Lecture Notes
Introduction to Stochastic Population Processes

Philippe CARMONA

March 18, 2019

Preface

These notes are based on lectures given at University of Nantes in 2018. Standard references for population dynamics are

- Stochastic models for Structured Populations Bansaye and Méléard [1]
- Markov Processes Ethier and Kurtz [2]
- Tom Britton, Etienne Pardoux, Stochastic epidemics in a homogeneous community, `https://arxiv.org/abs/1808.05350`
- Donald Dawson, Introductory Lectures on Stochastic Population Systems, `https://arxiv.org/abs/1705.03781`

The basic reference for convergence of Markov processes is the book by Ethier and Kurtz Ethier and Kurtz [2].

Contents

1
The Galton Watson process

1 Introduction and extinction probability

It is a dscter time Markov chain on the set $\mathbb{N}$. The discrete time parameter n is the generation number. In this population model, each individual produces, independently from the other individuals, a random number of descendant following the same offspring distribution, the law of an inteeger valeud random variable ξ. Given $(\xi^(k)_i, k \geq 1, i \geq 1)$ IID random variables distributed as ξ, the Markov hain is defined by induction : $X_0 = 1$ and

$$X_{n+1} = \sum_{i=1}^{X_n} \xi_i^{(n+1)} \tag{1.1}$$

We assume that the mean offspring is finite and not null

$$0 < m := \mathbb{E}[\xi] < +\infty\,. \tag{1.2}$$

Furthermore, we assume that the process is not deterministic, that is

$$\forall i \in \mathbb{N}, \mathbb{P}(\xi = i) < 1\,. \tag{1.3}$$

0 is an absorbing state, so $q_n = \mathbb{P}(X_n = 0)$ is increasing and

$$q := \lim \uparrow \mathbb{P}(X_n = 0) = \mathbb{P}(\exists n : X_n = 0) \quad \text{is the extinction probability.} \tag{1.4}$$

Let $f_n(s) = \mathbb{E}\left[s^{X_n}\right]$ be the generating function of X_n. We let $f(s) = f_1(s) = \mathbb{E}\left[s^{\xi}\right]$.

Lemma 1.1. *$f_{n+1}(s) = f_n(f(s))$ and therefore $f_n(s) = f \circ f \circ \cdots \circ f$ is the n-th composition of f.*

Proof. We have, $\xi_i^{(n+1)}$ independent from $\mathscr{F}_n := \sigma(\xi_j^{(k)}, k \leq n, j \geq 1) \supset \sigma(X_0, X_1, \ldots, X_n) =: \mathscr{F}_n^X$. Therefore

$$\mathbb{E}\left[s^{X_{n+1}} \mid X_n = k\right] = \mathbb{E}\left[s^{\xi_1^{(n+1)} + \ldots, \xi_k^{(n+1)}}\right] = f(s)^k\,. \tag{1.5}$$

Hence

$$f_{n+1}(s) = \mathbb{E}\left[\mathbb{E}\left[s^{X_{n+1}} \mid X_n\right]\right] = \mathbb{E}\left[f(s)^{X_n}\right] = f_n(f(s)). \tag{1.6}$$

□

Lemma 1.2. *The extinction probability satisfies $q \in [0,1]$ and*

$$q = f(q). \tag{1.7}$$

Proof. Consequently $q_n = \mathbb{P}(X_n = 0) = f_n(0)$ satisfies

$$q_{n+1} = f(q_n) \tag{1.8}$$

and taking limits yields the desired result. □

The study of the equation $f(s) = s$ for the function f non negative, increasing convex on $[0,1]$, with $f(1) = 1$, yields immediately the following dichotomy

Proposition 1.3. *If $m \leq 1$, then $q = 1$: there is almost sure extinction.*
If $m > 1$, then $q < 1$: there is a positive probability of non extinction.

We say that the process is subcritical if $m < 1$, critical if $m = 1$ and supercritical if $m > 1$.

The Galton–Watson process is a branching stochastic process arising from Francis Galton's statistical investigation of the extinction of family names. The process models family names as patrilineal (passed from father to son), while offspring are randomly either male or female, and names become extinct if the family name line dies out (holders of the family name die without male descendants) Assume that $\xi \sim \mathscr{B}(d,p)$ with $p = 1/2$. Then $f(s) = \left(\frac{s+1}{2}\right)^d$ and thus

- if $d = 3$, $1 - q = 0.77$ is the probability of survival of the name
- if $d = 5$, $1 - q = 0.96$.

Observe that if initially, $X_0 = 20$ then $q \to q^{20}$ and $1 - q^{20} \simeq 1$ for both $d = 3$ and $d = 5$.

2 The fundamental martingale and the actual growth of the population

The prcess $W_n = \frac{X_n}{m^n}$ is a positive martingale. Indeed, since $\xi_i^{(n+1)}$ are independent from $\mathscr{F}_n$

$$\mathbb{E}[X_{n+1} \mid \mathscr{F}_n] = \mathbb{E}[X_{n+1} \mid X_n]\mathbb{E}\left[\xi_1^{(n+1)} + \cdots + \xi_k^{(n+1)}\right]_{|k=X_n} = mX_n\,. \tag{1.9}$$ {eq:13}

Therefore, there exists a positive integrable finite rv W such that

$$W_n \to W \quad a.s. \tag{1.10}$$ {eq:14}

Hence, if $m < 1$, $X_n = m^n W_n$ converges as to 0 exponentially fast.
Trivially, $q = \mathbb{P}(\exists n_0, \forall n \geq 0, X_n = 0) \leq \mathbb{P}(W = 0)$ but we can say more.

Lemma 1.4.
$$\mathbb{P}(W = 0) \in \{q, 1\}.$$

Proof. We only need to prove that $s := \mathbb{P}(W = 0)$ satisfies $f(s) = s$. The i-th descendent from the firts generation has a martingale limit $W^{(i)}$. More precisely, if $X_n^{(i)}$ is the number of descendent in generation n of this i-th person, then the process $X^{(i)}$ are independent, distributed as X, independent from X_1 and for $n \geq 2$

$$X_n = \sum_{i=1}^{X_1} X_{n-1}^{(i)} \tag{1.11}$$ {eq:15}

Taking limit, we obtain that

$$W = \lim_{n \to +\infty} m^{-n} X_n = \frac{1}{m}(W^{(1)} + \cdots + W^{(X_1)}). \tag{1.12}$$ {eq:16}

Therefore

$$\begin{aligned}
s = \mathbb{P}(W = 0) = \mathbb{P}\big(\forall i \in \{1, \ldots, X_1\}, W^{(i)} = 0\big) \\
&= \sum_k \mathbb{P}\big(\forall i \in \{1, \ldots, X_1\}, W^{(i)} = 0 \mid X_1 = k\big)\mathbb{P}(X_1 = k) \\
&= \sum_k \mathbb{P}\big(W^{(i)} = 0, \forall i \in \{1, \ldots, k\}\big)\mathbb{P}(X_1 = k) \\
&= \sum_k s^k \mathbb{P}(X_1 = k) = f(s).
\end{aligned}$$

□

In the supercritical case, with mild integrability assumptions, either the process goes extinct, either the population grows a.s. at rate m^n.

Theorem 1.5 (Kesten-Stigum)**.** *Assume $m > 1$. If $\mathbb{E}\left[\xi \log^+ \xi\right] < +\infty$ then $(W_n)_n$ is UI, $\mathbb{E}[W] = 1$, $\mathbb{P}(W = 0) = q$ and $\{W > 0\} = \{\forall n, X_n > 0\}$ a.s. that is on the non extinction set the population grows exponentially fast.*
If $\mathbb{E}\left[\xi \log^+ \xi\right] = +\infty$, then W_n is not UI and $W = 0$ a.s.

Sketch. One can prove easily, exercise, that if $\mathbb{E}\left[\xi^2\right] < +\infty$ the W_n is UI. If W_n is UI, then by the optional stopping theorem at time $t = +\infty$, $\mathbb{E}[W] = \mathbb{E}[W_0] = 1$. Therefore one cannot have $\mathbb{P}(W = 0) = 1$. Hence $\mathbb{P}(W = 0) = q$.
Since $\{\exists n : X_n = 0\} \subset \{W = 0\}$ and these two sets have same probability, then these sets are equal a.s. □

In the subcritical case, we can determine the expected total size. In practice if the process models an infection with at time 0 exactly one infected person, then the total number of infected person is

$$\bar{X} = \sum_{n=0}^{+\infty} X_n$$

Lemma 1.6. *If $m < 1$ then $\mathbb{E}\left[\bar{X}\right] = \frac{1}{1-m}$.*

In the subcritical case, the asymptotics $X_n \sim m^n W$ suggests that the first hitting time of 0 has an exponential tail.

Lemma 1.7. *If $m < 1$ and $\mathbb{E}\left[\xi \log^+ \xi\right] < +\infty$, then there exists $K > 0$ such that*

$$\mathbb{P}(T_0 > n) = \mathbb{P}(X_n > 0) = K m^n (1 + o(1)) \quad (n \to +\infty).$$

We may be interested to study the mean time to extinction with an initial population of size N, with N large : it is of order $\ln(N)$.

Lemma 1.8. *If $m < 1$ and $\mathbb{E}\left[\xi \log^+ \xi\right] < +\infty$, then*

$$\mathbb{E}[T_0 \mid X_0 = N] \sim \frac{\ln N}{|\ln m|}. \tag{1.13}$$ {eq:17}

Proof. `https://www.math.uni-frankfurt.de/~ismi/vatutin/Lecture4.pdf` ☐

2

Birth and death processes

1 Definition and non explosion criteria

Definition 2.1. *A brth and death process is a pure jump Markov process, with values in $\mathbb{N}$, with jumps steps ± 1 and transition rates*

$$\begin{aligned} i \to i+1 \textit{ with rate} \qquad & \lambda_i \\ i \to i-1 \textit{ with rate} \qquad & \mu_i\,, \end{aligned}$$

with $\lambda_i \geq 0$, $\mu_i \geq 0$, $\lambda_0 = \mu_0 = 0$.

The Q matrix, or infinitesimal generator, is given by

$$Q_{i,i+1} = \lambda_i\,, \quad Q_{i,i-1} = \mu_i\,, \quad Q_{i,i} = -(\lambda_i + \mu_i) =: -q_i\,, \quad Q_{i,j} = 0 \text{ otherwise}\,. \tag{2.1}$$

When in state i the chain waits an $\mathcal{E}(q_i)$ time, then jumps to $i+1$ with probability $\frac{\lambda_i}{q_i}$, and to $i-1$ with probability $\frac{\mu_i}{q_i}$

Three important examples

1. The linear birth death process. It is a branching process with $\lambda_i = \lambda i$ and $\mu_i = \mu i$ ($\lambda, \mu > 0$ given).
2. The logistic birth death process: $\lambda_i = \lambda i$, $\mu_i = \mu i + c\, i(i-1)$ ($c > 0$).
3. The birth death process with immigration : $\lambda_i = \lambda i + \rho$, $\mu_i = \mu i$ ($\rho > 0$ given).

Le $(Z_n)_{n\geq 0}$ be the embedded Markov chain. It has transition matrix

$$p_{i,i+1} = 1 - p_{i,i-1} = \frac{\lambda_i}{q_i}\,. \tag{2.2}$$ {eq:1}

Recall that

Proposition 2.1. *There is non explosion iff*

$$\sum_{n\geq 0} \frac{1}{q_{Z_n}} = +\infty \quad a.s. \tag{2.3}$$ {eq:2}

Corollary 2.2. *Sufficient contions for non explosion are that*

- *either* $\sup_i q_i < +\infty$
- *either* $(Z_n)_{n\in\mathbb{N}}$ *is recurrent.*

Proposition 2.3. *The linear birth death process does not explode.*

Proof. Here we have for $i \geq 1$, $p_{i,i+1} = 1 - p_{i,i-1} = \frac{\lambda}{\lambda+\mu}$, and $p_{0,0} = 1$. So $Z_n = S_{n\wedge T_0}$ is a random walk $S_n = X_1 + \cdots + X_n$ with mean step $\mathbb{E}[X_1] = \frac{\lambda-\mu}{\lambda+\mu}$. If $\lambda > \mu$, then $S_n/n \to \mathbb{E}[X_1] > 0$, so $S_n \to +\infty$ and either $T_0 < +\infty$ and $\sum_{n\geq 0} \frac{1}{q_{Z_n}} = +\infty$, or

$$q_{Z_n} = \frac{1}{(\lambda+\mu)S_n} \sim \frac{C}{n}$$

and again $\sum_{n\geq 0} \frac{1}{q_{Z_n}} = +\infty$.
If $\lambda \leq \mu$ then $T_o < +\infty$ a.s, the chain is absorbed at 0, since $\liminf S_n = -\infty$ a.s. □

Theorem 2.4. *Let* $(X_t)_{t\geq 0}$ *be an integer valeud pure jump markov process with generator* Q. *Then* X *does not explodes a.s. iff the only non negative bounded solution of* $Q\phi(x) = \phi(x)$ *for* $x \geq 1$ *is* $\phi \equiv 0$.

Proof. We begin by showing that if $T_0 = 0 < T_1 < \cdots < T_n < \ldots$ are the jump times, with limit $T_\infty := \lim T_n$, then the function

$$\phi(x) := \mathbb{E}_x\left[e^{-T_\infty}\right] \tag{2.4}$$ {eq:3}

is non negative bounded and satisfies $Q\phi = \phi$. Indeed, $\phi(0) = 0$ and conditioning by T_1, thanks to the strong Markov property, if $x \geq 1$,

$$\mathbb{E}_x\left[e^{-T_\infty} \mid \mathscr{F}_{T_1}\right] = \mathbb{E}_x\left[e^{-T_1} e^{-(T_\infty - T_1)} \mid \mathscr{F}_{T_1}\right] = e^{-T_1}\mathbb{E}_{X_{T_1}}\left[e^{-T_\infty}\right] = e^{-T_1}\phi(Z_1).$$

Since T_1 and $Z_1 = X_{T_1}$ are independent,

$$\begin{aligned}\phi(x) &= \mathbb{E}_x\left[e^{-T_1}\right]\mathbb{E}_x\left[\phi(Z_1)\right] \\ &= \frac{q_x}{1+q_x}\sum_{y\neq x}\frac{q_{xy}}{q_x}\phi(y) \\ &= \frac{1}{1+q_x}\sum_{y\neq x} q_{xy}\phi(y).\end{aligned}$$

Thus,

$$Q\phi(x) = \sum_y q_{xy}\phi(y) = -q_x\phi(x) + \sum_{y\neq x} q_{xy}\phi(y) = \phi(x). \tag{2.5}$$ {eq:5}

Assume that the only non negative bounded solution of $Q\phi(x) = \phi(x)$ for $x \geq 1$ is $\phi \equiv 0$, then $\phi(x) = \mathbb{E}_x\left[e^{-T_\infty}\right] = 0$ so that $T_\infty = +\infty$, $\mathbb{P}_x$ a.s.
Reciprocaly if the process does not explodes a.s. there exists x such that $\phi(x) > 0$. □

Proposition 2.5. *Assume that $\lambda_i > 0$ for $i \geq 1$. Then, the birth death process does not explode a.s. iff*

$$\sum_{i\geq 1}\left(\frac{1}{\lambda_i} + \frac{\mu_i}{\lambda_i\lambda_{i-1}} + \cdots + \frac{\mu_i\cdots\mu_2}{\lambda_i\cdots\lambda_1}\right) = +\infty \tag{2.6}$$ {eq:8}

Corollary 2.6. *If there exists $\bar\lambda$ such that $\lambda_i \leq \bar\lambda i$, then the birth death process does not explode a.s.*

2 Extinction probabilities

Proposition 2.7. *The extinction probabilities $(u_i := \mathbb{P}_i(T_0 < +\infty), i \geq 1)$ satisfy the equation $Qu(i) = 0$ that is*

$$\lambda_i u_{i+1} - (\lambda_i + \mu_i)u_i + \mu_i u_{i-1} = 0. \tag{2.7}$$ {eq:9}

Proof. We condition by the value of the first jump $X_{T_1} = Z_1$. By the strong Markov property

$$\mathbb{P}_i\big(T_0 < +\infty \mid \mathscr{F}_{T_1}\big) = \mathbb{E}_{X_{T_1}}[T_0 < +\infty]. \tag{2.8}$$ {eq:4}

Therefore, taking expectations

$$u_i = \mathbb{E}_i[u(Z_1)] = \frac{\lambda_i}{\lambda_i + \mu_i}u_{i+1} + \frac{\mu_i}{\lambda_i + \mu_i}u_{i-1}. \tag{2.9}$$ {eq:10}

This is exactly the desired equation.

□

Proposition 2.8. *Given $N \geq 2$, let $u_i^{(N)} = \mathbb{P}_i(T_0 < T_N)$, for $0 \leq i \leq n$. Then $u_0^{(N)} = 1$, $u_N^{(N)} = 0$ and*

$$u_i^{(N)} = \frac{W_{N-1} - W_{i-1}}{W_{N-1}}, \quad \textit{with } W_n = 1 + \sum_{k=1}^n \frac{\mu_1\cdots\mu_k}{\lambda_1\cdots\lambda_k}. \tag{2.10}$$ {eq:11}

In particular

$$u_1^{(N)} = 1 - \frac{1}{W_{N-1}}. \tag{2.11}$$ {eq:12}

Proof. On can do the same proof as in the preceding proposition, or else apply the preceding proposition with rates $\lambda_i^{(N)} = \mu_i^{(N)} = 0$ for $i \geq N$, and get that $x_i^{(N)} := u_{i+1}^{(N)} - u_i^{(N)}$ satisfy $\lambda_i x_i^{(N)} = \mu_i x_{i-1}^{(N)}$, for $1 \leq i \leq N-1$. With $g_i = \frac{\mu_i}{\lambda_i}$ and $r_i = g_1 \cdots g_i$ we get $x_i^{(N)} = r_i x_0^{(N)}$,

$$u_i^{(N)} = 1 + x_0^{(N)} + \cdots x_{i-1}^{(N)} = 1 + x_0^{(N)} w_{i-1}\,.$$

The boundary condition $0 = u_N^{(N)}$ implies the value of $x_0^{(N)}$ and the formulas. □

By letting $N \to +\infty$, we have, when there is no explosion, $u_i^{(N)} \to u_i$

Theorem 2.9. *Let*

$$\alpha := 1 + \sum_{k=1}^{+\infty} \frac{\mu_1 \cdots \mu_k}{\lambda_1 \cdots \lambda_k}\,.$$

If $\alpha = +\infty$ then the extinction probabilities are all equal to 1. Otherwise, they are

$$u_i = \frac{1}{\alpha} \sum_{k=i}^{+\infty} \frac{\mu_1 \cdots \mu_k}{\lambda_1 \cdots \lambda_k}\,.$$

Example : the linear birth death process. $g_i = \frac{\mu}{\lambda}$, $r_i = \left(\frac{\mu}{\lambda}\right)^i$. There is a.s. extinction if $\frac{\mu}{\lambda}$. And if $\lambda > \mu$, the probability of extinction is

$$u_i = \left(\frac{\mu}{\lambda}\right)^i < 1\,.$$

Example : the logistic birth death process $g_i = \frac{\mu + c(i-1)}{\lambda}$,

$$r_i = \prod_{k=1}^{i} \frac{\mu + c(k-1)}{\lambda} \geq C k^i\,,$$

so $\sum_i r_i = +\infty$ and there is extinction a.s. In the mean the population stabilizes but the competition, stochastic, makes extinction inevitable.
Indeed if $f(x) = x$ and $x(t) = \mathbb{E}_1[X_t]$ then $Lf(x) = (\lambda - \mu + c)x - cx^2$. Therefore, by Kolmogorov forward equation, and Cauchy Schwarz inequality

$$x'(t) = \frac{d}{dt} P_t f(x) = P_t L f(x) = (\lambda - \mu + c)x(t) - c\mathbb{E}_1\left[X_t^2\right] \leq (\lambda - \mu + c)x(t) - c\,x(t)^2 \tag{2.12}$$

We write it

$$x'(t) = -C\,x(t)(x(t) - x_\infty), \quad \text{with} \quad x_\infty = \frac{c + \lambda - \mu}{c}\,. \tag{2.13}$$

Therefore if $x_\infty < 0$, then x is decreasing and if $x_\infty \geq 1$, then by comparison x stays below x_∞.

3

Linear Birth and Death process

1 The branching property

We say that the markov process $X = (X(t), t \geq 0; \mathcal{F}_t, t \geq 0; \mathbb{P}_x, x \in E)$ has the *branching property* if whenever X_1, X_2 are two independent copies of X starting from x_1, x_2 respectively, the process $X_1(t) + X_2(t)$ has the law of X starting from $x_1 + x_2$. Formally, this may be written as

$$\mathbb{P}_{x_1} * \mathbb{P}_{x_2} = \mathbb{P}_{x_1+x_2}. \tag{3.1}$$ {eq:defbran}

To establish 3.1 it suffices to prove equality of finite dimensional distributions. We let $\mathbb{P}_{x_1,x_2} = \mathbb{P}_{x_1} \otimes \mathbb{P}_{x_2}$ be the distribution of the couple of independent copies and $\mathbb{E}_{x_1,x_2}[.]$ be the corresponding expectation. We need to prove that for any $t_1 < t_2 < \ldots < t_n$, we have

$$\mathbb{E}_{x_1,x_2}\left[\prod_{i=1} f_i(X_1(t_i) + X_2(t_i))\right] = \mathbb{E}_{x_1+x_2}\left[\prod_i f_i(X(t_i))\right]. \tag{3.2}$$

It is easy to prove, by a monotone class theorem, that the process (X_1, X_2) is Markovian with respect to the filtration $\mathcal{G}_t := \sigma(X_1(s), X_2(s), s \leq t)$ with semi group

$$\mathbb{E}\big[f(X_1(t+s), X_2(t+s)) \mid \mathcal{G}_s\big] = \mathbb{E}_{X_1(s),X_2(s)}\big[f(X_1(t), X_2(t))\big]. \tag{3.3}$$

Therefore, by an easy induction, X is a branching process iff

$$\mathbb{E}_{x_1,x_2}\big[f(X_1(t) + X_2(t))\big] = \mathbb{E}_x\big[f(X(t))\big] \quad (\forall x_1, x_2, t). \tag{3.4}$$

Let us consider from now on, processes with values in $\mathbb{R}_+$ or $\mathbb{N}$. Then, since the Laplace tranform of positive rv's characterize their distributions, we let for $\theta \geq 0$, $f_\theta(x) := e^{-\theta x}$. By independence

$$\mathbb{E}_{x_1,x_2}\big[f_\theta(X_1(t) + X_2(t))\big] = \mathbb{E}_{x_1}\big[e^{-\theta X(t)}\big]\mathbb{E}_{x_2}\big[e^{-\theta X(t)}\big]. \tag{3.5}$$

Hence, X has the branching property, iff for any $\theta \geq 0, t \geq 0$ the function $h(x) = \mathbb{E}_x\left[e^{-\theta X(t)}\right]$ satisfies $h(x_1 + x_2) = h(x_1)h(x_2)$. This happens iff there exists $u(\theta, t)$ such that

$$P_t f_\theta(x) = \mathbb{E}_x\left[e^{-\theta X(t)}\right] = e^{-xu(\theta,t)}. \tag{3.6}$$

Let us specialize now to bd processes.

Proposition 3.1. *A birth and death process is a branching process iff it is a linear birth and death process.*

Proof. Recall that the generator is

$$Lf(x) = \lambda(x)(f(x+1) - f(x)) + \mu(x)(f(x-1) - f(x)),$$

and therefore the generator is

$$Lf_\theta(x) = \lim_{t\downarrow 0} \frac{1}{t}(P_t f_\theta(x) - f_\theta(x)) = -x\partial_t u(\theta, t)_{|t=0} e^{-xu(\theta,0)} = -x\partial_t u(\theta, t)_{|t=0} f_\theta(x). \tag{3.7}$$

On the other hand we have

$$Lf_\theta(x) = f_\theta(x)\big(\lambda(x)(e^{-\theta} - 1) + \mu(x)(e^\theta - 1)\big). \tag{3.8}$$

The only way for these two expressions to be equal for all x, θ is that $\lambda(x) = \lambda x$ and $\mu(x) = \mu x$.

□

2 Distribution at a fixed time

With a little bit extra work we can obtain the distribution of $X(t)$ at a fixed time, and deduc from it the extinction probability at a fixed time.

Proposition 3.2. *For a linear birth and death process starting from $x_0 = 1$, we have if $\lambda \neq \mu$,*

$$\mathbb{E}\left[e^{-\theta X_t}\right] = \frac{\mu(e^{-\theta} - 1)e^{(\lambda-\mu)t} - (\lambda e^{-\theta} - \mu)}{\lambda(e^{-\theta} - 1)e^{(\lambda-\mu)t} - (\lambda e^{-\theta} - \mu)}, \tag{3.9}$$

and if $\lambda = \mu$,

$$\mathbb{E}\left[e^{-\theta X_t}\right] = \frac{(\lambda t - 1)(e^{-\theta} - 1) - 1}{\lambda t(e^{-\theta} - 1) - 1}. \tag{3.10}$$

Of course, the branching property implies that

$$\mathbb{E}_x\left[e^{-\theta X_t}\right] = \left(\mathbb{E}_1\left[e^{-\theta X_t}\right]\right)^x. \tag{3.11}$$

Proof. We know from the branching property that $P_t f_\theta(x) = e^{-xu(\theta,t)}$. Therefore,

$$\mathbb{E}\left[X_t e^{-\theta X_t}\right] = -\partial_\theta P_t f_\theta(x), \tag{3.12}$$

and

$$\partial_t P_t f_\theta(x) = LP_t f_\theta(x) = (e^{-\theta}-1)\lambda\mathbb{E}\left[X_t e^{-\theta X_t}\right] + (e^{\theta}-1)\mu\mathbb{E}\left[X_t e^{-\theta X_t}\right] \tag{3.13}$$

$$= -(\lambda(e^{-\theta}-1)+\mu(e^{\theta}-1))\partial_\theta P_t f_\theta(x) \tag{3.14}$$

Therefore

$$\partial_t u(\theta,t) + (\lambda(e^{-\theta}-1)+\mu(e^{\theta}-1))\partial_\theta u(\theta,t) = 0. \tag{3.15}$$

We shall use the method of characteristics to solve this PDE. Let $(x_1(s), x_2(s))$ be a solution of

$$\frac{dx_2}{ds} = 1, \quad \frac{dx_1}{ds} = \lambda(e^{-x_1}-1)+\mu(e^{x_1}-1). \tag{3.16}$$

Then,

$$\frac{d}{ds}u(x_1(s), x_2(s)) = 0 \tag{3.17}$$

and $u(theta, t) = x_1(0)$ if we have the boundary conditions $x_2(0) = 0$, $x_2(t) = t$, $x_1(t) = \theta$.
For $\lambda \neq \mu$, we solve the ode for x_1

$$s + c = \int \frac{dx_1}{\lambda(e^{-x_1}-1)+\mu(e^{x_1}-1)} \tag{3.18}$$

$$= \int \frac{-dy}{\lambda(e^{y}-1)+\mu(e^{-y}-1)} \qquad (y = -x_1) \tag{3.19}$$

$$= -\frac{1}{\lambda-\mu}\ln\left(\frac{e^y-1}{\lambda e^y-\mu}\right) \tag{3.20}$$

That is

$$\frac{e^{(\lambda-\mu)s}(e^{-x_1(s)}-1)}{\lambda e^{-x_1(s)}-1} = constant. \tag{3.21}$$

Injecting the boundary conditions $x_1(t) = \theta, x_1(0) = u(\theta, t)$ yields the desired formula.

□

Letting $\theta \to +\infty$ in the preceding yileds the extinction probabilities

Corollary 3.3. *For a linear birth and death process starting from $x_0 = 1$, we have if $\lambda \neq \mu$,*

$$\mathbb{P}(X_t = 0) = \frac{\mu(1-e^{-(\lambda-\mu)t})}{\lambda-\mu e^{-(\lambda-\mu)t}}, \tag{3.22}$$

and if $\lambda = \mu$

$$\mathbb{P}(X_t = 0) = \frac{\lambda t}{1+\lambda t} \tag{3.23}$$

We can check that the (final) extinction probability is

$$q = \lim_{t\to+\infty} \mathbb{P}(X_t = 0) = 1 \wedge \frac{\mu}{\lambda} \tag{3.24}$$

3 The fundamental martingale

This linear BD process is the continuous analogue of the GW process. On non extinction, it grows exponentially fast.

Proposition 3.4. *Assume $\lambda > \mu$. The process $W_t = e^{-(\lambda-\mu)t} X_t$ is a positive martingale, Uniformly Integrable, converging a.s to a positive integrable random variable W_∞ and almost surely*

$$\{W_\infty > 0\} = \{\forall t, X_t > 0\} \tag{3.25}$$

Proof. Applying Kolmogorov forward equation to $f(x) = x$ yields

$$\partial_t P_t f(x) = P_t L f(x) = (\lambda - \mu) f(x)$$

and therefore $P_t f(x) = e^{(\lambda-\mu)t} f(x)$. Hence, if $s\ let$,

$$\mathbb{E}[W_t \mid \mathscr{F}_s] = e^{-(\lambda-\mu)t} P_{t-s} f(X_s) = W_s\,. \tag{3.26}$$

We have obviously

$$\{W_\infty > 0\} \subset \{\forall t, X_t > 0\} \tag{3.27}$$

and all we have to prove is that thes two sets have the same probability.
Similarly we can compute exactly $\mathbb{E}\left[X_t^2\right]$ and deduce that the martingale W_t is UI and thus $\mathbb{E}[W_\infty] = \mathbb{E}[W_0] = 1$.
On the other hand, conditionning by the first jump time, the strong Markov property yields that $s = \mathbb{P}(W_\infty = 0)$ satisfies

$$s = \frac{\lambda}{\lambda + \mu} s^2 + \frac{\mu}{\lambda + \mu} \tag{3.28}$$

Therefore $s \in \left\{1, \frac{\mu}{\lambda}\right\}$. and $\mathbb{E}[W_\infty] = 1$ imposes $s \neq 1$ therefore $s = \mu/\lambda$. □

4 Hitting times

Intuitively, the preceding results show that in the supercirtical case, it takes approximatively $logK$ unit of times to go from a population of 1 individual to a population of K individual. And in the subcritical case, it takes also $logK$ unit of time to go extinct starting from a polulation of order K.
Let $T_a = \inf\{t \geq 0 : X_t = a\}$ for $a \in \mathbb{N}$. Let $(t_K)_{K \geq 1}$ be a sequence of positive times such that $t_k >> \log K$.

{pro:hitlinbd}

Proposition 3.5. *1. Assume $\lambda < \mu$, subcritical case. Then for any $\epsilon > 0$*

$$\mathbb{P}_1\big(T_0 \le t_K \wedge T_{[\epsilon K]}\big) \to 1 \tag{3.29}$$ {eq:66}

and

$$\mathbb{P}_{[\epsilon K]}(T_0 \le t_K) \to 1. \tag{3.30}$$ {eq:67}

Moreover

$$\mathbb{P}_n(T_{kn} \le T_0) \le \frac{1}{k}. \quad (\forall n \ge 1, k \ge 1). \tag{3.31}$$ {eq:68}

2. Assume $\lambda > \mu$ (supercritical case). Then

$$\mathbb{P}_1\big(T_0 \le t_K \wedge T_{[\epsilon K]}\big) \to \frac{\mu}{\lambda}. \tag{3.32}$$ {eq:69}

and

$$\mathbb{P}_1\big(T_{[\epsilon K]} \le t_K\big) \to 1 - \frac{\mu}{\lambda} \tag{3.33}$$ {eq:70}

Proof. Since $T_a \ge a-1$, and $t_K \to +\infty$, we have

$$\mathbb{P}\big(T_0 \le t_K \wedge T_{[\epsilon K]}\big) \to \mathbb{P}(T_0 < +\infty)$$

and this yields (3.29) and (3.32)
The limit (3.30) follows from the exact computation of the extinction probability at time t_k

$$\mathbb{P}_{[\epsilon K]}(T_0 \le t_K) = \mathbb{P}\big(X_{t_K} = 0\big)^{[\epsilon K]} \tag{3.34}$$

The inequality (3.31) follows from Doob's stopping theorem applied to the UI martingale W_t and time $S = T_0 \wedge T_{kn}$

$$\mathbb{E}_n[W_S] = \mathbb{E}_n\big[W_{T_{kn}} \mathbf{1}_{(T_{kn} < T_0)}\big] = \mathbb{E}_n[W_0] = n \tag{3.35}$$

and sinc $\lambda < \mu$, $W_{T_{kn}} \ge kn$.
Eventually, (3.33) comes from the fact that on the extinction set, $\{W_\infty = 0\}$ of probability μ/λ, we have a finite progeny, to as $T_{[\epsilon K]} = +\infty$ for K large enough. On the survival set, $\{W_\infty > 0\} = \{T_0 = +\infty\}$ we have $X_t \sim W_\infty e^{(\lambda-\mu)t}$ and since $t_K >> log K$, we have both $T_0 = +\infty$ and $T_{[\epsilon K]} \le t_K$. □

4

Comparison of Markov Jump Processes

1 Motivation

Assume that X^1, X^2 are linear birth death processes with birth rates λ^i and death rates μ^i that satisfy

$$\lambda_1 \le \lambda_2\,, \quad \mu_1 \le \mu_2\,.$$

Our intuition tells us that if $X^1(0) \le X^2(0)$, then X^1 has more chances to be extinct at time t than X^2, that is

$$\text{if } x_1 \le x_2 \quad \text{then} \quad \mathbb{P}_{x_1}\big(X^1_t = 0\big) \ge \mathbb{P}_{x_2}\big(X^2_t = 0\big).$$

A first idea is to use exact computations that give

$$\mathbb{P}_{x_i}\big(X^i_t = 0\big) = \left(\frac{\mu_i(1 - e^{-(\lambda_i - \mu_i)t})}{\lambda_i - \mu_i e^{-(\lambda_i - \mu_i)t}}\right)^{x_i} \tag{4.1}$$

But even for $x_1 = x_2$, fixed t, checking that this function is decreasing in λ is not an easy task.

2 Stochastic Monotonicity

The state space $(E, \mathscr{E})$ is endowed with a measurable partial ordering $\prec$ such that

$$F := \{(x_1, x_2) : x_1 \prec x_2\} \in \mathscr{E}^2 \tag{4.2}$$

A measurable function $f : E \to \mathbb{R}$ is *monotone* if $x_1 \prec x_2 \implies f(x_1) \le f(x_2)$. A measurable set A is monotone if 1_A is monotone that is $x \in A$ and $x \prec y$ implies $y \in A$. We let $\mathscr{M}$ be the set of monotone functions (and the set of monotone sets). For two probability measures on E we say that $\mu_1 \prec \mu_2$ if for every non negative $f \in \mathscr{M}$, $\mu_1(f) \le \mu_2(f)$. For two random variables on $(E, \mathscr{E})$ we say that $X \prec Y$ if $P_X \prec P_Y$.

If F is closed and E polish theorem, then Strassen's theorem (see Lindvall [3]) states that it twho probabilities μ_1, μ_2 on $(E, \mathscr{E})$ satify $\mu_1 \prec \mu_2$ then there exists a probability measure P on $(E^2, \mathscr{E} \otimes \mathscr{E})$ with marginals μ_1, μ_2 such that $P(F) = 1$. In other words there exist random variables Y_1, Y_2 on $(E, \mathscr{E})$ such that $Y_1 \prec Y_2$ a.s. and $Y_i \sim \mu_i$.
Given two semigroups on $b\mathscr{E}$, we say that $P_1(t) \prec P_2(t)$ if

$$x_1 \prec x_2 \implies P_1(t)f(x_1) \le P_2(t)f(x_2)(t \ge 0, f \in b\mathscr{M}). \tag{4.3}$$

Lemma 4.1. *Let μ_1, μ_2 be two finite measures on $(E, \mathscr{E})$. The following are equivalent*

1. *$\mu_1(A) \le \mu_2(A)$ forall $A \in \mathscr{M}$*
2. *$\mu_1(f) \le \mu_2(f)$ forall $f \in b\mathscr{M}+$.*
3. *$\mu_1(f) \le \mu_2(f)$ forall $f \in \mathscr{M}+$.*

If furthermore $\mu_1(E) = \mu_2(E)$ then each of the above statement is also equivalent to

$$\mu_1(f) \le \mu_2(f) \tag{4.4}$$

for all $f \in \mathscr{M}$ such that the integrals exist.

We would like to compare generators and say that $L_1 \prec L_2$ implies $P_1(t) \prec P_2(t)$ forall t.
We say that the operators L_1, L_2 defined on $b\mathscr{E}$ satisfy $L_1 \prec L_2$ if $x_1 \prec x_2$ and A monotone, and either both $x_1, x_2 \in A$ or both $x_1, x_2 \in A^C$ imply $L_1 f(x_1) \le L_2 f(x_2)$.

Lemma 4.2. *If forall t $P_1(t) \prec P_2(t)$ then the corresponding generators satisfy $L_1 \prec L_2$.*

Proof. Assum $A \in \mathscr{M}$, $x_1 \prec x_2$. If $x_1 \in A, x_2 \in A$ then

$$L_i 1A(x_i) = \lim_{t\downarrow 0} \frac{1}{t}(P_i(t)1_A(x_i) - 1_A(x_i)) = \lim_{t\downarrow 0} \frac{1}{t}(P_i(t)1_A(x_i) - 1)$$

and since $P_1(t)1_A(x_1) \le P_2(t)1_A(x_2)$ we obtain $L1_A(x_1) \le L1_A(x_2)$. If $x_1 \not{i}nA$ and $x_2 \not{i}nA$ the proof is similar. □

We are going to prove that for processes with bounded rates, the condition $L_1 \prec L_2$ is also sufficient.
We have the following extension of Strassen's theorem.

Theorem 4.3. *Let X_1, X_2 be two Markov processes on the polish space $(E, \mathscr{E})$ with cadlag paths whose semigroups satisfy forall t $P_1(t) \prec P_2(t)$ and let μ, ν be two probabilities on E such that $\mu \prec \nu$. Then there exists a coupling that is two processes defined on the same probability space $(\hat{X}_1(t), \hat{X}_2(t), t \ge 0)$ such that $\hat{X}_1(0) \sim \mu$, $\hat{X}_2(0) \sim \nu$, $\hat{X}_1, \hat{X}_2$ have respective semigroups P_1, P_2 and*

$$a.s.\ \forall t \ge 0 \quad X_1(t) \le X_2(t). \tag{4.5}$$

Proof. This can be found in Kamae et al. [4, Theorem 5]. We say then that $X_1(t) \prec X_2(t)$ if we consider such a coupling. $\square$

3 Markov jump processes

Kernels and semigroups

Definition 4.1. *Let $(S, \mathscr{S})$ and $(T, \mathscr{T})$ be two measurable spaces. A function*

$$\kappa : S \times \mathscr{T} \to [0, +\infty] \tag{4.6}$$

is called a (transition) kernel if

1. *for any fixed $B \in \mathscr{T}$, the function $s \to \kappa_s(B)\kappa(s, B)$ is measurable.*
2. *for any fixed, $s \in S$, the function $B \to \kappa(s, B)$ is a measure on $(T, \mathscr{T})$.*

The kernel is said to be finite if all the measure κ_s are finite. It s a *Markov kernel*, or a probability kernel, if all the κ_s are probabilities.
To every finite kernel κ we associate the operator $A_\kappa : b\mathscr{T} \to b\mathscr{S}$ by:

$$A_\kappa f(s) = \int_T f(t)\kappa_s(dt) = \kappa_s(f). \tag{4.7}$$

If an operator $A : b\mathscr{T} \to b\mathscr{S}$ is positive in the sense that $f \geq 0$ implies $Af \geq 0$, then $\kappa(s, B) := A1_B(s)$ defines a finite kernel s.t. $A = A_\kappa$.
Let $(X(t), t \geq 0)$ be a stochastic process defined on a probability space with values in $(E, \mathscr{E})$, that is $(t, \omega) \to X(t, \omega)$ is a $\mathscr{B}([0, +\infty[) \otimes \mathscr{F} \to \mathbb{R}$ measurable function and let $\mathscr{F}_t^X := \sigma(X(s), s \leq t)$. Then X is a *Markov process* if

$$\mathbb{P}\left(X(t+s) \in A \mid \mathscr{F}_t^X\right) = \mathbb{P}(X(t+s) \in A \mid X(t)), \tag{4.8}$$

for all $s, t \geq 0$ and $A \in \mathscr{E}$.
If $(\mathscr{G}_t)_t$ is a filtration such that $\mathscr{F}_t^X \subset \mathscr{G}_t$ we say that X is a $(\mathscr{G}_t)$ Markov process if

$$\mathbb{P}(X(t+s) \in A \mid \mathscr{G}_t) = \mathbb{P}(X(t+s) \in A \mid X(t)), \tag{4.9}$$

for all $s, t \geq 0$ and $A \in \mathscr{E}$.

Proposition 4.4. *Assume that $(\mathbb{P}_x, x \in E)$ is a family of probability measure on $(\Omega, \mathscr{F})$ such that*

1. *X is a Markov process under each $\mathbb{P}_x$ such that*

$$\mathbb{P}_x(X(t+s) \in A \mid \mathscr{G}_t) = \mathbb{P}_{X(t)}(X(s) \in A) \tag{4.10}$$

for all $s, t \geq 0$, $x \in E$ and $A \in \mathscr{E}$.

2.

$$\mathbb{P}_x(X(0)=x)=1 \qquad (\forall x) \tag{4.11}$$

3. The operator $P_t f(x)=\mathbb{E}_x\left[f(X(t))\right]$ is a Markov kernel on E.

Then $(P_t, t\geq 0)$ is a Markov semigroup : $P_0 = id$ and $P_tP_s = P_{t+s}$.

Proof. If $f\in b\mathscr{E}$, then $P_0 f(x)=\mathbb{E}\left[f(X(0))\right]=f(x)$ so $P_0 = id$ and

$$P_{t+s}f(x)=\mathbb{E}_x\left[f(X(t+s))\right]=\mathbb{E}_x\left[\mathbb{E}_x\left[f(X(t+s))\mid \mathscr{G}_t\right]\right]=\mathbb{E}_x\left[P_s f(X_t)\right]=P_tP_sf(x). \tag{4.12}$$

□

To this semi-group we can associated transition kernels $P_t(x,dy)$ which are called transition functions associated to the markov process X, and the semi group equation is then called *Chapman-Kolmogorov equations*.

$$P_{t+s}(x,A)=\int P_t(x,dy)P_s(y,A). \tag{4.13}$$

It is worth observing that given such a Markovian semi-group, on a polish space $(E,\mathscr{E})$, then there exists a Markov process satisfying the assumptions of the proposition. It's distribution is uniquely determined (see e.g. [2, Theorem 1.1]). Therefore we shall identify Markov semi-groups with such Markov processes.

Definition and first properties

A pure jump Markov process defined on $(E,\mathscr{E})$ is a Markov process whose semigroup satisfies

$$\lim_{t\to 0} P_t 1_A(x)=1_A(x) \qquad (\forall x\in E, \forall A\in\mathscr{E}). \tag{4.14}$$

This is called the *continuity assumption* since it means that as $t\to 0$, $P_t 1_A(x)\to P_0 1_A(x)=1_A(x)$.
This is also called the *jump assumption* since this implies that the Markov process stays constant until the first jump.
For example a Brownian motion is not a pure jump process : it satisfies, for continuous bound f, $P_t f(x)\to f(x)$ but if $t>0$, since $p_t(x,dy)=\phi_t(y-x)\,dy$, the gaussian density, for $f=1_{\{x\}}$ we have $P_t f(x)=0$ and therefore $P_t f(x)\to 0\neq f(x)=1$.

Lemma 4.5. *Let X be a pure jump process and let $T:=\inf\{t>0 X_t\neq X_0\}$. Then, under $\mathbb{P}_x$ there exists $\alpha(x)\in[0,+\infty]$ such that $\mathbb{P}_x(T>t)=e^{-t\alpha(x)}$.*

Proof. By Markov property since $\{T>t\}\in\mathscr{F}_t$ and $\mathbf{1}_{(T>t+s)}=\mathbf{1}_{(T>t)}\mathbf{1}_{(T>s\circ\theta_t)}=\mathbf{1}_{(T>t)}\mathbf{1}_{(\tilde{T}>s)}$ with $\tilde{X}(s)=X(t+s)$ and $\tilde{T}=T(\tilde{X})$

$$\mathbb{P}_x(T>t+s\mid\mathscr{F}_t)=\mathbf{1}_{(T>t)}\mathbb{E}_{X_t}\left[\mathbf{1}_{(T>s)}\right]=\mathbf{1}_{(T>t)}\mathbb{P}_x(T>s) \tag{4.15}$$

□

If $\alpha(x) = 0$ (resp. $+\infty$, $\in (0, +\infty)$) we say that the state x is absorbing, instantaneous, stable.

Theorem 4.6. *(Chen [5, Theorem 1.4]) Let X be a Markov pure jump process with semigroup $(P_t)_{t\geq 0}$. Then there exists a measurable function $q : E \to [0, +\infty]$ such that*

$$\forall x, \lim_{t\to 0} \frac{1}{t}(1 - P_t 1_{\{x\}}(x)) = q(x). \tag{4.16}$$

We have $q(x) = \alpha(x)$, but this is not so simple to prove. The following applies in particular to processes with bounded rates,

Theorem 4.7. *(Chen [5, Theorem 1.11]) Let Let X be a markov pure jump process with semigroup $(P_t)_{t\geq 0}$ on E Polish, such that the set*

$$\{x : q(x) = +\infty\} \tag{4.17}$$

is at most countable. Then there exists a finite kernel q on E such that $q(x, \{x\}) = 0$ and for any $f \in b\mathcal{E}$

$$\lim_{t\to 0} \frac{1}{t} P_t f(x) - x = \int (f(y) - f(x)) q(x, dy) =: Lf(x). \tag{4.18}$$

L is the infinitesimal generator, and $q(x, E) = q(x)$ so we have, for x outside a countable set,

$$Lf(x) = \int f(y) q(x, dy) - q(x) f(x) \tag{4.19}$$

We shall assume from now on, except if otherwise stated, that forall x, $q(x) < +\infty$. The states for which $q(x) = 0$ are called absorbing. We have of course the Kolmogorov equations:

$$\frac{d}{dt} P_t f = P_t L f = L P_t f. \tag{4.20}$$

This of course applies to process on discrete state spaces whose Q matrix satisfy $q_i := \sum_{j\neq i} q_{ij} < +\infty$.

4 Bounded rate processes

We consider a jump process on $(E, \mathcal{E})$ with generator

$$Lf(x) = \int (f(y) - f(x)) q(x, dy) \tag{4.21}$$

with q a finite transition kernel that is a function $q : E \times \mathcal{E} \to \mathbb{R}_+$ such that

1. for each $x \in E$, $q(x, .)$ is a finite measure

2. for each $A \in \mathscr{E}$, $x \to q(x,A)$ is measurable. The total jump rate at state x is $q(x) := q(x,E)$.

Without loss in generality we shall assume that $q(x,\{x\}) = 0$. We say that the jup process has *bounded rates* if $\sup_{x\in E} q(x) < +\infty$.

Proposition 4.8. *Consider a Jump Markov process with bounded rates as above and let $b \geq \sup_x q(x)$. Then*

$$P_b = I + \frac{1}{b} L \tag{4.22}$$

is a Markov kernel that is $P_b f$ is positive bouded measurable if f is, and $P_b 1 = 1$. Let $Y = (Y_n)_{n\in\mathbb{N}}$ be a discrete time Markov chain with transition kernel P_b. Let $N = (N_t)_{t\geq 0}$ be a standard Poisson process on the line with rate b, independent from Y. Then

$$X_t = Y_{N_t} \tag{4.23}$$

is a Mrakov process with generator L.

Proof. Consider the filtration $\mathscr{F}_t = \sigma(Y_{n\wedge N_t}, n \in \mathbb{N}; N_s, s \leq t)$. The X_t is $\mathscr{F}_t$ measurable.
First we are going to prove that $(\tilde{N}_u = N_{t+u} - N_t, u \geq 0)$ is independent from $\mathscr{F}_t$. By the monotone class theorem it suffices to prove that

$$\mathbb{E}[UV] = \mathbb{E}[U]\mathbb{E}[V] \tag{4.24}$$

with $U = \prod_{j=1}^{K} h_j(\tilde{N}_{u_j})$, $V = \prod_{1\leq i\leq L} f_i(Y_{n_i\wedge t}) \prod_{1\leq l\leq M} g_l(N_{s_l})$, $u_1 < u_2 < u_M$, $n_1 < n_2 < \cdots < n_L$, $s_1 < \cdots < s_M \leq t$ and the functions f_i, g_l, h_j positive measurable bounded.
This is indeed true since

$$\begin{aligned}
\mathbb{E}[UV] &= \sum_{m_l, p_j} \mathbb{P}\Big(N_{s_l} = m_l, \tilde{N}_{u_j} = p_j\Big) \prod h_j(p_j) \prod g_l(m_l) \mathbb{E}\Big[\prod f_i(Y_{n_i\wedge t})\Big] \\
&= \sum_{m_l, p_j} \mathbb{P}\big(N_{s_l} = m_l\big)\mathbb{P}\Big(\tilde{N}_{u_j} = p_j\Big) \prod h_j(p_j) \prod g_l(m_l) \mathbb{E}\Big[\prod f_i(Y_{n_i\wedge t})\Big] \\
&= \mathbb{E}[U]\mathbb{E}[V]
\end{aligned}$$

Now we are going to prove that if we define

$$P_t f(x) := e^{tL} f(x) = \sum_{n\geq 0} \frac{t^n}{n!} L^n f(x)$$

which is well defined on $b\mathscr{E}$ since L is bounded, we have

$$\mathbb{E}\big[f(X_{t+s}) \mid \mathscr{F}_t\big] = P_s f(X_t), \tag{4.25}$$

for all positive bounded measurable f.

We decompose with respect to the values of $\tilde{N}_s = N_{t+s} - N_t$, which is independent of $\mathscr{F}_t$

$$
\begin{aligned}
\mathbb{E}\big[f(X_{t+s}) \mid \mathscr{F}_t\big] &= \mathbb{E}\big[f(Y_{N_t+\tilde{N}_s}) \mid \mathscr{F}_t\big] \\
&= \sum_k \mathbb{E}\big[f(Y_{N_t+k})\,\mathbf{1}_{(\tilde{N}_s=k)} \mid \mathscr{F}_t\big] \\
&= \sum_k \mathbb{P}\big(\tilde{N}_s = k\big)\mathbb{E}\big[f(Y_{N_t+k}) \mid \mathscr{F}_t\big] \\
&= \sum_k \mathbb{P}\big(\tilde{N}_s = k\big) p_b^k f(Y_{N_t}) \\
&= \sum_k e^{-bs}\frac{(bs)^k}{k!} p_b^k f(Y_{N_t}) \\
&= e^{-bs} e^{bsP_b} f(X_t) = e^{sL} f(X_t),
\end{aligned}
$$

since $e^{-bs}e^{bsP_b} = e^{-bsI+bs(I+\frac{1}{b}L)} = e^{sL}$.

□

The preceding construction dates back at least to Çinlar[6]

{thm:compbounded}

Theorem 4.9. *Assume that q_1, q_2 are finite transition kernels on $(E, \mathscr{E})$ a Polish space, with bounded rates, such that the associated generators satisfy $L_1 \prec L_2$. Then the associated semigroups satisfy forall t, $P_1(t) \prec P_2(t)$.*

Proof. Let $b > \sup_x q_1(x) + \sup_x q_2(x)$. Let Y_1, Y_2 be discrete time Markov chains associated to

$$
P_{b,i} := I + \frac{1}{b} L_i . \tag{4.26}
$$

Then $L_1 \prec L_2$ implies immediately that $P_{b,1} \prec P_{b,2}$. Indeed let $A \in \mathscr{M}$ and $x_1 \prec x_2$. If both x_1, x_2 are in A or A^C, since $L_1 1_A(x_1) \le L_2 1_A(x_2)$ we have $P_{b,1} 1_A(x_1) \le P_{b,2} 1_A(x_2)$. Since A is monotone the only case left to examine is $x_1 \notin A$ and $x_2 \in A$. We have

$$
L_1 1A(x_1) = q_1(x_1, A) \le q_1(x_1), \quad \text{and} \quad L_2 1_A(x_2) = q_2(x_2, A) - q_2(x_2) \ge -q_2(x_2) \tag{4.27}
$$

Therefore

$$
P_{b,1} 1_A(x_1) - P_2 1_A(x_2) = \frac{1}{b}(L_1 1_A(x_1) - L_2 1_A(x_2)) - 1 \le \frac{1}{b}(q_1(x_1) + q_2(x_2)) - 1 \le 0 .
$$

Let now $N = (N_t, t \ge 0)$ be a Poisson process with rate b, independent from Y_1 and Y_2. Let $X_i(t) = Y_i N_t$. Then, for any positive measurable f

$$
\begin{aligned}
P_1(t) f(x_1) = \mathbb{E}_{x_1}\big[f(Y_1(N_t))\big] &= \sum_n \mathbb{P}(N_t = n)\mathbb{E}_{x_1}\big[f(Y_n)\big] && \text{by independence} \\
&= \sum_n \mathbb{P}(N_t = n) P_{b,1}^n f(x_1) \\
&\le \sum_n \mathbb{P}(N_t = n) P_{b,2}^n f(x_2) && \text{since } P_{b,1} \prec P_{b,2} \\
&= P_2(t) f(x_2).
\end{aligned}
$$

□

5 Feller's construction of Markov jump process with unbounded rates

Assume that there exist borel subsets E_n o fE such that $B_n \uparrow E$ and $\sup_{n_n} q(x) \le n$. This is the case if q is locally bounded and E is locally compact separable. Then there exists on $E_\Delta = E \cup \{\Delta\}$ a Markov jump process $\bar{X}$ with generator $\bar{L}$ such that

- $\bar{q}(x,B) = q(x,B)$ if $B \in \mathscr{E}$
- if $\zeta = \left\{\inf t \ge 0 : X\bar{(}t) = \Delta\right\}$, then Δ is an absorbing point a.e. $\forall t \ge \zeta, X(t) = \Delta$.

We say that $X_t = \bar{X}_t \mathbf{1}_{(t<\zeta)}$ is a sub Markov jump process defined up its explosion time ζ. For bounded $f \in b\mathscr{E}$,

$$M_t^f := f(X_t) - f(X_0) - \int_0^t Lf(X_s)\,ds \qquad (t < \zeta) \tag{4.28}$$

is a local martingale with $Lf(x) = \int q(x,dy)(f(y)-f(x))$. Indeed, we know that $N_t = f(\bar{X}_t) - f(\bar{X}_0) - \int_0^t \bar{L}(X_s)\,ds$ is a martingale, hence $N_{t\wedge T_{E_n^C}}$ is, with f extended by $f(\Delta) = 0$, so $\bar{L}f(x) = Lf(x)$.

6 Unbounded rate jump Markov process

Assume that $(E,\mathscr{E})$ is Polish and that there exists G_δ's $E_n \uparrow E$ such that $\sup_{E_n} q(x) < +\infty$ and

$$H_n = \left\{y \in E \backslash E_n : \exists x \in E_n, x \prec y\right\} \text{ is monotone}, \tag{4.29}$$ {eq:comp27}

and if $H_n \ne \emptyset$, then there exist $b_n \in H_n$ such that

$$\forall x \in E_n, x \prec b_n. \tag{4.30}$$ {eq:comp28}

For example, if $E = \mathbb{R}^d, \mathbb{Z}^d, \ldots$ and q is locally bounded one could set

$$E_n = \{x \in E : -n \le x \le n\}, \quad b_n = (n+1,\ldots,n+1). \tag{4.31}$$

with $\le$ the classical lexical order.

Theorem 4.10. *Assume that q_1, q_2 are finite transition kernels on $(E,\mathscr{E})$ such that the associated generators satisfy $L_1 \prec L_2$ and such that $\sup_{x\in E_n} q_i(x) < +\infty$. Then the associated semigroups satisfy forall t, $P_1(t) \prec P_2(t)$.*

Proof. This is Chen [5, Theorem 5.47]. The idea is to build jump processes on $E_n + \{b_n\}$, apply the preceding results and taking limits. □

Definition 4.2. *We say that the semi group $P(t)$ is monotone if $P(t) \prec P(t)$ that is if for any $f \in b\mathscr{M}$*

$$x_1 \prec x_2 \implies \forall t, P(t)f(x_1) \leq P(t)f(x_2) \tag{4.32}$$

The comparison theorem is simpler when one of the processes is itself monotone.

Proposition 4.11. *Assume that q_1, q_2 are finite transition kernels on $(E, \mathscr{E})$, one of them monotone, such that the associated generators satisfy*

$$L_1 1_A(x) \leq L_2 1_A(x) \quad (\forall A \in \mathscr{M}, x \in E) \tag{4.33}$$

and $\sup_{x\in E_n} q_i(x) < +\infty$*. Then the associated semigroups satisfy forall t, $P_1(t) \prec P_2(t)$.*

Proof. Say that $P_1(t)$ is monotone. Then the assumption enables to prove as in Theorem 4.9 that for $f \in b\mathscr{M}$,

$$P_1(t)f(x) \leq P_2(t)f(x). \tag{4.34}$$

Therefore, if $x_1 \prec x_2$,

$$P_1(t)f(x_1) \leq P_1(t)f(x_2) \leq P_2(t)f(x_2) \tag{4.35}$$

Let us give a direct proof due to Rüschendorf [7]. Let us assume that for every $f \in b\mathscr{M}^+$, $L_1 f \leq L_2 f$ and that $P_2(t)$ is monotone. Fix $f \in b\mathscr{M}^+$ and consider

$$F(t,x) = P_2(t)f(x) - P_1(t)f(x). \tag{4.36}$$

Then

$$\partial_t F(t,x) = L_2 P_2(t)f(x) - L_1 P_1(t)f(x) = L_1 F(t,.)(x) + H(t,x) \tag{4.37}$$

with

$$H(t,x) = L_2 P_2(t)f(x) - L_1 P_2(t)f(x) = (L_2 - L_1)g(x) \tag{4.38}$$

with $g(x) = P_2(t)f(x)$ which by assumption is in $b\mathscr{M}^+$. Therefore, $H(t,x) \geq 0$ which is a crucial ingredient in the proof.
Observe now that

$$\frac{d}{ds} P_1(t-s)F(s,x) = P_1(t-s)\partial_s F(s,x) - P_1(t-s)L_1 F(s,x) = P_1(t-s)H(s,x) \geq 0 \tag{4.39}$$

Integrating this inequality between 0 and t yields then

$$F(t,x) - P_1(t)F(0,x) = \int_0^t P_1(t-s)H(s,x) \geq 0 \tag{4.40}$$

and since $F(0,x) = 0$ this yields $F(t,x) \geq 0$.

$\square$

7 Application

Comparison of Birth Death processes

As a warm up example we shall solve the problem of the introduction. q_i, $i = 1,2$ are BD processes with brth and death rates λ_i, μ_i such that $\lambda_1 \le \lambda_2$ and $\mu_1 \ge \mu_2$. The generators are thus

$$L_i f(x) = x(\lambda_i(f(x+1) - f(x)) + \mu_i(f(x-1) - f(x))). \tag{4.41}$$

We use the classical order on $\mathbb{N}:\ x \prec y$ if $x \le y$. Hence, a monotone set A is of the type $A = [n, +\infty)$ and we have

$$L_i 1_A(x) = x\lambda_i\, \mathbf{1}_{(x=n-1)} - x\mu_i\, \mathbf{1}_{(x=n)}. \tag{4.42}$$

Assume $x_1 \le x_2$. If $x_1 \in A$, is $n \le x_1$, then

$$L_1 1_A(x_1) = -x\mu_1\, \mathbf{1}_{(x_1=n)} \le L_2 1_A(x_2) = -x\mu_2\, \mathbf{1}_{(x_2=n)}. \tag{4.43}$$

If $x_2 \notin A$, i.e $x_2 < n$, then

$$L_1 1_A(x_1) = \lambda_1\, \mathbf{1}_{(x_1=n-1)} \le \lambda_2\, \mathbf{1}_{(x_2=n-1)} = L_2 1_A(x_2).$$

Therefore by the preceding theorem $P_1(t) \le P_2(t)$ and since $f(x) = -\mathbf{1}_{(x=0)}$ is monotone whenever $x_1 \le x_2$, $P_1(t)f(x_1) \le P_2(t)f(x_2)$ that is

$$\mathbb{P}_{x_1}\big(X_t^1 = 0\big) \ge \mathbb{P}_{x_2}\big(X_t^2 = 0\big). \tag{4.44}$$

Remark. *Observe that , taking $P_2 = P_1$ we have proved that birth death processes are monotone.*

Comparison of More General Jump Processes

We are now going to compare Markov jump processes that we consider in our law of large numbers. They have been introduced by Kurtz [8] as *density dependent Markov processes.* They have generators

$$L_i f(x) = \sum_{j=1}^{k} \beta_j^i(x)(f(x+h_j) - f(x)). \tag{4.45}$$

with locally bounded non negative rate functions β_j^i and $h_j \in \mathbb{Z}^d$.
We say that the vector h is quasi monotone if for any monotone set A and $x_1 \prec x_2$ if both x_1, x_2 are in A or in A^C then $1_A(x_1+h) - 1_A(x_1) \le 1_A(x_2+h) - 1_A(x_2)$.

Proposition 4.12. *If, for each j,*

1. *either h_j is quasi monotone and $x_1 \prec x_2$ implies $\beta^1(x_1) \le \beta^2(x_2)$*
2. *either $-h_j$ is quasi monotone and $x_1 \prec x_2$ implies $-\beta^1(x_1) \le -\beta^2(x_2)$*

Then $L_1 \prec L_2$ and thus forall $t \ge 0$, $P_1(t) \prec P_2(t)$.

The comparison of BD process is a simple application of this proposition with $h_1 = +1$, $\beta_1^1(x_1) = \lambda_1 x_1 \leq \lambda_2 x_2 = \beta_1^2(x_2)$ and $h_1 = +1$ is quasi monotone since if A is monotone, that is $A = [n, +\infty)$ then $1_A(x+1) - 1_A(x) = \mathbf{1}_{(x=n-1)} - \mathbf{1}_{(x=n)}$.
As before, everything is simpler when one of the processes is itself monotone.

Proposition 4.13. *Assume that either X_1 or X_2 is monotone. If for $A \in \mathscr{M}$ and forall x, $L_1 1_A(x) \leq L_2 1_A(x)$ then $P_1(t) \prec P_2(t)$.*

Comparison of SIR and BD processes

The SIR process has generator on Z^3:

$$L f(x) = \beta_1(x)(f(x+h_1) - f(x)) + \beta_2(x)(f(x+h_2) - f(x)) \tag{4.46}$$

with,if $x = (s, i, r) \in \mathbb{Z}_+^3$, $h_1 = (-1, 1, 0)$, $\beta_1(x) = \beta s i$, $h_2 = (0, -1, 1)$, $\beta_2(x) = \gamma i$.

Proposition 4.14. *Let $X_t = (S_t, I_t, R_t)$ be a SIR process with $N = \langle X_0, 1\rangle = S_0 + I_0 + R_0$ with parameters β, γ and let Z_t be a linear BD process with birth rate $\lambda = \beta N$ and $\mu = \gamma$ starting from $Z_0 \geq I_0$. Then* {pro:majsirbd}

$$I_t \prec Z_t \tag{4.47}$$

Proof. We introduce Y a Markov jump process on $\mathbb{Z}^3$ with generator

$$L^Y f(x) = \bar{\beta}_1(x)(f(x+h_1) - f(x)) + \beta_2(x)(f(x+h_2) - f(x)) \tag{4.48}$$

with $\bar{\beta}_1(x) = \beta N i$.
Our state space is $E = [0, N]^3 \cap \mathbb{Z}^3$ and $Y_0 \overset{d}{=} X_0$. The partial ordering we consider is $x = (s, i, r) \prec x' = (s', i', r')$ if and $i' \geq i$.
We check immediately that if $x \prec x'$ then $\bar{\beta}_1(x') = \beta N i' \geq \beta_1(x) = \beta s i$. A set A is monotone if for some n $A = \{(s, i, r) : i \geq n\}$. We check immediately that $h_1 = (-1, 1, 0)$ is quasi monotone since if $x \prec x'$ and both x, x' are in A or in A^C then

$$g(x) = 1_A(x+h_1) - 1_A(x) = \mathbf{1}_{(i+1 \leq n)} - \mathbf{1}_{(i \leq n)} \leq g(x')$$

Similarly $-h_2$ is quasi monotone. And thus if we let $Y_t = (S'_t, I'_t, R'_t)$ with $Y_0 \sim (S_0, Z_0, R_0)$ we have $X_0 \prec Y_0$ and thus $X_t \prec Y_t$. We now conclude since the I'_t is a BD process starting from Z_0. □

Corollary 4.15. *Let X be SIR process with parameters $\beta/N, \gamma$. Assume that the basic reproduction number $R_0 = \beta/\gamma \leq 1$, initial population $X_0 = (S_0, I_0, R_0)$ with $\langle X, 1\rangle = N$. Then the number of infected persons I_t goes to 0 in a time of order $\log I_0$ with a maximum of order $O(I_0)$.*

Proof. We have $I_t \prec Z_t$ and $W_t = Z_t e^{-(\beta-\gamma)t}$ is a UI martingale converging to W_∞ of expectation $\mathbb{E}[Z_0] = \mathbb{E}[I_0]$. Therefore $\max_t I_t \prec \max W_t$ is of order I_0. □

8 Killed processes and Markovian jump semigroups

Let $X = (X_t, t \geq 0)$ be a Markov jump process dénied on the polish space $(E, \mathscr{E})$ with generator

$$Lf(x) = \int_E (f(y) - f(x)) q(x, dy) \qquad (f \in b\mathscr{E}). \tag{4.49}$$

Define for $A \in \mathscr{E}$

$$T = T_{A^c} := \inf t > 0 : X_t \in A^C. \tag{4.50}$$

Let $A \in \mathscr{E}$ be such that $\forall x \in A, q(x) > 0$.
We let $\mathscr{A} = A \cap \mathscr{E}$ be the trace sigma field, and $b\mathscr{A}$ be the set of functions $f : A \to \mathbb{R}$, bounded, $\mathscr{E}$ measurable. For $f \in b\mathscr{A}$, and $t \geq 0$ we let

$$S_t f(x) := \mathbb{E}_x \left[f(X_t) \mathbf{1}_{(t<T)} \right] \qquad (x \in A). \tag{4.51}$$

Proposition 4.16. $(S_t, t \geq 0)$ *is a sub Markovian jump semigroup on* $b\mathscr{A}$*, that is*

1. *If* $f \geq 0$*, then* $S_t f \geq 0$ *and* $S_t 1 \leq 1$.
2. $S_{t+s} = S_t \circ S_s$
3. $S_0 f(x) = \lim_{t \downarrow 0} S_t f(x) = f(x)$.

Proof. 1. The first assertion is obvious.

2. For the second, we observe that

$$\mathbf{1}_{(t+s<T)} = \mathbf{1}_{(s<T)} \mathbf{1}_{(t<T)} \circ \theta_s. \tag{4.52}$$

Therefore, by Markov Property applied at time s

$$\begin{aligned} S_{t+s} f(x) &= \mathbb{E}_x \left[\mathbf{1}_{(s<T)} \mathbb{E}_x \left[(f(X_t) \mathbf{1}_{(t<T)}) \circ \theta_s \mid \mathscr{F}_s \right] \right] \\ &= \mathbb{E}_x \left[\mathbf{1}_{(s<T)} \mathbb{E}_{X_s} \left[f(X_t) \mathbf{1}_{(t<T)} \right] \right] = \mathbb{E}_x \left[\mathbf{1}_{(s<T)} S_t f(X_s) \right] \\ &= S_s (S_t f)(x). \end{aligned}$$

3. We have obviously for $x \in A$, $S_0 f(x) = f(x)$ since $T \geq T_1$ the first jump time of X. For $t > 0$, we decompose the expectation with respect to the value of $T_1 \sim \mathscr{E}(q(x))$ to obtain

$$\begin{aligned} S_t f(x) &= \mathbb{E}\left[f(X_t) \mathbf{1}_{(t \leq T_1)} \right] + \mathbb{E}\left[f(X_t) \mathbf{1}_{(T_1 < t)} \mathbf{1}_{(t<T)} \right] \\ &= f(x) e^{-tq(x)} + \int_0^t q(x) e^{-sq(x)} \mathbb{E}\left[f(X_t) \mathbf{1}_{(t<T)} \mid T_1 = s \right] ds \\ &= f(x) e^{-tq(x)} + \int_0^t e^{-sq(x)} \left(\int q(x, dy) f(y) \mathbf{1}_{(y \in A)} S_{t-s} f(y) \right) ds \end{aligned}$$

Indeed we know that X_{T_1} and T_1 are independent, with X_{T_1} with law $\frac{1}{q(x)}q(x,dy)$ so by the strons Markov property at time T_1

$$\begin{aligned}
\mathbb{E}\big[f(X_t)\mathbf{1}_{(t<T)} \mid T_1 = s\big] &= \int \mathbb{P}\big(X_{T_1} \in dy\big)\mathbb{E}\big[f(X_t)\mathbf{1}_{(t<T)} \mid T_1 = s, X_{T_1} = y\big] \\
&= \int \frac{1}{q(x)} q(x,dy)\mathbf{1}_{(y\in A)}\mathbb{E}\big[f(X_t)\mathbf{1}_{(t<T)} \mid T_1 = s, X_{T_1} = y\big] \\
&= \int \frac{1}{q(x)} q(x,dy)\mathbf{1}_{(y\in A)}\mathbb{E}\big[(f(X_{t-s})\mathbf{1}_{(t-s<T)})\circ\theta_{T_1} \mid T_1 = s, X_{T_1} = y\big] \\
&= \frac{1}{q(x)}\int q(x,dy)\mathbf{1}_{(y\in A)} S_{t-s} f(y).
\end{aligned}$$

Hence

$$S_t f(x) = f(x)e^{-tq(x)} + \int_0^t e^{-(t-s)q(x)}\bigg(\int q(x,dy)f(y)\mathbf{1}_{(y\in A)} S_s f(y)\bigg)ds \tag{4.53}$$

{eq:comp:27}

and we have $S_t f(x) \to f(x)$ as $t \to 0$ by dominated convergence.

$\square$

It is easy to determine the generator of this Markovian semigroup by using the last formula (4.53).

Lemma 4.17. *For every $f \in b\mathscr{A}$ and $x \in A$, the following limit exists*

$$Gf(x) = \lim_{t\downarrow 0} \frac{1}{t}\big(S_t f(x) - f(x)\big) \tag{4.54}$$

and we have

$$Gf(x) = q(x)f(x) - \int_A q_A(x,dy)f(y), \tag{4.55}$$

with $q_A(x,dy) = q(x,dy)\mathbf{1}_{(y\in A)}$. Furthermore, the semigroup property implies that for $t > 0$

$$\frac{d}{dt} S_t f(x) = GS_t f(x) = S_t G f(x). \tag{4.56}$$

We can always extend a submarkovian semigroup to a Markovian one by adding a cemetary point Δ. On $A_\Delta = A \cup \{\Delta\}$ we set $P_t f(\Delta) = f(\Delta)$ and for $x \in A$,

$$P_t f(x) = S_t f(x) + f(\Delta)(1 - \mathbb{P}_x(t < T)). \tag{4.57}$$

We have of course, $P_t f(x) = \mathbb{E}_x\big[f(Y_t)\big]$ with $Y_t = X_t\,\mathbf{1}_{(t<T)} + \Delta\,\mathbf{1}_{(t\geq T)}$. The generator is

$$\bar{L} f(x) = \int \bar{q}(x,dy)(f(y) - f(x)), \tag{4.58}$$

with $\bar{q}(x,B) = q(x,B)$ if $B \subset A$ and $\bar{q}(x,\{\delta\}) = q(x,A^C)$. Therefore if $f(\Delta) = 0$, $Gf(x) = \bar{L}f(x)$.
We now suppose that E is endowed with a measurable partial ordering. We extend the partial order to A_Δ by setting $\Delta \prec x$ forall x.
If $B \subset A_\Delta$ is monotone then either $\Delta \in B$ and then $B = A_\Delta$, either $\Delta \notin B$ and B is monotone in A.
Therefore we have $\bar{L}_1 \prec \bar{L}_2$ iff for any monotone B in A, and $x_1 \prec x_2$:

- either $x_1, x_2 \in B$ and $q_1(x,B) - q_1(x) \leq q_2(x,B) - q_2(x)$
- either $x_1, x_2 \notin B$ and $q_1(x,B) \leq q_2(x,B)$

In ohter words we only have to check that the conditions of $L_1 \prec L_2$ for $x_1, x_2 \in A$ and monotones $B \subset A$.

Proposition 4.18. *Assume $\bar{L}_1 \prec \bar{L}_2$. Then there exists a coupling (X_1, X_2) such that*

$$T_{A^C}(X_1) \leq T_{A^C}(X_2), \tag{4.59}$$

and

$$X_1(t) \prec X_2(t) \quad on\,[0, T_{A^C}(X_1)),. \tag{4.60}$$

Proof. We have $\bar{L}_1 \prec \bar{L}_2$, so we can construct the coupling for the killed processes $Y_1(t) \prec Y_2(t)$. If we have $T_{A^C}(X_1) > T_{A^C}(X_2)$, this implies that for some t, $Y_2(t) = \Delta$ and $Y_1(t) \neq \Delta$ which is contradictory. □

9 Another comparison between SIR and BD processes

Let $X_t = (S_t, I_t, R_t)$ be a SIR process with $N = \langle X_0, 1\rangle = S_0 + I_0 + R_0$ with parameters β, γ. Let $0 < \epsilon < 1$ and let Z_t be a linear BD process with birth rate $\lambda = \beta N(1-\epsilon)$ and $\mu = \gamma$ starting from $Z_0 = I_0$. Let B_t be the number of births in Z until time t and $T = \inf\{t > 0 : B_t < S_0 - N(1-\epsilon)\}$.

{pro:minsirbd}

Proposition 4.19. *There exists a coupling such that*

$$Z_t \leq I_t \quad on [0, T) \tag{4.61}$$

Proof. We consider $\bar{X}_t = (\bar{S}_t, \bar{I}_, \bar{R}_t)$ a jump Markov process with generator

$$\bar{L}f(x) = \bar{\beta}_1(x)(f(x+h_1) - f(x)) + \beta_2(x)(f(x+h_2) - f(x)), \tag{4.62}$$

with $\beta_1(x) = \beta N(1-\epsilon)i$. We consider the same order on $E = \left\{x = (s,i,r) \in (\mathbb{Z} \cap [0,N])^3\right\}$ as in Proposition 4.14
With $A = \{x \in E : s \geq N(1-\epsilon)\}$ we have $\bar{\beta}_1(x) \leq \beta_1(x)$ in A. We want to prove that we have $\bar{X} \prec X$ on $[0, T_{A^C}(\bar{X})$. Since $\bar{X}$ is monotone, we only need to prove that if $x \in A$ and B is monotone in A, then

$$\bar{L}1_B(x) \leq L1_B(x)$$

Since B is of the type $B = \{x \in A : i \geq n\}$ then the proof goes as in Proposition 4.14.
In the process $\bar{X}$ the process $\bar(I)$ is a linear BD process with birth rate $\beta N(1-\epsilon)$ and death rate γ, and $\bar{S}_t = S_0 - B_t$ since $\bar{S}_t$ decreases by 1 exactly when I_t increases by 1. Therefore $T_{A^C} = T$ defined above and we are done. □

Remark. *Observe that by the representation* (5.1)

$$Z_t = Z_0 + P_1(\lambda \int_0^t Z_s ds) - P_2(\int_0^t \mu Z_s ds)$$

and $B_t = P_1(\lambda \int_0^t Z_s ds)$ *We know that in the critical case, on the non explosion set of* Z, Z *grows exponentially fast, therefore with very high probability, if* S_0 *is of order* N, *then* T *is of order* $\log(N)$ *(since* $P(t)/t \sim 1$*).*

5

Law of Large numbers for Random Markov Epidemic Models

1 Another representation of Some Markov jump processes

{sec:anoth-repr-som

{pro:anoth-repr-som

Proposition 5.1. *Let $(h_i)_{1\le i\le k}$ be jump vectors in $\mathbb{Z}^d$ and $(P_i)_{1\le i\le k}$ be independent rate 1 Poisson processes independent from a random variable $X_0 \in \mathbb{Z}^d$. Let $\beta_j : \mathbb{Z}^d \to \mathbb{R}_+$ for $1 \le j \le k$.*
Then the equation

$$X_t = X_0 + \sum_{j=1}^{k} h_j P_j\left(\int_0^t \beta_j(X_s)\,ds\right), \tag{5.1}$$

{eq:defmarkepidemic

admits a.s. a unique solution which is a Markov jump process on $\mathbb{Z}^d$ with generator

$$Lf(x) = \sum_{j=1}^{k} \beta_j(x)(f(x+h_j) - f(x)) \qquad (f \text{ bounded.}) \tag{5.2}$$

{eq:genmarkovepidem

Remark. *Let us observe that this process may have a finite explosion time ζ. Furthermore, by the construction procedure if $X_0 \ge 0$ the forall t, $X_t \ge 0$.*

Example 5.1.
1. *The birth death process. $d = 1$, $h_1 = 1, \beta_1(x) = \lambda(x)$, $h_2 = -1, \beta_2(x) = \mu(x)$.*
2. *The SIR process $d = 3$, for $x = (s, i, r) \in \mathbb{Z}_+^d$,*

$$\begin{aligned} h_1 &= (-1, 1, 0), & \beta_1(x) &= \lambda s i \\ h_2 &= (0, -1, 1) & \beta_2(x) &= \gamma i \end{aligned}$$

λ is the percapita infectious contact rate and γ the percapita recovery rate. If $X_t = (S_t, I_t, R_t)$ then X is the solution of the SDE

$$S_t = S_0 - P_1\left(\int_0^t \lambda S_s I_s \, ds\right). \tag{5.3}$$

$$I_t = I_0 + P_1\left(\int_0^t \lambda S_s I_s \, ds\right) - P_2\left(\int_0^t \gamma I_s \, ds\right). \tag{5.4}$$

$$R_t = R_0 + P_2\left(\int_0^t \gamma I_s \, ds\right). \tag{5.5}$$

Proof. SInce independent Poisson processes do not jump at the same time a.s. we can do a pathwise construction of X_t, indectively along the jumps of X.
Let $Z_0 = X_0, \ldots, Z_n$ be the n first values of the jump chain, $S_1, \ldots, S_n$ the holding times, $T_N = S_1 + \cdots + S_n$ the n-th jump time. Then the next jup time is $T_{n+1} = T_n + S_{n+1}$ is the first time $t > T_n$ such that there exists j with

$$P_j\left(\int_0^{T_n} \beta_j(X_s)\, ds + \beta_j(Z_{n-1})(t - T_n)\right) - P_j\left(\int_0^{T_n} \beta_j(X_s)\, ds\right) \neq 0. \tag{5.6}$$

By the strong Markov property, these are independent Poisson processes of rates $\alpha_j = \beta_j(Z_{n-1})$. Let V_j be their respective first jump times. Then $V_j \sim \mathscr{E}(\alpha_j)$, $S_{n+1} = \inf_j V_j \sim \mathscr{E}(\sum_j \alpha_j)$ and $Z_n = Z_{n-1} + h_j$ with probability $\frac{\alpha_j}{\sum_i \alpha_i}$. This is exactly the usual construction of the Markov jump process with generator L given by (5.2). □

2 A non explosion criteria

We shall give a sufficient condition for non explosion for the process defined by Proposition 5.1. We shall exhibit a *Lyapunov function* if we make the following assumption on rates. for $x \in \mathbb{Z}^d$ we let $\langle x, 1\rangle = \sum_{i=1}^d x_i$ (if $x \geq 0$, the $\langle x, 1\rangle = \|x\|_1$.

Assumption A : rate control Let $J = \{j : \langle h_j, 1\rangle > 0\}$ Assume that for some $C_q < +\infty$,

$$\sup_{j\in J} \beta_j(x) \leq C_q(1 + \langle x, 1\rangle). \tag{5.7}$$

{pro:non-expl-crit}

Proposition 5.2. *Assume the rate control and that for some $p \geq 1$, $\mathbb{E}\left[\langle X_0, 1\rangle^p\right] < +\infty$ and $X_0 \geq 0$ a.s. Then, a.s. the process does not explodes and*

$$\forall T > 0, \quad \sup_{t\leq T} \mathbb{E}\left[\langle X_t, 1\rangle^p\right] < +\infty. \tag{5.8}$$

Proof. We let $Z_t = \langle X_t, 1\rangle$. Then given $a > 0$, we can consider bounded rates $\beta_j^a(x) = \beta_j(x)\mathbf{1}_{(\langle x,1\rangle \leq a)}$ and construct a process that has infinite life time X^a (for example by Cinlar construction we see that it does not explode) and the corresponding generator. We let

$$\tau_a = \inf t > 0 : \left\langle X_t^a, 1 \right\rangle \le a \tag{5.9}$$

and we define without ambiguity $X_t = X_t^a$ on $[0, \tau^a[$. We let $\zeta := \lim_{a \uparrow +\infty} \tau_a$. This is the lifetime of X.
Given a locally bounded function $f : \mathbb{Z}^d \to \mathbb{R}$, we define $f^a(x) = f(x)\mathbf{1}_{(\langle x,1\rangle \le a)}$ and

$$f^a(X_T^a) = f^a(X_0^a) + M_t^{f^a} + \int_0^t L^a f^a(X_s^a)\,ds, \tag{5.10}$$

with M^{f^a} a martingale. Since $L^a f^a(x) = Lf(x)$ for $\langle x, 1\rangle \le a$ we have, if we set $M_t^f = M_t^{f^a}$ on $[0, \tau_a[$,

$$f(X_t) = f(X_0) + M_t^f + \int_0^t Lf(X_s)\,ds \quad (t < \zeta).$$

From now on, we shall drop the superscript a, but keep in mind that on $[0, \tau_a[$ we are dealing with X^a, M^{f^a}, ...
Observe that

$$\begin{aligned}
Lf(x) &= \sum_j \beta_j(x)\Big(\big(\langle x,1\rangle + \left\langle h_j, 1\right\rangle\big)^p - \langle x,1\rangle^p\Big) \\
&\le \sum_{j\in J} \beta_j(x)\Big(\big(\langle x,1\rangle + \left\langle h_j, 1\right\rangle\big)^p - \langle x,1\rangle^p\Big) \\
&\le \sum_{j\in J} \beta_j(x) C_p \langle x,1\rangle^{p-1} \\
&\le C_p C_q k(1 + \langle x,1\rangle)\langle x,1\rangle^{p-1} \\
&\le C(1 + \langle x,1\rangle^p) = Cf(x).
\end{aligned}$$

Since $M_{t\wedge\tau_a}^f$ is a true martingale we get

$$\begin{aligned}
\mathbb{E}\big[f(X_{t\wedge\tau_a})\big] &= \mathbb{E}\big[f(X_0)\big] + \mathbb{E}\left[\int_0^{t\wedge\tau_a} Lf(X_s)\,ds\right] \\
&= \mathbb{E}\big[\langle X_0, 1\rangle^p\big] + C\int_0^t \mathbb{E}\big[1 + \left\langle X_{s\wedge\tau_a}, 1\right\rangle^p\big]\,ds.
\end{aligned}$$

By Gronwall's Lemma, there exists a constant C' that does not depend on a, but only on $\mathbb{E}\big[\langle X_0, 1\rangle^p\big]$ such that

$$1 + \mathbb{E}\Big[\left\langle X_{t\wedge\tau_a}, 1\right\rangle^p\Big] \le C'(1+t)e^{tC'} \tag{5.11}$$

{eq:1-+-espcrochetx

In particular $sup_{t\le T}\mathbb{E}\Big[\left\langle X_{t\wedge\tau_a}, 1\right\rangle^p\Big] < +\infty$.

Consequently, $\zeta := \lim_{a\uparrow+\infty} \tau_a = +\infty$ a.s. Indeed, otherwise thaere exists $T > 0$ such that $\mathbb{P}(\zeta \le T) > 0$. Then

$$C'(1+T)e^{TC'} \ge \mathbb{E}\left[\left\langle X_{T\wedge\tau_a}, 1\right\rangle^p\right] \ge \mathbb{E}\left[\left\langle X_{T\wedge\tau_a}, 1\right\rangle^p \mathbf{1}_{(\tau_a \le T)}\right]$$
$$\ge a^p \mathbb{P}(\tau_a \le T) \ge a^p \mathbb{P}(\zeta \le T) \quad \to +\infty (\text{as } a \to +\infty)$$

which is absurd. Hence $\zeta = +\infty$ a.s. and by Fatou's Lemma letting $a \uparrow +\infty$ in (5.11) we get

$$1 + \mathbb{E}\left[\langle X_t, 1\rangle^p\right] \le C'(1+t)e^{C't}. \tag{5.12}$$

□

With a little extra work we can get a maximal inequality.

{pro:non-expl-crit-

Proposition 5.3. *Under the same assumptions, for any $q \in \left[1, \frac{p+1}{2}\right]$ and any $T > 0$, we have*

$$\mathbb{E}\left[\sup_{t\le T} \langle X_t, 1\rangle^q\right] < +\infty. \tag{5.13}$$

Proof. With $f(x) = \langle x, 1\rangle^q$ we have $Lf(x) \le C(1+f(x))$ and thus

$$f(X_t) \le f(X_0) + M_t^f + C\int_0^t (1+f(X_s))\,ds. \tag{5.14}$$

Therefore, if $Y_t := \sup_{s\le t} f(X_s)$ we have for $t \in [0, T]$

$$Y_t \le Y_0 + \sup_{t\le T} M_t^f + Ct + C\int_0^t Y_s\,ds. \tag{5.15}$$

Hence, by Gronwall's Lemma

$$Y_T \le (Y_0 + \sup_{t\le T} M_T^f + CT)e^{CT}. \tag{5.16}$$

It remains to prove that $\mathbb{E}\left[\sup_{t\le T} M_t^f\right] < +\infty$. Remember that the predictable quadratic variation of the martingale M^f is given by the carré du champ operator

$$\left\langle M^f, M^f\right\rangle_t = \int_0^t (Lf^2 - 2fLf)(X_s)\,ds. \tag{5.17}$$

We have

$$\begin{aligned}
Lf^2(x) - 2f(x)Lf(x) &= \sum_j \beta_j(x)(f^2(x+h_j) - f^2(x) - 2f(x)(f(x+h_j) - f(x))) \\
&= \sum_j \beta_j(x)(f(x+h_j) - f(x))^2 \\
&= \sum_{j\in J} \beta_j(x)(f(x+h_j) - f(x))^2 && (f(x+h_j) = f(x) \text{ if } j \notin J) \\
&\le C(1+\langle x, 1\rangle)(1+\langle x, 1\rangle^{q-1})^2 \\
&\le C(1+\langle x, 1\rangle^p).
\end{aligned}$$

Hence, by Doob's maximal inequality

$$\mathbb{E}\left[\left(\sup_{t\leq T} M_t^f\right)^2\right] \leq C\mathbb{E}\left[(M_T^f)^2\right] = C\mathbb{E}\left[\left\langle M^f, M^f\right\rangle_T\right]$$
$$\leq C\int_0^T (1+\mathbb{E}\left[\langle X_t, 1\rangle^p\right])\,dt$$
$$\leq C\sup_{t\leq T}\mathbb{E}\left[\langle X_t, 1\rangle^p\right] < +\infty\,.$$

□

Corollary 5.4. *Assume thaht f locallly bounded satisfies for a constant C,*

$$\left|f(x)\right| + \left|Lf(x)\right| \leq C(1+\langle x,1\rangle^q), \tag{5.18}$$

with $q \in \left[1, \frac{1}{2}(p+1)\right]$.. Then the process

$$M_t^f := f(X_t) - f(X_0) - \int_0^t Lf(X_s)\,ds \tag{5.19}$$

is a true martingale.

Proof. $M = M^f$ is a local martingale, and thus $M_{t\wedge\tau_a}$ is a bounded local martingale, thus a true martingale. Observe that $Z_T := \sup_{t\leq T}\left|M_s^f\right| \in L^1(\mathbb{P})$ so if $0 \leq s \leq t \leq T$ and $A \in \mathscr{F}_s$ we can apply dominated convergence to the equality

$$\mathbb{E}\left[M_{t\wedge\tau_a}1_A\right] = \mathbb{E}\left[M_{s\wedge\tau_a}1_A\right]. \tag{5.20}$$

□

3 The law of large numbers

In a SIR model with an initial population N large, we are interested in the proportions of susceptibles, infected and recovered.
More generally, we consider Markov epidemic models $X^{(N)}$ with rates $\beta_j^{(N)}$ depending on a scale factor N that will goe to infinity. We are going to study the behaviour of

$$Z_t^N := \frac{X_t^{(N)}}{N}\,. \tag{5.21}$$

It's generator is

$$L^{Z^N} f(z) = L^{X^{(N)}} f(\frac{\cdot}{N})(Nz) = \sum_j \beta_j^{(N)}(z)(f(z+\frac{h_j}{N}) - f(z))\,. \tag{5.22}$$

We have the representation in terms of independent poisson processes p_J and thier compensated martingales $\tilde{P}_j(t) := P_j(t) - t$:

$$
\begin{aligned}
Z_t^N &= Z_0 + \sum_j \frac{h_j}{N} P_j\left(\int_0^t \beta_j^{(N)}(N Z_s^N)\,ds\right) \\
&= Z_0^N + \sum_j \frac{H_j}{N} \int_0^t \beta_j^{(N)}(N Z_s^N)\,ds + \sum_j \frac{h_j}{N} \tilde{P}_j\left(\int_0^t \beta_j^{(N)}(N Z_s^N)\,ds\right)
\end{aligned}
\tag{5.23}
$$ {eq:defznlaw-large-

If we neglect the martingale terms, and we impose that for functions $\beta_j : \mathbb{R}^d\ to \mathbb{R}_+$ smooth enough, we have

$$\beta_j^{(N)}(N z) := N \beta_j(z) \tag{5.24}$$ {eq:betanjnz-:=-n}

and we have $Z_0^N \to z_0 \in \mathbb{R}_+^d$, then Z^N will be close to the solution of the ODE

$$z(t) = z_0 + \sum_j h_j \int_0^t \beta_j(z(s))\,ds \tag{5.25}$$

i.e.

$$z'(t) = b(z(t)) \quad \text{with } b(z) := \sum_j h_j \beta_j(z), z(0) = z_0. \tag{5.26}$$ {eq:zt-=-bzt}

We also see that for smooth f, by a Taylor approximation,

$$L^{Z^N} f(x) \to \sum_j \beta_j(x) \nabla f(z).h_j = \nabla f.b(z). \tag{5.27}$$

Theorem 5.5. *Assume that the rate functions β_j are positive measurable and locally bounded. Assume that $b(z) = \sum_j h_j \beta_j(z)$ is locally Lipschitz. Assume that the sequence of positive rv's Z_0^N satisfy $\sup_N \mathbb{E}\left[< Z_0^N, 1 >^3\right] < +\infty$ and $Z_0^N \to z_0$ in distribution. Then the sequence of processes $(Z^N(t), 0 \leq t \leq T)$ defined by (5.23) converges in probability for the $L^\infty([0,T])$ norm to the continuous deterministic function z solution of (5.26).*

Proof. Assume first that the β_j are uniformly bounded and b globally Lipschitz

$$\sup_j \sup_z \beta_j(z) \leq M < +\infty \quad \text{and} \quad \sup_{y \neq z} \left|b(z) - b(y)\right| \leq M\left|y - z\right|. \tag{5.28}$$

Since we have

$$
\begin{aligned}
&Z_t^N = Z_0^N + \int_0^t b(Z_s^N)\,ds + \sum_k h_j \frac{1}{N} \tilde{P}_j\left(N \int_0^t \beta_j(Z_s^N)\,ds\right) \\
&z(t) = z_0 + \int_0^t b(z(s))\,ds
\end{aligned}
$$

we get for $t \in [0, T]$

$$\left|Z_t^N - z(t)\right| \le \left|Z_0^N - z_0\right| + M\int_0^t \left|Z_s^N - z(s)\right| ds + \frac{1}{N} C \sup_j \sup_{0\le s\le NMT} \left|\tilde{P}_j(s)\right|.$$

By Gronwall's Lemma, this implies for $t \in [0, T]$

$$\sup_{t\le T}\left|Z_t^N - z(t)\right| \le \left(\left|Z_0^N - z_0\right| + \frac{1}{N} C \sup_j \sup_{0\le s\le NMT} \left|\tilde{P}_j(s)\right|\right) e^{Mt}$$

We conclude that this quantity converges in probability to 0 thanks to the following Lemma

{lem:pttunif}

Lemma 5.6. *If $P(t)$ is a standard Poisson process and $\tilde{P}(t) = P(t) - t$ then for all $\alpha > \frac{1}{2}$,*

$$\frac{1}{n^\alpha} \sup_{t\in[0,n]} \left|\tilde{P}(t)\right| \to 0 \quad a.s. \tag{5.29}$$

Indeed first we have by assumption $\left|Z_0^N - z_0\right|$ Fix $\frac{1}{2} < \alpha < 1$. There exists $C_j(\omega) < +\infty$ such that a.s.

$$\forall n, \quad \sup_{t\le n}\left|\tilde{P}_j(t)\right| \le n^\alpha C_j \tag{5.30}$$

and thus

$$\frac{1}{N} \sup_j \sup_{0\le s\le NMT} \left|\tilde{P}_j(s)\right| \le \frac{(NMT)^\alpha}{N} \sup_j C_j(\omega) \to 0\,. \tag{5.31}$$

Let us consider now the general case. Looking closely at the proofs of Propositions (5.2) and (5.3), given $T > 0$ let $U_N := \sup_{t\le T}\left\langle Z_t^N, 1\right\rangle$. Since $\sup_N \mathbb{E}\left[\left\langle Z_0^N, 1\right\rangle^3\right] < +\infty$ we have

$$\sup_N \mathbb{E}\left[U_N^2\right] < +\infty\,. \tag{5.32}$$

The function b is locally Lipschitz and β_j is locally bounded. Therefore, for any $A > 0$ there exists $M_A < +\infty$ such that

$$\sup_j\left|\beta_j(x)\right| \le M_A\,, \quad \left|b(x) - b(y)\right| \le M_A \qquad (\text{if } \langle x, 1\rangle \le A, \left\langle y, 1\right\rangle \le A). \tag{5.33}$$

We choose $A > \sup_{t\le T} z(t)$. By the preceding arguments,on the event $\{U_A \le A\}$

$$\begin{aligned}\left|Z_t^N - z(t)\right| &\le C\left(\left|Z_0^N - z_0\right| + \frac{1}{N}\sup_{j,0\le n\le NTM_A}\left|\tilde{P}_j(s)\right|\right)e^{TM_A}\\ &\le C_{A,T}\left(\left|Z_0^N - z_0\right| + N^{\alpha-1}C(\omega)\right)\end{aligned}$$

with $C(\omega)$ a finite random variable. On the other hand we have

$$\mathbb{P}(U_N \ge A) \le \frac{1}{A^2}\mathbb{E}\left[U_N^2\right] \le \frac{C}{A^2}\,. \tag{5.34}$$

Combining twe two, we show easily that for any $\epsilon > 0$,

$$\lim_{N\to+\infty} \mathbb{P}\left(\sum_{t\le T} \left|Z_t^N - z(t)\right| \ge \epsilon\right) = 0. \tag{5.35}$$

□

Proof of Lemma 5.6. By Markov's inequality, for all $\gamma > 0, \epsilon > 0$

$$\mathbb{P}(P(t) - t > \epsilon) \le e^{-\gamma\epsilon}\mathbb{E}\left[e^{\gamma(P(t)-t)}\right] = exp\left(t(e^{\gamma} - 1 - \gamma) - \gamma\epsilon\right). \tag{5.36}$$

Taking the infimum, with respect to $\gamma > 0$, we get

$$\mathbb{P}(P(t) - t > \epsilon) \le \frac{e^{\epsilon}}{(1+\epsilon/t)^{t+\epsilon}} \tag{5.37}$$

Similarly,

$$\mathbb{P}(P(t) - t < -\epsilon) \le \frac{e^{-\epsilon}}{(1-\epsilon/t)^{t-\epsilon}} \tag{5.38}$$

Therefore, for $1/2 < \alpha < 1$ and $\epsilon = t^{\alpha}$ we get

$$\mathbb{P}\left(|P(t) - t| \ge t^{\alpha}\right) \le 2e^{-t^{2\alpha-1} + o(t^{3\alpha-2})} \tag{5.39}$$

By Borel Cantelli $\sup_n n^{-\alpha}|P(n) - n| < +\infty$ a.e. Since $t \to P(t)$ is increasing,

$$P(\lfloor t\rfloor) - \lfloor t\rfloor - 1 \le P(t) - t \le P(\lfloor t\rfloor + 1) - t \tag{5.40}$$

And thus

$$\sup_{t\ge1} \frac{|P(t)-t|}{t^{\alpha}} < +\infty \quad a.e. \tag{5.41}$$

Hence,for $\eta > 0$,

$$n^{-(\alpha+\eta)} \sup_{t\le n} |P(t) - t| \le n^{-(\alpha+\eta)} \sup_{t\le1} |P(t) - t| + n^{-\eta} \sup_{t\ge1} \frac{|P(t)-t|}{t^{\alpha}} \to 0 \quad a.e. \tag{5.42}$$

□

6 The duration of the basic stochastic epidemics

The stochastic SIR process is a pure jump Markov process X on $E=\mathbb{Z}^3$ (and a density dependent Markov process) with generator, for $x=(s,i,r)$ and bounded f

$$Lf(x)=\sum_j \beta_j(x)(f(x+h_j)-f(x))=\frac{\beta}{N}si(f(s-1,i+1,r)-f(s,i,r))+\gamma i(f(s,i-1,r+1)-f(s,i,r)), \tag{6.1}$$

We have seen that since when $x=(s,i,r)\in\mathbb{N}^3$ and $x+h_j\notin\mathbb{N}^3$, then $\beta_j(x)=0$. This ensures that starting from $X_0\in\mathbb{N}^3$ the process stays in $\mathbb{N}^3$: this we shall assume from now on. Letting $N_t=\|X_t\|_1=\left\langle X_,1\right\rangle=S_t+I_t+R_t$ we see that for every function g, if $f(x)=g(\langle x,1\rangle=$ we have $Lf(x)=0$ so $t\to P_tf(x)$ is constant and this implies that N_t stays constant a.e. (since it is cadlag).

1 The start of the epidemic

By proposition 4.14, let us assume that we start wuth $X_0=(N-1,1,0)$ that is $N-1$ susceptibles and one infected person. Then there exists a coupling with Z a linear branching process with rates (β,γ) starting from $Z_0=I_0=1$: a.e. forall t , $I_t\le Z_t$.
We define the *probability of a major outbreak* to be the supremum of the $\delta>0$ such that there exists $\epsilon>0$ and a N_0 such that for any $N\ge N_0$

$$\mathbb{P}\big(T_{[\epsilon N]}(I)<+\infty\mid X_0=(N-1,1,0)\big)\ge\delta\,, \tag{6.2}$$

with

$$T_a(I)=\inf\{t>0:I_t\ge a\}\,. \tag{6.3}$$

In words, the probability of a major outbreak is the probability that for large initial population N, starting with one infected individual, the population of infected reaches a macroscopic level, that is a positive fraction of the initial population.

Assume thus that $\beta < \gamma$. Then the linear BD process Z_t becomes a.s. extinct and $\sup_{t\geq 0} Z_t < +\infty$. Therefore, as $N \to +\infty$,

$$\mathbb{P}\big(T_{[\epsilon N]}(I) < +\infty \mid X_0 = (N-1,1,0)\big) \leq \mathbb{P}\left(\sup_{t\geq 0} Z_t \geq \epsilon N\right) \to 0. \tag{6.4}$$

An important ingredient of the preceding limit is that the random variable $\sup_{t\geq 0} Z_t$ does not depend on N. Therefore the probability of a major outbreak is 0.
Similarly, if $\beta > \gamma$, then Z becomes extinct with probability β/γ and therefore the probability of a major outbreak is less than $1-\beta/\gamma$.
On the other hand, thanks to Proposition 4.19, given $0 < \epsilon < 1$ small enough so that $\beta(1-\epsilon) > \gamma$, we can construct a coupling with a linear birth and death process $\bar{Z}$ with rates $(\beta(1-\epsilon), \gamma)$:

$$I_t \geq \bar{Z}_t \quad \text{on} \quad [0, T), \tag{6.5}$$

with $T = \inf\{t > 0 : B_t < N-1-N(1-\epsilon)\}$ and B_t the process of number of births of $\bar{Z}$.
Consider the martingale $\bar{W}_t = e^{-rt} Z_t$ with $r = \beta(1-\epsilon)-\gamma > 0$. On the set $\{\bar{W} > 0\}$, of probability $\frac{\gamma}{\beta(1-\epsilon)}$, the process $\bar{Z}_t$ grows exponentially fast, so T is of order $\log N$ and there is a $\eta > 0$ such that $Z_T \geq \eta N$ since we have $B_t \approx e^{rt}$ (Use again the comparison theorem to get a precise statement). Therefore, we have, for all N large enough,

$$\mathbb{P}\big(T_{[\eta N]}(I) < +\infty \mid X_0 = (N-1,1,0)\big) \geq \mathbb{P}\big(\bar{Z} \text{does not become extinct}\big) \geq \frac{\gamma}{\beta(1-\epsilon)}. \tag{6.6}$$

In conclusion, if $\beta > \gamma$, the probability of a major outbreak is $1-\gamma/\beta$, and the time for the infected population to reach a positive fraction of the initial population N is approximately $\log N$.

2 The deterministic SIR epidemic model

Assume now that we have a SIR process with initial population $X_0^{(N)} = (N(1-\epsilon), \epsilon N, 0)$. From the law of large numbers we now that $Z_t^N = \frac{1}{N} X_t^{(N)}$ is uniformly close to $z(t) = (s(t), i(t), r(t))$ the solution of the SIR ODE

$$s' = betasi \tag{6.7}$$

$$i' = \beta si - \gamma i \tag{6.8}$$

$$r' = \gamma i \tag{6.9}$$

with initial condition $z(0) = (1-\epsilon, \epsilon, 0)$. Let us do a brief study of this ODE of the type $z' = b(z)$ with b locally Lipschitz. Observe that $b(z) = 0$ for $z \in \partial K$ with K the positive orthant : the cone $K = \{z : s \geq 0, i \geq 0, r \geq 0\}$. Therefore since $z(0) \in K$, z stays in K for all times (this is a classical result on monotone dynamical systems : see e.g. Proposition 3.3 of [9])

Observe now that $n(t) = s(t) + r(t) + i(t)$ is constant : $n'(t) = 0$. thus, for all t, $n(t) = n(0) = 1$. We have $s'(t) = betasi \leq 0$, so it is a decreasing function, and $s(t) \in [0,1]$. Therefore it converges to $s(\infty)$. SImilarly, $r(t)$ is increasing and bound so converges to $r(\infty)$. Therefore $i(t) = 1 - r(t) - s(t)$ converges to $i(\infty)$. Since

$$r(t) = r(0) + \int_0^t r'(s)\,ds = 0 + \gamma \int_0^t i(s)\,ds \tag{6.10}$$

and $r(t)$ stays bounded and $i(t) \to i(\infty)$, we have $i(\infty) = 0$, and therefore $r(\infty) + s(\infty) = 1$.
Moreover, $\frac{s'}{s} = -\beta i$, thus

$$s(\infty) = s(0)\exp(-\beta \int_0^\infty i(s)\,ds) = \exp(-(\beta/\gamma)r(\infty)) \tag{6.11}$$

Combining all this yields that $z_\epsilon = 1 - s(\infty) = r(\infty)$ is a solution of

$$1 - z_\epsilon = (1-\epsilon)e^{-(\beta/\gamma)z_\epsilon} \tag{6.12}$$

and therefore $z = \lim_{\epsilon\to 0} z_\epsilon$ is the unique solution of

$$1 - z = e^{-R_0 z}, \tag{6.13}$$

with $R_0 = \beta/\gamma > 1$ the basic reproduction number. Let us rewrite down this equation, with $\sigma = 1 - s = \lim_{\epsilon\to 0} s(\infty)$

$$R_0 \sigma e^{-R_0\sigma} = R_0 e^{-R_0} \tag{6.14}$$

Therefore $\sigma = R_0\sigma \in (0,1)$ is the unique solution in $(0,1)$ of $xe^{-x} = R_0 e^{-R_0}$.

3 The end of the epidemic

Combining the law of large numbers, which we shall assume is an a.e. convergence, and the preceding results, let us choose ϵ small enough so that $R_0 s(\infty) < 1$ Let us choose $T > 0$ large enough so that $R_0 s(T) \leq 1 - 2\eta < 1$ and consider $X^{(N)}$ a SIR process starting from $X_0^{(N)} = (N(1-\epsilon), \epsilon N, 0)$.If N_0 is large enough, then for $N \geq N_0$, almost surely, $R_0 \frac{1}{N} S_T^{(N)} \leq 1 - \eta < 1$ and thus we shall use strong markov property at the finite time $T = \inf t > 0, R_0 S_T^{(N)} \leq N(1-\eta)$.
Using the comparison theorem, the extinction time of $X^{(N)}$ from this time on, is stochastically dominated by the extinction time of a linear birth and death process Z with rates $(\beta \frac{1-\eta}{R_0} = \gamma(1-\eta), \gamma)$ which is subcritical, and thus goes extinct in at most finite time that does not depends on N. Alas, the initial number Z_0 is less than N and we are left to prove that $T_0(Z) = \max(T_1, ..., T_n)$ the maximum of the hitting time of zero for N independent branching processes starting from 1, is of order $\log(N)$. This is an easy exercice because T_0 has an exponential tail.

$$\mathbb{P}(T_0 > t) = \mathbb{P}_1(Z_t \neq 0) \sim (1 - \lambda/\mu)e^{-(\mu-\lambda)t}. \tag{6.15}$$

x	0		σ		1		R_0
$f'(x)$		$+$		$+$	0	$-$	
$f(x)$	0	$\nearrow$	$f(R_0)$	$\nearrow$	$1/e$	$\searrow$	$f(R_0)$

7

Multi type Galtson Watson Processes

1 Motivation

Modelling the reproduction of bacteria in which a gene has two types of allele A and B. We assume $p_1, p_2, \alpha_1 \in (0,1)$ and consider two cases $\alpha_2 \in (0,1)$ and $\alpha_2 = 1$ (B alleles only yield B Alleles).
The questions we want to answer are the following :

- Do we have extinction, survival ? starting from all A's or all B's or a mixture ?
- When there is non extinction what is the growth rate of the total population ?
- Do we have relative asymptotic frequencies of A and B ?

2 The model

The population at generation n is a line vector $Z_n = (Z_{n1}, \dots, Z_{nd})$ of integer valued random variables, with d the number of different types. The type of an individual is an attribute that remains fixed throughout its lifetime. Individuals of the same type have the same offspring distribution. Different individuals reproduce independently.
The offspring of an individual of type i is distributed as $\xi_i = (\xi_{i1}, \dots, \xi_{id})$ and assumed to be integrable. The process Z_n satisfies the induction

$$Z_{n+1} = \sum_{j=1}^{d} \sum_{i=1}^{Z_{nj}} \xi_i^{(n+1),j}, \tag{7.1}$$ {eq:defgwmultitype}

with $(\xi_i^{(n+1),j}, n \geq 0, i \geq 1, 1 \leq j \leq d)$ independent and $\xi_i^{(n+1),j}$ distributed as ξ_j.

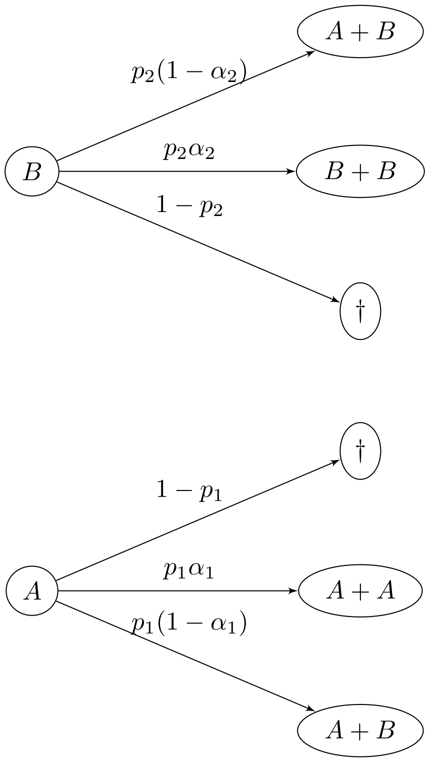

Figure 7.1: Reproduction of Bacteria

Example 7.1 (bacteria reproduction)**.** . $i=1$ *for type A,* $i=2$ *for type B. Then*

$$\begin{aligned} &\mathbb{P}(\xi_1=(0,0))=1-p_1\,, &\mathbb{P}(\xi_1=(2,0)) &=p_1\alpha_1\,, &\mathbb{P}(\xi_1=(1,1))=p_1(1-\alpha_1)\\ &\mathbb{P}(\xi_2=(0,0))=1-p_2\,, &\mathbb{P}(\xi_2=(0,2)) &=p_2\alpha_2\,, &\mathbb{P}(\xi_2=(1,1))=p_2(1-\alpha_2)\,. \end{aligned}$$

The mean matrix is $M=(m_{i,j})_{1\le i,j\le d}$ with

$$m_{i,j}=\mathbb{E}\big[\xi_{ij}\big]=\mathbb{E}\big[Z_{1j}\mid Z_0=e_i\big]. \tag{7.2}$$

The sigma-field is $\mathscr{F}_n=\sigma\Big(\xi_i^{(k),j},k\le n,j,i\Big)$ and Z_n is a $\mathscr{F}_n$ Markov chain.

Lemma 7.1. *If we take the conditional expectation of a vector to be the vector of its conditinal expectations we have*

$$\mathbb{E}[Z_{n+1}\mid\mathscr{F}_n]=Z_nM \tag{7.3}$$

Proof.

$$\mathbb{E}\Big[Z_{(n+1),k}\mid\mathscr{F}_n\Big]=\sum_{j=1}^{d}\sum_{i=1}^{Z_{nj}}\mathbb{E}\Big[\xi_{i,k}^{(n+1),j}\mid\mathscr{F}_n\Big]=\sum_{j=1}^{d}\sum_{i=1}^{Z_{nj}}m_{jk}=\sum_{j=1}^{d}Z_{n,j}m_{jk}=(Z_nM)_k\,. \tag{7.4}$$

$\square$

Corollary 7.2. *Let γ be a right eigenvector of M with eigenvalue $\lambda \in \mathbb{C}$. Then $Y_n = \lambda^{-n} Z_n \gamma$ is a martingale (possibly complex valued).*

Therefore it should not be surprising that the growth rate of Z_n is given, when there is no extinction, by the largest modulus eigenvalue, the *spectral radius*

$$\rho(M) := \sup\left\{|\lambda| : \lambda \in sp(M)\right\} \tag{7.5}$$

Gelfand's formula yields that for any matrix norm

$$\rho(M) = \lim_{n\to+\infty} \left\|M^n\right\|^{1/n}, \tag{7.6}$$

and we know that $\rho(M) < 1$ iff $M^n \to 0$.
In the baceria example,

$$m_{11} = \mathbb{E}[\xi_{11}] = 2p_1\alpha_1 + P_1(1-\alpha_1) = p_1(1+\alpha_1), m_{12} = p_1(1-\alpha_1), ... \tag{7.7}$$

If $\alpha_2 \in (0,1)$ we have $M >> 0$ that is $m_{ij} > 0$ for all entries i, j. On the other hand if $\alpha_2 = 1$ then M is triangular with spectrum $sp(M) = \left\{p_1(1+\alpha_2), 2p_2\right\}$.

Definition 7.1. *A matrix M with non negative entries, $M \geq 0$ is said to be* indecomposable *or* irreducible *if $\forall i, j \exists r (M^r)_{ij} > 0$. We say that the multitype branching process is indecomposable.*

This means that every type of individual may have eventually a progeny of any other type. We shall prove that then, when there is no extinction, for every starting mixture of types the growth rate is the same and given by ρ.
We see on the bacteria example that this is not the case for decomposable case $\alpha_2 = 1$ since we have two different rates.

3 Extinction probabilities

Proposition 7.3. *Let $f_i(s) = \mathbb{E}\left[s_1^{\xi_{i1}} \cdots s_d^{\xi_{id}}\right]$ and $f(s) = (f^i(s), 1 \leq i \leq d)$. Then the extinction probabilities*

$$q_i = \mathbb{P}(Z_n \to 0 \mid Z_0 = e_i) \tag{7.8}$$

satisfy $f(q) = q$.

Proof. As in dimension 1, condition on the first generation. There is extinction iff there is extinction for all the GW processes of the descendants of the ancestor, ξ_{ik} of type k, that have a probability q_k of extinction therefore. □

Assume that there exists a vector $u \in \mathbb{R}^d$ with $u_i > 0$ for all i, such that $Mu = \rho u$ with $\rho = \rho(M)$.

Proposition 7.4. *The process $W_n = \rho^{-n} Z_n u$ is a positive martingale such that $s_i = \mathbb{P}(W = 0 \mid Z_0 = e_i)$ satisfies $f(s) = s$. Therefore if $\rho < 1$, then there is almost sure extinction. Assume $\rho > 1$, if forall i, k, $\mathbb{E}\left[\xi_{ik} \log^+ \xi_{ik}\right] < +\infty$ then W_n is an UI martingale. If, moreover, 0 is the only absorbing point of the chain Z, then $s = q$: on non extinction the process grows exponentially.*

Proof. Condition on the first generation : if $w^{l,j}$ is the variable corresponding to the multitype GW process $Z_n^{l,j}$ of the l-th child of type j of $\emptyset$, then

$$Z_n = \sum_{j=1}^{d} \sum_{l=1}^{\xi_{ij}} Z_{n-1}^{l,j} \tag{7.9}$$

So taking limits in $\frac{1}{\rho^n} Z_n u$ yields

$$W = \frac{1}{\rho} \sum_{j=1}^{d} \sum_{l=1}^{\xi_{ij}} W^{l,j} \tag{7.10}$$

And thus, by independence,

$$\mathbb{P}(W = 0 \mid Z_0 = e_i) = \mathbb{E}\left[\mathbb{P}\left(\sum_{j=1}^{d} \sum_{l=1}^{\xi_{ij}} W^{l,j} = 0 \mid \xi_{ij}, 1 \le i \le d\right)\right] = \mathbb{E}\left[\prod s_j^{\xi_{ij}}\right] = f^i(s). \tag{7.11}$$

SInce $u >> 0$, we have $\{W > 0\} \subset \{\forall n, Z_n > 0\}$. By monotonicity

$$q_i = \mathbb{P}(\exists n : Z_n = 0 \mid Z_0 = e_i) \tag{7.12}$$

is the smallest root of $f(q) = q$ since $q_i = \lim \uparrow q_{i,n} = \mathbb{P}(Z_n = 0 \mid Z_0 = e_i)$ and $q_{i,0} = 0$. So if s is another solution $s_i \ge 0 = q_{i,0}$ implies $s_i = f^i(s) \ge f^i(0) = q_{i,1}$ and by induction $s_i \ge q_{i,n} \to q_i$.
Since W_n i a UI martingale, we have $\mathbb{E}[W \mid Z_0 = e_i] = \mathbb{E}[W_0 \mid Z_0 = e_i] = u_i > 0$ and thus for all i, $s_i < 1$.
Observe that $M_n = s^{Z_n} = \prod s_i^{Z_{n,i}}$ is a martingale since

$$\mathbb{E}[M_{n+1} \mid \mathscr{F}_n] = \mathbb{E}_{Z_n}\left[s^{Z_1}\right] = \prod f_i(s)^{Z_{n,i}} = M_n \tag{7.13}$$

Since $M_n \in [0,1]$ it is a UI martingale, and thus converges to M_∞. The range of M_n is discrete : $E = \left\{\prod_{1 \le i \le d} s_i^{n_i}, n_i \in \mathbb{N}\right\}$ and therefore $M_\infty \in \{0\} \cup E$. When $M_\infty \in E$, then M_n is a stationary sequence that is ther exist $n_0(\omega)$ such that for all i,fora all $n \ge n_0$, $Z_{n,i} = Z_{\infty,i}$ Since the only absorbing point of Z is 0, this implies that $Z_\infty = 0$ is that the process goes extinct, and therefore $M_\infty = 1$. Therefore

$$\begin{aligned} s_i &= \mathbb{E}[M_0 \mid Z_0 = e_i] = \mathbb{E}[M_\infty \mid Z_0 = e_i] = \mathbb{E}\left[M_\infty \mathbf{1}_{(M_\infty \in E)} \mid Z_0 = e_i\right] && (7.14) \\ &\le \mathbb{P}(Z_n \to 0 \mid Z_0 = e_i) = q_i\,. && (7.15) \end{aligned}$$

Combining everything we do get $s = q$. $\square$

4 Perron Frobenius Theorem

This theorem gives sufficient conditions for matrices with positive entries to have their spectral radius as the only eignevalue for a vecteur with strictly positive entries.
For $x \in \mathbb{R}^d$ we say that $x \geq 0$ if forall i, $x_i \geq 0$. We say $x >$ if $x \geq 0$ and $x \neq 0$, and $x >> 0$ if forall i, $x_i > 0$. We have the same notations for a real matrix $A \in \mathcal{M}_{d\times d}$. Observe that if $x > 0$ and $A >> 0$ then $Ax >> 0$.

Theorem 7.5. *Assume $A >> 0$ and let $\rho = \rho(A) > 0$ be its spectral radius. Then*

1. *There exists $u >> 0$ such that $Au = \rho u$.*
2. *If $\lambda \in sp(A)$ and $\lambda \neq \rho$ then, $|\lambda| < \rho$.*
3. *The dimension of the eigenspace associated to ρ is 1.*

Proof. Let $x \neq 0$ be an eigenvector of A with eignvalue λ. Then $y = |x|$ defined by $y_i = |x_i|$ satisfies

$$|\lambda| y_i = \left| \sum_j a_{ij} x_j \right| \leq \sum_j a_{ij} y_j \tag{7.16}$$ {eq:pfeqmod}

Assume that $|\lambda| = \rho$. And let $z = Ay - \rho y$. Assume $z > 0$ and so $Az >> 0$. Let $a = \sup\{c > 0 : Az \geq c y\}$; We have $a > 0$ so $Ay \geq (\rho + a) y$ and $A^n y \geq (a + \rho)^n y$ and so $\|A^n\| \geq (a + \rho)^n$ and by Gelfand's formula, $\rho(A) \geq a + \rho$, which is absurd. Therefore $z = 0$, that is $Ay = \rho y$.
Recall that if $w_1, \ldots, w_n$ are complex numbers such that $|w_1 + \cdots + w_n| = \sum |w_i|$ then they have the same argument : there exists $\theta \in \mathbb{R}$ and $\lambda_i \geq 0$ such that $w_k = \lambda_k e^{i\theta}$.
Applying this to (7.16) yields that the x_j have tha same argument, and therefore $x_j = e^{i\theta} y_j$ so $\lambda = \rho$. Since $y > 0$ and $A >> 0$ we have $y = \frac{1}{\lambda} Ay >> 0$ and this proves the theorem. □

{theo:perronfrobeni

Theorem 7.6. *If $A \geq 0$ and there exists $m \in \mathbb{N}^*$ such that $A^m >> 0$ then the conclusions of the preceding theorem hold.*

Proof. We have $\rho(A^m) = \rho(A)^m$. Let $n \neq 0$ and λ such that $Ax = \lambda x$. Then $A^m x = \lambda^m x$ If $|\lambda| = \rho$ then $|\lambda^m| = \rho(A^m)$ so by the preceding theorem $\lambda^m = \rho$.
Observe that since $A^m >> 0$, $0 \notin sp(A)$ so every line and every column of A is > 0. Therefore $A^{m+1} = A^m A >> 0$ and we have also $\lambda^{m+1} = \rho$. This yields $\lambda = \rho$ so $A^m x = \rho^m x$ and by the preceding theorem $x = \eta y$ with $y >> 0$, $\eta \in \mathbb{C}$. Therefore $Ay = y$ and we are done. □

Assume $A \geq 0$ and $A^m >> 0$ for some integer m. We saw that A has a right eigenvector $u >> 0$ such that $Au = \rho u$. We normalize u so that $\sum_i u_i = 1$.
Then A has a left eigenvector with eigenvalue ρ (consider the transpose): $vA = \rho v$. And we can normalize v so that $vu = u.v^T = \sum_i v_i u_= 1$.

Prove as an exercice that the operator $Px = (x.v^T)u = v x u$ is a projector $P^2 = P$ that commutes with A : $AP = PA = \rho P$. (It is a projector on the eigenspace with eigenvalue ρ).

Lemma 7.7. *Let $B = A - \rho P$. Then $\rho(B) < \rho$ and $\frac{A^n}{\rho^n} \to P$.*

Proof. Assume $Bx = \lambda x$ with $\lambda \neq 0$, $x \neq 0$. Then

$$\lambda P x = PBx = PAx - \rho P^2 x = 0, \tag{7.17}$$

o $Px = 0$ and $Ax = \lambda x$. If $|\lambda| = \rho$ then $\lambda = \rho$ and $Px = x$ so $x = 0$, absurd. Therefore $|\lambda| < \rho$ and we have proved that $\rho(B) < \rho$.By induction

$$A^n = B^n + \rho^n P, \tag{7.18}$$

and therefore $\rho^{-n} A^n = P + \rho^{-n} B^n$. But if $\delta > 0$, then there exists by Gelfand formula a n_0 such that for $n \geq n_0$

$$\left\| \rho^{-n} B^n \right\| \leq \rho - n(\rho B + \delta)^n \to 0 \tag{7.19}$$

if we have chosen δ small enough. And thus $\rho^{-n} B^n \to 0$.

□

Prove as an exercice that if $x > 0$ and $Ax \geq \lambda x$ for some $\lambda \geq 0$ then $\lambda = \rho$ and so x is a multiple of u.

Fortunately, we know exactly when to apply Theorem 7.6. A matrix $A \geq 0$ is *irreducible* if

$$\forall x, y \quad \exists m \quad (A^m)_{xy} > 0. \tag{7.20}$$

The *period* of an element x is $d(x) = gcd\{n \geq 1 : (A^n)_{xx} > 0\}$. If A is irreducible then all states have the same period, forall x $d(x) = d(A)$. We say then that A is *aperiodic* if $d(A) = 1$.

Proposition 7.8. *Let A be a matrix such that $A \geq 0$. Then there exists $m \in \mathbb{N}^*$ such that $A^m >> 0$ iff A is irreducible and aperiodic.*

5 The supercritical case and geometric growth

We assume that the mean matrix $M = (m_{ij})$ is irreducible and aperiodic. Thanks to Perron Frobenius theory, if ρ is the spectral radius of M then there exists $u >> 0$, $v >> 0$ with $1 = \sum_i u_i v_i$ such that $Mu = \rho u$ and $vM = \rho M$.
We know then that $W_n = \rho^{-n} Z_n u$ is a positive martingale converging to a finite rv W.

{theo:supercritical

Theorem 7.9. *Assume $\rho > 1$ and $\sup_{i,k} \mathbb{E}\left[\xi_{ik} \log^+ \xi_{ik}\right] < +\infty$. Then*

$$\rho^{-n} Z_n \to W v \qquad a.e. \tag{7.21}$$

Corollary 7.10. *Under the preceding assumptions, on the non extinction set the assymptotic proportions of each type converge a.e. to a deterministic number. If* $|Z_n| = \sum_j Z_{nj}$ *then, a.e. on* $\{W > 0\}$

$$\frac{Z_{ni}}{|Z_n|} \to \frac{v_i}{|v|}. \tag{7.22}$$

Proof of Theorem 7.9. To simplify the proofs we shall assume that $\sup_{i,k} \mathbb{E}\left[\xi_{ik}^2\right] < +\infty$ and follow the arguments of Kesten and Stigum [10].

Lemma 7.11. *There exists a constant* $C > 0$ *such that for all* $a \in \mathbb{R}^d$, $z \in \mathbb{R}_+^d$

$$\mathbb{E}_z\left[((Z_1 - zM)a)^2\right] \le C\|a\|_2^2 |z|. \tag{7.23}$$

Proof. Remember that $\mathbb{E}[Z_{n+1} \mid \mathscr{F}_n] = Z_n M$, hence if $X = Z_1 a = \sum_i Z_{1i} a_i$, we have

$$\mathbb{E}_z[X] = \mathbb{E}_x[\mathbb{E}[Z_1 a \mid \mathscr{F}_0]] = zMa. \tag{7.24}$$

Therefore,

$$\begin{aligned}
\mathbb{E}_z\left[((Z_1 - zM)a)^2\right] = \operatorname{Var}_z(X) &= \operatorname{Var}\left(\left(\sum_j \sum_{i=1}^{z_j} \xi_i^{(1),j}\right) a\right) \\
&= \sum_j \sum_{i=1}^{z_j} \operatorname{Var}\left(\xi_j a\right) && \text{by independence,} \\
&= \sum_j z_j \operatorname{Var}\left(\xi_j a\right) \le C\|a\|_2^2 |z|.
\end{aligned}$$

□

Lemma 7.12. *The series* $\sum_n (Z_{n+1} - Z_n M)\rho^{-n}$ *converges as in* $\mathbb{R}^d$.

Proof. We shall prove that for every $a \in \mathbb{R}^d$ the series $\sum_n U_n$ converges a.e. with $U_n := \rho^{-n}(Z_{n+1} - Z_n M)a$. First observe that by induction $Z_n \in L^2$ so $U_n \in L^2$. Then $\mathbb{E}[U_n \mid \mathscr{F}_n] = 0$ hence $M_n = U_1 + \cdots + U_n$ is an L^2 martingale and it converges a.e. as soon as $\sum_n \mathbb{E}\left[U_n^2\right] < +\infty$. Indeed,

$$\mathbb{E}[U_n \mid \mathscr{F}_n] = \rho^{-n}(\mathbb{E}[Z_{n+1} a \mid \mathscr{F}_n] - Z_n M a) = 0. \tag{7.25}$$

Moreover, by Markov property,

$$\mathbb{E}\left[U_n^2 \mid \mathscr{F}_n\right] = \rho^{-2n} \mathbb{E}_{Z_n}\left[((Z_1 - Z_0 M)a)^2\right] \le C\rho^{-2n}\|a\|_2^2 |Z_n|. \tag{7.26}$$

Remember that since $u >> 0$, $W_n = \rho^{-n} Z_n = \rho^{-n} \sum_i u_i Z_{ni} \ge C\rho^{-n}|Z_n|$, and thus

$$\mathbb{E}[|Z_n|] \le C\rho^n \mathbb{E}[W_n] \le C\rho^n \mathbb{E}[W_0] \tag{7.27}$$

Hence,

$$\mathbb{E}\left[U_n^2\right] = \mathbb{E}\left[\mathbb{E}\left[U_n^2 \mid \mathscr{F}_n\right]\right] \le C\rho^{-2n}\mathbb{E}[|Z_n|] \le C'\rho^{-n} \tag{7.28}$$

and thus $\sum_n \mathbb{E}\left[U_n^2\right] < +\infty$. □

{lem:supercr-det}

Lemma 7.13. *Let $u, v, x_n \in \mathbb{R}^d$ be such that $u.v = 1$ and $\lim_{n,p\to+\infty} x_{n+p} - (x_n.u)v = 0$. Then there exists $\lambda \in \mathbb{R}$ such that $x_n \to \lambda v$.*

The proof is left as an exercise.

{lem:supercr-case-g

Lemma 7.14.

$$a.e. \quad \lim_{r_0\to+\infty, r_1-r_0\to+\infty} Z_{r_1+1}\rho^{-(r_1+1)} - \left(\rho^{-r_0} Z_{r_0} u\right) v = 0 \tag{7.29}$$

Proof. From Perron Frobenius theory we know that if $zP = z\,u\,v$ is the projection then $B = M - \rho P$ has spetral radius $spr(B) < \rho$ and thus $\|B^n\| \le \rho_1^n$ for some $0 < \rho_1 < \rho$. SInce $\rho^{-n}M^n = P + \rho^{-n}B$, we have

$$\begin{aligned}
I(r_0, r_1) &:= \sum_{r=r_0}^{r_1} \left(Z_{r+1}M^{r_1-r} - Z_r M^{r_1-r+1}\right)\rho^{-1-r_1} \\
&= \sum_{r=r_0}^{r_1} (Z_{r+1} - Z_r M)\rho^{-(r+1)}\left(\rho^{r-r_1} M^{r_1-r}\right) \\
&= \sum_{r=r_0}^{r_1} (Z_{r+1} - Z_r M)\rho^{-(r+1)} P + \sum_{r=r_0}^{r_1} (Z_{r+1} - Z_r M)\rho^{-(r+1)}\rho^{r-r_1} B^{r_1-r}
\end{aligned}$$

Therefore, with $U_r = (Z_{r+1} - Z_r M)\rho^{-(r+1)} P$,

$$\|I(r_0, r_1)\| \le \left\| \sum_{r=r_0}^{r_1} U_r \right\| + \sup_{r\ge r_0} \|U_r\| \sum_{r=r_0}^{r_1} \rho^{r-r_1} \left\| B^{r_1-r} \right\|, \tag{7.30}$$

and $\lim_{r_0\to+\infty, r_1-r_0\to+\infty} I(r_0, r_1) = 0$ since the first term is a remainder for a convergent series $\sum_r U_r$, and the second term is bounded by

$$C \sup_{r\ge r_0} \|U\|_r. \tag{7.31}$$

Now observe that $I(r_0, r_1)$ is a telescopic sum:

$$I(r_0, r_1) = \rho^{-(r_1+1)} Z_{r_1+1} - \rho^{-r_0} Z_{r_0} \rho^{-(r_1-r_0+1)} M^{r_1-r_0+1} \tag{7.32}$$

Since $\rho^{-n}M^n \to P$ this yields

$$0 = \lim_{r_0\to+\infty, r_1-r_0\to+\infty} I(r_0, r_1) = \lim_{r_0\to+\infty, r_1-r_0\to+\infty} \rho^{-(r_1+1)} Z_{r_1+1} - \rho^{-r_0} Z_{r_0} P, \tag{7.33}$$

and this ends our proof of the Lemma. □

We now resume the proof of the Theorem. Combining Lemmas 7.13 and 7.14 we obtain the existence of a random variable T such that a.e. $\rho^{-n} Z_n \to T v$. Since $W_n = \rho^{-n} Z_n u \to W$, we have a.e. $W = T v u = T$. □

8 Exercises

1 Exercises on Galton Watson processes

Exercice 1.1

Prove that for a subcritical GW process($m < 1$) the mean total progeny is

$$\mathbb{E}\left[\bar{X}\right] = \frac{1}{1-m}$$

Exercice 1.2

Assume that $\sigma^2 := \mathrm{Var}(\xi) < +\infty$. Show that

$$\mathrm{Var}(X_{n+1}) = m^n \sigma^2 + m^2 \mathrm{Var}(X_n), \qquad (8.1) \quad \texttt{\{eq:18\}}$$

and then that

$$\mathrm{Var}(X_n) = \begin{cases} \frac{\sigma^2 m^n (m^n - 1)}{m^2 - m} & \text{if } m \neq 1. \\ n\sigma^2 & \text{if } m = 1. \end{cases} \qquad (8.2) \quad \texttt{\{eq:19\}}$$

Show that if $m > 1$ then the martingale $W_n = \frac{X_n}{m^n}$ is UI.

Exercice 1.3

The Galton Watson process with immigration is defined by the recurrence

$$X_{n+1} = \sum_{i=1}^{X_n} \xi_i^{(n+1)} + Y_{n+1}$$

where the $(\xi_i^{(k)}, k \geq 1, i \geq 1)$ are IID distributed as ξ and are independent from $(Y_k, k \geq 1)$ IID distributed as Y. In this model $\xi_i^{(n+1)}$ is the number of children of the i-th individual of the n-th generation, and Y_n is the number of immigrants in the n-th generation. We assume that $0 < m = \mathbb{E}[\xi] < +\infty$, $\forall j \mathbb{P}\big(\xi = j\big) < 1$ and $0 < \lambda = \mathbb{E}[Y] < +\infty$.

1. Prove that one has
$$X_n = Z_n + U_n^{(1)} + \cdots + U_n^{(n)}, \tag{8.3}$$ {eq:21}

   with Z_n the number of descendants at generation n of the initial individual, $U_n^{(i)}$ is the number of descendants at generation n of immigrants that arrived at generation i, and all these processes are independent.

2. Let $V_n = m^{-n} X_n$. Show that
$$\mathbb{E}[X_{n+1} \mid \mathscr{F}_n] = m X_n + \lambda, \tag{8.4}$$ {eq:22}

   after defining precisely $\mathscr{F}_n$.

3. Show that V_n is a positive submartingale.

4. Assume from now on that $m > 1$. Show that
$$\mathbb{E}[X_n] = \frac{m^n(m+\lambda-1)-\lambda}{m-1}, \tag{8.5}$$ {eq:23}

   and infer that $C := \sup \mathbb{E}[V_n] < +\infty$.

5. Show that there exists a rv V, $0 \le V < +\infty$ a.e. and $V_n \to V$ a.e.

6. Recall that $W_n = m^{-n} Z_n \to W$ a.e. for a positive finite rv W. Let $\beta = \frac{m+\lambda-1}{m-1}$ and let
$$T = \sum_{k=1}^{Y} W_k, \tag{8.6}$$ {eq:24}

   with $(W_k, k \ge 1)$ IID distributed as W. Show that
$$m^{-n} U_n^{(i)} \to m^{-i} T^{(i)} a.e. \quad \text{with} \quad T^{(i)} \overset{d}{=} T. \tag{8.7}$$ {eq:25}

   Deduce that
$$V \ge U := W + \sum_{i=1}^{+\infty} m^{-i} T^{(i)} \quad a.e. \tag{8.8}$$ {eq:26}

7. Using independence in (8.3), compute for $\lambda > 0$, $\mathbb{E}\big[e^{-\lambda V_n}\big]$ Combining the inequality for a positive random variable
$$-\log \mathbb{E}\big[e^{-X}\big] \le \mathbb{E}[X]$$

   with the fact that
$$\mathbb{E}\big[m^{-n} U_n^{(i)}\big] = m^{-n} \mathbb{E}[\mathbb{E}[Z_{n-i} \mid Z_0 = Y]] = m^{-n} m^{n-i} \mathbb{E}[Y] = \lambda m^{-i}, \tag{8.9}$$

and show that

$$\mathbb{E}\big[e^{-\lambda V}\big] = \lim_{n\to+\infty} \mathbb{E}\big[e^{-\lambda V_n}\big] = \mathbb{E}\big[e^{-\lambda U}\big], \tag{8.10}$$

and deduce from it that $V = U$ a.e.

8. Show that if $\mathbb{E}\big[\xi \log^+ \xi\big] < +\infty$ the $V > 0$ a.e. and that if $\mathbb{E}\big[\xi \log^+ \xi\big] = +\infty$ then $V = 0$ a.e.

2 Exercises on birth and death processes

Exercice 2.1

Let ϕ be a non identically null non negative solution of $Q\phi = \phi$, with $\phi(0) = 0$. Let $\Delta_n = \phi_n - \phi_{n-1}$, $f_n = \frac{1}{\lambda_n}$, $g_n = \frac{\mu_n}{\lambda_n}$.

1. Show that $\Delta_1 = \phi_1$, $\Delta_{n+1} = \Delta_n g_n + f_n \phi_n$.
2. Show that ϕ_n is increasing
3. Let $r_n = f_n + \sum_{k=1}^{n-1} f_k g_{k+1} \dots g_n + g_1 \dots g_n$. Show that

$$r_n \phi_1 \le \Delta_{n+1} \le r_n \phi_n$$

and deduce that

$$\phi_1(1 + r_1 + \dots + r_n) \le \phi_{n+1} \le \phi_1 \prod_{k=1}^{n} (1 + r_k)$$

4. Show that $\sum_k r_k$ converges iff ϕ is bounded (and relate this to the non explosion criterion).

Exercice 2.2

Consider a linear birth and death process with $\lambda = \mu$. Let $q(t) := \mathbb{P}_1(X_t = 0)$ be the extinction probability at time t, when starting with one individual. Explain why $\mathbb{P}_x(X_t = 0) = q(t)^x$. Condition by the first jump time and show that

$$q(t) = \int_0^t e^{-2\lambda s} (\lambda q(t-s)^2 + \lambda). \tag{8.11}$$

Deduce that $q(t)$ satisfies the ode (Ricatti)

$$\frac{d}{dt} q = \lambda (q-1)^2, \tag{8.12}$$

and establish the formula $q(t) = \frac{\lambda t}{1 + \lambda t}$.

Exercice 2.3

Consider a linear birth and death process. Apply Kolmogorov forward equation to $f(x)=x^2$ to show that $u(t)=P_t f(x)$ satisfies the ode

$$u' = (\lambda+\mu)(2u+1) \tag{8.13}$$

Deduce from it that with $W_t = e^{-(\lambda-\mu)t}X_t$ we have if $\lambda > \mu$, $\sup_t \mathbb{E}\left[W_t^2\right] < +\infty$, and thus the martingale W is UI.

3 Exercises on stochastic comparison of Markov Processes

Exercice 3.1

Show that if X is a branching process on $\mathbb{N}^2$, then it is monotone.

Exercice 3.2

Let X be a pure jump process on $\mathbb{N}^2$ with semigroup $(P_t)_{t\geq 0}$ and generator L. For every $\theta = (\theta_1, \theta_2) \in (0, +\infty)^2$ we define the function $f_\theta(x) = e^{-\theta . x} = e^{-(\theta_1 x_1 + \theta_2 x_2)}$. Then the following assertions are equivalent

1. X has the branching property
2. for every θ and every x, $P_t f_\theta(x+y) = P_t f_\theta(x) P_t f_\theta(y)$.
3. for every θ and every x,
$$L f_\theta(x+y) = f_\theta(x) L f_\theta(y) + f_\theta(y) L f_\theta(x) \tag{8.19}$$
4. for every θ, there exists a constant vector C_θ such that
$$L f_\theta(x) = C_\theta . x f_\theta(x) \tag{8.20}$$

Exercice 3.3

(see [11]) Let $\{N(t) = (N_1(t), N_2(t))\}$ be a $\mathbb{N}\times\mathbb{N}$-valued pure jump Markov process with the following transition rates.

$$\begin{array}{rl}
m b_1 & \text{from } (m,n) \text{ to } (m+1,n) \\
n b_2 & \text{from } (m,n) \text{ to } (m,n+1) \\
m d_1(m,n) & \text{from } (m,n) \text{ to } (m-1,n) \\
n d_2(m,n) & \text{from } (m,n) \text{ to } (m,n-1).
\end{array}$$

Here b_1, b_2 are positive constants and d_1, d_2 are functions from $\mathbb{N}\times\mathbb{N}$ to $\mathbb{R}_+$. Suppose that there is a set $\mathscr{S} \subset \mathbb{R}_+ \times \mathbb{R}_+$ and constants $d_1^+, d_1^-, d_2^+, d_2^- \in [0,\infty]$ such that

$$d_1^- \le \inf d_1(\mathscr{S}) \le \sup d_1(\mathscr{S}) \le d_1^+ \text{ and}$$
$$d_2^- \le \inf d_2(\mathscr{S}) \le \sup d_2(\mathscr{S}) \le d_2^+$$

Assume that $(N_1(0), N_2(0)) \in \mathscr{S}$ and let $T_{\mathscr{S}}$ be the random time defined by

$$T_{\mathscr{S}} = \inf\{t \ge 0 : N(t) \notin \mathscr{S}\}.$$

Let $z_1^+, z_1^-, z_2^+, z_1^-$ be positive integers satisfying $z_1^- \le N_1(0) \le z_1^+$ and $z_2^- \le N_2(0) \le z_2^+$. Then on the same probability space as N, we can construct four $\mathbb{N}$-valued processes B_1^+, B_1^-, B_2^+ and B_2^- with laws $\mathbf{P}(b_1, d_1^-, z_1^+)$, $\mathbf{P}(b_1, d_1^+, z_1^-)$, $\mathbf{P}(b_2, d_2^-, z_2^+)$, $\mathbf{P}(b_2, d_2^+, z_1^-)$ such for all $t \le T_{\mathscr{S}}$ the following relations are satisfied almost surely,

$$B_1^-(t) \le N_1(t) \le B_1^+(t) \quad \text{and} \quad B_2^-(t) \le N_2(t) \le B_2^+(t).$$

(Hint : prove that when the function sd_1, d_2 are constant we have a branching process, which is therefore monotone, and use the comparison theorems)

4 Exercises on Multitype branching processes

Exercice 4.1

Let $u, v, x_n \in \mathbb{R}^d$ be such that $u.v = 1$ and $\lim_{n,p\to+\infty} x_{n+p} - (x_n.u)v = 0$. Then there exists $\lambda \in \mathbb{R}$ such that $x_n \to \lambda v$.

Exercice 4.2

Consider the bacteria example of the lecture notes. Determine the growth rate of the population when there is no extinction, and show that there is an asymptotic proportion of type A cells. Take $\alpha_1 = 0.9998, p_1 = 0.8$, $\alpha_2 = 0.999, p_2 = 0.9$.

5 Hints and Solutions

Solution de l'Exercice 1.2

By the conditional variance formula

$$\begin{aligned}
\mathrm{Var}(X_{n+1}) &= \mathrm{Var}(\mathbb{E}[X_{n+1} \mid \mathscr{F}_n]) + \mathbb{E}\left[(X_{n+1} - \mathbb{E}[X_{n+1} \mid \mathscr{F}_n])^2\right] \\
&= m^2 \mathrm{Var}(X_n) + \mathbb{E}\left[\mathbb{E}\left[(X_{n+1} - mX_n)^2 \mid X_n\right]\right] \\
&= m^2 \mathrm{Var}(X_n) + \mathbb{E}\left[X_n \sigma^2\right].
\end{aligned}$$

Solution de l'Exercice 2.3

if $f(x) = x^2$ then

$$Lf(x) = \lambda x((x+1)^2 - x^2) + \mu x((x-1)^2 - x^2) = 2(\lambda - \mu)x^2 + (\lambda + \mu)x \tag{8.14}$$

so if $u(t) = P_t f(x)$, taking into accoutn $\mathbb{E}_x[X_t] = x e^{(\lambda-\mu)t}$

$$u'(t) = P_t Lf(x) = 2(\lambda - \mu)u(t) + (\lambda + \mu)x e^{(\lambda-\mu)t} \tag{8.15}$$

and since $u(0) = x^2$, if $\lambda \neq \mu$

$$u(t) = e^{2(\lambda_\mu)t}\left(x^2 + (\lambda+\mu)x\frac{1 - e^{-(\lambda-\mu)t}}{\lambda - \mu}\right) \tag{8.16}$$

Therefore, if $\lambda > \mu$, $\sup_t \mathbb{E}\left[W_t^2\right] < +\infty$.

Solution de l'Exercice 3.1

It is the same proof as in dimension 1 with the partial order $x \leq y$ if $x_1 \leq y_1$ and $x_2 \leq y_2$. Then if $x \leq y$, there exists $z \in \mathbb{N} \times \mathbb{N}$ such that $y = x + z$ and we have, for a monotone f:

$$\begin{aligned}
P_t f(y) = \mathbb{E}_y\left[f(X(t))\right] &= \mathbb{E}_{x,z}\left[f(X^x(t) + X^z(t))\right] \qquad (8.17)\\
&\geq \mathbb{E}_{x,z}\left[f(X^x(t))\right] = \mathbb{E}_x\left[f(X(t))\right] = P_t f(x) \qquad (8.18)
\end{aligned}$$

Solution de l'Exercice 3.2

In the lecture notes we saw 1 $\iff$ 2.
For 2 $\implies$ 3 we only have to take derivatives at time $t = 0$
For 3 $\implies$ 4 it is trivial : let $\gamma(x) = Lf(x)/f(x)$ it satisfies $\gamma(x+y) = \gamma(x) + \gamma(y)$.
Eventually, 4 $\implies$ 1 comes from the method of characteristics. Let $C_1(\theta), C_2(\theta)$ denote the components of C_θ. Let ϕ_t be the flow of the ODE

$$\begin{aligned}
\theta_1' &= +C_1(\theta_1, \theta_2) \qquad (8.21)\\
\theta_2' &= +C_2(\theta_1, \theta_2) \qquad (8.22)
\end{aligned}$$

Then, by Kolmogorov's equation the function $u : t \to P_t f_{\phi_t(\theta)}(x)$ is constant: indeed let $v(t,\theta_1,\theta_2) = P_t f_\theta(x)$

$$\begin{aligned} u'(t) &= \partial_t v(t,\phi_t(\theta)) + \theta_1' \partial_{\theta_1} v(t,\phi_t(\theta)) + \theta_2' \partial_{\theta_2} v(t,\phi_t(\theta)) \\ &= \partial_t P_t f_\eta(x) - P_t L f_\eta(x) \qquad\qquad \text{for } \eta = \phi_t(\theta) \\ &= 0 \end{aligned}$$

so $P_t f_{\phi_t(\theta)}(x) = u(0) = f_\theta(x)$ and thus

$$P_t f_\theta(x) = f_{\phi_{-t}(\theta)}(x) = e^{\phi_{-t}(\theta).x}. \tag{8.23}$$

This is true at least for t in $(0,\delta)$, since on $(-\delta,\delta)$ the flow exists by Cauchy Lipschitz theory. Therefore, by semi group property, this is true for all t.

Solution de l'Exercice 3.3

The generator is

$$Lf(x) = \sum_i x_i(b_i(f(x+e_i) - f(x)) + d_i(x)(f(x-e_i) - f(x))) \tag{8.24}$$

Therefore $Lf_\theta(x) = f_\theta(x)\sum_i x_i(b_i(e^{\theta_i} - 1) + d_i(x)(e^{-\theta_i} - 1))$ and the process is branching when the death rates are constant.

Therefore, to prove that we can produce the coupling on $[0, T_{\mathscr{S}}(B^-))$ we only have to prove that for every monotone set $A \subset \mathscr{S}$

$$L^- 1_A(x) \le L 1_A(x) \le L^+ 1_A(x) \qquad (x \in \mathscr{S}). \tag{8.25}$$

SInce the monotone sets are of the type $A = [a, +\infty) \times [b, +\infty)$ this is fairly easy to prove.

Solution de l'Exercice 4.1

Let $\lambda_n := x_n.u$. Given $\epsilon > 0$, there exists n_0, p_0 such that forall $n \ge n_0$, $p \ge p_0$, $\left\| x_{n+p} - \lambda_n v \right\| \le \epsilon$.

Therefore, since $u.v = 1$, for $n \ge n_0$, $p \ge p_0$

$$\left|\lambda_{n+p} - \lambda_n\right| = \left|x_{n+p}.u - \lambda_n v.u\right| \le \|u\|_\infty \left\|x_{n+p} - \lambda_n v\right\| \le C\epsilon. \tag{8.26}$$

Hence, for $n \ge n_0$, $p \ge p_0$

$$\lambda_n - C\epsilon \le \lambda_{n+p} \le \lambda_n + C\epsilon. \tag{8.27}$$

Letting $p \to +\infty$, we obtain with $a = \liminf \lambda_k$ and $b = \limsup_k \lambda_k$,

$$\lambda_n - C\epsilon \le a \le b \le \lambda_n + C\epsilon. \tag{8.28}$$

Taking $n \to +\infty$ in $\lambda_n \le a + C\epsilon$ yields $b \le a + C\epsilon$. Letting $\epsilon \to 0$ yileds $b = a$ so there exists $\lambda \in \mathbb{R}$ such that $\lambda_n \to \lambda$. Let ϵ, n_0, p_0 be as above. Then, for $n \ge n_0$ and $p \ge p_0$

$$\left\|x_{n+p} - \lambda v\right\| \le \left\|x_{n+p} - \lambda_n v\right\| + |\lambda - \lambda_n| \|v\| \le \epsilon + |\lambda - \lambda_n| \|v\|. \tag{8.29}$$

Letting $n \to +\infty$ yields

$$\limsup_k \|x_k - \lambda v\| \le \epsilon . \tag{8.30}$$

Then, letting $\epsilon \to 0$ yields $x_k \to \lambda v$.

Solution de l'Exercice 4.2

The mean matrix is $M >> 0$

$$M = \begin{pmatrix} p_1(1+\alpha_1) & p_1(1-\alpha_1) \\ p_2(1-\alpha_2) & p_2(1+\alpha_2) \end{pmatrix} \tag{8.31}$$

with $det(M) = 2p_1p_2\alpha_1\alpha_2 > 0$ and $tr(M) > 0$. We can apply Perron Frobenius theorem, and the important thing is to find a left eigenvector. Numerically, we get $\rho = 1.79910717$ and a left eigenvector, normalized, with $vM = \rho M$ is $v = [0.00446413, 0.99553587]$. Therefore the asymptotic proportion of type A cells is $v[0] = 4.410^{-3}$ very small.

Bibliography

[1] Vincent Bansaye and Sylvie Méléard. *Stochastic models for structured populations*, volume 1 of *Mathematical Biosciences Institute Lecture Series. Stochastics in Biological Systems*. Springer, Cham; MBI Mathematical Biosciences Institute, Ohio State University, Columbus, OH, 2015. ISBN 978-3-319-21710-9; 978-3-319-21711-6. doi: 10.1007/978-3-319-21711-6. URL `https://doi.org/10.1007/978-3-319-21711-6`. Scaling limits and long time behavior.

[2] Stewart N. Ethier and Thomas G. Kurtz. *Markov processes*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York, 1986. ISBN 0-471-08186-8. doi: 10.1002/9780470316658. URL `https://doi.org/10.1002/9780470316658`. Characterization and convergence.

[3] Torgny Lindvall. On Strassen's theorem on stochastic domination. *Electron. Comm. Probab.*, 4:51–59, 1999. ISSN 1083-589X. doi: 10.1214/ECP.v4-1005. URL `https://doi.org/10.1214/ECP.v4-1005`.

[4] T. Kamae, U. Krengel, and G. L. O'Brien. Stochastic inequalities on partially ordered spaces. *Ann. Probability*, 5(6):899–912, 1977.

[5] Mu-Fa Chen. *From Markov chains to non-equilibrium particle systems*. World Scientific Publishing Co., Inc., River Edge, NJ, second edition, 2004. ISBN 981-238-811-7. doi: 10.1142/9789812562456. URL `https://doi.org/10.1142/9789812562456`.

[6] Erhan Çinlar. *Introduction to stochastic processes*. Prentice-Hall, Inc., Englewood Cliffs, N.J., 1975.

[7] Ludger Rüschendorf. On a comparison result for Markov processes. *J. Appl. Probab.*, 45(1):279–286, 2008. ISSN 0021-9002. doi: 10.1239/jap/1208358968. URL `https://doi.org/10.1239/jap/1208358968`.

[8] Thomas G. Kurtz. Strong approximation theorems for density dependent Markov chains. *Stochastic Processes Appl.*, 6(3):223–240, 1977/78. ISSN 0304-4149. doi: 10.1016/0304-4149(78)90020-0. URL `https://doi.org/10.1016/0304-4149(78)90020-0`.

[9] M. W. Hirsch and Hal Smith. Monotone dynamical systems. In *Handbook of differential equations: ordinary differential equations. Vol. II*, pages 239–357. Elsevier B. V., Amsterdam, 2005.

[10] H. Kesten and B. P. Stigum. A limit theorem for multidimensional Galton-Watson processes. *Ann. Math. Statist.*, 37:1211–1223, 1966. ISSN 0003-4851. doi: 10.1214/aoms/1177699266. URL `https://doi.org/10.1214/aoms/1177699266`.

[11] Ankit Gupta, J. A. J. Metz, and Viet Chi Tran. A new proof for the convergence of an individual based model to the trait substitution sequence. *Acta Appl. Math.*, 131:1–27, 2014. ISSN 0167-8019. doi: 10.1007/s10440-013-9847-y. URL `https://doi.org/10.1007/s10440-013-9847-y`.