IBMPopSim: An R package for simulating Individual-Based Models

Sarah Kaakaï

Laboratoire Manceau de Mathématiques (LMM), Le Mans Université

Joint work with D. Giorgi, and V. Lemaire

Rencontre de la chaire MMB, 22 septembre 2021

IBMPopSim

- Provide a unified and user-friendly R framework for simulating stochastic individual-based population models.
- Many applications in biology, ecology, epidemiolgy, simulation of human populations...
- Large class of population models:
  - Including interactions between individuals.
  - Description of non-stationary phenomena (reduction of mortality over time, shocks, ...).

- Package **IBMPoPSim** available on CRAN
- Documentation and examples:
  `https://daphnegiorgi.github.io/IBMPopSim/`

Simulation challenges

- Large class of IBMs while staying user-friendly.
- Low computational cost:
  - Efficient algorithm, adapted from Fournier and Méléard (2004), Tran (2008).
  - Use of Rcpp,
  - Appropriate data structures.
- Various outputs:
  - Age-pyramids, number of deaths, exposure-to-risk...

1. Population model

2. Algorithm

Outline

The population

- An **individual** I is characterized by date of birth τ and characteristics $x \in \mathcal{X}$ at a given time:
  - Gender, place of living, wealth, patch, strategy, infected, size ...
- A **population** Z_t at time t is a collection of individuals:

$$\{(\tau_i, x_t^i)\ ;\ i = 1...N_t\}, \quad \text{or } \{(a_t^i, x_t^i) = (t - \tau_i, x_t^i)\ ;\ i = 1...N_t\}.$$

Population composition is modified due to different **types of events**:
↳ Describe how frequently events happens and how they modify the population composition.

Events

The population composition can changes following 4 main types of events:

- Birth at time t An individual gives birth to a new individual (t, x_t^{new}).
- Death/Exit An individual is removed from the population.
- Change of characteristics (swap) An individual changes characteristics from (τ_i, x_t^i) to (τ_i, y).
- Entry/"Immigration" An individual (τ_i, x_t^i) is added to the population.

Events rates (I)

Frequency of events are described by individual rates:

$$\lambda^{e}(I, t, Z)dt \simeq \mathrm{P}(\text{event e occuring to individual } I \in [t, t+dt] | \mathcal{F}_t).$$

The event intensity λ^e can depend on:

- Age and characteristics of I.
- The time t.
- Interactions with other individuals in the population.

Event rates (II)

For an event e, we consider event rates of form:

$$\lambda(I, t, Z) = f^e(I, t) + \underbrace{\sum_{J \in Z} U^e(I, J, t)}_{\text{Interaction kernel}}$$

Event rates (II)

For an event e, we consider event rates of form:

$$\lambda^e(I, t, Z) = {\color{blue} f^e(I, t)} + \underbrace{\sum_{J \in Z} U^e(I, J, t)}_{\text{Interaction kernel}}$$

- **Individual term**
  - Age-dependent birth rate: $\lambda^b(I, t) = b(a_I(t), t)$.
  - Joint mortality forecast for socioeconomic subgroups, $x = i \in [\![1; n]\!]$:

$$\log(\lambda^d(i, a, t)) = \alpha^i_{\lfloor a \rfloor} + \beta^i_{\lfloor a \rfloor}\kappa_{\lfloor t \rfloor}, \quad \kappa_n = d + \kappa_{n-1} + \xi_n.$$

Event rates (II)

For an event e, we consider event rates of form:

$$\lambda^e(I,t,Z) = f^e(I,t) + \underbrace{\sum_{J\in Z} U^e(I,J,t)}_{\text{Interaction kernel}}$$

- **Individual term**
  - Age-dependent birth rate: $\lambda^b(I,t) = b(a_I(t),t)$.
  - Joint mortality forecast for socioeconomic subgroups, $x = i \in [\![1;n]\!]$:

$$\log(\lambda^d(i,a,t)) = \alpha^i_{\lfloor a\rfloor} + \beta^i_{\lfloor a\rfloor}\kappa_{\lfloor t\rfloor}, \quad \kappa_n = d + \kappa_{n-1} + \xi_n.$$

- **Interaction term**
  - SIR model: transition from susceptible to infected.

$$\lambda^s(I,t) = \sum_{J\in Z} \beta \mathbb{1}_{\text{susceptible}}(I)\mathbb{1}_{\text{infected}}(J)$$

  - Competition between individuals (Tran (2008)): body size $x_0 + ga$.

$$\lambda^d(x_0,a,Z) = \sum_{(x_0^i,a^i)\in Z} U(x_0+ga, x_0^i+ga^i), \; U(x,y) = C\left(1-\frac{1}{1+e^{-\nu(x-y)}}\right)$$

Global intensities

- For an event e, the total event intensity is the sum of individual intensities:
$$\Lambda^e(t) = \sum_{I \in Z_t} \lambda^e(I, t, Z).$$
- Particular case: Entry event with (inhomogeneous) Poisson intensity: $\Lambda^i(t) = \lambda_t$.
- Total intensity: $\Lambda(t) = \sum_{e=b,d,s,i} \Lambda^e(t)$.

Global intensities

- For an event e, the total event intensity is the sum of individual intensities:
$$\Lambda^e(t) = \sum_{I \in Z_t} \lambda^e(I, t, Z).$$
- Particular case: Entry event with (inhomogeneous) Poisson intensity: $\Lambda^i(t) = \lambda_t$.
- Total intensity: $\Lambda(t) = \sum_{e=b,d,s,i} \Lambda^e(t)$.

Simple example: $\lambda^d(I, t) = d(a_I(t))$,

$$\Lambda^d(t) = \sum_{I \in Z_t} d(a_I(t)) = \sum_{I \in Z_t} d(t - \tau_I)$$

Long computation time!

Outline

Assumption on intensities

Individual event rate:

$$\lambda^e(I,t) = f^e(I,t) + \underbrace{\sum_{J \in Z} U^e(I,J,t)}_{\text{Interaction kernel}}$$

- Individuals rates are assumed to be bounded:
  - Individual term: $f^e(I,t) \leqslant \bar{e}$.
  - Interaction term: $U^e(I,J,t) \leqslant C_e$.
- **Event rate bound** $e = b, d, s$:

$$e(I,t) \leqslant \bar{e} + C_e N_t, \quad N_t = \text{nb of individuals in the population.}$$

Acceptance/Rejection algo

- Goal : Simulate $(N^e(ds, di))_e = ((T_n^e, I_n^e))_e$ of intensity $e(i,s)1_{\{i \leqslant N_s\}} ds \otimes di$.
- Befor the first event :

$$e(i,s)1_{\{i \leqslant N_0\}} \leqslant (\bar{e} + C_e N_0)1_{\{i \leqslant N_0\}}$$

- Coupling Idea :
  1. Simulate population $(\bar{N}^e)_e$ with more events, at times easy to simulate.
  2. Acceptance/rejection procedure.

Algorithm without interactions

Consider a population with only births and deaths: $\Lambda(t) \leqslant (\bar{b} + \bar{d})N_t$.

Algorithm without interactions

Consider a population with only births and deaths: $\Lambda(t) \leqslant (\bar{b} + \bar{d}) N_t$.

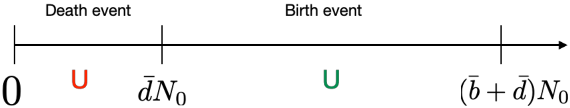
Step 1 Simulate a candidate event time $T_1 \sim \mathcal{E}\left((\bar{b} + \bar{d}) N_0\right)$.

Algorithm without interactions

Consider a population with only births and deaths: $\Lambda(t) \leqslant (\bar{b} + \bar{d})N_t$.

Step 1 Simulate a candidate event time $T_1 \sim \mathcal{E}\left((\bar{b} + \bar{d})N_0\right)$.

Step 2 Choose individual $I \sim \mathcal{U}(\{1, \dots, N_0\})$ and event type $e = b$ or d: let $U \sim \mathcal{U}([0, (\bar{b} + \bar{d})N_0])$

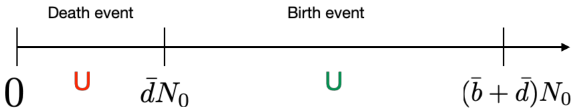


Algorithm without interactions

Consider a population with only births and deaths: $\Lambda(t) \leqslant (\bar{b}+\bar{d})N_t$.

Step 1 Simulate a candidate event time $T_1 \sim \mathcal{E}\left((\bar{b}+\bar{d})N_0\right)$.

Step 2 Choose individual $I \sim \mathcal{U}(\{1,\dots,N_0\})$ and event type $e = b$ or d: let $U \sim \mathcal{U}([0,(\bar{b}+\bar{d})N_0])$



Step 3 Accept/reject event using $\theta \sim \mathcal{U}([0,1])$:

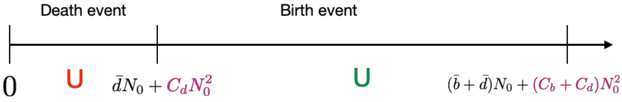
1. If $\theta \leqslant \dfrac{\lambda^e(I,T_1)}{\bar{e}}$ then: event e occurs to I at time T_1.
   Step 1 from $T_2 = T_1 + \mathcal{E}\left((\bar{b}+\bar{d})(N_0 \pm 1)\right)$
2. Else, nothing happen at T_1.
   Step 1 from $T_2 = T_1 + \mathcal{E}\left((\bar{b}+\bar{d})N_0\right)$

Algorithm with interactions

$$\Lambda(t) \leqslant (\bar{b} + \bar{d})N_t + (C_d + C_b)N_t^2$$

Step 1 Candidate event time $T_1 \sim \mathcal{E}\left((\bar{b} + \bar{d})N_0 + (C_d + C_b)N_0^2\right)$.

Step 2 Choose event type e and individual I :



Step 3 Accept/reject event with $\theta \sim \mathcal{U}([0,1])$.

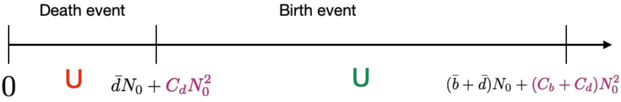
1. If $\theta \leqslant \dfrac{\lambda^e(I, T_1, Z_{T_1})}{C_e N_0}$ then: event e occurs to J at time T_1.
2. Else, nothing happen at T_1. BUT

Algorithm with interactions

$$\Lambda(t) \leqslant (\bar{b} + \bar{d})N_t + (C_d + C_b)N_t^2$$

Step 1 Candidate event time $T_1 \sim \mathcal{E}\left((\bar{b} + \bar{d})N_0 + (C_d + C_b)N_0^2\right)$.

Step 2 Choose event type e and individual I :



Step 3 Accept/reject event with $\theta \sim \mathcal{U}([0,1])$.

1. If $\theta \leqslant \dfrac{\lambda^e(I, T_1, Z_{T_1})}{C_e N_0}$ then: event e occurs to J at time T_1.
2. Else, nothing happen at T_1. BUT

$$\lambda^e(I, T_1, Z_{T_1}) = \sum_{J \in Z_{T_1}} U^e(I, J, T_1)$$

Need information on all the population to compute individual rate $\Rightarrow$ high cost

Interaction between two individuals

Idea See U^e as the interaction rate between two individuals taken at random in the population

- Step 3 becomes: Choose individual I and J at random and compute interaction intensity $U^e(I, J, T_1)$.
- Accept/Reject :
  1. If If $\theta \leqslant \frac{U^e(I, J, T_1)}{C_e}$ then: event e occurs to I at time T_1.
  2. Else, nothing happen at T_1.

A word on data management

- Removing someone from the population is costly but not picking someone at random.
- **Idea** Let "dead people" in the population and occasionally remove all dead individuals.
- Simple but very efficient : quadratic to linear simulation time (when they are no interactions).
- +Automatic parallelization when no interactions.

Example

Thank you!