot, s, t)\right)(x)}{m(x, s, t)}$$

Examples

- In the neutral case, $B \equiv b$, we find again the process of Bansaye et al.:

$$\mathcal{A}f(x) = \mathcal{G}f(x) + 2b \int_{\mathcal{X}} (f(y) - f(x)) \frac{m(x, dy)}{2}$$

Examples

- In the neutral case, $B \equiv b$, we find again the process of Bansaye et al.:

$$\mathcal{A}f(x) = \mathcal{G}f(x) + 2b \int_{\mathcal{X}} (f(y) - f(x)) \frac{m(x, dy)}{2}$$

- In the case of a size-structured population with exponential growth at rate a and binary division at rate $B(x) = x$, we have:

$$m(x, s, t) = 1 + \frac{x}{a} \left(e^{a(t-s)} - 1 \right),$$

and:

$$\mathcal{A}_s^{(t)} f(x) = axf'(x) + x \left(1 + \frac{1}{1 + \frac{x}{a} \left(e^{a(t-s)} - 1 \right)} \right) \left(f\left(\frac{x}{2} \right) - f(x) \right).$$

Large population approximation

Let $\nu \in \mathcal{M}_P(\mathcal{X})$ satisfying:

$$\mathbb{P}_\nu(N_t > 0) > 0.$$

Let

$$\nu_n := \sum_{i=1}^{n} \delta_{X_i},$$

with X_i independent identically distributed according to ν.

Theorem

Under Assumption A,B, +technical assumptions, we have in law in $\mathbb{D}([0,t],\mathcal{X})$:

$$X_{[0,t]}^{U(t),\nu_n} \mathbf{1}_{A_n^\nu(t)} \xrightarrow[n\to+\infty]{} Y_{[0,t]}^{(t),\pi_t}, \text{ avec } \pi_t(dx) = \frac{m(x,0,t)\nu(dx)}{m(\nu,0,t)}.$$

Future prospects

Future prospects

- Long time behaviour of the process, law of large numbers,
- Statistical estimation of the division rate.