Poisson Process

Birth and Death Processes

Références

[1] Karlin, S. and Taylor, H. M. (75), *A First Course in Stochastic Processes*, Academic Press : New-York.

[2] Karlin, S. and Taylor, H. M. (81), *A Second Course in Stochastic Processes*, Academic Press : New-York.

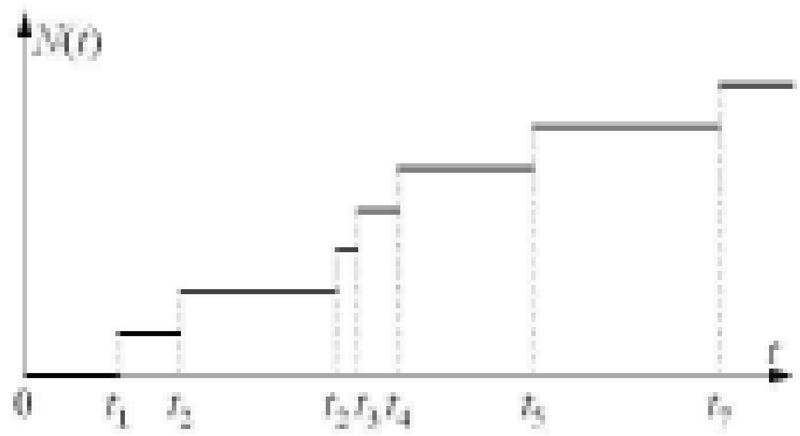
[3] Renshaw, E (91), *Modeling biological populations in space and time,* Cambridge University Press

The process $N(t)$ counts the number of events which occurred between time 0 and t.

One assumes that $N(0) = 0$.

Denoting $T_1, T_2, \ldots$ the times at which the event occurred, we have that the process $N(t)$ increases with 1 at each time.

$$
\begin{aligned}
N(t) &= 0 \quad \text{if } t < t_1, \\
&= 1 \quad \text{if } t_1 \le t \le t_2, \\
&\vdots \\
&= k \quad \text{if } t_k \le t < t_{k+1}, \\
& etc.
\end{aligned}
$$



Poisson Process

Examples.

- Arrivals of customers at a ticket office.
- Bus arrivals at a station.

Definition. $\{N(t), t > 0\}$ is a Poisson process with intensity λ if it satisfies the two following hypotheses :

Markov Process : Events occur independently from each other.

The future only depends on the past via the current value of $N(t)$.

Homogeneity : The probability for an event to occur between t and $t + \Delta t$ proportional to Δt (for Δt small) :

$$\Pr\{N(t+\Delta t) - N(t) = 1\} = \lambda \Delta t + o(\Delta t).$$

where λ is constant$\Rightarrow$ the process is homogeneous in time.

Differential equation system. For Δt small, we have

$$\begin{cases} \Pr\{N(t+\Delta t)-N(t)=1\} = \lambda\Delta t + o(\Delta t), \\ \Pr\{N(t+\Delta t)-N(t)=0\} = 1-\lambda\Delta t + o(\Delta t), \\ \Pr\{N(t+\Delta t)-N(t)\geq 2\} = o(\Delta t). \end{cases}$$

$\Rightarrow$ This last equation means that two events (or more) do not occurred in a same time.

Consequence. Denoting $p_n(t) = \Pr\{N(t)=n\}$, we get

$$p_n(t+\Delta t) = p_n(t) + \lambda\Delta t\,[p_{n-1}(t) - p_n(t)] + o(\Delta t)$$

which implies

$$p_n'(t) = \lambda\,[p_{n-1}(t) - p_n(t)]\,,$$

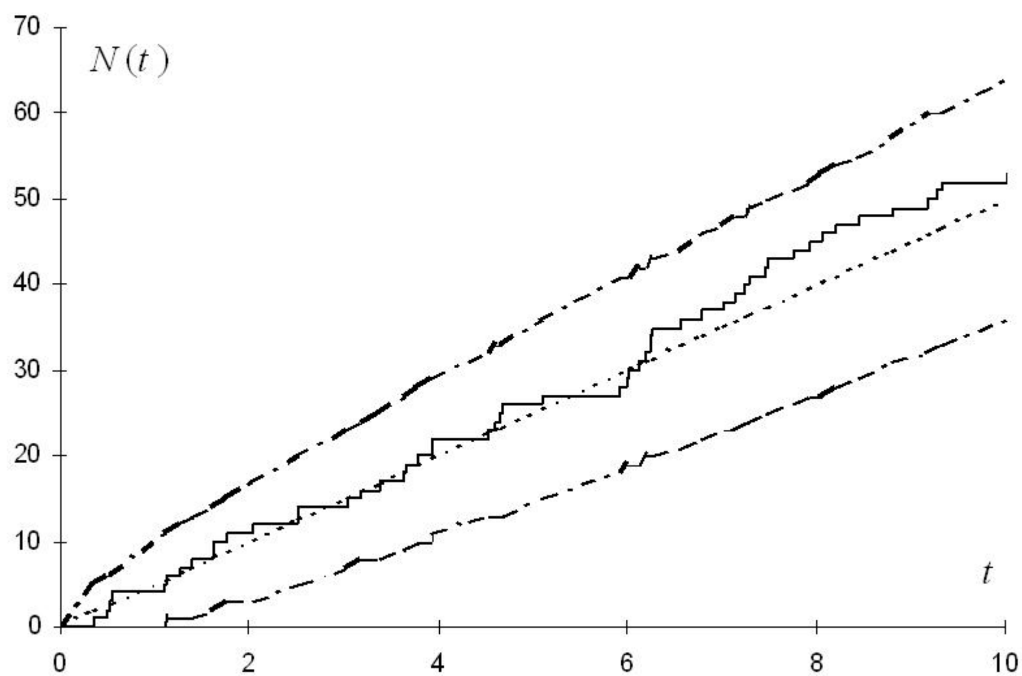
and

$$p_0'(t) = -\lambda p_0(t).$$

Distribution of the count $N(t)$. The solution of the preceding systems implies that $N(t)$ has a Poisson distribution :

$$p_n(t) = e^{-\lambda t}\frac{(\lambda t)^n}{n!} \qquad \Rightarrow \qquad N(t) \sim \mathcal{P}(\lambda t).$$

As a consequence : $\mathbb{E}[N(t)] = \lambda t, \quad \mathbb{V}[N(t)] = \lambda t$.



Interpretation of the intensity λ.

By taking $t = 1$, we have that $N(1) \sim \mathcal{P}(\lambda)$ and

$$\mathbb{E}[N(1)] = \lambda,$$

So λ is the number of events in average which occurred in a time unit.

We have that $N(t)$ tends to infinity

$$\Pr\{N(t) = \infty\} \longrightarrow 1$$

Poisson Process. Main properties

Property of the times $T_1, \ldots, T_n$.

Suppose that n events occur between $[0, t]$.

Recall that T_k is the time at which the event k occurs, then we have that

the distribution of the times $T_1 < \ldots < T_n$

is an uniform on the interval $[0, t]$.

Waiting times. Let us denote $\Delta T_i = (T_{i+1} - T_i)$ the time between two events,

$$\Delta T_i = (T_{i+1} - T_i) \sim \mathcal{E}(\lambda) \qquad \Rightarrow \qquad \Pr\{\Delta T_i > t\} = \exp(-\lambda t)$$

As a consequence : $\mathbb{E}(\Delta T_i) = 1/\lambda$, $\quad \mathbb{V}(\Delta T_i) = 1/\lambda^2$ so $1/\lambda$ is the average time between two events.

Distribution of T_i. We can derive that

$$T_i \sim \gamma(i, \lambda) \quad \mathbb{E}(T_i) = i/\lambda$$

Property of the exponential distribution.

Conditional distribution :

$$\Pr\{T > s + t \mid T > s\} = \Pr\{T > t\}\,.$$

The absence of memory (Markov) implies the 'Bus stop' paradox :

$$\mathbb{E}[T - s \mid T > s] = \mathbb{E}[T]$$

$\Rightarrow$ Whatever the time at which we arrive at the bus station, the mean waiting time is the mean time.

Poisson Process.

Estimator of λ ?

$$\hat{\lambda} = \frac{N(t)}{t},$$

i.e. the number of events on the interval $[0, t]$ over the considered time interval.

Likelihood.

$$V(N(t); \lambda | N(t) = n) = \lambda^n exp\left[\lambda(\sum_{i=1}^{n} -(t_i - t_{i-1}))\right] exp\left[\lambda - (T - t_n))\right]$$

$\Rightarrow$ $\hat{\lambda}$ is the maximum likelihood estimator.

Application to Genetic Distance

Statistical model

Crossing-over occur along the chromosome according to an (homogeneous) Poisson process with intensity λ.

$N(t)$ = number of crossing over occurring in a portion of length t :

$$N(t) \sim \mathcal{P}(\lambda t).$$

Probability of common origin for 2 loci (at distance t).

$$p(t) = \Pr\{N(t) \text{ is even}\} = \sum_k \Pr\{N(t) = 2k\}.$$

Remarking that $\mathrm{e}^{\lambda t} + \mathrm{e}^{-\lambda t} = 2 \sum_{k \geq 0} \left[(\lambda t)^{2k} / 2k! \right]$ we get

$$p(t) = \left(1 + \mathrm{e}^{-2\lambda t} \right) / 2$$

which goes to $1/2$ when t goes to infinity.

Application to Genetic Distance

Recombination probability. The 'recombination probability' is the probability for two loci (at distance t) to be issued from different parents : $q(t) = 1 - p(t)$.

For a small t, we have

$$q(t) = \frac{1}{2}\left(1 - \mathrm{e}^{-2\lambda t}\right) \simeq \lambda t$$

CentiMorgan (cM) definition. One cM is the distance d such as $q(d) = 1\%$, i.e.

$$d = -\frac{\log(0.98)}{2\lambda}.$$

If t is measured in cM, owe get

$$\begin{aligned} q(t) &= \frac{1}{2}\left\{1 - \exp\left[2\lambda t \frac{\log(0.98)}{2\lambda}\right]\right\} \\ &= \frac{1}{2}\left(1 - 0.98^{t}\right). \end{aligned}$$

which also goes to $1/2$ when t goes to infinity.

Pure Birth Process

Model The $N(t)$ process counts the number of births between 0 and t. The intensity of the process depends on the population size at time t :

$$\Pr\{N(t+\Delta t) = N(t)+1\} = \lambda[N(t)]\Delta t + o(\Delta t),$$

and

$$\Pr\{N(t+\Delta t) = N(t)+2\} = o(\Delta t), \text{ no more births at the same time}$$

where $\lambda(n)$ is some given function.

- The process is always without memory : the number of birth before t does not affect the number of births after t,
- The process is not homogeneous in time ($\lambda(n)$ could depends on n)

Pure Birth Process

Different function $\lambda(n)$:

Poisson process :	$\lambda(n) = \lambda.$	
Linear birth process :	$\lambda(n) = \lambda n$	proportional to the population size.
Quadratic birth process :	$\lambda(n) = \lambda n^2$	proportional to the number of couples in the population.
Density dependent :	$\lambda(n) = \lambda n \left(1 - \frac{n}{n_{n_{\max}}}\right)$	$n_{n_{\max}}$ is the maximal capacity of the environment.

Case of linear birth process

Distribution of $N(t)$?

Solving a differential equation system, we get

$$p_n(t) = \binom{n-1}{n_0-1} \left(e^{-\lambda t}\right)^{n_0} \left(1 - e^{-\lambda t}\right)^{n-n_0},$$

which means that the population size at time t has a binomial negative distribution : $N(t) - n_0 \sim \mathcal{NB}\left(n_0, e^{-\lambda t}\right)$,

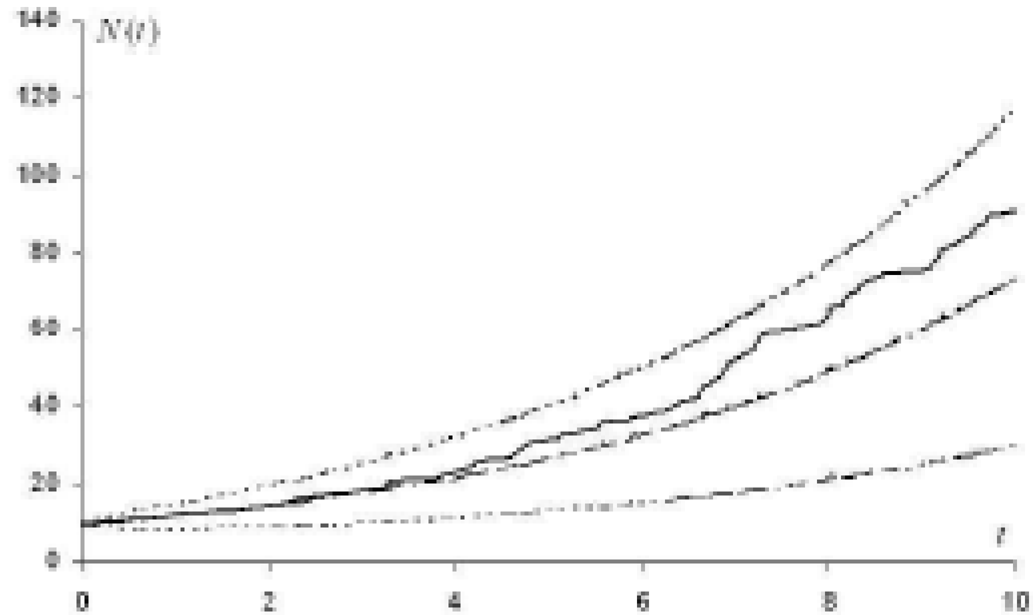
with $1 - e^{-\lambda t}$ is the probability for an individual to be born in $[0, t]$

Pure Birth Process

$$\mathbb{E}[N(t)] = n_0 e^{\lambda t}$$

consistent with the exponential growth deterministic model.

$$\mathbb{V}[N(t)] = n_0 e^{\lambda t}(e^{\lambda t} - 1)$$



Waiting times ?

Property of the exponential distribution : distribution of the minimum.

Let $X \sim \mathcal{E}(\lambda)$ and $Y \sim \mathcal{E}(\mu)$ be 2 independent random variables and Z their minimum :

$$Z = \inf(X, Y),$$

we have

$$\begin{array}{ll} (i) & Z \sim \mathcal{E}(\lambda + \mu); \\ (ii) & \Pr\{Z = X\} = \dfrac{\lambda}{\lambda + \mu}, \qquad \Pr\{Z = Y\} = \dfrac{\mu}{\lambda + \mu}. \end{array}$$

Pure Birth Process

Waiting time

- $N(t)$ individuals are present at time t.
- Each of them (numbered $i = 1..N(t)$) will give birth to a new individual in random time T_i with exponential $\mathcal{E}(\lambda)$ distribution.
- Next birth will occur at time

$$t + \min_{i=1..N(t)} T_i$$

- The waiting time until the next birth is distributed as the minimum of $N(t)$ independent exponential $\mathcal{E}(\lambda)$ random times. It is hence distributed as :

$$\mathcal{E}[\lambda N(t)] \qquad \Rightarrow \qquad \text{Mean waiting time :} \frac{1}{\lambda N(t)}$$

Counting process. At each time, the population size increases with 1.

Estimator of λ

The maximum likelihood estimator is

$$\hat{\lambda} = \frac{N(t) - n_0}{n_0 t + \sum_{i=1}^{N(t)-n_0} (t - t_i)}$$

$N(t) - n_0$ events on $[0, t]$ and

- n_0 live on $[0, t]$,
- 1 lives on $[t_1, t]$,
- $\ldots$
- 1 lives on $[t_{N(t)-n_0}, t]$

Pure Death Process

Model : Linear Death intensity $\mu(n) = \mu \times n$

Starting with n_0, the death rate is proportional to the population size :

$$\Pr\{N(t+\Delta t) = N(t) - 1\} = \mu N(t)\Delta t + o(\Delta t),$$

Survival time. Each individual stays alive an exponential $\mathcal{E}(\mu)$ time so it has probability $e^{-\mu t}$ to be still alive at time t.

Population size has a binomial distribution :

$$N(t) \sim \mathcal{B}(n_0, e^{-\mu t}) \qquad \Rightarrow \qquad p_n(t) = \binom{n_0}{n} e^{-n\mu t} \left(1 - e^{-\mu t}\right)^{n_0 - n}.$$

Waiting times. The waiting time until the next has an exponential $\mathcal{E}[\mu N(t)]$ distribution.

Counting process. At each time, the population size decreases with 1.

Pure Death Process

Time to extinction

Let denote T_0 the time to extinction of the population, we have that

$$\Pr\{T_0 \leq t\} = (1 - e^{-\mu t})^{n_0}$$

Mean Time to extinction

$$\mathbb{E}[T_0] = \mathbb{E}[\Delta T_0 + \Delta T_1 + \ldots + \Delta T_{n_0 - 1}] = (1/\mu) \times (0.577 + \log(n_0))$$

Estimation of μ

$$\hat{\mu} = \frac{n_0 - N(t)}{(n_0 - N(t))t + \sum_{i=1}^{N(t)} t_i}$$

Birth and Death Process

Model. Linear intensities.

Both probabilities to observe either a death or a birth between t and $t + \Delta t$ are proportional to Δt and to the population size $N(t)$

Differential equation system. Following the preceding models, we get

$$\begin{aligned} p_n(t + \Delta t) &= p_n(t) \times [1 - n(\lambda + \mu)\Delta t] \\ &\quad + p_{n-1}(t) \times (n-1)\lambda\Delta t \\ &\quad + p_{n+1}(t) \times (n+1)\mu\Delta t \\ &\quad + o(\Delta t). \end{aligned}$$

which implies

$$p'_n(t) = -n(\lambda + \mu)p_n(t) + (n-1)\lambda p_{n-1}(t) + (n+1)\mu p_{n+1}(t)$$

.... difficult, but solvable.

Remark : other intensities

Other intensities $\lambda(n)$ and $\mu(n)$ (quadratic, density-dependent, etc.) can be considered.

We then get the general differential equation system

$$p_n'(t) = -[\lambda(n) + \mu(n)]p_n(t) + \lambda(n-1)p_{n-1}(t) + \mu(n+1)p_{n+1}(t)$$

which is not solvable in a close form in general.

$\Rightarrow$ The system can yet be studied using computer simulations.

Case of an initial population $n_0 = 1$. Solving the differential equation system, we get

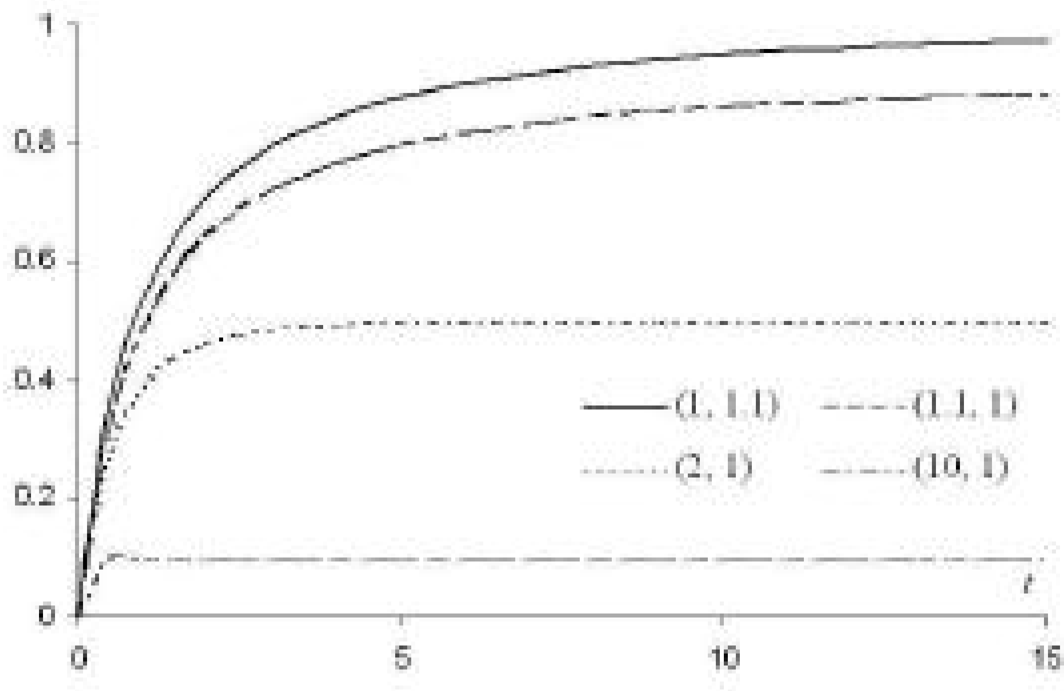
$$\begin{aligned} p_0(t) &= \mu g(t) = \text{probability of extinction before } t \\ p_n(t) &= (1-\mu g(t))(1-\lambda g(t))(\lambda g(t))^{n-1} \end{aligned}$$

with $g(t) = (\exp((\lambda-\mu)t))/(\lambda \exp((\lambda-\mu)t) - \mu)$ and

$$\mathbb{E}[N(t)] = \exp((\lambda-\mu)t)$$

$\Rightarrow$ depends on the sign of $\lambda - \mu$.

Probability of the time to extinction of the population for $n_0 = 1$ for different values of (λ, μ)



When $\mu > \lambda$, the extinction is certain.

When $\lambda > \mu$, even $\lambda >> \mu$, the probability of the extinction of the population is not zero.

Important properties. At time t, $N(t) = n$ individual are present.

Next birth : The waiting time until the next birth is $\mathcal{E}[\lambda(n)]$.

Next death : The waiting time until the next death is $\mathcal{E}[\mu(n)]$.

Next event : The waiting time until the next event is

$$\min\{\mathcal{E}[\lambda(n)], \mathcal{E}[\mu(n)]\} = \mathcal{E}[\lambda(n) + \mu(n)].$$

Birth or death ? The next event will be

a birth	with probability :	$\lambda(n)/[\lambda(n) + \mu(n)]$
a death	with probability :	$\mu(n)/[\lambda(n) + \mu(n)]$

This process can be seen as two processes :

Population size. The sequence of the population size (regardless of the time) is a Markov chain with transition matrix $\mathbf{\Pi} =$

$$\begin{pmatrix} \ddots & \ddots & \ddots & & & \\ & \frac{\mu(n-1)}{\mu(n-1)+\lambda(n-1)} & 0 & \frac{\lambda(n-1)}{\mu(n-1)+\lambda(n-1)} & & \\ & & \frac{\mu(n)}{\mu(n)+\lambda(n)} & 0 & \frac{\lambda(n)}{\mu(n)+\lambda(n)} & \\ & & & \frac{\mu(n+1)}{\mu(n+1)+\lambda(n+1)} & 0 & \frac{\lambda(n+1)}{\mu(n+1)+\lambda(n+1)} \\ & & & & \ddots & \ddots \end{pmatrix}$$

Waiting times between events are random ($X \sim \mathcal{E}[\lambda]$ then $X/n \sim \mathcal{E}[\lambda n]$)

$\Rightarrow$ The waiting times between events, and the sequences of the population sizes can be simulated independently.

Transition rate matrix. Both transition probabilities ($\mathbf{\Pi}$) and waiting times (exponential) can be summarized in the transition rate matrix $\mathbf{R} =$

$$\begin{pmatrix} \ddots & \ddots & \ddots & & & \\ & \mu(n-1) & -[\mu(n-1)+\lambda(n-1)] & \lambda(n-1) & & \\ & & \mu(n) & -[\mu(n)+\lambda(n)] & \lambda(n) & \\ & & & \mu(n+1) & -[\mu(n+1)+\lambda(n+1)] & \lambda(n+1) \\ & & & & \ddots & \ddots \end{pmatrix}$$

Distribution at time t. The general form of the differential equation system is

$$\mathbf{p}'(t) = \mathbf{p}(t)\mathbf{R} \qquad \text{where } \mathbf{p}(t) = [\ p_0(t) \quad p_1(t) \quad \ldots \quad p_n(t) \quad \ldots\ ]$$

and its solution is

$$\mathbf{p}(t) = \mathbf{p}(0)\exp(\mathbf{R}t) = \mathbf{p}(0)\exp(\mathbf{R})^t.$$

Stationary distribution. Stationary distributions are eigenvectors of $\mathbf{R}$ associated with a null eigenvalue.

Example of density-dependent birth process

Initial size : $N(0) = n_0$

Birth rate : $\lambda(n) = \lambda \left(1 - \frac{n}{n n_{\max}}\right)$

Death rate : $\mu(n) = \mu n$

Immigration rate $\gamma(n) = \gamma$

In presence of immigration, the state $N(t) = 0$, is not absorbing.

Parameters : $n_0 = 1, n_{\max} = 5, \lambda = 1, \mu = 0, \gamma = 0$

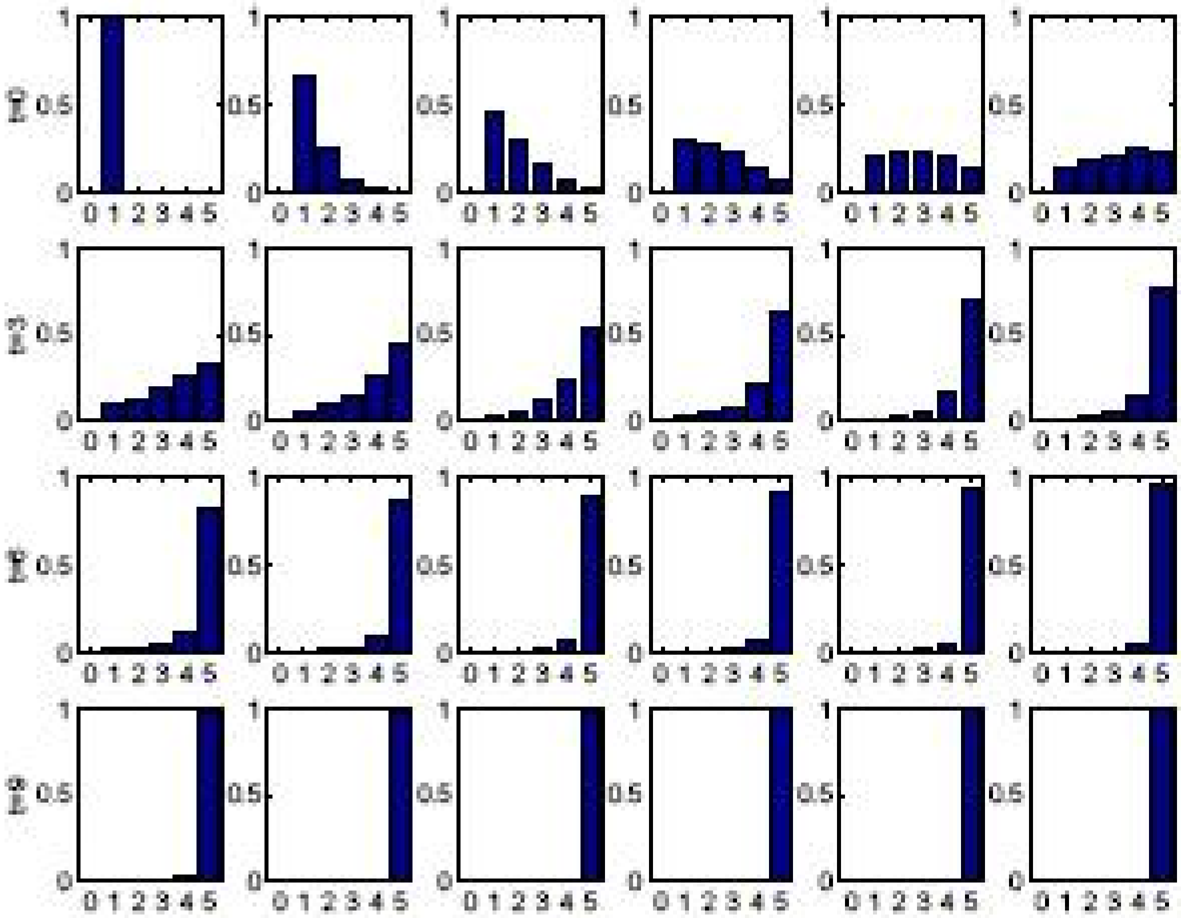
$$\mathbf{R} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -0.8 & 0.8 & 0 & 0 & 0 \\ 0 & 0 & -1.2 & 1.2 & 0 & 0 \\ 0 & 0 & 0 & -1.2 & 1.2 & 0 \\ 0 & 0 & 0 & 0 & -0.8 & 0.8 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Two stationary distributions :

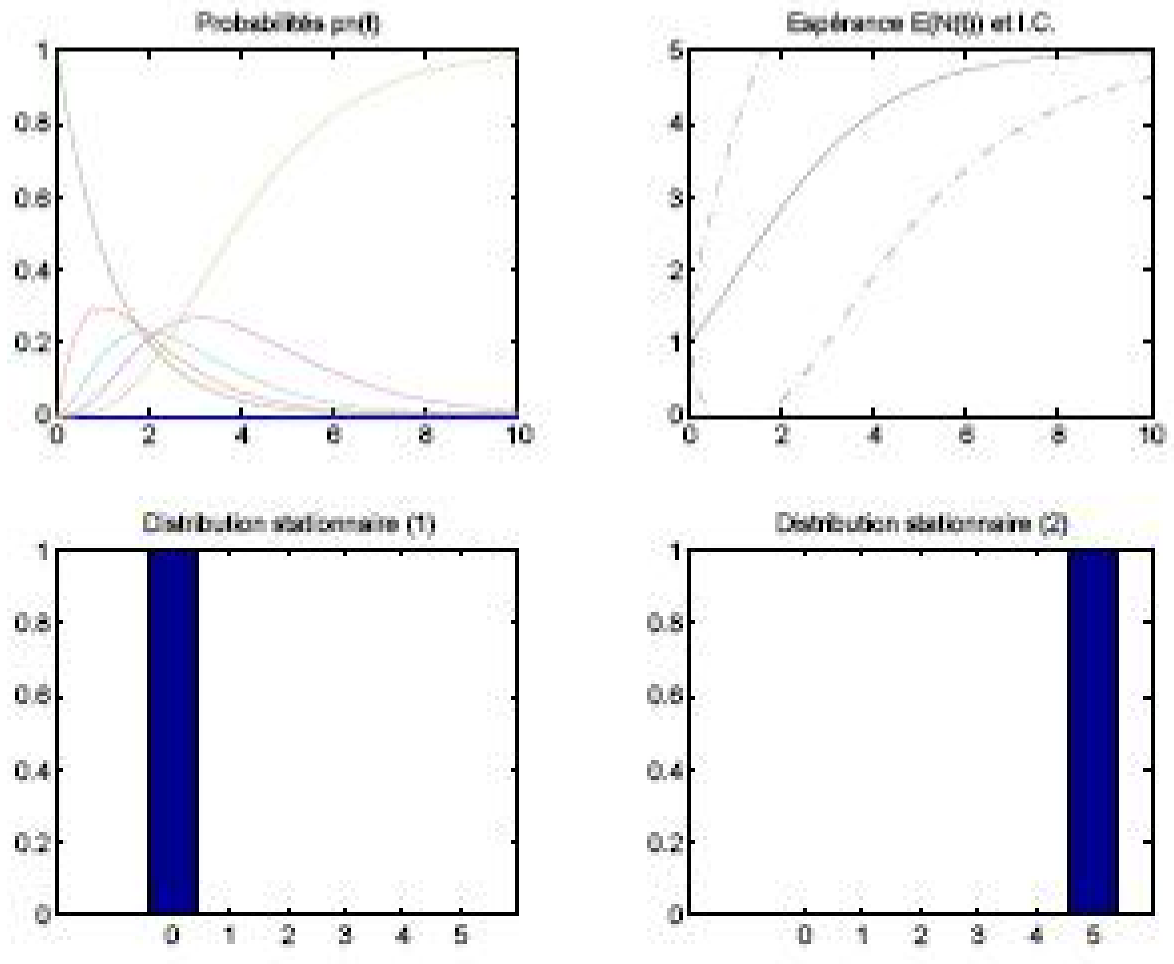
$$\begin{aligned} \boldsymbol{\mu} &= [\ 1 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0\ ] \quad \text{i.e. } N(t) = 0 \\ \boldsymbol{\mu}' &= [\ 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 1\ ] \quad \text{i.e. } N(t) = n_{\max} \end{aligned}$$

so the chain is reducible. ($\boldsymbol{\mu}$ can not be reached from $n_0 = 1$.)

Distribution as a function of time.



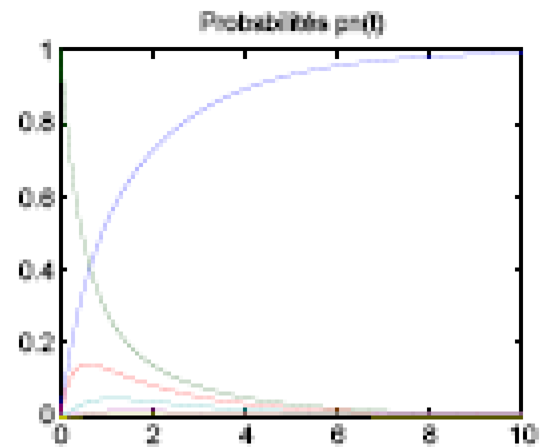
Mean, confidence intervals and stationary distributions :

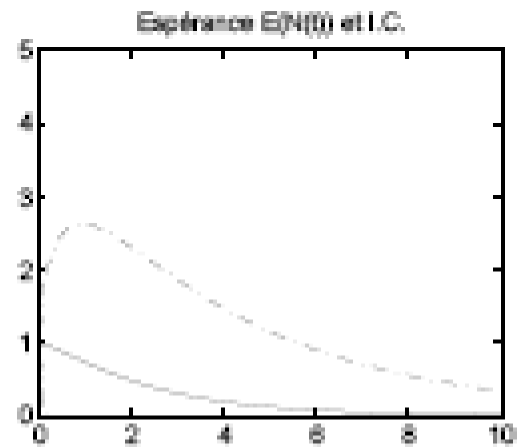


The population grows until it reaches $n_{\max}$.

Birth and death (1) $n_0 = 1, n_{\max} = 5, \lambda = 1, \mu = 1, \gamma = 0$

$$\mathbf{R} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & -1.8 & 0.8 & 0 & 0 & 0 \\ 0 & 2 & -3.2 & 1.2 & 0 & 0 \\ 0 & 0 & 3 & -4.2 & 1.2 & 0 \\ 0 & 0 & 0 & 4 & -4.8 & 0.8 \\ 0 & 0 & 0 & 0 & 5 & -5 \end{pmatrix}$$

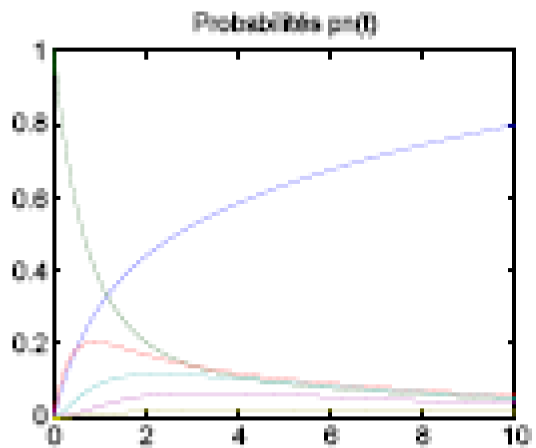


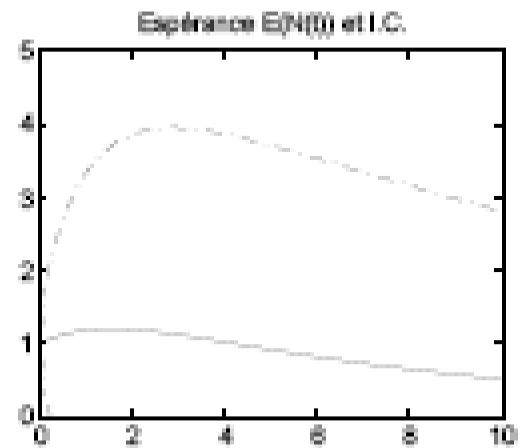


$N = 0$ is an absorbing state toward which the process converges.

Birth and death (2) $n_0 = 1, n_{\max} = 5, \lambda = 1, \mu = 0.5, \gamma = 0$

$$\mathbf{R} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0.5 & -1.3 & 0.8 & 0 & 0 & 0 \\ 0 & 1 & -2.2 & 1.2 & 0 & 0 \\ 0 & 0 & 1.5 & -2.7 & 1.2 & 0 \\ 0 & 0 & 0 & 2 & -2.8 & 0.8 \\ 0 & 0 & 0 & 0 & 2.5 & -2.5 \end{pmatrix}$$

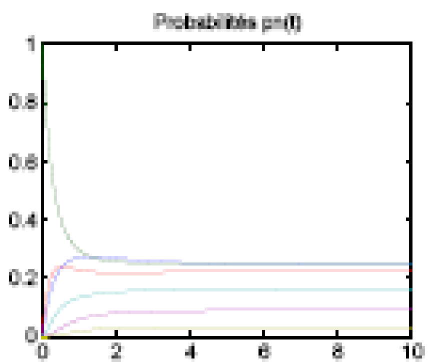


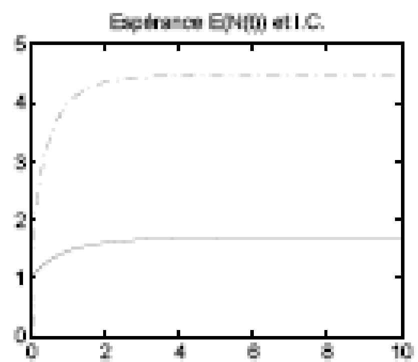


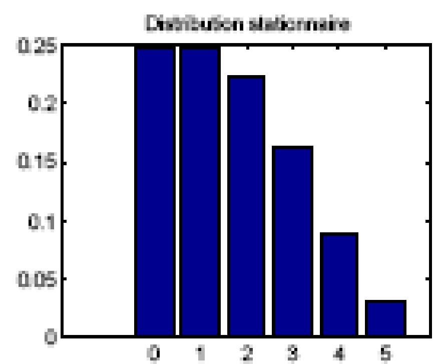
Reducing the death rate only delays the extinction of the population.

Birth, death and immigration $n_0 = 1, n_{\max} = 5, \lambda = 1, \mu = 1, \gamma = 1$

$$\mathbf{R} = \begin{pmatrix} -1 & 1 & 0 & 0 & 0 & 0 \\ 1 & -2.8 & 1.8 & 0 & 0 & 0 \\ 0 & 2 & -4.2 & 2.2 & 0 & 0 \\ 0 & 0 & 3 & -5.2 & 2.2 & 0 \\ 0 & 0 & 0 & 4 & -5.8 & 1.8 \\ 0 & 0 & 0 & 0 & 5 & -5 \end{pmatrix}$$







$$\boldsymbol{\mu} = [\ 0.247 \quad 0.247 \quad 0.22 \quad 0.16 \quad 0.09 \quad 0.03\ ]$$

Aim

- Estimate times since divergence between species
- Reconstruct the phylogenetic tree

Data

- For each specie $i = 1..n$, we have one nucleotide sequence

$$\mathbf{S}_i = (S_{i1}, \dots S_{i\ell})$$

S_{ix} = nucleotide in position $x = 1..\ell$

- The n sequences $(\mathbf{S}_1, \dots \mathbf{S}_n)$ are aligned : $\begin{bmatrix} S_{11} & \dots & S_{1x} & \dots & S_{1\ell} \\ \vdots & & \vdots & & \vdots \\ S_{i1} & \dots & S_{ix} & \dots & S_{i\ell} \\ \vdots & & \vdots & & \vdots \\ S_{n1} & \dots & S_{nx} & \dots & S_{n\ell} \end{bmatrix}$

Model

- Nucleotides are supposed to evolve (i.e. mutate) independently
- according to a continuous time Markov process with state space

$$\mathcal{A} = \{\mathtt{a}, \mathtt{c}, \mathtt{g}, \mathtt{t}\}$$

- and transition rates

$$\mathbf{R} = \begin{pmatrix} - & r(\mathtt{a}, \mathtt{c}) & r(\mathtt{a}, \mathtt{g}) & r(\mathtt{a}, \mathtt{t}) \\ r(\mathtt{c}, \mathtt{a}) & - & r(\mathtt{c}, \mathtt{g}) & r(\mathtt{c}, \mathtt{t}) \\ r(\mathtt{g}, \mathtt{a}) & r(\mathtt{g}, \mathtt{c}) & - & r(\mathtt{g}, \mathtt{t}) \\ r(\mathtt{t}, \mathtt{a}) & r(\mathtt{t}, \mathtt{c}) & r(\mathtt{t}, \mathtt{g}) & - \end{pmatrix}$$

$\pi_{ab}(t) =$ transition probability from nucleotide a to nucleotide b in a time t.

Example : Juke and Cantor Model

Hypothesis. All transition rates are equal :

$$\mathbf{R} = \begin{pmatrix} - & \alpha & \alpha & \alpha \\ \alpha & - & \alpha & \alpha \\ \alpha & \alpha & - & \alpha \\ \alpha & \alpha & \alpha & - \end{pmatrix}$$

Property.

$$\pi_{aa}(t) = \frac{1}{4} + \frac{3}{4}e^{-4\alpha t}, \qquad \pi_{ab}(t) = \frac{1}{4}\left(1 - e^{-4\alpha t}\right)$$

Estimation of the divergence time.

$$\hat{t}_{ij} = -\frac{3}{4}\ln\left(1 - \frac{4}{3}p_{ij}\right)$$

where p_{ij} is the proportion of different nucleotides between sequences i and j