POISSON PROCESSES

Cyril Labbé
Stefano Olla
Université Paris-Dauphine
olla@ceremade.dauphine.fr
www.ceremade.dauphine.fr/olla/

January 6, 2023

Contents

Introduction

In this course, we will introduce stochastic processes in *continuous time*. They form a very rich class of processes that arise in many different situations. The aim of this course is to introduce a few important continuous-time processes (counting processes, Poisson processes, compound Poisson processes, jump Markov processes) . We will start with counting processes, that are the simplest example of continuous-time processes: they are processes that are non-decreasing, right continuous with left limits, and take values in **N**. The most important such process is the Poisson process that we will study thoroughly in Chapter 1. In Chapter 2 we will introduce elements of the theory of Markov processes on a countable state-space. Chapter 3 we will introduce queueing theory and branching processes. Finally, in Chapter 4 we will move to renewal processes and renewal theory, which is used in risk theory.

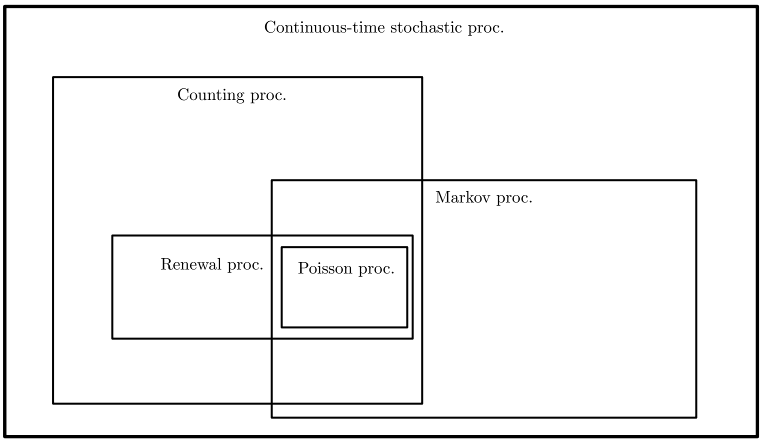

Figure 1: Continuous-time stochastic processes include Counting processes and Markov processes. Among Counting processes, one finds the class of Renewal processes. The intersection of all these classes of processes consists of Poisson processes.

We conclude this introduction with general definitions on continuous-time stochastic processes. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space.

A random variable Y is a measurable map from $(\Omega, \mathcal{F})$ into some measurable space $(E, \mathcal{E})$. Most of the time, the measurable space is $(\mathbf{R}, \mathcal{B}(\mathbf{R}))$, that is, the set of real values endowed with its Borel sigma-field, or $(\mathbf{N}, \mathcal{P}(\mathbf{N}))$ the set of integers endowed with its natural sigma-field. From time to time (and implicitly), the measurable space is taken to be $\mathbf{R} \cup \{+\infty\}$ endowed with its natural sigma-field. The law of a random variable Y is the pushforward (=mesure image) of $\mathbb{P}$ through the map Y: it is a probability measure μ on the space $(E, \mathcal{E})$ defined by

$$\mu(A) = \mathbb{P}(Y^{-1}(A)) \ ,$$

for all A in $\mathcal{E}$. Recall that $Y^{-1}(A) = \{\omega \in \Omega : Y(\omega) \in A\}$.

At the first semester, were introduced stochastic processes in *discrete* time, that is, collections of real-valued random variables indexed by the set $\mathbb{N}$. For many different reasons, it is natural to deal with stochastic processes in *continuous* time.

Definition 0.0.1. *A stochastic process $X = (X_t, t \geq 0)$ is a collection of real-valued random variables indexed by the set of nonnegative real numbers $[0, \infty)$.*

In discrete time, the trajectories of the process, that is the maps $n \mapsto X_n(\omega)$ for $\omega \in \Omega$, are random sequences. In continuous time, they become “random functions” $t \mapsto X_t(\omega)$ and one can ask about the regularity of these functions.

Fix $\omega \in \Omega$. The map $t \mapsto X_t(\omega)$ is said to be right continuous if for all $t \geq 0$, $X_t(\omega) = \lim_{s\downarrow 0} X_{t+s}(\omega)$. Similarly, it is said to be left continuous if for all $t > 0$, $X_t(\omega) = \lim_{s\downarrow 0} X_{t-s}(\omega)$. It is said to admit left limits if for all $t > 0$, $\lim_{s\downarrow 0} X_{t-s}(\omega)$ exists: in that case, we denote by $X_{t-}(\omega)$ the left limit at t.

Definition 0.0.2 (Continuity)**.** *We say that a stochastic process $(X_t, t \in \mathbf{R}_+)$ is almost surely right continuous, respectively left continuous, if the event*

$$\{\omega \in \Omega : \forall t \in \mathbf{R}_+, X_t(\omega) = \lim_{s\downarrow 0} X_{t+s}(\omega)\} \ ,$$

respectively the event

$$\{\omega \in \Omega : \forall t \in \mathbf{R}_+ \backslash \{0\}, X_t(\omega) = \lim_{s\downarrow 0} X_{t-s}(\omega)\} \ ,$$

belongs to $\mathcal{F}$ and has probability 1.
We say that $(X_t, t \in \mathbf{R}_+)$ is almost surely continuous if it is right and left continuous.

Definition 0.0.3 (Càdlàg)**.** *We say that $(X_t, t \in \mathbf{R}_+)$ admits almost surely left limits if*

$$\{\omega \in \Omega : \forall t > 0, \lim_{s\downarrow 0} X_{t-s}(\omega) \textit{ exists}\} \ ,$$

belongs to $\mathcal{F}$ and has probability 1.
We say that $(X_t, t \in \mathbf{R}_+)$ is almost surely càdlàg if it is right continuous and if it admits left limits.

In English, non-decreasing means "croissant" while increasing means "strictement croissant".

Definition 0.0.4 (Non-decreasing)**.** *We say that $(X_t, t \in \mathbf{R}_+)$ is almost surely non-decreasing if the event*

$$\{\omega \in \Omega : \forall t > s \geq 0, X_t(\omega) \geq X_s(\omega)\} \ ,$$

belongs to $\mathcal{F}$ and has probability 1.

Frequently, we will simply write "continuous" or "càdlàg" for "almost surely continuous" or "almost surely càdlàg".

Remark 0.0.5. *The measurability of the events above is not granted by the mere fact that* $(X_t, t \in \mathbf{R}_+)$ *is a stochastic process. Indeed, these events depend on uncountably many random variables* X_t*,* $t \in \mathbf{R}_+$*.*

Recall that

$$\mathbb{P}(\omega \in \Omega : \forall t \in \mathbf{R}_+, X_t(\omega) = \lim_{s \downarrow 0} X_{t+s}(\omega)) = 1 \ ,$$

is equivalent to saying that $\mathbb{P}$-almost surely

$$\forall t \in \mathbf{R}_+, X_t = \lim_{s \downarrow 0} X_{t+s} \ .$$

$\mathbb{P}$-almost surely is often abbreviated $\mathbb{P}$-a.s.

Chapter 1

Counting processes and the Poisson process

1.1 Counting processes

Definition 1.1.1. *A process $N = (N_t, t \geq 0)$ is a counting process if $\mathbb{P}$-a.s.:*

- $N_0 = 0$,
- *N is non-decreasing and right continuous,*
- *$N_t \in \mathbf{N} \cup \{+\infty\}$ for all $t \geq 0$.*

The jump times of a counting process N are defined as the sequence of random variables

$$T_n := \inf\{t \geq 0 : N_t \geq n\} , \quad n \geq 0 ,$$

with the convention $\inf \emptyset = +\infty$. Notice that $T_0 = 0$ and $(T_n, n \geq 0)$ is non-decreasing, $\mathbb{P}$-a.s.

Example 1.1.2. *The total number of claims to an insurance company can be modelled by a counting process $(N_t, t \geq 0)$. The r.v. T_n is the time of the n-th claim.*

Lemma 1.1.3. *We have $\mathbb{P}$-a.s.*

$$N_t = \#\{k \geq 1 : T_k \leq t\} = \sum_{k \geq 1} \mathbf{1}_{\{T_k \leq t\}} , \quad t \geq 0 .$$

Proof. By definition, N takes values in $\mathbf{N} \cup \{+\infty\}$ and is non-decreasing. Consequently, $\mathbb{P}$-a.s. we have for all $t \geq 0$ and all $n \geq 0$:

$$N_t = n \Leftrightarrow n \leq N_t < n + 1 \Leftrightarrow T_n \leq t < T_{n+1} \Leftrightarrow \#\{k \geq 1 : T_k \leq t\} = n ,$$

thus yielding the first identity. The second identity is immediate. □

Proposition 1.1.4. *The law of a counting process N is completely characterised by the law of its sequence of jump times $(T_n)_{n \geq 1}$.*

Remark 1.1.5. *Before talking about the notion of law of a stochastic process, we need to specify the measurable space it lives in: we consider the product space* $\mathbf{R}^{[0,\infty)}$ *endowed with the product sigma-algebra. The law of a stochastic process* X *is the pushforward of* $\mathbb{P}$ *through the map* $\omega \mapsto (X_t(\omega), t \in [0,\infty))$.

Proof. The law of N is characterised by the law of its marginals $(N_{t_1}, \ldots, N_{t_n})$ for all $0 \le t_1 < t_2 < \ldots < t_n$ and all $n \ge 1$. For any $k_1 \le k_2 \le \ldots \le k_n$, the events

$$\{N_{t_1} = k_1, \ldots, N_{t_n} = k_n\} ,$$

and

$$\{T_{k_1} \le t_1 < T_{k_1+1}, T_{k_2} \le t_2 < T_{k_2+1}, \ldots, T_{k_n} \le t_n < T_{k_n+1}\} ,$$

coincide up to $\mathbb{P}$-negligible sets. As a consequence, the law of $(N_{t_1}, \ldots, N_{t_n})$ is completely characterised by the law of the sequence $(T_k)_{k\ge1}$, thus concluding the proof. □

Definition 1.1.6. *We say that* N *is a standard counting process if it is a counting process and*

1. $N_t \to \infty$ *as* $t \to \infty$, $\mathbb{P}$*-a.s.,*
2. $N_t - N_{t-} \in \{0,1\}$ *for all* $t \ge 0$, $\mathbb{P}$*-a.s.,*
3. $N_t < \infty$ *for all* $t \ge 0$, $\mathbb{P}$*-a.s.*

Recall that increasing means "strictement croissant".

Proposition 1.1.7. *Let* N *be a counting process and* $(T_n)_{n\ge0}$ *its jump times.* N *is standard if and only if*

(a) $T_n < \infty$ *for all* $n \ge 0$, $\mathbb{P}$*-a.s.,*

(b) $(T_n)_{n\ge0}$ *is increasing,* $\mathbb{P}$*-a.s.,*

(c) $T_n \to \infty$ *as* $n \to \infty$, $\mathbb{P}$*-a.s.*

Proof. We prove the equivalence (c) ⇔ 3. Almost surely we have

$$\begin{aligned}\{T_n \to \infty\} &= \{\forall t \ge 0, \exists n_t \in \mathbf{N} : T_{n_t} > t\} \\ &= \{\forall t \ge 0, \exists n_t \in \mathbf{N} : N_t < n_t\} \\ &= \{\forall t \ge 0, N_t < \infty\} .\end{aligned}$$

We turn to the equivalence (a) ⇔ 1. Almost surely we have

$$\{N_t \to \infty\}^c = \{\exists n_0 : \forall t \ge 0, N_t < n_0\} = \{\exists n_0 : T_{n_0} = \infty\} = \{\forall n \ge 0 : T_n < \infty\}^c .$$

We now prove that (b) implies 2. If there exists $t > 0$ such that $N_t - N_{t-} \ge 2$, then $T_{N_t} = T_{N_t - 1}$. Consequently $\mathbb{P}$-a.s.

$$\{(T_n)_{n\ge0} \text{ is increasing }\} \subset \{\forall t > 0 : N_t - N_{t-} \in \{0,1\}\} .$$

Finally we prove that 2 implies (b). If there exists $n \ge 0$ such that $T_n = T_{n+1}$, then $N_{T_n} = N_{T_{n+1}} \ge n+1$. If $n = 0$, this implies that $N_0 \ge 1$ but this holds with null probability since N is a counting process. If $n \ge 1$, we find $N_t - N_{t-} \ge 2$ for $t = T_n$ but this holds with null probability by 2. □

Definition 1.1.8 (Stationarity and independence). *A counting process N has independent increments if for all $n \geq 1$ and all $0 = t_0 < t_1 < \ldots < t_n$, $(N_{t_i} - N_{t_{i-1}}, i = 1, \ldots, n)$ is a vector of independent random variables.*
A counting process N has stationary increments if for all $0 \leq s < t$, $N_t - N_s$ and N_{t-s} have the same law.

Definition 1.1.9. *Given a counting process N and its sequence of jump times $(T_n)_{n\geq 0}$, we call*

$$\delta_n := T_n - T_{n-1} \ , \quad n \geq 1 \ ,$$

the inter-arrival times of N.

Notice that $(\delta_n)_{n\geq 1}$ is not necessarily a sequence of i.i.d. r.v.

Definition 1.1.10 (Renewal process). *Let $(\delta_n, n \geq 1)$ be a sequence of i.i.d. random variables taking values in $(0, \infty)$. Set $T_0 = 0$ and $T_n = \delta_1 + \ldots + \delta_n$ for all $n \geq 0$. Then, $(T_n)_{n\geq 0}$ is called a renewal process.*

1.2 Some classical probability distributions

Recall that the exponential distribution $\mathcal{E}(\lambda)$ with parameter $\lambda > 0$ is the probability distribution on $[0, \infty)$ associated to the density $\lambda e^{-\lambda x}\mathbf{1}_{x>0}$. If X has law $\mathcal{E}(\lambda)$, then

$$\mathbb{E}X = \frac{1}{\lambda} \ , \quad \mathbb{V}X = \frac{1}{\lambda^2} \ .$$

Its characteristic function is given by:

$$\mathbb{E}[e^{itX}] = \frac{1}{(1 - \frac{it}{\lambda})} \ , \quad t \in \mathbf{R} \ .$$

The exponential distribution is memoryless: the law of $X - t$ given $X > t$ is $\mathcal{E}(\lambda)$. More precisely, we have the following.

Lemma 1.2.1. *Let X be distributed as $\mathcal{E}(\lambda)$. For all $t, s > 0$:*

$$\mathbb{P}(X > t + s \,|\, X > t) = \mathbb{P}(X > s) \ .$$

Conversely, let X be a (strictly) positive random variable. If for all $s, t > 0$ we have

$$\mathbb{P}(X > t + s \,|\, X > t) = \mathbb{P}(X > s) \ ,$$

then there exists $\lambda > 0$ such that X has law $\mathcal{E}(\lambda)$.

Proof. The first part of the proof comes from a simple computation based on the identity $\mathbb{P}(X > t) = e^{-\lambda t}$ for all $t \geq 0$:

$$\mathbb{P}(X > t + s \,|\, X > t) = \frac{\mathbb{P}(X > t + s; X > t)}{\mathbb{P}(X > t)} = \frac{\mathbb{P}(X > t + s)}{\mathbb{P}(X > t)} = e^{-\lambda(t+s)+\lambda t} = \mathbb{P}(X > s) \ .$$

We turn to the converse property. First assume that $\mathbb{P}(X > t) > 0$ for all $t \geq 0$ and set $\varphi(t) = \mathbb{P}(X > t)$. We have for all $t, s > 0$

$$\begin{aligned}\varphi(t+s) &= \mathbb{P}(X > t+s) = \mathbb{P}(X > t+s\ ;\ X > t) = \mathbb{P}(X > t+s \,|\, X > t)\mathbb{P}(X > t)\\ &= \mathbb{P}(X > s)\mathbb{P}(X > t) = \varphi(t)\varphi(s)\ .\end{aligned}$$

Consequently, a simple recursion shows that for all $k, p \in \mathbf{N}^*$ we have

$$\varphi(k) = \varphi(1)^k\ , \quad \varphi(k) = \varphi(k/p)^p\ ,$$

and therefore

$$\varphi(k/p) = \varphi(1)^{k/p}\ .$$

Notice that $0 < \varphi(1) \leq 1$ so that we can set $\lambda = -\log\varphi(1) \geq 0$. We have shown that $\varphi(t) = e^{-\lambda t}$ for all $t \in \mathbf{Q}_+^*$. From the right continuity of distribution functions, we deduce that $t \mapsto \mathbb{P}(X > t) = 1 - \mathbb{P}(X \leq t)$ is right continuous itself. The density of $\mathbf{Q}_+^*$ in $\mathbf{R}_+$ ensures that $\varphi(t) = e^{-\lambda t}$ holds for all $t \geq 0$. To show that λ is actually strictly positive, it suffices to observe that, if it were equal to 0 then $\varphi(t) = 1$ for all $t \geq 0$ and so $\mathbb{P}(X = \infty) = 1$ thus raising a contradiction.
Let us now show that $\mathbb{P}(X > t) > 0$ for all $t \geq 0$. Since X is a positive random variable, there exists $\epsilon > 0$ such that $\mathbb{P}(X > \epsilon) > 0$. Thus, $\mathbb{P}(X > 2\epsilon) = \mathbb{P}(X > 2\epsilon \,|\, X > \epsilon)\mathbb{P}(X > \epsilon) = \mathbb{P}(X > \epsilon)^2$. A simple recursion yields $\mathbb{P}(X > n\epsilon) = \mathbb{P}(X > \epsilon)^n$ for all $n \in \mathbf{N}$. Consequently, $\mathbb{P}(X > t) > 0$ for all $t > 0$. □

Lemma 1.2.2. *Let $X_1, \ldots, X_n$ be n independent random variables with law $\mathcal{E}(\lambda_i)$, $\lambda_i > 0$. If we set $M_n = \min(X_1, \ldots, X_n)$ then*

$$M_n \overset{\text{(law)}}{=} \mathcal{E}(\lambda_1 + \ldots + \lambda_n)\ .$$

Proof. For all $t \geq 0$, using the independence of the X_i's we get:

$$\mathbb{P}(M_n > t) = \mathbb{P}(\cap_{i=1}^n \{X_i > t\}) = \prod_{i=1}^n \mathbb{P}(X_i > t) = \prod_{i=1}^n e^{-\lambda_i t} = e^{-\sum_{i=1}^n \lambda_i t}\ ,$$

thus identifying completely the law of M_n. □

The Gamma distribution $\Gamma(\alpha, \beta)$ with parameter $\alpha, \beta > 0$ is the probability distribution on $[0, \infty)$ associated to the density

$$x^{\alpha-1}\frac{\beta^\alpha}{\Gamma(\alpha)}e^{-\beta x}\mathbf{1}_{x>0}\ ,$$

where $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1}e^{-x}dx$. Recall that $\Gamma(n) = (n-1)!$ for all $n \in \mathbf{N}^*$. If X has law $\Gamma(\alpha, \beta)$ then

$$\mathbb{E}X = \frac{\alpha}{\beta}\ , \quad \mathbb{V}X = \frac{\alpha}{\beta^2}\ .$$

Its characteristic function is given by:

$$\mathbb{E}[e^{itX}] = \frac{1}{(1 - \frac{it}{\beta})^\alpha}\ , \quad t \in \mathbf{R}\ .$$

It turns out that the distribution $\Gamma(1, \lambda)$ coincides with the distribution $\mathcal{E}(\lambda)$. Furthermore, if $X_1, \ldots, X_n$ are independent r.v. with distribution $\Gamma(\alpha_i, \beta)$ for $\alpha_1, \ldots, \alpha_n > 0$, then their sum $X_1 + \ldots + X_n$ has law $\Gamma(\alpha_1 + \ldots + \alpha_n, \beta)$.

A random variable X has a Poisson distribution $\mathcal{P}(\mu)$ with parameter $\mu > 0$ if for all $n \in \mathbf{N}$, we have $\mathbb{P}(X = n) = e^{-\mu}\frac{\mu^n}{n!}$. Then,

$$\mathbb{E}X = \mu \ , \quad \mathbb{V}X = \mu \ .$$

Its characteristic function is given by:

$$\mathbb{E}[e^{itX}] = e^{\lambda(e^{it}-1)} \ , \quad t \in \mathbf{R} \ .$$

1.3 The Poisson process

Definition 1.3.1 (First definition of the Poisson process)**.** *Let $\lambda > 0$. A Poisson process with intensity λ is a counting process with stationary and independent increments such that*

$$N_t - N_s \overset{\text{(law)}}{=} \mathcal{P}(\lambda(t-s)) \ , \quad \forall t \geq s \geq 0 \ .$$

The existence of this process will be proven later on. Let us check that this definition characterises the law of a unique process. To that end, it suffices to show that the definition characterises completely the law of the vector $(N_{t_1}, \ldots, N_{t_n})$ for any $0 = t_0 < t_1 < \ldots < t_n$ and any $n \geq 1$. But this vector is a linear transformation of the vector $(N_{t_1}, N_{t_2} - N_{t_1}, \ldots, N_{t_n} - N_{t_{n-1}})$ which is distributed as n independent Poisson random variables with parameters $t_i - t_{i-1}$.
Notice that if N is a Poisson process, then $N_0 = 0$ a.s. and therefore $N_t \overset{\text{(law)}}{=} \mathcal{P}(\lambda t)$ for all $t \geq 0$. Furthermore, we have

$$\mathbb{E}[N_t - N_s] = \lambda(t-s) \ , \quad \mathbb{V}[N_t - N_s] = \lambda(t-s) \ .$$

Theorem 1. *A counting process N is a Poisson process of intensity λ if and only if its sequence of inter-arrival times $(\delta_n)_{n\geq 1}$ is i.i.d. with law $\mathcal{E}(\lambda)$.*

This result yields a second, equivalent definition of the Poisson process.

Definition 1.3.2 (Second definition of the Poisson process)**.** *Let $\lambda > 0$. A Poisson process with intensity λ is a counting process whose sequence of inter-arrival times is i.i.d. with law $\mathcal{E}(\lambda)$.*

Then, the existence of the Poisson process is immediate as it can be constructed from a sequence of i.i.d. random variables with law $\mathcal{E}(\lambda)$. This second definition ensures that the law of T_n is $\Gamma(n, \lambda)$ since this is the sum of n independent $\mathcal{E}(\lambda)$ r.v.

Proof. By Proposition 1.1.4, the law of a counting process is completely characterised by the law of its sequence of jump times. Consequently, it suffices to check that the law of $(T_1, \ldots, T_n)$ is the same in both cases.

Case 1: The law of $(T_1, \ldots, T_n)$ when N is a Poisson process. Suppose that N is a Poisson process of intensity λ. Fix $n \geq 1$ and $0 < s_1 < \ldots < s_n$. For any $h_1, \ldots, h_n > 0$ such that $s_1 + h_1 < s_2 < s_2 + h_2 < \ldots < s_n < s_n + h_n$ we use the independence and stationarity of the increments of N to derive the following identities:

$$
\begin{aligned}
&\mathbb{P}\big(\cap_{i=1}^n \{T_i \in (s_i, s_i + h_i]\}\big) \\
&= \mathbb{P}\big(N_{s_1} = 0, N_{s_1+h_1} - N_{s_1} = 1, N_{s_2} - N_{s_1+h_1} = 0, \dots \\
&\qquad , N_{s_n} - N_{s_{n-1}+h_{n_1}} = 0, N_{s_n+h_n} - N_{s_n} \geq 1\big) \\
&= \mathbb{P}\big(N_{s_1} = 0\big)\mathbb{P}\big(N_{s_1+h_1} - N_{s_1} = 1\big)\mathbb{P}\big(N_{s_2} - N_{s_1+h_1} = 0\big) \dots \\
&\qquad \mathbb{P}\big(N_{s_n} - N_{s_{n-1}+h_{n_1}} = 0\big)\mathbb{P}\big(N_{s_n+h_n} - N_{s_n} \geq 1\big) \\
&= \mathbb{P}\big(N_{s_1} = 0\big)\mathbb{P}\big(N_{s_n+h_n} - N_{s_n} \geq 1\big) \prod_{i=1}^{n-1} \mathbb{P}\big(N_{s_i+h_i} - N_{s_i} = 1\big)\mathbb{P}\big(N_{s_{i+1}} - N_{s_i+h_i} = 0\big) \\
&= e^{-\lambda s_1}(1 - e^{-\lambda h_n}) \prod_{i=1}^{n-1} e^{-\lambda h_i} \lambda h_i e^{-\lambda(s_{i+1} - s_i - h_i)} \ .
\end{aligned}
$$

This computation suggests to define for all $0 < s_1 < \ldots < s_n$:

$$
f_n(s_1, \ldots, s_n) = \lim_{h_1,\ldots,h_n \downarrow 0} \frac{\mathbb{P}\big(\cap_{i=1}^n \{T_i \in (s_i, s_i + h_i]\}\big)}{\prod_{i=1}^n h_i} = \lambda^n e^{-\lambda s_n} \ ,
$$

and to set this function to 0 for all other values of $(s_1, \ldots, s_n)$. In other words

$$
f_n(s_1, \ldots, s_n) = \mathbf{1}_{\{0 < s_1 < \ldots < s_n\}} \lambda^n e^{-s_n} \ .
$$

Let us show recursively that

$$
\int_{\mathbf{R}^n} f_n(s_1, \ldots, s_n) \prod_{i=1}^n ds_i = 1 \ .
$$

For any $n \geq 2$, we have:

$$
\begin{aligned}
\int_{\mathbf{R}^n} f_n(s_1, \ldots, s_n) \prod_{i=1}^n ds_i &= \int_{0 < s_1 < \ldots < s_{n-1}} \lambda^{n-1} e^{-s_{n-1}} \int_{s_n \in (s_{n-1}, \infty)} \lambda e^{-(s_n - s_{n-1})} ds_n \prod_{i=1}^{n-1} ds_i \\
&= \int_{0 < s_1 < \ldots < s_{n-1}} \lambda^{n-1} e^{-s_{n-1}} \prod_{i=1}^{n-1} ds_i \\
&= \int_{\mathbf{R}^{n-1}} f_{n-1}(s_1, \ldots, s_{n-1}) \prod_{i=1}^{n-1} ds_i \ .
\end{aligned}
$$

Consequently, a simple recursion shows that this last expression equals

$$
\int_{(0,\infty)} \lambda e^{-\lambda s} ds = 1 \ .
$$

Hence f_n is the density of a probability distribution on $\mathbf{R}^n$ supported in $\{0 < s_1 < \ldots < s_n\}$. Moreover, for all s_i, h_i as above we have

$$
\mathbb{P}\big(\cap_{i=1}^n \{T_i \in (s_i, s_i + h_i]\}\big) = \int_{\prod_{i=1}^n (s_i, s_i + h_i]} f_n(t_1, \ldots, t_n) \prod_{i=1}^n dt_i \ .
$$

Consequently, $(T_1, \ldots, T_n)$ admits a density on $\mathbf{R}^n$ given by the function f_n.
Case 2: The law of $(T_1, \ldots, T_n)$ *when* N *is built from a sequence of i.i.d.* $\mathcal{E}(\lambda)$. Assume that $(\delta_n)_{n\geq 1}$ is a sequence of i.i.d. r.v. with law $\mathcal{E}(\lambda)$ and set $T_n = \sum_{i=1}^n \delta_i$ for all $n \geq 1$. Observe that $\delta_i = T_i - T_{i-1}$ for every $i \geq 1$.
Fix $n \geq 1$. If we denote by δ the vector $(\delta_1, \ldots, \delta_n)$ and similarly, we denote by T the vector $(T_1, \ldots, T_n)$ then $\delta = AT$ where

$$A = \begin{pmatrix} 1 & 0 & \ldots & 0 & 0 \\ -1 & 1 & \ldots & 0 & 0 \\ 0 & -1 & \ldots & 0 & 0 \\ \vdots & \vdots & \ldots & 0 & 0 \\ 0 & 0 & \ldots & 1 & 0 \\ 0 & 0 & \ldots & -1 & 1 \end{pmatrix}$$

A is an invertible matrix so that the map $t \mapsto At$ is a diffeomorphism from $\mathbf{R}^n$ into itself, and its determinant is equal to 1. Let $\varphi : \mathbf{R}^n \to \mathbf{R}^n$ be a bounded measurable map, and denote by $g(s_1, \ldots, s_n)$ the density of the vector δ:

$$g_n(s_1, \ldots, s_n) = \prod_{i=1}^n \lambda e^{-\lambda s_i} \mathbf{1}_{s_i > 0} \ .$$

By the change-of-variable formula and the fact that $\sum_{i=1}^n (At)_i = t_n$, we find

$$\begin{aligned} \mathbb{E}[\varphi(T_1, \ldots, T_n)] = \mathbb{E}[\varphi(A^{-1}\delta)] &= \int_{s\in\mathbf{R}^n} \varphi(A^{-1}s) g_n(s_1, \ldots, s_n) \prod_{i=1}^n ds_i \\ &= \int_{t\in\mathbf{R}^n} \varphi(t) g_n(At) |\det A| \prod_{i=1}^n dt_i \\ &= \int_{t\in\mathbf{R}^n} \varphi(t) \lambda^n e^{-\lambda t_n} \prod_{i=1}^n \mathbf{1}_{\{0<t_1<\ldots<t_n\}} dt_i \ , \end{aligned}$$

thus identifying the density of $(T_1, \ldots, T_n)$ as being given by f_n. □

1.4 A reminder on conditional expectations

Let $\mathcal{B}$ be a sigma-field included in $\mathcal{F}$. Let X be a real-valued random variable in $L^1(\Omega, \mathbb{P})$.

Definition 1.4.1. *The conditional expectation* $\mathbb{E}[X|\mathcal{B}]$ *of* X *given* $\mathcal{B}$ *is the unique (up to almost sure equivalence)* $\mathcal{B}$*-measurable random variable that satisfies for all* $B \in \mathcal{B}$

$$\mathbb{E}[X\mathbf{1}_B] = \mathbb{E}[\mathbb{E}[X|\mathcal{B}]\mathbf{1}_B] \ .$$

Theorem 2. *Let* $\mathcal{B}$ *be a sigma-field, and let* X, Y *be two real valued random variables, such that* X *is* $\mathcal{B}$*-measurable and* Y *is independent from* $\mathcal{B}$*. Then for any bounded measurable map* $f : \mathbf{R}^2 \to \mathbf{R}$*, we have the following almost sure identity:*

$$\mathbb{E}\big[f(X,Y)\,|\,\mathcal{B}\big] = F(X) \ ,$$

where

$$F(x) := \mathbb{E}[f(x,Y)] \ , \quad x \in \mathbf{R} \ .$$

Proof. Fix $B \in \mathcal{B}$. We aim at showing that

$$\mathbb{E}[f(X,Y)\mathbf{1}_B] = \mathbb{E}[F(X)\mathbf{1}_B] \ .$$

It suffices to show that it holds true for all maps f of the form $f = \mathbf{1}_A$ with $A \in \mathcal{B}(\mathbf{R}^2)$. Indeed, by linearity of this identity, we then deduce that it holds true for all linear combination of such indicators, and by the Dominated Convergence Theorem for any map f which is non-negative. The general case is then obtained by splitting f into its positive and negative parts.
Let us introduce the class $\mathcal{M} := \{A \in \mathcal{B}(\mathbf{R}^2) : \mathbb{E}[\mathbf{1}_A(X,Y)\mathbf{1}_B] = \mathbb{E}[F_A(X)\mathbf{1}_B]\}$ where $F_A(x) = \mathbb{E}[\mathbf{1}_A(x,Y)]$. If $A \subset A'$ are in $\mathcal{M}$, then by linearity $A' \backslash A$ is in $\mathcal{M}$ as well. Furthermore, $\mathcal{M}$ is stable under increasing limit. Finally, Ω belongs to $\mathcal{M}$ since $F_\Omega = 1$. Hence $\mathcal{M}$ is a monotone class that contains Ω. Now if we let $\mathcal{C}$ be the class of all product sets $A_1 \times A_2$ with $A_1, A_2 \in \mathcal{B}(\mathbf{R})$, then $\mathcal{C}$ is stable under finite intersection and belongs to $\mathcal{M}$. By the Monotone Class Theorem, we deduce that $\sigma(\mathcal{C})$ also belongs to $\mathcal{M}$: since $\sigma(\mathcal{C}) = \mathcal{B}(\mathbf{R}^2)$ this concludes the proof. □

1.5 Some properties of the Poisson process

From now on, N is a Poisson process of intensity $\lambda > 0$.

Proposition 1.5.1 (Law of large numbers)**.** *As $t \to \infty$, we have the following almost sure convergence:*

$$\frac{N_t}{t} \to \lambda \ .$$

Proof. For any $n \geq 1$, observe that $N_n = \sum_{i=1}^n (N_i - N_{i-1})$. Since the increments of a Poisson process are stationary and independents, we can apply the strong law of large numbers to the sequence $(N_i - N_{i-1})_{i \geq 1}$ and get the following almost sure convergence as $n \to \infty$:

$$\frac{N_n}{n} \to \mathbb{E}[N_1 - N_0] = \mathbb{E}[N_1] = \lambda \ .$$

Let us now write $n_t = \lfloor t \rfloor$ for the integer part of $t \geq 0$. Since N has non-decreasing paths, we have for all $t \geq 1$:

$$\frac{n_t}{n_t + 1}\frac{N_{n_t}}{n_t} = \frac{N_{n_t}}{n_t+1} \leq \frac{N_t}{t} \leq \frac{N_{n_t+1}}{n_t} = \frac{n_t+1}{n_t}\frac{N_{n_t+1}}{n_t+1} \ .$$

Observe that $n_t/(n_t+1) \to 1$ as $t \to \infty$. Consequently, the almost sure convergence $N_n/n \to \lambda$ ensures that the leftmost and rightmost terms of the previous inequalities converge to λ almost surely, thus yielding the asserted result. □

We introduce the filtration generated by the process N:

$$\mathcal{F}_t := \sigma(N_s, s \leq t) \ , \quad t \geq 0 \ .$$

Lemma 1.5.2. *For all $t, s \geq 0$, the random variable $N_{t+s} - N_t$ is independent of $\mathcal{F}_t$.*

Proof. Let $\mathcal{M}$ be the collection of all events $E \in \mathcal{F}_t$ such that E is independent from $N_{t+s} - N_t$. It is simple to check that $\Omega \in \mathcal{M}$. Furthermore, if $E \subset E'$ are in $\mathcal{M}$, then for all $f : \mathbf{R} \to \mathbf{R}$ measurable and bounded we have

$$\begin{aligned}\mathbb{E}[\mathbf{1}_{E'\setminus E} f(N_{t+s}-N_t)] &= \mathbb{E}[(\mathbf{1}_{E'}-\mathbf{1}_E) f(N_{t+s}-N_t)] \\ &= \mathbb{E}[\mathbf{1}_{E'} f(N_{t+s}-N_t)] - \mathbb{E}[\mathbf{1}_E f(N_{t+s}-N_t)] \\ &= \mathbb{P}(E')\mathbb{E}[f(N_{t+s}-N_t)] - \mathbb{P}(E)\mathbb{E}[f(N_{t+s}-N_t)] \\ &= \mathbb{P}(E'\setminus E)\mathbb{E}[f(N_{t+s}-N_t)] \ ,\end{aligned}$$

so that $E'\setminus E$ is in $\mathcal{M}$ as well. Similarly, one can check that if E_n is an increasing sequence of events in $\mathcal{M}$, then $\cup_n E_n$ is also in $\mathcal{M}$. Consequently, $\mathcal{M}$ is a monotone class that contains Ω. Let us introduce $\mathcal{C}$ as the collection of all events of the form

$$\{N_{s_1}\in A_1,\ldots,N_{s_n}\in A_n\} \ ,$$

for $0\le s_1<\ldots<s_n\le t$, $A_1,\ldots,A_n$ some Borel sets of **R** and $n\ge 1$. Notice that $\mathcal{C}$ is stable under finite intersections. From the independence of the increments of a Poisson process, we deduce that $\mathcal{C}$ is a subset of $\mathcal{M}$. The Monotone Class Theorem yields that $\sigma(\mathcal{C})\subset\mathcal{M}$. By definition, $\mathcal{F}_t$ coincides with $\sigma(\mathcal{C})$ so this concludes the proof. □

Proposition 1.5.3. *If N, M are two independent Poisson processes of intensity λ,μ then their sum is a Poisson process of intensity $\lambda+\mu$.*

Proof. Left as an exercise. □

Our next proposition describes the law of the jump times $(T_1,\ldots,T_n)$ conditionally given $N_t = n$. In order to state the result, we need to introduce the following.

Definition 1.5.4 (Order statistics)**.** *Let $(X_i, i=1,\ldots,n)$ be a collection of real valued random variables which are almost surely pairwise distinct. We define $X_{(1)} := \inf\{X_i, i\in\{1,\ldots,n\}\}$ and recursively*

$$X_{(i)} = \inf\{X_j : X_j > X_{(i-1)}\} \ .$$

The random variables $(X_{(i)}, i=1,\ldots,n)$ are called the order statistics of $(X_i, i=1,\ldots,n)$.

Notice that this definition makes sense almost surely since the r.v. $(X_i, i=1,\ldots,n)$ are pairwise distinct with probability one. Notice that if the X_i are independent and have a density, then they are pairwise distinct with probability one. We now state a simple property of order statistics.

Lemma 1.5.5. *If $(X_i, i=1,\ldots,n)$ is a sequence of i.i.d., real valued r.v. with common density f, then the vector $(X_{(1)},\ldots,X_{(n)})$ admits the following density:*

$$f_n(y_1,\ldots,y_n) = n!\prod_{i=1}^n f(y_i)\mathbf{1}_{y_1<\ldots<y_n} \ .$$

Proof. Let Σ_n be the set of all permutations of $\{1,\ldots,n\}$. We have for every measurable and bounded map $g:\mathbf{R}^n\to\mathbf{R}$

$$\mathbb{E}[g(X_{(1)},\ldots,X_{(n)})] = \sum_{\sigma\in\Sigma_n}\mathbb{E}\Big[g(X_{\sigma(1)},\ldots,X_{\sigma(n)})\mathbf{1}_{\{X_{\sigma(1)}<\ldots<X_{\sigma(n)}\}}\Big] \ .$$

We define $D := \{x \in \mathbf{R}^n : x_1 < \ldots < x_n\}$. For every permutation σ of $\{1,\ldots,n\}$, we set $D_\sigma := \{x \in \mathbf{R}^n : x_{\sigma(1)} < \ldots < x_{\sigma(n)}\}$. The mapping $\varphi_\sigma : x \in D_\sigma \mapsto (x_{\sigma(1)},\ldots,x_{\sigma(n)}) \in D$ is a diffeomorphism. Consequently, by the change of variables formula we get

$$\begin{aligned}\mathbb{E}\Big[g(X_{\sigma(1)},\ldots,X_{\sigma(n)})\mathbf{1}_{\{X_{\sigma(1)}<\ldots<X_{\sigma(n)}\}}\Big] &= \int_{x\in D_\sigma} g(x_{\sigma(1)},\ldots,x_{\sigma(n)})f(x_1)\ldots f(x_n)dx_1\ldots dx_n\\ &= \int_{y\in D} g(y_1,\ldots,y_n)\big|J\varphi_\sigma^{-1}\big|f(y_1)\ldots f(y_n)dy_1\ldots dy_n\ .\end{aligned}$$

Since $\big|J\varphi_\sigma^{-1}\big| = 1$, this last expression does not depend on σ. Since Σ_n contains $n!$ elements, we deduce that

$$\mathbb{E}[g(X_{(1)},\ldots,X_{(n)})] = n!\int_{y\in D} g(y_1,\ldots,y_n)f(y_1)\ldots f(y_n)dy_1\ldots dy_n\ ,$$

thus identifying uniquely the density of the distribution of $(X_{(1)},\ldots,X_{(n)})$. □

As usual, we let N be a Poisson process of intensity λ and we denote by $T_1, T_2, \ldots$ its jump times.

Theorem 3. *Fix $t > 0$ and $n \in \mathbf{N}$. Let $U_i, i = 1\ldots n$ be a sequence of i.i.d. uniform r.v. on $[0,t]$. Then, the law of $(T_1,\ldots,T_n)$ given $\{N_t = n\}$ coincides with the law of $(U_{(1)},\ldots,U_{(n)})$, the order statistics of $(U_1,\ldots,U_n)$.*

Proof. Let f be a measurable and bounded map from $\mathbf{R}^n$ into $\mathbf{R}$. We have

$$\begin{aligned}&\mathbb{E}[f(T_1,\ldots,T_n)\mathbf{1}_{N_t=n}]\\ &= \mathbb{E}\Big[f(\delta_1,\delta_1+\delta_2,\ldots,\delta_1+\ldots+\delta_n)\mathbf{1}_{\{\delta_1+\ldots+\delta_n\le t<\delta_1+\ldots+\delta_{n+1}\}}\Big]\\ &= \int_{s_1,\ldots,s_{n+1}\ge 0} f(s_1,\ldots,s_1+\ldots s_n)\lambda^{n+1}\mathbf{1}_{\{s_1+\ldots+s_n\le t<s_1+\ldots+s_{n+1}\}}\prod_{i=1}^{n+1} e^{-\lambda s_i}ds_i\\ &= \int_{s_1,\ldots,s_n\ge 0} f(s_1,\ldots,s_1+\ldots s_n)\mathbf{1}_{\{s_1+\ldots+s_n\le t\}}\prod_{i=1}^{n} e^{-\lambda s_i}ds_i\lambda^n e^{-\lambda(t-s_1-s_2-\ldots-s_n)}\\ &= \int_{s_1,\ldots,s_n\ge 0} f(s_1,\ldots,s_1+\ldots s_n)\mathbf{1}_{\{s_1+\ldots+s_n\le t\}}\prod_{i=1}^{n} ds_i\lambda^n e^{-\lambda t}\\ &= \int_{t_1,\ldots,t_n\ge 0} f(t_1,\ldots,t_n)\mathbf{1}_{\{t_1\le t_2\le\ldots\le t_n\le t\}}\prod_{i=1}^{n} dt_i\lambda^n e^{-\lambda t}\ .\end{aligned}$$

Since in addition $\mathbb{P}(N_t = n) = e^{-\lambda t}(\lambda t)^n/n!$ we obtain

$$\begin{aligned}&\mathbb{E}[f(T_1,\ldots,T_n)\,\big|\,N_t = n]\\ &= \frac{\mathbb{E}[f(T_1,\ldots,T_n)\mathbf{1}_{N_t=n}]}{\mathbb{P}(N_t=n)}\\ &= n!\int_{t_1,\ldots,t_n\ge 0} f(t_1,\ldots,t_n)\mathbf{1}_{\{0<t_1<\ldots<t_n\le t\}}t^{-n}dt_1\ldots dt_n\ ,\end{aligned}$$

We deduce that the law of $(T_1,\ldots,T_n)$ given $\{N_t = n\}$ admits the following density: $n!\mathbf{1}_{\{0<t_1<\ldots<t_n\le t\}}t^{-n}$. By Lemma 1.5.5, we recognise the density of the order statistics of n i.i.d. uniform r.v. on $[0,t]$. □

1.6 The mixed Poisson process

Definition 1.6.1. *Let N be a Poisson process of intensity 1, and let Θ be a positive random variable, independent of N. The process*

$$\tilde{N}_t := N_{\Theta t} \ , \quad t \geq 0 \ ,$$

is called a mixed Poisson process with random intensity Θ.

Remark 1.6.2. *It is not completely immediate to check that the marginals of $\tilde{N}$ are well-defined random variables.*

To make computations on the law of a mixed Poisson process, one can condition first on its random intensity. Indeed, conditionally given Θ, $\tilde{N}$ is a Poisson process of intensity Θ as the following result shows.

Proposition 1.6.3. *Conditionally given Θ, $\tilde{N}$ has the law of a Poisson process of intensity Θ.*

Proof. First of all, $t \mapsto \tilde{N}_t$ is an **N**-valued, non-decreasing, right-continuous process starting from 0: it is therefore a counting process. It makes jumps of size 1 a.s. and its inter-arrival times $\tilde{\delta}_n$ satisfy:

$$\tilde{\delta}_n = \delta_n / \Theta \ , \quad n \geq 1 \ .$$

Consequently, for all bounded measurable maps $f_1, \dots, f_n$ we have by Theorem 2

$$\mathbb{E}[f_1(\tilde{\delta}_1) \dots f_n(\tilde{\delta}_n) \,|\, \Theta] = \mathbb{E}[f_1(\delta_1/\Theta) \dots f_n(\delta_n/\Theta) \,|\, \Theta] = F(\Theta) \ ,$$

where

$$F(x) = \mathbb{E}[f_1(\delta_1/x) \dots f_n(\delta_n/x)] \ .$$

Since for all $x > 0$, $(\delta_1/x, \dots, \delta_n/x)$ has the law of n independent $\mathcal{E}(x)$ r.v., we deduce that the conditional law of $(\tilde{\delta}_1, \dots, \tilde{\delta}_n)$ given Θ is the law of n independent $\mathcal{E}(\Theta)$ r.v. This yields the statement of the proposition. □

Lemma 1.6.4. *Let $\tilde{N}$ be a mixed Poisson process with random intensity Θ. For all $t > 0$, we have*

$$\mathbb{E}[\tilde{N}_t] = \mathbb{E}[\Theta] t \ , \quad \textit{Var}\ [\tilde{N}_t] = \mathbb{E}[\Theta] t + \textit{Var}\ [\Theta] t^2 \ .$$

Proof. Since $\tilde{N}_t$, conditionally given Θ, is a Poisson r.v. with parameter Θt, we get

$$\mathbb{E}[\tilde{N}_t] = \mathbb{E}[\mathbb{E}[\tilde{N}_t \,|\, \Theta]] = \mathbb{E}[\Theta t] \ ,$$

as well as

$$\mathbb{E}[\tilde{N}_t^2] = \mathbb{E}[\mathbb{E}[\tilde{N}_t^2 \,|\, \Theta]] = \mathbb{E}[(\Theta t)^2 + \Theta t] \ .$$

Consequently,

$$\text{Var}\ [\tilde{N}_t] = \mathbb{E}[\tilde{N}_t^2] - \mathbb{E}[\tilde{N}_t]^2 = \mathbb{E}[(\Theta t)^2 + \Theta t] - \mathbb{E}[\Theta t]^2 = \text{Var}\ [\Theta] t^2 + \mathbb{E}[\Theta] t \ .$$

□

Proposition 1.6.5. *Let $\tilde{N}$ be a mixed Poisson process with random intensity Θ. Then:*

(i) $\tilde{N}_t$ does not follow a Poisson distribution, except if Θ is deterministic,

(ii) $\tilde{N}$ *does not have independent increments, except if* Θ *is deterministic,*

(iii) $\tilde{N}$ *has stationary increments.*

Proof. Recall that the expectation and variance of a Poisson r.v. coincide. By the previous lemma, the expectation and variance of $\tilde{N}_t$ coincide if and only if Var $\Theta = 0$. But

$$\text{Var } \Theta = 0 \Leftrightarrow \mathbb{E}[(\Theta - \mathbb{E}\Theta)^2] = 0 \Leftrightarrow \Theta - \mathbb{E}\Theta = 0 \text{ a.s. },$$

which is equivalent to saying that Θ is deterministic. Consequently, if Θ is not deterministic then $\tilde{N}_t$ does not follow a Poisson distribution. On the other hand, if Θ is deterministic, then $\tilde{N}_t$ is a Poisson random variable with intensity Θ. This proves (i).
For all $t > s \geq 0$, we have

$$\mathbb{E}[\tilde{N}_s(\tilde{N}_t - \tilde{N}_s)] = \mathbb{E}[\mathbb{E}[\tilde{N}_s(\tilde{N}_t - \tilde{N}_s) \,|\, \Theta]] = \mathbb{E}[\Theta s \Theta(t-s)] = \mathbb{E}[\Theta^2] s(t-s) \;,$$

while

$$\mathbb{E}[\tilde{N}_s]\mathbb{E}[\tilde{N}_t - \tilde{N}_s] = \mathbb{E}[\Theta]^2 s(t-s) \;.$$

Consequently, as soon as $\mathbb{E}[\Theta^2] \neq \mathbb{E}[\Theta]^2$, the random variables $\tilde{N}_s$ and $\tilde{N}_t - \tilde{N}_s$ are not independent so that the increment of $\tilde{N}$ are not independent. Notice that $\mathbb{E}[\Theta^2] \neq \mathbb{E}[\Theta]^2$ is equivalent with Var $\Theta \neq 0$, which is itself equivalent with Θ is not deterministic. On the other hand, if Θ is deterministic, then $\tilde{N}$ is a Poisson process of intensity Θ and its increments are independent. This proves (ii).
To prove the stationarity, we observe that for all $y \in \mathbf{R}$ and all $t \geq s \geq 0$ we have

$$\mathbb{E}[e^{iy(\tilde{N}_t - \tilde{N}_s)}] = \mathbb{E}[\mathbb{E}[e^{iy(\tilde{N}_t - \tilde{N}_s)} \,|\, \Theta]] = \mathbb{E}[\mathbb{E}[e^{iy\tilde{N}_{t-s}} \,|\, \Theta]] = \mathbb{E}[e^{iy\tilde{N}_{t-s}}] \;.$$

This proves (iii). □

Proposition 1.6.6. *Fix* $t > 0$ *and* $n \in \mathbf{N}$*. Let* $0 < \tilde{T}_1 < \ldots < \tilde{T}_n < \ldots$ *be the jump times of a mixed Poisson process* $\tilde{N}$ *of random intensity* Θ*. The law of* $(\tilde{T}_1, \ldots, \tilde{T}_n)$ *conditionally given* $\tilde{N}_t = n$ *is independent of* Θ *and coincides with the law of* $(U_{(1)}, \ldots, U_{(n)})$*, the order statistics of* $(U_1, \ldots, U_n)$*, taken to be* n *i.i.d. uniform r.v. over* $[0, t]$

Proof. Recall that $\tilde{N}$, conditionally given Θ, is a Poisson process of intensity Θ. Similarly as in the proof of Theorem 3, we find for every bounded and measurable function $g : \mathbf{R}^n \to \mathbf{R}$:

$$\begin{aligned}
\mathbb{E}[g(\tilde{T}_1, \ldots, \tilde{T}_n)\mathbf{1}_{\{\tilde{T}_n \leq t < \tilde{T}_{n+1}\}} \,|\, \Theta] &= \Theta^{n+1} \int_{0<t_1<\ldots<t_{n+1}} e^{-\Theta t_{n+1}} g(t_1, \ldots, t_n)\mathbf{1}_{\{t_n \leq t < t_{n+1}\}} dt_1 \ldots dt_{n+1} \\
&= \Theta^n e^{-\Theta t} \int_{0<t_1<\ldots<t_n} g(t_1, \ldots, t_n)\mathbf{1}_{\{t_n \leq t\}} dt_1 \ldots dt_n \\
&= \frac{(\Theta t)^n}{n!} e^{-\Theta t} \mathbb{E}\big[g(U_{(1)}, \ldots, U_{(n)})\big] \\
&= \mathbb{P}(\tilde{N}_t = n \,|\, \Theta)\mathbb{E}\big[g(U_{(1)}, \ldots, U_{(n)})\big] \;,
\end{aligned}$$

where the equality before the last comes from Lemma 1.5.5 applied to $(U_1, \ldots, U_n)$.

We now take $h : \mathbf{R} \to \mathbf{R}$ bounded measurable. We have

$$
\begin{aligned}
\mathbb{E}[g(\tilde{T}_1, \ldots, \tilde{T}_n)h(\Theta) \,|\, \tilde{N}_t = n] &= \frac{\mathbb{E}[g(\tilde{T}_1, \ldots, \tilde{T}_n)h(\Theta)\mathbf{1}_{\{\tilde{N}_t = n\}}]}{\mathbb{P}(\tilde{N}_t = n)} \\
&= \frac{\mathbb{E}[g(\tilde{T}_1, \ldots, \tilde{T}_n)h(\Theta)\mathbf{1}_{\{\tilde{T}_n \le t < \tilde{T}_{n+1}\}}]}{\mathbb{E}[\mathbf{1}_{\{\tilde{T}_n \le t < \tilde{T}_{n+1}\}}]} \\
&= \frac{\mathbb{E}[h(\Theta)\mathbb{E}[g(\tilde{T}_1, \ldots, \tilde{T}_n)\mathbf{1}_{\{\tilde{T}_n \le t < \tilde{T}_{n+1}\}} \,|\, \Theta]]}{\mathbb{E}[\mathbb{E}[\mathbf{1}_{\{\tilde{T}_n \le t < \tilde{T}_{n+1}\}} \,|\, \Theta]]}
\end{aligned}
$$

so that using the previous calculation, we get

$$
\begin{aligned}
&= \frac{\mathbb{E}[\frac{(\Theta t)^n}{n!}e^{-\Theta t}h(\Theta)]}{\mathbb{E}[\frac{(\Theta t)^n}{n!}e^{-\Theta t}]}\mathbb{E}\big[g(U_{(1)}, \ldots, U_{(n)})\big] \\
&= \frac{\mathbb{E}[\Theta^n e^{-\Theta t}h(\Theta)]}{\mathbb{E}[\Theta^n e^{-\Theta t}]}\mathbb{E}\big[g(U_{(1)}, \ldots, U_{(n)})\big]
\end{aligned}
$$

Taking $h = 1$, we deduce that

$$
\mathbb{E}[g(\tilde{T}_1, \ldots, \tilde{T}_n) \,|\, \tilde{N}_t = n] = \mathbb{E}\big[g(U_{(1)}, \ldots, U_{(n)})\big] \ ,
$$

and taking $g = 1$ we find

$$
\mathbb{E}[h(\Theta) \,|\, \tilde{N}_t = n] = \frac{\mathbb{E}[\Theta^n e^{-\Theta t}h(\Theta)]}{\mathbb{E}[\Theta^n e^{-\Theta t}]} \ .
$$

Therefore, for any g and h bounded and measurable, we have the identity

$$
\mathbb{E}[g(\tilde{T}_1, \ldots, \tilde{T}_n)h(\Theta) \,|\, \tilde{N}_t = n] = \mathbb{E}[g(\tilde{T}_1, \ldots, \tilde{T}_n) \,|\, \tilde{N}_t = n]\mathbb{E}[h(\Theta) \,|\, \tilde{N}_t = n] \ ,
$$

which proves the statement of the proposition. □

Corollary 1.6.7. *Let $n \ge 1$ and $0 \le t_1 \le \ldots \le t_n$. The law of $(\tilde{N}_{t_1}, \ldots, \tilde{N}_{t_n})$ conditionally given $\tilde{N}_{t_n} = m$ is independent of Θ and therefore coincides with the law of $(N_{t_1}, \ldots, N_{t_n})$ conditionally given $N_{t_n} = m$, where N is a Poisson process of arbitrary intensity $\lambda > 0$.*

Proof. For all $0 \le k_1 \le \ldots \le k_n = m$, we have

$$
\begin{aligned}
&\mathbb{P}(\tilde{N}_{t_1} = k_1, \ldots, \tilde{N}_{t_n} = k_n \,|\, \tilde{N}_{t_n} = m) \\
&= \mathbb{P}(\tilde{T}_{k_1} \le k_1 < \tilde{T}_{k_1+1} \ldots \tilde{T}_{k_2} \le t_2 < \ldots \tilde{T}_{k_{n-1}} \le t_{n-1} < \tilde{T}_{k_{n-1}+1} \,|\, \tilde{N}_{t_n} = m) \ .
\end{aligned}
$$

By the previous proposition, we know that the law of $(\tilde{T}_1, \ldots, \tilde{T}_{k_n})$ conditionally given $\tilde{N}_{t_n} = k_n$ is independent of Θ and coincides with the same quantity for a Poisson process of arbitrary intensity $\lambda > 0$. Consequently, the last quantity equals

$$
\begin{aligned}
&= \mathbb{P}(T_{k_1} \le t_1 < T_{k_1+1} \ldots T_{k_2} \le t_2 < \ldots T_{k_{n-1}} \le t_{n-1} < T_{k_{n-1}+1} \,|\, N_{t_n} = m) \\
&= \mathbb{P}(N_{t_1} = k_1, \ldots, N_{t_n} = k_n \,|\, N_{t_n} = m) \ .
\end{aligned}
$$

□

1.7 The compound Poisson process

Definition 1.7.1. *The process $S_t = \sum_{i=1}^{N_t} X_i$, $t \geq 0$, is called a compound Poisson process if the sequence $(X_i)_{i\geq 1}$ is i.i.d. with values in $\mathbf{R}$, and N is a Poisson process independent of $(X_i)_{i\geq 1}$.*

As usual, we will denote by $\lambda > 0$ the intensity of the Poisson process N.

Proposition 1.7.2. *Let S be a compound Poisson process. Then:*

(i) S is almost surely càdlàg and has stationary and independent increments,

(ii) The distribution function of S_t satisfies for all $x \in \mathbf{R}$

$$\mathbb{P}(S_t \leq x) = \sum_{n\geq 0} \mathbb{P}(N_t = n)\mathbb{P}(X_1 + \ldots + X_n \leq x) \ ,$$

(iii) If X is integrable, then

$$\mathbb{E}[S_t] = \mathbb{E}[N_t]\mathbb{E}[X] \ ,$$

and if furthermore X has a finite second moment, then

$$\mathit{Var}\,[S_t] = \mathit{Var}\,[N_t]\mathbb{E}[X]^2 + \mathit{Var}\,[X]\mathbb{E}[N_t] \ .$$

(iv) For all $t \geq 0$ and all $q \in \mathbf{R}$ such that $\mathbb{E}[e^{qX}] < \infty$ we have

$$\mathbb{E}[e^{qS_t}] = \exp\left(\lambda t(\mathbb{E}[e^{qX}] - 1)\right) \ .$$

Proof. To prove that S is almost surely càdlàg, it suffices to use the fact that N is itself almost surely càdlàg so that almost surely for all $t \geq 0$:

$$\lim_{s\downarrow 0} S_{t+s} = \lim_{s\downarrow 0} \sum_{i=1}^{N_{t+s}} X_i = \sum_{i=1}^{N_t} X_i = S_t \ ,$$

and for all $t > 0$:

$$\lim_{s\downarrow 0} S_{t-s} = \lim_{s\downarrow 0} \sum_{i=1}^{N_{t-s}} X_i = \sum_{i=1}^{N_{t-}} X_i = S_{t-} \ .$$

Let us show that S has stationary and independent increments. Fix $n \geq 1$, and let $q_1, \ldots, q_n \in \mathbf{R}$ and $0 = t_0 < t_1 < \ldots < t_n$. Denote by $\Phi_X(q) = \mathbb{E}[e^{iqX}]$ the characteristic function of the law of X. We have

$$\begin{aligned}\mathbb{E}\Big[e^{i\sum_{k=1}^n q_k(S_{t_k} - S_{t_{k-1}})}\Big] &= \mathbb{E}\Big[\prod_{k=1}^n e^{iq_k \sum_{j=N_{t_{k-1}}+1}^{N_{t_k}} X_j}\Big] \\ &= \mathbb{E}\Big[\mathbb{E}\Big[\prod_{k=1}^n e^{iq_k \sum_{j=N_{t_{k-1}}+1}^{N_{t_k}} X_j} \,|\, N\Big]\Big] \ .\end{aligned}$$

Since the random variables $(X_j)_{j\geq 1}$ are independent of N, their conditional law given N is the same: in particular, they are i.i.d. Consequently, we have

$$= \mathbb{E}\Big[\prod_{k=1}^{n} \mathbb{E}\Big[e^{iq_k \sum_{j=N_{t_{k-1}}+1}^{N_{t_k}} X_j} \mid N\Big]\Big]$$
$$= \mathbb{E}\Big[\prod_{k=1}^{n} \Phi_X(q_k)^{N_{t_k}-N_{t_{k-1}}}\Big] .$$

Since the process N has stationary and independent increments, we get

$$= \prod_{k=1}^{n} \mathbb{E}\Big[\Phi_X(q_k)^{N_{t_k - t_{k-1}}}\Big]$$
$$= \prod_{k=1}^{n} \mathbb{E}\Big[e^{iq_k S_{t_k - t_{k-1}}}\Big] .$$

This concludes the proof of (i).
Regarding (ii), by independence we have

$$\begin{aligned}
\mathbb{P}(S_t \leq x) &= \sum_{n\geq 0} \mathbb{P}(S_t \leq x, N_t = n) \\
&= \sum_{n\geq 0} \mathbb{P}(X_1 + \ldots + X_n \leq x, N_t = n) \\
&= \sum_{n\geq 0} \mathbb{P}(X_1 + \ldots + X_n \leq x)\mathbb{P}(N_t = n) .
\end{aligned}$$

To prove (iii), we first show that $|S_t|$ is integrable:

$$\begin{aligned}
\mathbb{E}[|S_t|] &= \sum_{n\geq 0} \mathbb{E}[|S_t|\mathbf{1}_{N_t=n}] \\
&= \sum_{n\geq 0} \mathbb{E}[|X_1 + \ldots X_n|]\mathbb{P}(N_t = n) \\
&\leq \sum_{n\geq 0} n\mathbb{E}[|X_1|]\mathbb{P}(N_t = n) \\
&\leq \mathbb{E}[N_t]\mathbb{E}[|X_1|] < \infty .
\end{aligned}$$

Then, a similar computation shows that $\mathbb{E}[S_t] = \mathbb{E}[N_t]\mathbb{E}[X_1]$. If X has finite second moment, a similar argument shows that S_t also has a finite second moment and we get

$$\begin{aligned}
\mathbb{E}[S_t^2] &= \sum_{n\geq 0} \mathbb{E}[(X_1 + \ldots + X_n)^2]\mathbb{P}(N_t = n) \\
&= \sum_{n\geq 0} \big(n\mathbb{E}[X_1^2] + n(n-1)\mathbb{E}[X_1]^2\big)\mathbb{P}(N_t = n) \\
&= \mathbb{E}[N_t]\mathbb{E}[X_1^2] + \mathbb{E}[N_t^2 - N_t]\mathbb{E}[X_1]^2 \\
&= \mathbb{E}[N_t](\mathbb{E}[X_1^2] - \mathbb{E}[X_1]^2) + \mathbb{E}[N_t^2]\mathbb{E}[X_1]^2 .
\end{aligned}$$

Consequently,

$$\text{Var } S_t = \mathbb{E}[S_t^2] - \mathbb{E}[S_t]^2 = \mathbb{E}[N_t]\text{Var } X_1 + \mathbb{E}[N_t^2]\text{Var } X_1 \ .$$

Property (iv) is a consequence of the following computation:

$$\begin{aligned}
\mathbb{E}[e^{qS_t}] &= \sum_{n\geq 0} \mathbb{E}[e^{qS_t}\mathbf{1}_{N_t=n}] \\
&= \sum_{n\geq 0} \mathbb{E}[e^{q(X_1+\ldots+X_n)}]\mathbb{P}(N_t=n) \\
&= \sum_{n\geq 0} \mathbb{E}[e^{qX_1}]^n\mathbb{P}(N_t=n) \\
&= \sum_{n\geq 0} \mathbb{E}[e^{qX_1}]^n\frac{(\lambda t)^n}{n!}e^{-\lambda t} \\
&= \exp\big(\lambda t(\mathbb{E}[e^{qX_1}]-1)\big) \ .
\end{aligned}$$

□

Proposition 1.7.3. *Let S be a compound Poisson process and assume that $\mathbb{E}[|X|] < \infty$. Then, as $t \to \infty$ we have the following almost sure convergence*

$$\frac{S_t}{t} \to \lambda\,\mathbb{E}[X] \ .$$

Proof. Set

$$\Omega_1 := \{\omega \in \Omega : \frac{N_t(\omega)}{t} \to \lambda\} \ , \quad \Omega_2 := \{\omega \in \Omega : \frac{1}{n}\sum_{i=1}^{n} X_i(\omega) \to E[X_1]\} \ .$$

By Proposition 1.5.1 for the first event, and the classical law of large numbers for the second event, we deduce that $\mathbb{P}(\Omega_1 \cap \Omega_2) = 1$. Then for every $\omega \in \Omega_1 \cap \Omega_2$ we have as $t \to \infty$

$$\frac{S_t(\omega)}{t} = \frac{N_t(\omega)}{t}\frac{1}{N_t(\omega)}\sum_{i=1}^{N_t(\omega)} X_i(\omega) \to \lambda\,\mathbb{E}[X_1] \ .$$

□

1.8 An exercise

Insects fall into a soup bowl according to a Poisson process N with intensity $\lambda > 0$ (the event $\{N_t = n\}$ means that there are n insects in the bowl at time t). Assume that every insect is green with probability $p \in (0,1)$ and that its colour is independent of the colours of all other insects. Show that the number of green insects that fall into the bowl, as a function of time, is a Poisson process with intensity λp.

<u>Correction:</u> First of all, we need to introduce mathematical objects to model the problem. We let N be the Poisson process of intensity λ whose jump times T_n are the times at which an insect falls into the bowl. We assume that the insects are numbered from 1 to ∞ according to their order of appearance in the bowl. We then consider a sequence $(\epsilon_n)_{n\geq 1}$ of i.i.d. Bernoulli(p) r.v., independent of N, that models

the colours of the insects: the colour of the n-th insect is green if and only if $\epsilon_n = 1$.

It is clear then that the number of green insects at time t in the bowl is given by

$$N_t^g = \sum_{j=1}^{N_t} \epsilon_j \ , \quad t \geq 0 \ ,$$

that is a compound Poisson process. As such it has independent and stationary increments, and all we have to do is to compute the distribution at time t.

$$\mathbb{E}\left(e^{qN_t^g}\right) = \exp\left(\lambda t\left(\mathbb{E}\left[e^{q\epsilon_1}\right] - 1\right)\right) = \exp\left(\lambda t\left(\left[pe^{q\epsilon_1} + 1 - p\right] - 1\right)\right) = e^{\lambda pt(e^q - 1)},$$

that implies that N_t^g has a Poisson distribution of parameter λpt.

Chapter 2

Continuous-time Markov processes with countable state-space

In this chapter, we study continuous-time processes that take values in a countable set E and satisfy the Markov property. Recall that E is a countable set if it is either a finite set or a set in bijection with **N**. We will see that continuous-time Markov processes are intimately related with Markov chains, which were studied at the first semester.

2.1 The Markov property: from discrete-time to continuous-time processes

Let us recall the definition of Markov chains from the first semester.

Definition 2.1.1 (Markov chain - first version)**.** *Let $(X_n)_{n\geq 0}$ be a discrete-time process taking values in E. We say that $(X_n)_{n\geq 0}$ is a time-homogeneous Markov chain if:*

1. *(Markov property): For all $n \geq 0$ and all $(x_0, \ldots, x_{n-1}, x, y) \in E^{n+2}$ such that $\mathbb{P}(X_0 = x_0, \ldots, X_{n-1} = x_{n-1}, X_n = x) > 0$, we have*

$$\mathbb{P}(X_{n+1} = y \,|\, X_0 = x_0, \ldots, X_{n-1} = x_{n-1}, X_n = x) = \mathbb{P}(X_{n+1} = y \,|\, X_n = x) \ .$$

2. *(Time-homogeneity): There exists a matrix $\Pi : E \times E \to \mathbf{R}$ such that*

$$\mathbb{P}(X_{n+1} = y \,|\, X_n = x) = \Pi(x, y) \ .$$

Recall that Π is often called the *transition matrix* of the chain, although it is an "infinite" matrix as soon as E is an infinite set. Recall also that Π satisfies $\Pi(x, y) \geq 0$ for all $x, y \in E$ and

$$\sum_{y \in E} \Pi(x, y) = 1 \ ,$$

for all $x \in E$.

Our first task in this chapter is to extend this definition to the setting of continuous-time processes. While the intuition behind the Markov property (the relevant information to determine the future evolution of the process consists of the current state of the process) remains the same in continuous-time, the

precise definition is a priori unclear: the time parameter being continuous, there is no obvious way to generalise the conditioning by $X_0 = x_0, \ldots, X_{n-1} = x_{n-1}, X_n = x$ from the definition.
At this point, we actually observe that one can simplify the above definition using the notion of filtration. Let $\mathcal{F}_n := \sigma(X_0, \ldots, X_n)$, $n \geq 0$, be the natural filtration associated to the process $(X_n, n \geq 0)$.

Definition 2.1.2 (Markov chain - second version)**.** *Let $(X_n)_{n\geq 0}$ be a discrete-time process taking values in E. We say that $(X_n)_{n\geq 0}$ is a time-homogeneous Markov chain if:*

1. *(Markov property): For all $n \geq 0$ and all $y \in E$, we have*

$$\mathbb{P}(X_{n+1} = y \,|\, \mathcal{F}_n) = \mathbb{P}(X_{n+1} = y \,|\, X_n) \; .$$

2. *(Time-homogeneity): There exists a matrix $\Pi : E \times E \to \mathbf{R}$ such that almost surely*

$$\mathbb{P}(X_{n+1} = y \,|\, X_n) = \Pi(X_n, y) \; .$$

Lemma 2.1.3. *These two definitions are equivalent.*

Proof. Let us start from the second definition. Under the assumption that $\mathbb{P}(X_0 = x_0, \ldots, X_{n-1} = x_{n-1}, X_n = x) > 0$, we have

$$\begin{aligned}
\mathbb{P}(X_{n+1} = y, X_0 = x_0, \ldots, X_{n-1} = x_{n-1}, X_n = x) &= \mathbb{E}[\mathbf{1}_{X_{n+1}=y}\mathbf{1}_{X_0=x_0,\ldots,X_{n-1}=x_{n-1},X_n=x}] \\
&= \mathbb{E}[\mathbb{E}[\mathbf{1}_{X_{n+1}=y}\mathbf{1}_{X_0=x_0,\ldots,X_{n-1}=x_{n-1},X_n=x} \,|\, \mathcal{F}_n]] \\
&= \mathbb{E}[\mathbf{1}_{X_0=x_0,\ldots,X_{n-1}=x_{n-1},X_n=x}\mathbb{E}[\mathbf{1}_{X_{n+1}=y} \,|\, \mathcal{F}_n]] \\
&= \mathbb{E}[\mathbf{1}_{X_0=x_0,\ldots,X_{n-1}=x_{n-1},X_n=x}\mathbb{E}[\mathbf{1}_{X_{n+1}=y} \,|\, X_n]] \\
&= \mathbb{E}[\mathbf{1}_{X_0=x_0,\ldots,X_{n-1}=x_{n-1},X_n=x}\, \varphi_n(X_n)] \\
&= \varphi_n(x)\mathbb{E}[\mathbf{1}_{X_0=x_0,\ldots,X_{n-1}=x_{n-1},X_n=x}] \; ,
\end{aligned}$$

where $\varphi_n(z) := \mathbb{E}[\mathbf{1}_{X_{n+1}=y} \,|\, X_n = z]$ for all $z \in E$. Hence

$$\begin{aligned}
&\mathbb{P}(X_{n+1} = y \,|\, X_0 = x_0, \ldots, X_{n-1} = x_{n-1}, X_n = x) \\
&= \frac{\mathbb{P}(X_{n+1} = y, X_0 = x_0, \ldots, X_{n-1} = x_{n-1}, X_n = x)}{\mathbb{P}(X_0 = x_0, \ldots, X_{n-1} = x_{n-1}, X_n = x)} \\
&= \varphi_n(x) = \mathbb{P}(X_{n+1} = y \,|\, X_n = x) \; .
\end{aligned}$$

Let us now prove the converse implication. We start from the first definition. We aim at showing that for all $A \in \mathcal{F}_n$, we have

$$\mathbb{E}[\mathbf{1}_{X_{n+1}=y}\mathbf{1}_A] = \mathbb{E}[\mathbb{E}[\mathbf{1}_{X_{n+1}=y} \,|\, X_n]\mathbf{1}_A] \; . \tag{2.1}$$

We claim that this is true whenever A is of the form $A = \{X_0 = x_0, \ldots, X_{n-1} = x_{n-1}, X_n = x\}$. Indeed, if $\mathbb{P}(A) = 0$ then both sides of the identity vanish. If $\mathbb{P}(A) > 0$, then the assertion of the first definition yields

$$\begin{aligned}
\mathbb{E}[\mathbf{1}_{X_{n+1}=y}\mathbf{1}_A] &= \mathbb{P}(X_{n+1} = y \,|\, X_0 = x_0, \ldots, X_{n-1} = x_{n-1}, X_n = x) \\
&\qquad \times \mathbb{P}(X_0 = x_0, \ldots, X_{n-1} = x_{n-1}, X_n = x) \\
&= \mathbb{P}(X_{n+1} = y \,|\, X_n = x)\mathbb{P}(X_0 = x_0, \ldots, X_{n-1} = x_{n-1}, X_n = x) \\
&= \varphi_n(x)\mathbb{P}(X_0 = x_0, \ldots, X_{n-1} = x_{n-1}, X_n = x) \; ,
\end{aligned}$$

where φ_n is the (measurable) map from E to **R** such that $\varphi_n(X_n) = \mathbb{P}(X_{n+1} = y \,|\, X_n)$ almost surely. Hence

$$\begin{aligned}\mathbb{E}[\mathbf{1}_{X_{n+1}=y}\mathbf{1}_A] &= \varphi_n(x)\mathbb{P}(X_0 = x_0, \ldots, X_{n-1} = x_{n-1}, X_n = x) \\ &= \varphi_n(x)\mathbb{E}[\mathbf{1}_A] \\ &= \mathbb{E}[\varphi_n(X_n)\mathbf{1}_A] \;,\end{aligned}$$

which concludes the proof of the claim. Since the class $\mathcal{C}_n$ of all events A of the above form is stable under finite intersections and is such that $\sigma(\mathcal{C}_n) = \mathcal{F}_n$, the Monotone Class Theorem allows to deduce (2.1). □

Given the second definition above, we can easily pass to continuous-time processes.

Definition 2.1.4. *A càdlàg process $(X_t, t \geq 0)$ that takes values in E is a time-homogeneous Markov process if:*

1. *(Markov property): For all $t, s \geq 0$ we have almost surely*

$$\mathbb{P}[X_{t+s} = y \,|\, \mathcal{F}_t] = \mathbb{P}[X_{t+s} = y \,|\, X_t] \;,$$

   where $\mathcal{F}_t := \sigma(X_s, s \in [0,t])$.

2. *(Time-homogeneity): For every $s \geq 0$ there exists a matrix $P_s : E \times E \to$ **R** such that for all $y \in E$ we have almost surely*

$$\mathbb{P}[X_{t+s} = y \,|\, X_t] = P_s(X_t, y) \;.$$

It will be convenient to write $\mathbb{P}_x$ and $\mathbb{E}_x$ when the process X starts from the deterministic initial condition $X_0 = x$.

Observe that we impose X to be càdlàg. Since E is countable, any càdlàg function with values in E is necessarily piecewise constant.

Remark 2.1.5. *We have not specified the topology on E while we have been talking of continuity for E-valued functions. Here E is endowed with the discrete topology, that is, the topology induced by the metric*

$$d(x,y) := \begin{cases} 1 \text{ if } x \neq y \;, \\ 0 \text{ if } x = y \;. \end{cases}$$

From now on, all our Markov processes will be time-homogeneous so we will simply write "Markov process" for "time-homogeneous Markov process".

The next result shows that conditionally given $\mathcal{F}_t$, $(X_{t+s}, s \geq 0)$ is a Markov process starting from X_t.

Lemma 2.1.6. *Let X be a Markov process. Then for all $t > 0$, $0 < t_1 < \ldots < t_n$ and $y_1, \ldots, y_n \in E^n$ we have*

$$\mathbb{P}(X_{t+t_1} = y_1, \ldots, X_{t+t_n} = y_n \,|\, \mathcal{F}_t) = P_{t_1}(X_t, y_1)P_{t_2-t_1}(y_1, y_2)\ldots P_{t_n-t_{n-1}}(y_{n-1}, y_n) \;.$$

Proof. The proof consists of a recursion on n. The case $n = 1$ comes from the definition. Assume that the identity holds true at rank n. At rank $n + 1$, for all $A \in \mathcal{F}_t$, we have

$$\begin{aligned}
&\mathbb{E}[\mathbf{1}_{X(t+t_1)=y_1,\dots,X(t+t_{n+1})=y_{n+1}}\mathbf{1}_A] \\
&= \mathbb{E}[\mathbb{E}[\mathbf{1}_{X(t+t_1)=y_1,\dots,X(t+t_{n+1})=y_{n+1}}\mathbf{1}_A \,|\, \mathcal{F}_{t+t_n}]] \\
&= \mathbb{E}[\mathbf{1}_{X(t+t_1)=y_1,\dots,X(t+t_n)=y_n}\mathbf{1}_A\mathbb{E}[\mathbf{1}_{X(t+t_{n+1})=y_{n+1}} \,|\, \mathcal{F}_{t+t_n}]] \;, \\
&= \mathbb{E}[\mathbf{1}_{X(t+t_1)=y_1,\dots,X(t+t_n)=y_n}\mathbf{1}_A P_{t_{n+1}-t_n}(y_n, y_{n+1})] \\
&= \mathbb{E}[\mathbf{1}_{X(t+t_1)=y_1,\dots,X(t+t_n)=y_n}\mathbf{1}_A] P_{t_{n+1}-t_n}(y_n, y_{n+1}) \\
&= \mathbb{E}[\mathbf{1}_A P_{t_1}(X_t, y_1)] P_{t_2-t_1}(y_1, y_2) \dots P_{t_n-t_{n-1}}(y_{n-1}, y_n) P_{t_{n+1}-t_n}(y_n, y_{n+1}) \;,
\end{aligned}$$

which ensures that the conditional probability of $X(t + t_1) = y_1, \dots, X(t + t_{n+1}) = y_{n+1}$ given $\mathcal{F}_t$ is $P_{t_1}(X_t, y_1)P_{t_2-t_1}(y_1, y_2) \dots P_{t_{n+1}-t_n}(y_n, y_{n+1})$. □

Similarly as in discrete-time, we say that a non-negative r.v. T is an $\mathcal{F}$-stopping time if $\{T \le t\}$ is $\mathcal{F}_t$-measurable for all $t \ge 0$. We then define the sigma-field $\mathcal{F}_T$ as the set of events A such that $A \cap \{T \le t\} \in \mathcal{F}_t$ for all $t \ge 0$.

Proposition 2.1.7 (Strong Markov Property)**.** *Let T be an $\mathcal{F}$-stopping time such that $T < \infty$ almost surely. Then, the process $(X_{t+T}, t \ge 0)$ is a Markov process such that for all $y \in E$ and all $s \ge 0$ we have almost surely*

$$\mathbb{P}(X_{s+T} = y \,|\, \mathcal{F}_T) = P_s(X_T, y) \;.$$

Proof. It suffices to prove the identity of the statement: then, the fact that the process starting at time T is a Markov process is proven similarly as in Lemma 2.1.6. When T is deterministic, this is the definition. Let us now assume that T takes values in a countable set, say $\mathbf{Q}_+$. Then, for every event $A \in \mathcal{F}_T$, we have

$$\begin{aligned}
\mathbb{P}(X_{s+T} = y, A) &= \sum_{q\in\mathbf{Q}_+} \mathbb{P}(X_{s+T} = y; A \cap \{T = q\}) \\
&= \sum_{q\in\mathbf{Q}_+} \mathbb{E}[\mathbb{E}[\mathbf{1}_{\{X_{s+q}=y\}\cap A\cap\{T=q\}} \,|\, \mathcal{F}_q]] \\
&= \sum_{q\in\mathbf{Q}_+} \mathbb{E}[\mathbf{1}_{A\cap\{T=q\}}\mathbb{E}[\mathbf{1}_{\{X_{s+q}=y\}} \,|\, \mathcal{F}_q]] \;,
\end{aligned}$$

At this point, we apply the Markov property at time q and obtain:

$$\begin{aligned}
&= \sum_{q\in\mathbf{Q}_+} \mathbb{E}[\mathbf{1}_{A\cap\{T=q\}}\mathbb{E}[\mathbf{1}_{\{X_{s+q}=y\}} \,|\, X_q]] \\
&= \sum_{q\in\mathbf{Q}_+} \mathbb{E}[\mathbf{1}_{A\cap\{T=q\}} P_s(X_q, y)] \\
&= \mathbb{E}[\mathbf{1}_A P_s(X_T, y)] \;.
\end{aligned}$$

Henceforth, the conditional expectation $\mathbb{E}[\mathbf{1}_{\{X_{s+T}=y\}} \,|\, \mathcal{F}_T]$ coincides with $P_s(X_T, y)$ almost surely, thus establishing the proposition in this case.
We now consider a general stopping time T. We consider its dyadic approximation:

$$T_n := \sum_{k=0}^{\infty} \frac{k+1}{2^n} \mathbf{1}_{\{\frac{k}{2^n} \le T < \frac{k+1}{2^n}\}} \;.$$

It is not difficult to check that $T_n \geq T$ almost surely, and $T_n \downarrow T$ almost surely as $n \to \infty$. Furthermore, for every $n \geq 1$, T_n is an $\mathcal{F}$-stopping time. Finally, the right continuity of X and the Dominated Convergence Theorem show that for all $A \in \mathcal{F}_T$ we have

$$\mathbb{P}(X_{s+T} = y, A) = \lim_{n\to\infty} \mathbb{P}(X_{s+T_n} = y, A) \; .$$

Since T_n takes values in a countable set, we can apply the result proved right before and deduce that

$$\mathbb{P}(X_{s+T_n} = y, A) = \mathbb{E}[\mathbf{1}_A P_s(X_{T_n}, y)] \; .$$

This readily shows that

$$\mathbb{P}(X_{s+T} = y, A) = \lim_{n\to\infty} \mathbb{E}[\mathbf{1}_A P_s(X_{T_n}, y)] = \mathbb{E}[\mathbf{1}_A P_s(X_T, y)] \; .$$

□

2.2 The embedded Markov chain

In this section, we establish a deep connection between Markov processes and Markov chains. The key observation is the following result.

Proposition 2.2.1. *Let X be a Markov process starting from $X_0 = x$ almost surely, for some $x \in E$. Denote by $\tau := \inf\{s \geq 0 : X_s \neq X_0\}$. Then the random variable τ has an exponential law whose parameter will be denoted λ_x or $\lambda(x)$.*
If $\lambda_x > 0$ then $\tau < \infty$ a.s. and the r.v. X_τ is independent of τ. We denote by $\Pi(x, \cdot)$ the corresponding probability measure on E.
If $\lambda_x = 0$ then $\tau = \infty$. In that case, we set $\Pi(x, x) = 1$ and $\Pi(x, y) = 0$ for all $y \neq x$.

This result shows the following fact. A continuous-time Markov process starting from some point $x \in E$ at time 0 stays at this point for a random time τ_1 of exponential law with parameter λ_x: at that time, the process "jumps" to a random point y independent of τ_1 and chosen according to a probability law $\Pi(x, \cdot)$. Then, the process stays at point y for a random time τ_2 of exponential law with parameter λ_y: at time $\tau_1 + \tau_2$ it jumps to a point z chosen according to the probability law $\Pi(y, \cdot)$ and independent of τ_1, τ_2. And so on.

This implies that the *law* of a Markov process is completely characterised by λ_x and $\Pi(x, \cdot)$ for all $x \in E$.

Proof. We work under $\mathbb{P}_x$. Fix $t, s > 0$ and introduce $\tau' := \inf\{r \geq 0 : X_{t+r} \neq X_t\}$. We have

$$\mathbb{P}_x(\tau > t + s) = \mathbb{E}_x[\mathbb{E}_x[\mathbf{1}_{\{\tau > t+s\}} \,|\, \mathcal{F}_t]] = \mathbb{E}_x[\mathbb{E}_x[\mathbf{1}_{\{\tau > t\} \cap \{\tau' > s\}} \,|\, \mathcal{F}_t]] = \mathbb{E}_x[\mathbf{1}_{\{\tau > t\}} \mathbb{E}_x[\mathbf{1}_{\{\tau' > s\}} \,|\, \mathcal{F}_t]] \; .$$

At this point, we observe that

$$\mathbb{E}_x[\mathbf{1}_{\{\tau' > s\}} \,|\, \mathcal{F}_t] = \mathbb{E}_x[\mathbf{1}_{\{\forall r \in [0,s], X_{t+r} = X_t\}} \,|\, \mathcal{F}_t] \; .$$

From the Markov property, we know that the process $(X_{t+r}, r \geq 0)$ is a Markov process starting from the random initial condition X_t, and that its conditional expectation given $\mathcal{F}_t$ is the same as its conditional expectation given X_t. Therefore, there exists a measurable map φ such that

$$\mathbb{E}_x[\mathbf{1}_{\{\forall r\in[0,s], X_{t+r}=X_t\}} \,|\, \mathcal{F}_t] = \mathbb{E}_x[\mathbf{1}_{\{\forall r\in[0,s], X_{t+r}=X_t\}} \,|\, X_t] = \varphi(X_t) \ .$$

Let us determine φ. Conditionally given $X_t = y$, the process $(X_{t+r}, r \geq 0)$ has the same law as $(X_r, r \geq 0)$ starting from y. Consequently

$$\varphi(y) = \mathbb{E}_y[\mathbf{1}_{\{\forall r\in[0,s], X_r=y\}}] = \mathbb{P}_y(\tau > s) \ .$$

Furthermore, on the event $\tau > t$ we have $X_t = X_0$. Hence,

$$\begin{aligned}\mathbb{E}_x[\mathbf{1}_{\{\tau>t\}}\mathbb{E}_x[\mathbf{1}_{\{\tau'>s\}} \,|\, \mathcal{F}_t]] &= \mathbb{E}_x[\mathbf{1}_{\{\tau>t\}}\varphi(X_t)] \\ &= \mathbb{E}_x[\mathbf{1}_{\{\tau>t\}}\varphi(x)] \\ &= \mathbb{E}_x[\mathbf{1}_{\{\tau>t\}}\mathbb{P}_x(\tau > s)] = \mathbb{P}_x(\tau > s)\mathbb{P}_x(\tau > t) \ .\end{aligned}$$

By Lemma 1.2.1, we deduce that the r.v. τ has an exponential distribution. Regarding the independence of $X(\tau)$ with τ, we observe that on the event $\{\tau > t\}$ we have $\tau = t + \tau'$. Consequently

$$\begin{aligned}\mathbb{E}_x[\mathbf{1}_{X(\tau)=y}\mathbf{1}_{\tau>t}] &= \mathbb{E}_x[\mathbf{1}_{X(t+\tau')=y}\mathbf{1}_{\tau>t}] \\ &= \mathbb{E}_x[\mathbb{E}_x[\mathbf{1}_{X(t+\tau')=y}\mathbf{1}_{\tau>t} \,|\, \mathcal{F}_t]] \\ &= \mathbb{E}_x[\mathbf{1}_{\tau>t}\mathbb{E}_x[\mathbf{1}_{X(t+\tau')=y} \,|\, \mathcal{F}_t]] \\ &= \mathbb{E}_x[\mathbf{1}_{\tau>t}\mathbb{E}_x[\mathbf{1}_{X(t+\tau')=y} \,|\, X_t]] \\ &= \mathbb{E}_x[\mathbf{1}_{\tau>t}\mathbb{E}_x[\mathbf{1}_{X(\tau)=y}]] \\ &= \mathbb{P}_x(\tau > t)\mathbb{P}_x(X(\tau) = y) \ .\end{aligned}$$

This is sufficient to conclude to the asserted independence. □

Let us now introduce T_n as the n-th jump time of the process X:

$$T_n := \inf\{t > T_{n-1} : X(t) \neq X(T_{n-1})\} \ , \quad T_0 := 0 \ ,$$

then the discrete-time process $Y_n := X(T_n)$, $n \geq 0$ is a Markov chain.

Proposition 2.2.2. *The process $Y_n := X(T_n)$ is a discrete-time Markov chain of transition matrix Π.*

Proof. The proof consists of a recursion of the previous result. □

It is also possible to show that the process $(Y_n, n \geq 0)$ is independent of $(\tau_n \lambda_{Y_{n-1}}, n \geq 1)$, where $\tau_n := T_n - T_{n-1}$ for every $n \geq 1$.

2.3 Construction of Markov processes

Given a transition matrix Π and a collection of non-negative parameters λ_x, $x \in E$ we can construct a Markov process starting from some initial condition X_0 as follows. Assume that we are given:

1. a collection of IID r.v. E_n, $n \geq 1$ of exponential law with parameter 1,

2. an independent Markov chain $(Y_n, n \geq 0)$ of transition matrix Π, starting from $Y_0 := X_0$.

Remark 2.3.1. *The careful reader will notice that, to be consistent with the notations introduced in Proposition 2.2.1, one needs to impose a little restriction on Π and λ, namely: for all $x \in E$, we have*

$$\Pi(x,x) \neq 0 \Leftrightarrow \Pi(x,x) = 1 \Leftrightarrow \lambda_x = 0 \ .$$

Remark 2.3.2. *We do not address the construction of the chain $(Y_n, n \geq 0)$. However this can be achieved using random recursions, see Section 3.1.2 of the lecture notes of "Processus Discrets".*

We set $\tau_n := \frac{E_n}{\lambda_{Y_{n-1}}}$ for every $n \geq 1$, and we define $T_0 := 0$ and

$$T_n := \tau_1 + \ldots + \tau_n \ .$$

We define $\zeta := \sup_{n\geq 0} T_n$. For all $t \in [0, \zeta[$, we let $n \geq 0$ be the unique (random) integer such that $t \in [T_n, T_{n+1}[$ and we set

$$X_t := Y_n \ .$$

Lemma 2.3.3. *We have $\mathbb{P}(\zeta < \infty) = 0$ in any of the following situations:*

1. *E is a finite set,*
2. *$\sup_{x\in E} \lambda_x < \infty$,*
3. *the chain Y is irreducible and recurrent.*

Proof. If E is finite then $\sup_{x\in E} \lambda_x < \infty$. Consequently, the first case is a particular instance of the second. Let us therefore prove the second case: assume that $\bar{\lambda} := \sup_{x\in E} \lambda_x < \infty$. Since $\bar{\lambda}\tau_k \geq E_k$ almost surely for all $k \geq 1$, we deduce that

$$\bar{\lambda}\zeta \geq \sum_{k\geq 1} E_k \ .$$

Since the E_k's are IID $\mathcal{E}(1)$ r.v., the law of large numbers ensures that their sum is infinite almost surely. Consequently $\zeta = \infty$ almost surely.
Let us now assume that Y is irreducible and recurrent, and that E is not finite (if E is finite then we can apply the previous arguments). Necessarily for any $x \in E$, Y visits x infinitely many times almost surely. Moreover $\Pi(x,x) \neq 1$: indeed, if it were equal to 1 then Y would not be irreducible. Consequently $\lambda_x > 0$. Let's call $N_1, N_2, \ldots$ the successive discrete times n at which $Y_n = x$. We have

$$\lambda_x \zeta \geq \lambda_x \sum_{k=1}^{\infty} \tau_{N_k+1} = \sum_{k\geq 1} E_{N_k+1} \ .$$

Since $(N_k)_{k\geq 1}$ only depends on the chain Y, it is independent of $(E_n)_{n\geq 1}$. Hence, $(E_{N_k})_{k\geq 1}$ is a sequence of IID $\mathcal{E}(1)$ r.v. Its sum is infinite almost surely. □

Proposition 2.3.4. *If $\zeta = \infty$ almost surely, then the process $(X_t, t \geq 0)$ is a continuous-time Markov process.*

The proof of this proposition is delicate and will not be presented in the lectures.

Proof. Let us show that for all $n \geq 1$, all $0 < t_1 < \ldots < t_n < t < t+s$ and all $x_1, \ldots, x_n, x, y, z \in E$, we have

$$\mathbb{P}_x(X_{t_1} = x_1, \ldots, X_{t_n} = x_n, X_t = y, X_{t+s} = z) = \mathbb{P}_x(X_{t_1} = x_1, \ldots, X_{t_n} = x_n, X_t = y)\mathbb{P}_y(X_s = z) \ .$$

This will be enough to deduce that X satisfies the Markov property and is time-homogeneous.

First of all, observe that it suffices to show that for any $m \geq 0$

$$\begin{aligned} &\mathbb{P}_x(X_{t_1} = x_1, \ldots, X_{t_n} = x_n, X_t = y, X_{t+s} = z, T_m \leq t < T_{m+1}) \\ &= \mathbb{P}_x(X_{t_1} = x_1, \ldots, X_{t_n} = x_n, X_t = y, T_m \leq t < T_{m+1})\mathbb{P}_y(X_s = z) \ . \end{aligned} \tag{2.2}$$

From now on, $m \geq 0$ is fixed. By construction, there exists a map Φ such that $X = \Phi((E_n)_{n\geq 1}, (Y_n)_{n\geq 0})$. We do not give a precise definition of this map but it is intuitive that X is completely characterised by $(E_n)_{n\geq 1}, (Y_n)_{n\geq 0}$. In particular, note that for all $r \geq 0$, we have $X_r = Y_n$ where $n \geq 0$ is the unique integer such that $T_n \leq r < T_{n+1}$ and where $T_n = \sum_{j=1}^n \frac{E_j}{\lambda_{Y_{j-1}}}$.
On the event $\{T_m \leq t < T_{m+1}\}$, set

$$E_1' := \lambda_{Y_m}(T_{m+1} - t) \ , \quad E_2' := \lambda_{Y_{m+1}}(T_{m+2} - T_{m+1}) = E_{m+2} \ , \quad \ldots$$

and for every $k \geq 1$

$$T_k' := \sum_{j=1}^k \frac{E_j'}{\lambda_{Y_{m+j-1}}} \ .$$

Define $X_r' := X_{t+r}$ for all $r \geq 0$. On the event $\{T_m \leq t < T_{m+1}\}$, we see that $X_r' = Y_{m+n}$ where $n \geq 0$ is the unique integer such that $T_n' \leq r < T_{n+1}'$. We can then deduce that $X' = \Phi((E_n')_{n\geq 1}, (Y_{m+n})_{n\geq 0})$.
For $\bar{i} := (i_1, i_2, \ldots, i_n)$ with $0 \leq i_1 \leq \ldots \leq i_n \leq m$ and $\bar{y} := (y_1, \ldots, y_m) \in E^m$ define the event

$$A_{\bar{i},\bar{y}} := \cap_{j=1}^n \{T_{i_j} \leq t_j < T_{i_j+1}\} \cap \{T_m \leq t < T_{m+1}\} \cap \{(Y_1, \ldots, Y_m) = (y_1, \ldots, y_m)\} \ .$$

This event only depends on $E_1, \ldots, E_{m+1}$ and $Y_1, \ldots, Y_m$. By construction, conditionally given $A_{\bar{i},\bar{y}}$ the process $(Y_{m+n})_{n\geq 0}$ is a Markov chain starting from y_m which is independent from $(E_n')_{n\geq 1}$. Note that $E_1' = E_{m+1} - \lambda_{Y_m}(t - T_m)$. By the absence of memory of exponential r.v. we see that E_1', conditionally given $A_{\bar{i},\bar{y}}$ is still an exponential r.v. of parameter 1. Finally, one can check that conditionally given this same event, the sequence $(E_n')_{n\geq 1}$ is made of i.i.d. $\mathcal{E}(1)$ r.v.
We therefore have

$$\begin{aligned} &\mathbb{P}_x(X_{t_1} = x_1, \ldots, X_{t_n} = x_n, X_t = y, X_{t+s} = z \mid A_{\bar{i},\bar{y}}) \\ &= \mathbf{1}_{\cap_{j=1}^n \{x_j = y_{i_j}\}} \mathbf{1}_{\{y = y_m\}} \mathbb{P}_x(X_s' = z \mid A_{\bar{i},\bar{y}}) \\ &= \mathbf{1}_{\cap_{j=1}^n \{x_j = y_{i_j}\}} \mathbf{1}_{\{y = y_m\}} \mathbb{P}_y(X_s = z) \ . \end{aligned}$$

Moreover

$$\begin{aligned} &\mathbb{P}_x(X_{t_1} = x_1, \ldots, X_{t_n} = x_n, X_t = y, T_m \leq t < T_{m+1}) \\ &= \sum_{\bar{i},\bar{y}} \mathbb{P}_x(X_{t_1} = x_1, \ldots, X_{t_n} = x_n, X_t = y, A_{\bar{i},\bar{y}}) \\ &= \sum_{\bar{i},\bar{y}} \mathbf{1}_{\cap_{j=1}^n \{x_j = y_{i_j}\}} \mathbf{1}_{\{y = y_m\}} \mathbb{P}(A_{\bar{i},\bar{y}}) \ . \end{aligned}$$

Consequently

$$\begin{aligned}
&\mathbb{P}_x(X_{t_1} = x_1, \ldots, X_{t_n} = x_n, X_t = y, X_{t+s} = z, T_m \le t < T_{m+1}) \\
&= \sum_{\bar{i},\bar{y}} \mathbb{P}_x(X_{t_1} = x_1, \ldots, X_{t_n} = x_n, X_t = y, X_{t+s} = z, A_{\bar{i},\bar{y}}) \\
&= \sum_{\bar{i},\bar{y}} \mathbb{P}_x(X_{t_1} = x_1, \ldots, X_{t_n} = x_n, X_t = y, X_{t+s} = z \,|\, A_{\bar{i},\bar{y}}) \mathbb{P}(A_{\bar{i},\bar{y}}) \\
&= \sum_{\bar{i},\bar{y}} \mathbf{1}_{\cap_{j=1}^n \{x_j = y_{i_j}\}} \mathbf{1}_{\{y=y_m\}} \mathbb{P}_y(X_s = z) \mathbb{P}(A_{\bar{i},\bar{y}}) \\
&= \mathbb{P}_x(X_{t_1} = x_1, \ldots, X_{t_n} = x_n, X_t = y, T_m \le t < T_{m+1}) \mathbb{P}_y(X_s = z) \;,
\end{aligned}$$

so that (2.2) is proved. □

Remark 2.3.5. *Our definition of Markov processes excludes the so-called explosions, that is, the case where $\zeta < \infty$ with positive probability. It is possible to introduce a more general definition that encompasses this case, see Section 2.7.*

2.4 Semigroup and Kolmogorov equations

Definition 2.4.1. *We call* semigroup *a collection of matrices $(P_t, t \ge 0)$ that satisfies the following properties:*

1. $P_0 = Id$ *,*
2. *For every $t \ge 0$, P_t is a stochastic matrix, that is, for all $x, y \in E$ we have $P_t(x,y) \ge 0$ and $\sum_{y \in E} P_t(x,y) = 1$,*
3. *For every $t, s \ge 0$, we have $P_{t+s} = P_t P_s$, that is, for all x, y*

$$P_{t+s}(x,y) = \sum_{z \in E} P_t(x,z) P_s(z,y) \;.$$

Lemma 2.4.2. *Let X be a Markov process. The collection of matrices $(P_t, t \ge 0)$ defined by*

$$P_t(x,y) := \mathbb{P}_x(X_t = y) \;,$$

is a semigroup.

Proof. Left as an exercise. □

Given a semigroup $(P_t, t \ge 0)$ and a measure μ, one can define a collection of measures μP_t, $t \ge 0$ by setting

$$(\mu P_t)(y) = \sum_x \mu(x) P_t(x,y) \;, \quad y \in E \;.$$

Note that the mass of μP_t is the same as the mass of μ.

Let μ be a probability measure on E, and let X be a Markov process starting from the probability measure μ, that is, such that X_0 has law μ. Then the law of X_t is given by μP_t: indeed we have

$$\mathbb{P}_\mu(X_t = y) = \sum_{x \in E} \mu(x)\mathbb{P}_x(X_t = y) = \sum_{x \in E} \mu(x)P_t(x, y) = \mu P_t(y) \ .$$

More generally, for all $0 < t_1 < \ldots < t_n$ and all $x_1, \ldots, x_n \in E$ we have

$$\mathbb{P}_\mu(X(t_1) = x_1, \ldots, X(t_n) = x_n) = \sum_{x_0 \in E} \mu(x_0)P_{t_1}(x_0, x_1)P_{t_2-t_1}(x_1, x_2) \ldots P_{t_n-t_{n-1}}(x_{n-1}, x_n) \ .$$

Recall that λ_x is the rate of the exponential distribution of the time the process X stays at x, and Π is the transition matrix of the embedded Markov chain Y. We define a matrix Q by setting

$$Q(x, y) := \begin{cases} \lambda(x)\Pi(x, y) & \text{if } x \neq y \ , \\ -\lambda(x) & \text{if } x = y \ . \end{cases}$$

The matrix Q is called the *generator* of the Markov process X. Roughly speaking, $Q(x, y)$ is the "rate" at which the Markov process jumps from x to y while $Q(x, x)$ is the opposite of the rate at which the Markov process leaves x, that is, the sum over $y \neq x$ of the rates from x to y.

Remark 2.4.3. *We saw that if $X_0 = x$ then X stays at x an exponential time T_1 of parameter λ_x and then jumps to a random point drawn according to the probability measure $\Pi(x, \cdot)$. An equivalent point of view is the following. Consider independent exponential r.v. $S_{x,y}$ of parameters $\lambda_x\Pi(x, y)$ for every $y \neq x$. Then $\min_y S_{x,y}$ is an exponential r.v. of parameter $\sum_y \lambda_x\Pi(x, y) = \lambda_x$. Furthermore $\arg\min_y S_{x,y}$ has law $\Pi(x, \cdot)$. As a consequence, there is a competition between independent exponential r.v. of parameters $Q(x, y)$ for $y \neq x$: the next transition of the Markov process corresponds to the minimum of these r.v.*

The generator is intimately related to the semigroup of X as the following theorem shows. In the following f is a bounded function on X, and we denote $\|f\|_\infty = \sup_{x \in X} |f(x)|$.

Theorem 4. *For any $x \in X$ we have*

$$\lim_{t \to 0} \frac{1}{t} \left[P_t f(x) - f(x)\right] \ = \ Qf(x), \tag{2.3}$$

furthermore there exists a constant $C < \infty$ such that

$$\sup_x \left| \frac{1}{t} \left[P_t f(x) - f(x)\right] - Qf(x) \right| \leq Ct\|f\|_\infty. \tag{2.4}$$

Proof. Recall that by 2.2.1 T_1 and X_{T_1} are independent. Then

$$\begin{aligned}
&P_t f(x) = \mathbb{P}_x(T_1 > t)f(x) + \mathbb{E}_x(1_{\{T_1 \leq t\}} f(X_t)) \\
&= e^{-\lambda_x t} f(x) + \mathbb{E}_x(1_{\{T_1 \leq t\}} f(X_{T_1})) + \mathbb{E}_x \left(1_{\{T_1 \leq t\}} \left(f(X_t) - f(X_{T_1})\right)\right) \\
&= f(x) - \left(1 - e^{-\lambda_x t}\right) f(x) + \mathbb{P}_x(T_1 \leq t)E_x(f(X_{T_1})) + \mathbb{E}_x \left(1_{\{T_2 \leq t\}} \left(f(X_t) - f(X_{T_1})\right)\right) \\
&= f(x) - \left(1 - e^{-\lambda_x t}\right) f(x) + \left(1 - e^{-\lambda_x t}\right) \sum_{y \neq x} \Pi(x, y)f(y) + \mathbb{E}_x \left(1_{\{T_2 \leq t\}} \left(f(X_t) - f(X_{T_1})\right)\right) \\
&= f(x) + tQf(x) + \left(1 - e^{-\lambda_x t} - \lambda_x t\right) \left(\sum_{y \neq x} \Pi(x, y)f(y) - f(x)\right) + \mathbb{E}_x \left(1_{\{T_2 \leq t\}} \left(f(X_t) - f(X_{T_1})\right)\right),
\end{aligned} \tag{2.5}$$

that gives

$$\frac{1}{t}\left[P_t f(x) - f(x)\right] + Qf(x) = \frac{\left(1 - e^{-\lambda_x t} - \lambda_x t\right)}{t}\left(\sum_{y\neq x}\Pi(x,y)f(y) - f(x)\right) + \frac{1}{t}\mathbb{E}_x\left(1_{\{T_2\leq t\}}\left(f(X_t) - f(X_{T_1})\right)\right) \tag{2.6}$$

that gives the bound

$$\left|\frac{1}{t}\left[P_t f(x) - f(x)\right] + Qf(x)\right| \leq c\lambda_x^2 t\|f\|_\infty + \frac{2}{t}\mathbb{P}_x\left(T_2 \leq t\right)\|f\|_\infty. \tag{2.7}$$

Since $\sup_x \lambda_x$ is assumed finite, we have also that $\frac{2}{t}\mathbb{P}_x\left(T_2 \leq t\right) \leq Ct$ (see the following Lemma 2.4.4) and (2.4) follows. □

Lemma 2.4.4. *Given two independent r.v,* $\tau_1 \sim \mathcal{E}(\lambda_1)$*,* $\tau_2 \sim \mathcal{E}(\lambda_2)$*, then* $\mathbb{P}\left(\tau_1 + \tau_2 \leq t\right) \leq Ct^2$*.*

Proof of 2.4.4. Exercise. □

Theorem 5. *The semigroup* $(P_t, t \geq 0)$ *satisfies:*

1. *The* Chapman-Kolmogorov backward equation*:*

$$\partial_t P_t(x,y) = (QP_t)(x,y)\ , \quad P_0(x,y) = \mathbf{1}_x(y)\ , \tag{2.8}$$

2. *The* Chapman-Kolmogorov forward equation*:*

$$\partial_t P_t(x,y) = (P_t Q)(x,y)\ , \quad P_0(x,y) = \mathbf{1}_x(y)\ . \tag{2.9}$$

Furthermore, there exists a unique semigroup satisfying any of the two equations.

Proof of Chapman-Kolmogorov backward equation (2.8).

$$\partial_t(P_t f)(x) = \lim_{s\to 0}\frac{1}{s}(P_{t+s} - P_t)f(x) = \lim_{s\to 0}\frac{1}{s}\left(P_s(P_t f)(x) - P_t f(x)\right) = QP_t f(x). \tag{2.10}$$

□

Proof of Chapman-Kolmogorov forward equation (2.9).

$$\partial_t(P_t f)(x) = \lim_{s\to 0}\frac{1}{s}(P_{t+s} - P_t)f(x) = \lim_{s\to 0}P_t\left(\frac{(P_s f)(x) - f(x)}{s}\right), \tag{2.11}$$

and the bound (2.4) justify the exchange of the $\lim_{s\to 0}$ with P_t and (2.9) follows. □

2.5 Recurrence and transience of Markov processes

The notions of recurrence and transience, already introduced for Markov chains, find natural counterparts in continuous time.

Definition 2.5.1. *Fix $x \in E$. We say that x is transient for X if $\mathbb{P}_x(\{t > 0 : X(t) = x\}$ is unbounded$) = 0$. We say that x is recurrent for X if $\mathbb{P}_x(\{t > 0 : X(t) = x\}$ is unbounded$) = 1$.*
If x is recurrent for X, we say that x is positive recurrent for X if $\lambda_x = 0$ or if $\lambda_x > 0$ and

$$\mathbb{E}_x[R_x] < \infty \ , \qquad \textit{where} \qquad R_x := \inf\{t \geq T_1 : X(t) = x\} \ ,$$

otherwise we say that x is null recurrent.

As we will see, any state x is either recurrent or transient.

Remark 2.5.2. *An equivalent definition of recurrence/transience would be: x is recurrent if $\lambda_x = 0$ or if $\lambda_x > 0$ and $\mathbb{P}_x(R_x < \infty) = 1$; otherwise x is transient.*

Remark 2.5.3. *Note that if $\lambda_x = 0$, then the process X stays at x at all times if $X(0) = x$.*

Recall that Y is the Markov chain embedded in X.

Proposition 2.5.4. *Any state x is either transient or recurrent for X. Furthermore:*

1. *x is transient for X if and only x is transient for Y,*
2. *x is recurrent for X if and only x is recurrent for Y.*

In addition, if $0 < \inf_x \lambda_x \leq \sup_x \lambda_x < \infty$, then x is positive recurrent (respectively null recurrent) for X if and only if x is positive recurrent (respectively null recurrent) for Y.

Proof. Suppose x is transient for Y. Then there exists a random integer N such that almost surely $Y_n \neq x$ for all $n \geq N$. Consequently, almost surely $X(t) \neq x$ for all $t \geq T_N$, and therefore almost surely $\{t > 0 : X(t) = x\}$ is bounded. This shows that x is transient for X.
Suppose x is recurrent for Y. Then there exists an infinite random sequence $N_1 < N_2 < \ldots$ such that almost surely $Y_{N_k} = x$ for all $k \geq 1$. Consequently, almost surely $X(T_{N_k}) = x$ for all $k \geq 1$. From the construction of Markov processes, we know that almost surely $T_{N_k} \to \infty$ as $k \to \infty$. Therefore almost surely $\{t > 0 : X(t) = x\}$ is unbounded: x is recurrent for X.
Since any state x is either transient or recurrent for Y, we thus deduce that the same holds for X and that there is an equivalence between transience/recurrence for X and Y.

Let us introduce $R_x^Y := \inf\{n > 0 : Y_n = x\}$. Recall that x is positive recurrent for Y if $\mathbb{E}_x[R_x^Y] < \infty$, and x is null recurrent for Y if $\mathbb{E}_x[R_x^Y] = \infty$. From the construction of Markov processes, we have almost surely

$$R_x := \sum_{n=1}^{R_x^Y} \tau_n \ .$$

Now set $\bar{\lambda} := \sup_x \lambda_x$ and $\underline{\lambda} := \inf_x \lambda_x$. We have almost surely

$$\tau_n \underline{\lambda} \leq E_n \leq \tau_n \bar{\lambda} \ .$$

Consequently,

$$R_x\underline{\lambda} \leq \sum_{n=1}^{R_x^Y} E_n \leq R_x\bar{\lambda} \ .$$

Since R_x^Y is mesurable w.r.t. $\sigma(Y)$ and since $(E_n)_{n\geq 1}$ is independent of $\sigma(Y)$, we deduce that

$$\mathbb{E}_x[\sum_{n=1}^{R_x^Y} E_n] = \mathbb{E}_x[\mathbb{E}_x[\sum_{n=1}^{R_x^Y} E_n \,|\, Y]] = \mathbb{E}_x[R_x^Y \mathbb{E}[E_1]] = \mathbb{E}_x[R_x^Y] \ .$$

We therefore obtain

$$\mathbb{E}_x[R_x]\underline{\lambda} \leq \mathbb{E}_x[R_x^Y] \leq \mathbb{E}_x[R_x]\bar{\lambda} \ .$$

This shows that $\mathbb{E}_x[R_x]$ is finite if and only if $\mathbb{E}_x[R_x^Y]$ is finite. □

Let us finally introduce the notion of irreducibility.

Definition 2.5.5. *We say that X is irreducible if for all $x \neq y \in E$ there exists $t > 0$ such that*

$$\mathbb{P}_x(X_t = y) > 0 \ .$$

We have the following result whose proof is admitted.

Proposition 2.5.6. *If X is irreducible then for all $x \neq y \in E$ and all $t > 0$ we have*

$$\mathbb{P}_x(X_t = y) > 0 \ .$$

Furthermore, X is irreducible if and only if Y is irreducible.

We then deduce that if X is irreducible then all states are either recurrent or transient. Indeed, if X is irreducible then Y is irreducible too: all states of Y are either recurrent or transient. From the previous proposition, we deduce the asserted property.
Therefore we say that an irreducible Markov process X is recurrent if one state (and therefore all states) is (are) transient.

2.6 Invariant measure

From now on, we will assume that $\lambda_x > 0$ for all $x \in E$ in order to avoid "pathological" cases. Of course, all the results presented below can be adapted to encompass the general setting but at the price of complexifying the statements and the proofs.

Definition 2.6.1. *Let μ be a positive measure on E and let X be a Markov process. We say that μ is invariant for X if $\mu Q = 0$, that is, if for all $x \in E$*

$$\mu(x)\lambda_x = \sum_{y\neq x} \mu(y)\lambda_y \Pi(y,x) \ .$$

The next result shows that any invariant measure is invariant for the embedded Markov chain Y and vice versa.

Lemma 2.6.2. *Let μ and ν be two positive measures on E satisfying for all $x \in E$*

$$\lambda_x \mu(x) = \nu(x) \ .$$

Then $\mu Q = 0$ if and only if $\nu\Pi = \nu$.

Proof. We have for any $y \in E$

$$\mu Q(y) = \sum_x \mu(x) Q(x,y) = -\mu(y)\lambda_y + \sum_{x \neq y} \mu(x)\lambda_x \Pi(x,y) = -\nu(y) + \sum_{x\neq y} \nu(x)\Pi(x,y) \ .$$

From the above identity we deduce that $\mu Q(y) = 0$ if and only if $\nu(y) = \nu\Pi(y)$, this proves the lemma. □

We now present results on existence and uniqueness of invariant measure in the irreducible and recurrent case.

Theorem 6. *Assume that X is an irreducible and recurrent Markov process. Then for any given $x \in E$, the measure $\mu^{(x)}$ defined by*

$$\mu^{(x)}(y) = \mathbb{E}_x\Big[\int_0^{R_x} \mathbf{1}_{\{X_t = y\}} dt\Big] \ , \quad y \in E \ , \tag{2.12}$$

is an invariant measure. Furthermore, any invariant measure μ' satisfies $\mu' = c\mu^{(x)}$ for some $c > 0$.

This theorem implies that if X is irreducible and recurrent, then it admits at most one invariant probability measure: indeed, either $\mu^{(x)}$ has infinite mass and there is no invariant probability (even finite) measure, or $\mu^{(x)}$ has finite mass and there exists exactly one $c > 0$ that makes μ' a probability measure.

Proof. We recall that the measure

$$\nu^{(x)}(y) = \mathbb{E}_x\Big[\sum_{n=0}^{R_x^Y - 1} \mathbf{1}_{\{Y_n = y\}}\Big] \ , \quad y \in E \ ,$$

is invariant for the Markov chain Y and that any measure ν' that is invariant for Y is of the form $\nu' = c\nu^{(x)}$ for some $c > 0$ (see Theorem 3.9 from the course "Processus discrets"). If we show that $\mu^{(x)}$ defined in (2.14) satisfies $\mu^{(x)}(y) := \nu^{(x)}(y)/\lambda_y$ for all $y \in E$, then we will deduce from Lemma 2.6.2 the statement of the proposition.
Recall that τ_n is the random time that X spends in state Y_{n-1} before jumping to Y_n. We have

$$\int_0^{R_x} \mathbf{1}_{\{X_t = y\}} dt = \sum_{n=0}^{R_x^Y - 1} \tau_{n+1} \mathbf{1}_{\{Y_n = y\}} = \sum_{n \geq 0} \tau_{n+1} \mathbf{1}_{\{Y_n = y \ ; \ n < R_x^Y\}} \ ,$$

so that Fubini's Theorem yields

$$\mathbb{E}_x\Big[\int_0^{R_x} \mathbf{1}_{\{X_t = y\}} dt\Big] = \sum_{n \geq 0} \mathbb{E}_x\big[\tau_{n+1} \mathbf{1}_{\{Y_n = y \ ; \ n < R_x^Y\}}\big] \ .$$

From the construction of Markov processes, we know that conditionally given the process Y the r.v. $(\tau_n)_{n\geq 1}$ are IID $\mathcal{E}(\lambda_{Y_{n-1}})$. Consequently

$$\begin{aligned}\mathbb{E}_x\big[\tau_{n+1}\mathbf{1}_{\{Y_n=y\,;\,n<R_x^Y\}}\big] &= \mathbb{E}_x\big[\mathbb{E}_x\big[\tau_{n+1}\mathbf{1}_{\{Y_n=y\,;\,n<R_x^Y\}}\,|\,Y\big]\big] \\ &= \mathbb{E}_x\big[\mathbf{1}_{\{Y_n=y\,;\,n<R_x^Y\}}\mathbb{E}_x\big[\tau_{n+1}\,|\,Y\big]\big] \\ &= \mathbb{E}_x\big[\frac{1}{\lambda_{Y_n}}\mathbf{1}_{\{Y_n=y\,;\,n<R_x^Y\}}\big] \\ &= \frac{1}{\lambda_y}\mathbb{E}_x\big[\mathbf{1}_{\{Y_n=y\,;\,n<R_x^Y\}}\big]\ .\end{aligned}$$

Consequently

$$\begin{aligned}\mu^{(x)}(y) = \mathbb{E}_x\Big[\int_0^{R_x}\mathbf{1}_{\{X_t=y\}}dt\Big] &= \frac{1}{\lambda_y}\sum_{n\geq 0}\mathbb{E}_x\big[\mathbf{1}_{\{Y_n=y\,;\,n<R_x^Y\}}\big] \\ = \frac{1}{\lambda_y}\mathbb{E}_x\Big[\sum_{n=0}^{R_x^Y-1}\mathbf{1}_{\{Y_n=y\}}\Big] & \\ = \frac{1}{\lambda_y}\nu^{(x)}(y)\ . &\end{aligned}$$

□

Proposition 2.6.3. *Let X be irreducible and recurrent. Let μ' be a measure on E. Then $\mu' Q = 0$ if and only if $\mu' P_t = \mu'$ for some $t > 0$ if and only if $\mu' P_t = \mu'$ for all $t > 0$.*

This proposition gives a more intuitive meaning to the notion of invariant measure: in the irreducible and recurrent case, a measure is invariant if and only if the law of the process starting from this measure is the same at all times.

Proof. Assume that $\mu' Q = 0$. By Theorem 6, there exists $c > 0$ such that $\mu' = c\mu^{(x)}$ where $\mu^{(x)}$ is defined in (2.14). Therefore, it suffices to show $\mu^{(x)} P_t = \mu^{(x)}$ for all $t > 0$. Fix $t > 0$. By the strong Markov property applied at time R_x, we have

$$\mathbb{E}_x\Big[\int_0^t\mathbf{1}_{\{X_s=y\}}ds\Big] = \mathbb{E}_x\Big[\int_{R_x}^{R_x+t}\mathbf{1}_{\{X_s=y\}}ds\Big]\ .$$

Consequently

$$\begin{aligned}\mathbb{E}_x\Big[\int_0^{R_x}\mathbf{1}_{\{X_s=y\}}ds\Big] &= \mathbb{E}_x\Big[\int_0^t\mathbf{1}_{\{X_s=y\}}ds\Big] + \mathbb{E}_x\Big[\int_t^{R_x}\mathbf{1}_{\{X_s=y\}}ds\Big] \\ &= \mathbb{E}_x\Big[\int_{R_x}^{R_x+t}\mathbf{1}_{\{X_s=y\}}ds\Big] + \mathbb{E}_x\Big[\int_t^{R_x}\mathbf{1}_{\{X_s=y\}}ds\Big] \\ &= \mathbb{E}_x\Big[\int_t^{R_x+t}\mathbf{1}_{\{X_s=y\}}ds\Big] \\ &= \mathbb{E}_x\Big[\int_0^{\infty}\mathbf{1}_{s<R_x}\mathbf{1}_{\{X_{s+t}=y\}}ds\Big]\ .\end{aligned}$$

We then compute

$$
\begin{aligned}
\mu^{(x)}(y) &= \mathbb{E}_x\Big[\int_0^\infty \mathbf{1}_{s<R_x}\mathbf{1}_{\{X_{s+t}=y\}}ds\Big] \\
&= \int_{\mathbf{R}_+} \mathbb{P}_x(X_{s+t}=y, s<R_x)ds \\
&= \int_{\mathbf{R}_+} \sum_{z\in E} \mathbb{P}_x(X_s=z, s<R_x)\mathbb{P}_z(X_t=y)ds \\
&= \sum_{z\in E} \mathbb{E}_x\Big[\int_0^{R_x} \mathbf{1}_{\{X_s=z\}}ds\Big]P_t(z,y) \\
&= \sum_{z\in E} \mu^{(x)}(z)P_t(z,y) \\
&= \mu^{(x)}P_t(y) \ .
\end{aligned}
$$

We have therefore proved that $\mu'Q=0$ implies $\mu'P_t=\mu'$ for all $t>0$.
If $\mu'P_t=\mu'$ for all $t>0$, then of course, $\mu'P_t=\mu'$ for some $t>0$.
It remains to prove that if $\mu'P_t=\mu'$ for some $t>0$ then $\mu'Q=0$: this will be proven in the exercises. □

Proposition 2.6.4. *Let X be irreducible. The following are equivalent:*

1. *All states x are positive recurrent,*
2. *One state x is positive recurrent,*
3. *There exists an invariant probability measure μ'.*

If one these conditions hold, then the invariant probability measure μ' is unique and is given by

$$
\mu'(x) := \frac{1}{\lambda_x\mathbb{E}_x[R_x]} \ , \quad x\in\mathbb{E} \ .
$$

To prove the proposition, we will need an intermediate fact on Markov chains, whose proof will be given in the exercises.

Lemma 2.6.5. *Fix $x\in E$. Let us introduce the measure*

$$
\nu^{(x)}(y) := \mathbb{E}_x[\sum_{n=0}^{R_x^Y-1} \mathbf{1}_{\{Y_n=y\}}] \ , \quad y\in E \ .
$$

If ν' satisfies $\nu'\Pi=\nu'$ and $\nu'(x)=1$ then $\nu'\geq\nu^{(x)}$.

Proof of Proposition 2.6.4. Recall the measure $\mu^{(x)}$ of (2.14). In the proof of Theorem 6 we have shown that for all $y\in E$

$$
\mu^{(x)}(y) = \frac{1}{\lambda_y}\nu^{(x)}(y) \ .
$$

Let us now assume that there exists an invariant probability measure μ'. Fix $x \in E$. We claim that $\mu'(x)\lambda_x > 0$. We postpone the proof of this claim and carry on the proof. We define

$$\nu'(y) := \frac{\mu'(y)\lambda_y}{\mu'(x)\lambda_x} \;, \quad y \in E \;.$$

Observe that the measure $y \mapsto \mu'(y)/(\mu'(x)\lambda_x)$ is invariant for X. By Lemma 2.6.2, we deduce that $\nu'\Pi = \nu'$. Moreover, $\nu'(x) = 1$. Consequently by Lemma 2.6.5

$$\nu' \geq \nu^{(x)} \;.$$

Using the identity recalled at the beginning of the proof, we obtain

$$\mathbb{E}_x[R_x] = \sum_y \mu^{(x)}(y) = \sum_y \frac{\nu^{(x)}(y)}{\lambda_y} \leq \sum_y \frac{\nu'(y)}{\lambda_y} = \sum_y \frac{\mu'(y)}{\mu'(x)\lambda_x} = \frac{1}{\mu'(x)\lambda_x} \;.$$

Given the claim, this last quantity is finite and therefore x is positive recurrent. Since x was arbitrary, we deduce that all states are positive recurrent.
Let us prove the claim. First of all, for any $y \in E$ we have $\lambda_y > 0$ by assumption (this is the standing assumption of this section). If $\mu'(x) = 0$, then the fact that $\mu' Q(x) = 0$ implies that

$$\mu'(x)Q(x,x) + \sum_{y \neq x} \mu'(y)Q(y,x) = \sum_{y\neq x} \mu'(y)Q(y,x) = 0 \;,$$

and then $\mu'(y) = 0$ whenever $Q(y,x) > 0$, that is, whenever $\Pi(y,x) > 0$. Iterating this argument, we see that $\mu'(y) = 0$ whenever $\Pi^n(y,x) = 0$ for some $n \geq 0$. Since Y is irreducible

$$\cup_{n\geq 0}\{y \in E : \Pi^n(y,x) > 0\} = E \;.$$

Consequently, $\mu'(y) = 0$ for all $y \in E$ thus contradicting the fact that μ' is a probability measure.

If all states are positive recurrent, then obviously there exists one state which is positive recurrent.

Let us now assume that some x is positive recurrent. Then the measure $\mu^{(x)}$ of (2.14) satisfies

$$\begin{aligned}
\mu^{(x)}(E) &= \sum_{y\in E} \mu^{(x)}(y) \\
&= \sum_{y\in E} \mathbb{E}_x\Big[\int_0^{R_x} \mathbf{1}_{\{X_t = y\}} dt\Big] \\
&= \mathbb{E}_x\Big[\int_0^{R_x} \sum_{y \in E} \mathbf{1}_{\{X_t=y\}} dt\Big] \\
&= \mathbb{E}_x[R_x] \;,
\end{aligned}$$

where we used Fubini's Theorem to go from the second to the third line. Since x is positive recurrent we deduce that $\mu^{(x)}$ is a finite measure. By Theorem 6, we deduce that there exists a *unique* invariant probability measure μ', which is given by $\mu' = c\mu^{(x)}$ with $c = 1/\mathbb{E}_x[R_x]$. The identity recalled at the beginning of the proof shows that $\mu^{(x)}(x) = \nu^{(x)}(x)/\lambda_x$. Since $\nu^{(x)}(x) = 1$, we deduce that

$$\mu'(x) = \frac{1}{\lambda_x \mathbb{E}_x[R_x]} \;.$$

Since we showed that the existence of an invariant probability measure implies that all states are positive recurrent, the above identity holds for *all* x, thus concluding the proof. □

A consequence of the last result is that an irreducible and transient Markov process does never admit a finite invariant measure.

2.7 General remarks on Markov processes

Generally speaking, the theory of Markov processes is delicate. Here we have concentrated on the case where the state-space E is countable, but uncountable state-spaces are relevant too (for instance, the Brownian motion takes values in **R** and satisfies the Markov property).

Our presentation did not encompass *all* Markov processes taking values in countable spaces: indeed, we assumed that our Markov processes are càdlàg and this assumption is actually restrictive. In particular, Markov processes that explode in finite time, that is, Markov processes for which $\zeta = \lim_{n\to\infty} T_n < \infty$, do not admit a left limit at ζ and therefore do not satisfy this assumption. However, explosive Markov processes appear in many different situations. If one replaces càdlàg by simply right-continuous in Definition 2.1.4, then explosive processes are allowed.

More generally, one can suppress the regularity assumption on the trajectories: in that case, the definition of Markov processes allow for "monsters". For instance, if one considers a collection X_t, $t \geq 0$ of IID random variables then X is a Markov process but its trajectories do not have any regularity (except if the law of X_t is trivial).

Chapter 3

Some examples of Markov processes

3.1 Queueing theory

In this section, we will study stochastic processes that model a *queue* in a service unit. A service unit is made of one or several servers. Customers arrive at random times in the queue. If a server is available at the time a customer arrives, then the customer goes to that server - otherwise it waits in the queue. Each customer requires some *service time*: this is the duration of a time required for the customer to be served. The service unit has a maximal capacity (possibly infinite). The quantity of interest in this model is the total number of customers that are either being served or are waiting in the queue.

There is a standard terminology to specify the parameters/characteristics of a queue. It consists of [A] / [S] / [s] / [c] / [Discipline] where :

- A indicates the distribution of the interarrival times of customers. It can be G (general), nothing is specified ; GI (general independent), that is, the interarrival times are IID; M (Markov), the interarrival times are IID with exponential distribution; D (deterministic), the interarrival times are deterministic.

- S indicates the distribution of the service time. The possible values are the same as for A.

- s is the total number of servers in the service unit: it's either an integer or $+\infty$,

- c is capacity of the service unit, that is, the maximal number of customers in the queue it's either an integer or $+\infty$,

- Discipline is the service discipline: usually it is either FIFO (first in first out), which means that among all customers that are waiting in the queue the first one who arrived will be the first one to be served, or LIFO (last in first out), which means that the last customer who arrived is served first.

By abuse of notation, the word "queue" is often used insead of "service unit".

Let us give an example. M/GI/1/∞/FIFO denotes the queue where: customers arrive according to a Poisson process, service times are IID, there is only one server, the service unit has infinite capacity, the service discipline is FIFO.

If we do not specify the last two parameters, then they are implicitly taken to be ∞ and FIFO.

The total number of customer in the service unit at time t is denoted $X(t)$.

3.1.1 M/M/1

We first examine the queue M/M/1. We let $\lambda > 0$ be the parameter of the exponential r.v. associated to the arrival of new customers and $\gamma > 0$ the parameter of the exponential r.v. for the duration of the service times. If there are n customers in the queue at time t, that is, if $X(t) = n$ then:

- Either $n \geq 1$. Then, the process X will jump by 1 after an exponential time of parameter λ or by -1 after an exponential time of parameter γ. Consequently, the next jump of X occurs at the minimum of two independent exponential r.v. of parameters λ and γ, which is itself an exponential r.v. of parameter $\lambda + \gamma$.
- Or $n = 0$. Then, the process X will jump by $+1$ after an exponential time of parameter λ.

This discussion suffices to deduce that X is a Markov process with values in $E = \mathbf{N}$. Its transition matrix Π and transition rates λ are given by

$$\lambda_n = \lambda + \gamma \ , \qquad \Pi(n, n+1) = \frac{\lambda}{\lambda + \gamma} \ , \qquad \Pi(n, n-1) = \frac{\gamma}{\lambda + \gamma} \ , \qquad n \geq 1 \ ,$$

and

$$\lambda_0 = \lambda \ , \qquad \Pi(0, 1) = 1 \ .$$

(For all other values of n, m, $\Pi(n, m) = 0$.)

Remark 3.1.1. *The process X is "almost" a compound Poisson process of intensity $\lambda + \gamma$ and jump law $\lambda\delta_{+1} + \mu\delta_{-1}$. This is true when the process is strictly positive.*

It is clear that X is irreducible.

We turn to the investigation of invariant measures. Let us set $\rho = \lambda/\gamma$. Given Lemma 2.6.2, we start with the invariant measure for the embedded Markov chain Y.

Proposition 3.1.2. *A measure ν is invariant for Y if and only if it satisfies*

$$\nu(n) = \rho^{n-1}(1 + \rho)\nu(0) \ , \quad n \geq 1 \ .$$

Proof. The measure ν is invariant for Y if and only if $\nu = \nu\Pi$. The later identity is equivalent to

$$\nu(0) = \nu(1)\Pi(1, 0) \ , \quad \nu(n) = \nu(n-1)\Pi(n-1, n) + \nu(n+1)\Pi(n+1, n) \ , \quad n \geq 1 \ .$$

This can be rewritten as

$$(1 + \rho)\nu(0) = \nu(1) \ , \quad \nu(n) = \nu(n-1)\frac{\rho}{1 + \rho} + \nu(n+1)\frac{1}{1 + \rho} \ , \quad n \geq 1 \ . \tag{3.1}$$

Consider a measure ν that satisfies the condition of the statement of the proposition. Then, it is easy to check that (3.1) is satisfied. Conversely, assume that (3.1) is satisfied, and let us prove by recursion that

$$\nu(n) = \rho^{n-1}(1 + \rho)\nu(0) \ , \quad n \geq 0 \ .$$

At rank $n = 1$ this holds true. Assume that this is true up to some rank $n \geq 1$. Then, at rank $n + 1$ we have

$$\nu(n + 1) = (1 + \rho)\nu(n) - \rho\nu(n - 1) = \rho^2\nu(n - 1) = \rho^n(1 + \rho)\nu(0) \ .$$

This concludes the proof. □

We therefore deduce from Lemma 2.6.2 that the only invariant measures for X are the measures μ that satisfy

$$\mu(n) := \rho^n \mu(0) , \quad n \geq 1 .$$

We deduce that X admits an invariant probability measure if and only if $\rho < 1$. If this condition holds, then the invariant probability measure is Geometric with parameter ρ. From Proposition 2.6.4, we deduce that X is positive recurrent if and only if $\rho < 1$.

Proposition 3.1.3. *The following holds:*

1. *If $\rho < 1$, then X is positive recurrent,*
2. *If $\rho = 1$, then X is null recurrent,*
3. *If $\rho > 1$, then X is transient.*

To prove this proposition, we start with an auxiliary lemma. Let $(Z_n, n \geq 0)$ be a discrete-time Markov chain with transition matrix Π^Z given by

$$\Pi^Z(n, n+1) = \lambda/(\lambda + \gamma) , \quad \Pi^Z(n, n-1) = \gamma/(\lambda + \gamma) , \qquad \forall n \in \mathbf{Z} ,$$

and $\Pi^Z(n, m) = 0$ for all other values of m, n.

Lemma 3.1.4. *For any $q \in \mathbf{R}$, let $c_q := e^q \rho(1+\rho)^{-1} + e^{-q}(1+\rho)^{-1}$. The process $M_n, n \geq 0$ is a martingale where*

$$M_n := e^{qZ_n} c_q^{-n} , \quad n \geq 0 .$$

Furthermore, if we let $T_0 := \inf\{n \geq 1 : Z_n = 0\}$ we have for all $q \in (-\infty, \ln(\rho^{-1} \wedge 1))$

$$\mathbb{E}_1[\exp(-T_0 \ln c_q)] = e^q .$$

As a consequence $\mathbb{P}(T_0 < \infty) = 1 \wedge \rho^{-1}$ and $\mathbb{E}[T_0] < \infty$ if and only if $\rho < 1$.

Proof. Note that we have the deterministic bound $|Z_n - Z_0| \leq n$ for all $n \geq 0$ so that M_n is integrable for all $n \geq 0$. Regarding the martingale property, we have

$$\begin{aligned}
\mathbb{E}[e^{qZ_{n+1}} \,|\, \mathcal{F}_n] &= e^{qZ_n}\big(e^q \lambda/(\lambda+\gamma) + e^{-q}\gamma/(\lambda+\gamma)\big) \\
&= e^{qZ_n}(e^q \rho(1+\rho)^{-1} + e^{-q}(1+\rho)^{-1}) .
\end{aligned}$$

Consequently $\mathbb{E}[M_{n+1} \,|\, \mathcal{F}_n] = M_n$.
A simple computation shows that $(-\infty, \ln(\rho^{-1} \wedge 1)] \ni q \mapsto c_q$ is decreasing and equals 1 at $\ln(\rho^{-1} \wedge 1)$. Consequently, $(M_{n \wedge T_0}, n \geq 0)$ is a bounded martingale for all $q \in (-\infty, \ln(\rho^{-1} \wedge 1)]$. By the Stopping Theorem, we deduce that

$$\mathbb{E}_1[M_{n \wedge T_0}] = \mathbb{E}_1[M_0] = e^q .$$

On the event $\{T_0 = +\infty\}$, we have $Z_n \geq 1$ for all $n \geq 1$. Since $q \leq 0$, we have $M_n \leq c_q^{-n}$ so that $M_n \to \infty$ as $n \to \infty$ on this same event. Consequently, almost surely

$$M_{n \wedge T_0} \to M_{T_0} \mathbf{1}_{\{T_0 < \infty\}} , \quad n \to \infty .$$

By the Dominated Convergence Theorem, we thus deduce that

$$\mathbb{E}[M_{T_0} \mathbf{1}_{\{T_0 < \infty\}}] = e^q .$$

This yields for $q < \ln(\rho^{-1} \wedge 1)$

$$\mathbb{E}[c_q^{-T_0}] = e^q \ ,$$

and for $q = \ln(\rho^{-1} \wedge 1)$

$$\mathbb{P}(T_0 < \infty) = 1 \wedge \rho^{-1} \ .$$

This immediately ensures that for $\rho > 1$, $\mathbb{E}[T_0] = +\infty$. Let us now assume that $\rho \leq 1$, in which case $\ln(\rho^{-1} \wedge 1) = 0$. For any $q < 0$, by the Differentiation Theorem under the integral we have

$$\mathbb{E}[-T_0 \frac{c_q'}{c_q} c_q^{-T_0}] = e^q \ .$$

Rearranging terms, we get

$$\mathbb{E}[T_0 c_q^{-T_0}] = -\frac{c_q}{c_q'} e^q \ .$$

By the Monotone Convergence Theorem we have

$$\lim_{q \uparrow 0} \mathbb{E}[T_0 c_q^{-T_0}] = \mathbb{E}[T_0] \ .$$

On the other hand

$$\lim_{q \uparrow 0} -\frac{c_q}{c_q'} e^q = \frac{1+\rho}{1-\rho} \ .$$

We thus deduce that $\mathbb{E}[T_0] < \infty$ if and only if $\rho < 1$. □

Proof of Proposition 3.1.3. Since the jump rates are bounded from above and below, Proposition 2.5.4 ensures that it suffices to determine the recurrence/transience nature of the irreducible Markov chain Y. Let us determine the recurrence/transience of the state $x = 0$: by irreducibility, this suffices to deduce the recurrence/transience of the whole chain. Let us observe that, if $Y_0 = 0$ then $Y_1 = 1$. Consequently R_0^Y under $\mathbb{P}_0$ has the same law as $1 + R_0^Y$ under $\mathbb{P}_1$. Furthermore, since the processes Y and Z have the same law up to their first hitting time of 0, we deduce that R_0^Y under $\mathbb{P}_1$ has the same law as T_0 for Z starting from 1. From the previous lemma, we deduce that $\mathbb{P}_1(R_0^Y < \infty) = 1$ if and only if $\rho \leq 1$. Consequently 0 is recurrent if and only if $\rho \leq 1$, and therefore Y is recurrent if and only if $\rho \leq 1$. Furthermore, $\mathbb{E}_1[R_0^Y] < \infty$ if and only if $\rho < 1$. Consequently 0 is positive recurrent for Y if and only if $\rho < 1$. By Proposition 2.5.4, 0 is positive recurrent for X if and only if $\rho < 1$. By Proposition 2.6.4 we deduce that X is positive recurrent if and only if $\rho < 1$. □

Hence we see that, if $\rho > 1$ then almost surely the queue is never empty (X never hits 0) after some random time, while if $\rho \leq 1$ then the set of times at which the queue is empty is unbounded.

3.1.2 M/M/s and M/M/∞

Let us examine the M/M/s queue with $s \in \mathbf{N} \cup \{+\infty\}$. If there are n customers in the queue at time t, that is, if $X(t) = n$ then:

- Either $n \geq 1$. Then, the process X will jump by 1 after an exponential time of parameter λ or by -1 after an exponential time of parameter $\min(n, s)\gamma$. Consequently, the next jump of X occurs at the minimum of two independent exponential r.v. of parameters λ and $\min(n, s)\gamma$, which is itself an exponential r.v. of parameter $\lambda + \min(n, s)\gamma$.

- Or $n = 0$. Then, the process X will jump by $+1$ after an exponential time of parameter λ.

This discussion suffices to deduce that X is a Markov process with values in $E = \mathbf{N}$. Its transition matrix Π and transition rates λ are given by

$$\lambda_n = \lambda + \min(n,s)\gamma \ , \qquad \Pi(n,n+1) = \frac{\lambda}{\lambda + \min(n,s)\gamma} \ , \qquad \Pi(n,n-1) = \frac{\min(n,s)\gamma}{\lambda + \min(n,s)\gamma} \ , \qquad n \geq 1 \ ,$$

and

$$\lambda_0 = \lambda \ , \qquad \Pi(0,1) = 1 \ .$$

(For all other values of n, m, $\Pi(n,m) = 0$.)

We will concentrate on the case $s = \infty$. Recall that $\rho = \lambda/\gamma$.

Proposition 3.1.5. *When $s = \infty$, the process X admits an invariant probability measure given by*

$$\mu(n) = \frac{\rho^n}{n!} e^{-\rho} \ , \quad n \geq 0 \ .$$

As a consequence X is recurrent positive.

The proof will be the content of an exercise.

3.2 Branching processes

Let ξ be a probability law on $\mathbf{N}$, called the *offspring distribution* and denote by $\phi(r) := \sum_{k \geq 0} r^k \xi(k)$ its generating function. We start with the following Markov chain Z, usually called a *Galton-Watson* process. If $Z_n = k$, then each of the k individuals is replaced at time $n+1$ by a random number of children distributed according to ξ. The associated transition matrix is given by

$$\Pi^Z(1,k) = \xi(k) \ , \quad k \in \mathbf{N} \ ,$$

and

$$\Pi^Z(n,k) = \sum_{k_1, k_2, \ldots, k_n : k_1 + \ldots + k_n = k} \xi(k_1) \ldots \xi(k_n) \ , \quad n \geq 2, k \geq 0 \ ,$$

and $\Pi^Z(0,0) = 1$.

Proposition 3.2.1. *We have for all $r \in (0,1]$*

$$\mathbb{E}_x[r^{Z_n}] = (\phi_n(r))^x \ ,$$

where $\phi_n(r) = \phi \circ \phi_{n-1}(r)$ and $\phi_0(r) = r$.

We now consider a continuous-time process associated to the above model. Fix $c > 0$. To each individual we associate a random liftetime distributed according to an exponential law of parameter $c > 0$. We call $X(t), t \geq 0$ the corresponding process.

The following result is admitted.

Proposition 3.2.2. *For all* $r \in (0,1)$ *and all* $t \geq 0$*, we have*

$$\mathbb{E}_x[r^{X(t)}] = F(t,r)^x \ ,$$

where $F(t,r)$ *is the unique value satisfying*

$$\int_r^{F(t,r)} \frac{dy}{c(\phi(y) - y)} = t \ .$$

Introduce $\tau_\infty = \inf\{t \geq 0 : X(t) = \infty\}$ and $\tau_0 = \inf\{t \geq 0 : X(t) = 0\}$.

Lemma 3.2.3. *We have for all* $t > 0$

$$\mathbb{P}_x(\tau_0 \leq t) = F(t, 0+)^x \ , \quad \mathbb{P}_x(\tau_\infty \leq t) = 1 - F(t, 1-)^x \ .$$

At least intuitively, X is a Markov process with generator

$$Q(n,n) = -cn \ , \quad Q(n, n-1+k) = cn\xi(k) \ .$$

However, our definition of Markov process does not allow for explosion in finite time and therefore, only in the case where $\tau_\infty = \infty$ almost surely we can apply our definition.

Chapter 4

Introduction to Renewal Theory

4.1 Renewal processes

Recall that a renewal process $(T_n)_{n\geq 1}$ is a non-decreasing sequence defined by setting:

$$T_n = \delta_1 + \ldots + \delta_n \;, \quad n \geq 1 \;,$$

where $(\delta_n)_{n\geq 1}$ is a sequence of i.i.d. positive random variables.

Proposition 4.1.1. *Let $(T_n)_{n\geq 1}$ be a renewal process and set*

$$N_t = \sum_{n\geq 1} \mathbf{1}_{\{T_n \leq t\}} \;, \quad t \geq 0 \;.$$

Then, N is a standard counting process.

Proof. Since $\delta_n \in (0,\infty)$ for all $n \geq 1$, $\mathbb{P}$-a.s., we deduce that $(T_n)_{n\geq 0}$ is increasing and $T_n < \infty$ for all $n \geq 0$, $\mathbb{P}$-a.s.
By the law of large numbers applied to the sequence $(\min(\delta_n, 1))_{n\geq 1}$, we have $\mathbb{P}$-a.s.

$$\frac{1}{n}\sum_{i=1}^{n} \min(\delta_i, 1) \to \mathbb{E}[\min(\delta_1, 1)] \;.$$

Notice that $\mathbb{E}[\min(\delta_1, 1)] > 0$ since, otherwise, we would have $\delta_1 = 0$ almost surely. Since $T_n \geq \sum_{i=1}^{n} \min(\delta_i, 1)$, we deduce that $T_n \to \infty$ as $n \to \infty$, $\mathbb{P}$-a.s. Finally, given the definition of N, it is simple to check that it is a counting process (left as an exercise). □

Lemma 4.1.2 (Wald's identity)**.** *Let X_n be a sequence of i.i.d. random variables. Let R be a stopping time in the filtration $\mathcal{F}_n := \sigma(X_1, \ldots, X_n)$. Assume that X_1 and R are integrable. Then, if we set $S_n = X_1 + \ldots + X_n$ for every $n \geq 0$, we have*

$$\mathbb{E}[S_R] = \mathbb{E}[R]\mathbb{E}[X_1] \;.$$

Proof. We set

$$Z := \sum_{n\geq 1} |X_n| \mathbf{1}_{n\leq R} \;.$$

The event $\{n \leq R\}$ is the complement of $\{R \leq n-1\}$ which belongs to $\mathcal{F}_{n-1}$ by assumption. Since X_n is independent from $X_1, \ldots, X_{n-1}$, it is independent from $\mathcal{F}_{n-1}$. Consequently

$$\mathbb{E}[|X_n|\mathbf{1}_{n\leq R}] = \mathbb{E}[|X_n|]\mathbb{P}(n \leq R) = \mathbb{E}[|X_1|]\mathbb{P}(n \leq R) \ .$$

Since R is integrable, we have $\mathbb{E}[R] = \sum_{n\geq 1} \mathbb{P}(n \leq R) < \infty$, and therefore $\mathbb{E}[Z] < \infty$. As Z is integrable, we can apply Fubini's Theorem (the version for non-necessarily positive functions) and deduce that

$$\mathbb{E}[\sum_{n\geq 1} X_n \mathbf{1}_{n\leq R}] = \sum_{n\geq 1} \mathbb{E}[X_n \mathbf{1}_{n\leq R}] \ .$$

By the same argument as above $\mathbb{E}[X_n\mathbf{1}_{n\leq R}] = \mathbb{E}[X_1]\mathbb{P}(n \leq R)$ and therefore

$$\mathbb{E}[\sum_{n\geq 1} X_n \mathbf{1}_{n\leq R}] = \mathbb{E}[X_1]\mathbb{E}[R] \ .$$

□

Corollary 4.1.3. *Let N be a counting process associated with a renewal process. Assume that δ_1 is integrable. Then for every $t \geq 0$, we have*

$$\mathbb{E}[T_{N_t+1}] = \mathbb{E}[N_t + 1]\mathbb{E}[\delta_1] \ .$$

Proof. Fix $m \geq 1$ and take $R = N_t \wedge m + 1$. The event $\{R \leq n\}$ belongs to $\mathcal{F}_n = \sigma(\delta_1, \ldots, \delta_n)$. Indeed, either $n \geq m+1$ in which case $\{R \leq n\} = \Omega$. Or $n < m+1$ in which case this event coincides with $\{N_t \leq n-1\} = \{t < T_n\} \in \mathcal{F}_n$.
Since R is bounded by $m+1$, it is integrable. Applying Wald's identity, we get

$$\mathbb{E}[T_{N_t\wedge m+1}] = \mathbb{E}[N_t \wedge m + 1]\mathbb{E}[\delta_1] \ .$$

Since N is standard, $N_t < \infty$ almost surely. Applying the Monotone Convergence Theorem, we can pass to the limit on $m \to \infty$ and get

$$\mathbb{E}[T_{N_t+1}] = \mathbb{E}[N_t + 1]\mathbb{E}[\delta_1] \ .$$

□

Proposition 4.1.4. *Let N be a counting process associated with a renewal process. Assume that $\mathbb{E}[\delta_1] < \infty$. As $t \to \infty$ we have*

$$\frac{N_t}{t} \to \frac{1}{\mathbb{E}[\delta_1]} \ , \quad \textit{a.s.}$$

and

$$\frac{\mathbb{E}N_t}{t} \to \frac{1}{\mathbb{E}[\delta_1]} \ .$$

Proof. By the strong law of large numbers we know that almost surely

$$\frac{T_n}{n} \to \mathbb{E}[\delta_1] \ ,$$

as $n \to \infty$. By Proposition 1.1.7, we know that N is a standard counting process so that $N_t \to \infty$ as $t \to \infty$ almost surely. Consequently we have the almost sure convergence $T_{N_t}/N_t \to \mathbb{E}[\delta_1]$ as $t \to \infty$. Then we write

$$\frac{N_t + 1}{T_{N_t+1}} \frac{N_t}{N_t + 1} = \frac{N_t}{T_{N_t+1}} \le \frac{N_t}{t} \le \frac{N_t}{T_{N_t}} \ .$$

The leftmost and rightmost terms converge to $1/\mathbb{E}[\delta_1]$ almost surely. This yields the first convergence of the statement.
Applying Fatou's Lemma, we get

$$\liminf_{t\to\infty} \mathbb{E}[N_t/t] \ge \mathbb{E}[\liminf_{t\to\infty} N_t/t] = \frac{1}{\mathbb{E}[\delta_1]} \ .$$

To bound the $\limsup$, we proceed as follows. Let $\delta_i' := \min(\delta_i, K)$ for some $K > 0$ and let N' be the associated counting process. Since the interarrival times of N' are shorter than those of N, we have the almost sure bound $N_t \le N_t'$ for all $t \ge 0$. Thus

$$\limsup_{t\to\infty} \mathbb{E}[N_t/t] \le \limsup_{t\to\infty} \mathbb{E}[N_t'/t] \ .$$

By Corollary 4.1.3, we have

$$\mathbb{E}[T'_{N_t'+1}] = \mathbb{E}[\delta_1']\mathbb{E}[N_t' + 1] \ .$$

Since $T'_{N_t'+1} = T'_{N_t'} + \delta_{N_t'+1} \le t + K$ almost surely, we get

$$\mathbb{E}[N_t'/t] = \frac{1}{t}\Big(\frac{\mathbb{E}[T'_{N_t'+1}]}{\mathbb{E}[\delta_1']} - 1\Big) \le \frac{t+K}{t\mathbb{E}[\delta_1']} \ .$$

Therefore

$$\limsup_{t\to\infty} \mathbb{E}[N_t/t] \le \limsup_{t\to\infty} \mathbb{E}[N_t'/t] \le \frac{1}{\mathbb{E}[\delta_1']} \ .$$

Since $\mathbb{E}[\delta_1'] \uparrow \mathbb{E}[\delta_1]$ as $K \to \infty$, we obtain the desired result. □

We now introduce the forward recurrence time process $(B_t, t \ge 0)$ as follows:

$$B_t := T_{N_t+1} - t \ , \quad t \ge 0 \ .$$

At any time t, the random variable B_t measures the time remaining until the next jump of the counting process N.
The process B takes values in $(0, \infty)$ and its evolution is as follows. If it starts from $b > 0$ at time 0, then it decreases linearly like $b - t$ on $(0, b)$. At time b, it makes a jump of random size distributed like δ_1: from there, it decreases linearly again until its next jump.

4.2 A reminder on convolution of measures

Consider the map $\sigma : \mathbf{R}^2 \to \mathbf{R}$ defined by $\sigma(x, y) = x + y$. Let μ and ν be two finite measures on $\mathbf{R}$. Recall that the product measure $\mu \otimes \nu$ is the unique measure on $\mathcal{B}(\mathbf{R}^2)$ such that for all Borel sets A, B of $\mathbf{R}$, we have

$$\mu \otimes \nu(A \times B) = \mu(A)\nu(B) \ .$$

The convolution $\mu * \nu$ of μ and ν is defined as the pushforward measure (=mesure image) on $\mathbf{R}$ of $\mu \otimes \nu$ through the map σ. In other words, we have for all Borel sets A in $\mathbf{R}$:

$$\mu * \nu(A) = \mu \otimes \nu(\sigma^{-1}(A)) .$$

More generally, for all bounded and measurable function φ we have

$$\int_{x \in \mathbf{R}} \varphi(x) \mu * \nu(dx) = \int_{y \in \mathbf{R}} \int_{z \in \mathbf{R}} \varphi(y+z) \mu(dy) \nu(dz) , \tag{4.1}$$

In the particular case where μ and ν are probability measures, we have the following probabilistic interpretation of the convolution $\mu * \nu$. Let X and Y be two independent r.v. with laws μ and ν. Then the sum $X + Y$ has law $\mu * \nu$.

In the general case, let us mention that if μ or ν has a density, then $\mu * \nu$ has a density as well.

For convenience, we will denote by μ^{*n} the measure obtained by convolving n times μ with itself: in the particular case where $n = 0$, by convention we set $\mu^{*0}(dx) = \delta_0(dx)$ where δ_0 is the Dirac measure at 0: for all Borel set A in $\mathbf{R}$, $\delta_0(A) = 1$ if $0 \in A$, $\delta_0(A) = 0$ otherwise.

4.3 The renewal equation

Let F be a finite measure on $[0, \infty)$ that does not charge 0. Let z be a function on $[0, \infty)$. We introduce the *renewal equation* associated to F and z:

$$Z(t) = z(t) + \int_{[0,t]} Z(t-u) F(du) , \quad t \geq 0 . \tag{4.2}$$

The main theorem of this section ensures existence and uniqueness of the solution of this equation (under some hypothesis). Before we come to this result, let us introduce some further notations.
The measure

$$U(dx) = \sum_{n=0}^{\infty} F^{*n}(dx) ,$$

is called the *renewal measure* and its distribution function

$$U(t) := \int_{[0,t]} U(dx) , \quad t \geq 0 ,$$

is called the *renewal function*.

Remark 4.3.1. *Any measure on $[0, \infty)$ can be uniquely extended into a measure on $\mathbf{R}$ by setting to 0 its total mass on $(-\infty, 0)$: the notion of distribution function that we use here thus coincides with the classical one.*

Let us collect the following simple fact:

Lemma 4.3.2. *The renewal function $U(t)$ is finite for all $t \geq 0$. Furthermore, if the total mass $\int_{[0,\infty)} F(ds)$ of the measure F equals 1, then*

$$U(t) = 1 + \mathbb{E} N_t , \quad t \geq 0 ,$$

where N is a counting process associated with a renewal sequence $(\delta_i)_{i \geq 1}$ with law F (the law of δ_1 is F). Furthermore, we have

$$U(t+a) - U(t) \leq U(a) , \quad \forall t, a \geq 0 . \tag{4.3}$$

Proof. Recall that the measure F does not charge $(-\infty, 0)$. Consequently, it admits a Laplace transform:

$$\int_{[0,\infty)} e^{-qs} F(ds) < \infty \ , \quad \forall q \geq 0 \ .$$

Observe that $e^{-qs} \to 0$ as $q \to \infty$ for all $s > 0$, and that $e^{-qs} \leq 1$. Since F is a finite measure on $(0, \infty)$, the constant 1 is integrable. By the Dominated Convergence Theorem:

$$\int_{[0,\infty)} e^{-qs} F(ds) \to 0 \ , \quad \text{as } q \to \infty \ .$$

Therefore, there exists $q_0 > 0$ such that $\int_{[0,\infty)} e^{-q_0 s} F(ds) \in [0, 1)$. We let δ be the latter quantity.
We claim that for every $n \geq 0$ and every $q \geq 0$, we have the identity

$$\int_{[0,\infty)} e^{-qs} F^{*n}(ds) = \Big(\int_{[0,\infty)} e^{-qs} F(ds) \Big)^n \ .$$

We postpone the proof of the claim, and proceed to the proof of the first part of the lemma. For every $t \geq 0$ and $n \geq 0$, we have

$$\begin{aligned}
F^{*n}(t) = \int_{[0,t]} F^{*n}(ds) &\leq e^{q_0 t} \int_{[0,t]} e^{-q_0 s} F^{*n}(ds) \\
&\leq e^{q_0 t} \int_{[0,\infty)} e^{-q_0 s} F^{*n}(ds) \\
&\leq e^{q_0 t} \Big(\int_{[0,\infty)} e^{-q_0 s} F(ds) \Big)^n \\
&\leq e^{q_0 t} \delta^n \ ,
\end{aligned}$$

so that

$$U(t) = \sum_{n \geq 0} F^{*n}(t) \leq \sum_{n \geq 0} e^{q_0 t} \delta^n < \infty \ .$$

It remains to prove the claim. The case $n = 0$ is trivial since $F^{*n} = \delta_0$. In the case $n = 1$, there is nothing to prove. Let us consider the case $n \geq 2$. By the defining property (4.1) of the convolution, we have

$$\begin{aligned}
\int_{[0,\infty)} e^{-qs} F^{*n}(ds) &= \int_{\mathbf{R}} \mathbf{1}_{[0,\infty)}(s) e^{-qs} F^{*n}(ds) \\
&= \int_{\mathbf{R}} \dots \int_{\mathbf{R}} \mathbf{1}_{[0,\infty)}(s_1 + \ldots + s_n) e^{-q(s_1 + \ldots + s_n)} F(ds_1) \ldots F(ds_n) \\
&= \int_{[0,\infty)} \dots \int_{[0,\infty)} e^{-q(s_1 + \ldots + s_n)} F(ds_1) \ldots F(ds_n) \\
&= \Big(\int_{[0,\infty)} e^{-qs_1} F(ds_1) \Big)^n \ ,
\end{aligned}$$

thus concluding the proof of the first part of the lemma.
We now assume that $\int_{[0,\infty)} F(ds) = 1$. Let $(\delta_i)_{i \geq 1}$ be an i.i.d. sequence with law F, let $T_n = \delta_1 + \ldots + \delta_n$

for every $n \geq 1$ and let N be the associated counting process. Notice that the law of T_n is F^{*n}. We have for all $t \geq 0$

$$\begin{aligned}\mathbb{E}N_t &= \sum_{n\geq 1}\mathbb{P}(N_t \geq n) = \sum_{n\geq 1}\mathbb{P}(T_n \leq t) = \sum_{n\geq 1}F^{*n}(t)\\ &= U(t) - 1 \ ,\end{aligned}$$

as asserted.
Fix $t, a > 0$. We have

$$N_{t+a} - N_t = \sum_{n\geq 0}\mathbf{1}_{T_{N_t+1+n}\leq t+a} = \mathbf{1}_{T_{N_t+1}\leq t+a} + \sum_{n\geq 1}\mathbf{1}_{T_{N_t+1+n}\leq t+a} \ .$$

We introduce the r.v.

$$T'_n = T_{N_t+1+n} - T_{N_t+1} \ , \quad n \geq 1 \ .$$

Notice that $T_{N_t+1} = t + B_t$ so that

$$N_{t+a} - N_t = \mathbf{1}_{0\leq a-B_t} + \sum_{n\geq 1}\mathbf{1}_{T'_n\leq a-B_t} \ .$$

It is possible to show that $(T'_n)_{n\geq 1}$ is independent of B_t and has the same law as $(T_n)_{n\geq 1}$. Since F^{*0} is a Dirac mass at 0, and F^{*n} is the law of T_n, we deduce that

$$\mathbb{E}[N_{t+a} - N_t] = \mathbb{E}[\mathbb{E}[N_{t+a} - N_t | B_t]] = \mathbb{E}[\sum_{n\geq 0}F^{*n}(a - B_t)] = \mathbb{E}[U(a - B_t)] \ .$$

Since U is non-decreasing, the latter quantity is bounded by $U(a)$. □

Theorem 7. *If the function z is bounded on finite intervals then the renewal equation* (4.2) *admits a unique solution which is bounded on finite intervals and this solution is given by*

$$Z(t) = \int_{[0,t]} z(t-x)U(dx) \ , \quad t \geq 0 \ .$$

Notice that U has a Dirac mass at 0 so it is very important to specify whether we integrate over $[0, t]$ or $(0, t]$.

Proof. For every $t \geq 0$, we set

$$Z(t) = \int_{[0,t]} z(t-x)U(dx) \ ,$$

and we observe that

$$|Z(t)| \leq \sup_{x\in[0,t]}|z(x)|\int_{[0,t]}U(dx) = \sup_{x\in[0,t]}|z(x)|U(t) < \infty \ ,$$

by assumption on z and by the previous lemma. Consequently, Z is bounded on finite intervals. Let us now check that Z indeed satisfies the renewal equation. We compute

$$\begin{aligned}\int_{[0,t]} Z(t-s)F(ds) &= \int_{[0,t]}\int_{[0,t-s]} z(t-s-x)U(dx)F(ds)\\ &= \int_{\mathbf{R}}\int_{\mathbf{R}} \mathbf{1}_{s+x\in[0,t]} z(t-s-x)U(dx)F(ds)\\ &= \int_{[0,t]} z(t-y)U*F(dy)\\ &= \int_{[0,t]} z(t-y)\sum_{n\geq 1} F^{*n}(dy)\ .\end{aligned}$$

Since $z(t) = \int_{[0,t]} z(t-y)F^{*0}(dy)$ we deduce that

$$z(t) + \int_{[0,t]} Z(t-s)F(ds) = Z(t)\ ,$$

as required.

We thus showed that Z solves the renewal equation. It remains to show that this is the unique solution among all functions that are bounded on finite intervals. If Z_1 and Z_2 are two solutions, then their difference $D = Z_1 - Z_2$ solves

$$D(t) = \int_{[0,t]} D(t-s)F(ds)\ , \quad t \geq 0\ .$$

Iterating this equation and by the definition of the convolution of two measures, we get

$$D(t) = \int_{[0,t]}\int_{[0,t-s]} D(t-s-u)F(du)F(ds) = \int_{[0,t]} D(t-x)F^{*2}(dx)\ .$$

By induction, we then show that for all $n \geq 1$

$$D(t) = \int_{[0,t]} D(t-x)F^{*n}(dx)\ , \quad t \geq 0\ .$$

Consequently

$$|D(t)| \leq \int_{[0,t]} |D(t-x)|F^{*n}(dx) \leq \sup_{u\in[0,t]} |D(u)|F^{*n}(t)\ ,$$

which goes to 0 as $n \to \infty$ as we showed in the proof of the previous lemma. □

4.4 Blackwell's Renewal Theorem and Key Renewal Theorem

From now on, we assume that F is non-lattice, that is, F is not supported on a set of the form $\{0, r, 2r, 3r, \ldots\}$ for some $r > 0$. We also assume that F is a probability measure: $\int_{[0,\infty)} F(ds) = 1$. Finally, we set $\mu = \int_0^\infty xF(dx)$.

Theorem 8 (Blackwell's Renewal Theorem)**.** *Assume that F is a non-lattice probability measure and that $\mu < \infty$. Then for all $a > 0$ we have*

$$U(t+a) - U(t) \to \frac{a}{\mu} \;, \quad t \to \infty \;.$$

Theorem 9 (Key Renewal Theorem)**.** *Assume that F is a non-lattice probability measure and that $\mu < \infty$. Consider the renewal equation (4.2) and assume that the function z appearing therein is non-increasing, non-negative and Lebesgue integrable. Then*

$$Z(t) = \int_{[0,t]} z(t-s)U(ds) \to \frac{1}{\mu}\int_{[0,\infty)} z(x)dx \;, \quad t \to \infty \;.$$

Remark 4.4.1. *The theorem remains true under much less restrictive assumptions on z.*

Although the Key Renewal Theorem seems to be stronger, these two theorems are equivalent.

Proposition 4.4.2. *The Key Renewal Theorem and Blackwell's Renewal Theorem are equivalent.*

Proof. Assume that the statement of the Key Renewal Theorem holds true. Take $z(s) = \mathbf{1}_{[0,a)}(s)$, it is plain that z is integrable, non-negative and non-increasing so that we have as $t \to \infty$

$$\int_{[0,t+a]} z(t+a-s)U(ds) \to \frac{1}{\mu}\int_{[0,\infty)} z(x)dx = \frac{a}{\mu} \;.$$

On the other hand, we have for all $t > a$

$$\int_{[0,t+a]} z(t+a-s)U(ds) = U(t+a) - U(t) \;,$$

so that the conclusion of Blackwell's Renewal Theorem follows.

Conversely, let us assume that Blackwell's Renewal Theorem holds true. Assume that z is integrable, non-negative and non-increasing and define $I_k(x) = [x-(k+1)h, x-kh)$ for some $h > 0$. Then for any $x \in (nh, (n+1)h]$ we have

$$Z(x) = \int_0^x z(x-y)U(dy) = \int_0^{x-nh} z(x-y)U(dy) + \sum_{k=0}^{n-1}\int_{I_k} z(x-y)U(dy) \;.$$

From the assumptions on z, we deduce that z goes to 0 at infinity. Notice that $x - nh < h$ for all x and h so that

$$\Big|\int_0^{x-nh} z(x-y)U(dy)\Big| \le \int_0^h \sup_{r\in[x-h,x]} |z(r)|U(dy) = \sup_{r\in[x-h,x]} |z(r)|U(h) \;,$$

goes to 0 as $x \to \infty$.
Using the fact that z is non-increasing we obtain for all $k \in \{0,\dots,n-1\}$

$$z((k+1)h)(U((k+1)h) - U(kh)) \le \int_{I_k} z(x-y)U(dy) \le z(kh)(U((k+1)h) - U(kh)) \;.$$

For some fixed n_0, we then write

$$\sum_{k=0}^{n-1}\int_{I_k} z(x-y)U(dy) \le \sum_{k=0}^{n-1} z(kh)\big(U(x-kh)-U(x-(k+1)h)\big)$$
$$\le \sum_{k=0}^{n_0} z(kh)\big(U(x-kh)-U(x-(k+1)h)\big) + \sum_{k=n_0+1}^{n-1} z(kh)U(h) \; .$$

Therefore for any fixed n_0 we obtain

$$\limsup_{x\to\infty} Z(x) \le \frac{h}{\mu}\sum_{k=0}^{n_0} z(kh) + U(h)\sum_{k=n_0+1}^{\infty} z(kh) \; ,$$

and then taking the limit $n_0 \to \infty$ and then $h \downarrow 0$

$$\limsup_{x\to\infty} Z(x) \le \lim_{h\downarrow 0}\frac{1}{\mu}\int z(t)dt = \frac{1}{\mu}\int z(t)dt \; .$$

On the other hand, we have for all $n > n_0$

$$\sum_{k=0}^{n-1}\int_{I_k} z(x-y)U(dy) \ge \sum_{k=0}^{n_0} z((k+1)h)\big(U(x-kh)-U(x-(k+1)h)\big) \; .$$

so that

$$\liminf_{x\to\infty} Z(x) \ge \frac{h}{\mu}\sum_{k=0}^{n_0} z((k+1)h) \; ,$$

and taking the limit $n_0 \to \infty$ and then $h \downarrow 0$ one gets

$$\liminf_{x\to\infty} Z(x) \ge \lim_{h\downarrow 0}\frac{1}{\mu}\int z(t)dt = \frac{1}{\mu}\int z(t)dt \; .$$

Consequently, $Z(x)$ admits a limit as $x \to \infty$ and this limit coincides with $\frac{1}{\mu}\int z(t)dt$ as required. □

The proof of Blackwell's Renewal Theorem is delicate so we do not present it here, and refer the interested reader to [Asm03, Section V.5].

4.5 The exponential case

Assume that δ_1 is distributed as $\mathcal{E}(\lambda)$ for some $\lambda > 0$. Then the renewal process are the jump times of a Poisson process N of intensity λ. The parameter μ equals and $1/\lambda$.
The function U is then explicit:

$$U(t) = 1 + \mathbb{E}[N_t] = 1 + \lambda t \; , \quad t \ge 0 \; .$$

In particular, the conclusion of Blackwell's Theorem holds not only for $t \to \infty$ but for any $t \ge 0$:

$$U(t+a) - U(t) = \lambda a = \frac{a}{\mu} \; .$$

The forward recurrence process B_t has a very special behaviour.

Lemma 4.5.1. *The law of B_t does not depend on t and is $\mathcal{E}(\lambda)$.*

Proof.

$$\begin{aligned}\mathbb{P}(B_t > x) &= \mathbb{P}(T_{N_t+1} - t > x) = \sum_{k\geq 0} \mathbb{P}(T_{N_t+1} - t > x; N_t = k) \\ &= \sum_{k\geq 0} \mathbb{P}(T_k \leq t < t + x < T_{k+1}) \ .\end{aligned}$$

For every $k \geq 0$, recall that T_k has a $\Gamma(k, \lambda)$ distribution and is independent from δ_{k+1} so that we have:

$$\begin{aligned}\mathbb{P}(T_k \leq t < t + x < T_{k+1}) &= \mathbb{P}(T_k \leq t < t + x < T_k + \delta_{k+1}) \\ &= \int_{s\leq t<t+x<s+r} s^{k-1} \frac{\lambda^k}{(k-1)!} e^{-\lambda s} \lambda e^{-\lambda r} dsdr \\ &= \int_{s\leq t} s^{k-1} \frac{\lambda^k}{(k-1)!} e^{-\lambda s} \int_{r>t+x-s} \lambda e^{-\lambda r} dsdr \\ &= \int_{s\leq t} s^{k-1} \frac{\lambda^k}{(k-1)!} e^{-\lambda s} e^{-\lambda(t+x-s)} \\ &= e^{-\lambda(t+x)} \int_{s\leq t} s^{k-1} \frac{\lambda^k}{(k-1)!} \\ &= e^{-\lambda(t+x)} t^k \frac{\lambda^k}{k!} \ .\end{aligned}$$

Consequently,

$$\begin{aligned}\mathbb{P}(B_t > x) &= \sum_{k\geq 0} \mathbb{P}(T_k \leq t < t + x < T_{k+1}) = e^{-\lambda(t+x)} \sum_{k\geq 0} t^k \frac{\lambda^k}{k!} \\ &= e^{-\lambda x} \ ,\end{aligned}$$

thus yielding the asserted result. ☐

Bibliography

[Asm03] S. R. ASMUSSEN. *Applied probability and queues*, vol. 51 of *Applications of Mathematics (New York)*. Springer-Verlag, New York, second ed., 2003, xii+438. Stochastic Modelling and Applied Probability.

[Nor98] J. R. NORRIS. *Markov chains*, vol. 2 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge, 1998, xvi+237. Reprint of 1997 original.