Chapter 2

Stochastic Processes

In this chapter we summarize some definitions and properties of stochastic processes that are used in different parts of this thesis. The material included here has been extracted from different sources, such as Doob (1953), Gikhman and Skorokhod (1963), Karlin and Taylor (1975), Sato (1999) and Daley and Vere-Jones (2002), while trying to keep consistency in the notation used.

2.1 Introduction

The theory of stochastic processes has two important applications in this thesis. On the one hand, it will provide us with the theoretical framework to define and study the stochastic models we are going to develop. Additionally, in a two fold purpose, it will provide us with the sufficient insights to understand Bayesian nonparametric procedures, through the manipulation of random probability measures.

Let us anticipate here that Bayesian nonparametric procedures will be used with two different purposes: As a probabilistic framework to work out marginal and conditional probabilities involving random probability measures, with the purpose of model construction; and, as the traditional framework for making statistical inference and prediction.

We organize this chapter as follows. In Section 2.2 we review the general defini-

tion and properties of stochastic processes. Section 2.3 focuses on discrete time process, in particular Markov processes with general state space. Markov processes provide us with the theoretical background needed for the time-series model construction we shall develop. A relevant extension theorem for probability measures is also reviewed. Section 2.4 reviews two types of continuous-time stochastic processes: point processes and processes with independent and stationary increments, more specifically Lévy process. Continuous-time processes play a fundamental role for the understanding of Bayesian nonparametric methods related to random distribution functions with support on continuous unidimensional spaces, as we shall review in Chapter 3.

2.2 Definitions and properties

Before continuing, let us state some notation and provide the general definition of a stochastic process. An abstract probability space is denoted by $(\Omega, \mathscr{B}_\Omega, \mathbb{P})$, i.e. Ω is a sample space and $\mathscr{B}_\Omega$ is its corresponding Borel σ-field, with $\mathbb{P}$ a probability measure on $\mathscr{B}_\Omega$. Let us take $(\mathcal{Y}, \mathscr{B}_\mathcal{Y})$ to be a general observation space. It is also assumed that $\mathcal{Y}$ is a subset of $\mathbb{R}^d$, for some integer $d \geq 1$; hence, it follows that $(\mathcal{Y}, \mathscr{B}_\mathcal{Y})$ is also a Polish space [i.e. a complete metrizable topological space]. The mathematical definition of a stochastic process is given as follows.

Definition 2.2.1. (Stochastic process). A *stochastic process* with common state-space $(\mathcal{Y}, \mathscr{B}_\mathcal{Y})$ is a collection of random variables $\{Y(t,\omega); t \in \mathcal{T}, \omega \in \Omega)$ defined on a probability space $(\Omega, \mathscr{B}_\Omega, \mathbb{P})$ with indexed set $\mathcal{T}$, such that:

i) for a fixed ω, $Y(\cdot, \omega)$ is a function from $\mathcal{T}$ into $\mathcal{Y}$,

ii) for a fixed t, $Y(t, \cdot)$ is a function from Ω into $\mathcal{Y}$.

In the definition, (i) is reference to the realized path of the stochastic process, while (ii) defines a random variable for any t in $\mathcal{T}$.

According to the above definition, it is usual to take the index set $\mathcal{T}$ to be either countable or continuous, giving rise to stochastic processes in discrete or continuous time, respectively. A *discrete-time stochastic process* will be denoted by $\{Y_t\}_{t\in\mathcal{T}}$, with $\mathcal{T}$ being a countable subset of $\mathbb{Z}$. Without lost of generality, we shall assume that $\mathcal{T}$ coincides with $\mathbb{N}$. Whereas, a *continuous-time stochastic process* will be denoted by $\{Y(t); t \in \mathcal{T}\}$, with $\mathcal{T}$ being a dense subset of the real line.

A stochastic process is usually completed with what is known as the filtration of the process, the accumulated information generated by the process itself at each time period. Mathematically speaking, the filtration is a collection of σ-algebras generated by the past history of the process. That is, the natural filtration of the process at time t is defined as $\mathscr{F}_t^{\mathcal{Y}} = \sigma\left(\{Y(s)\} : s \leq t\right)$. So, the filtration of the stochastic process $\{Y(t); t \in \mathcal{T}\}$ is defined as $\{\mathscr{F}_t^{\mathcal{Y}} : t \in \mathcal{T}\}$, and it is said that $\{Y(t)\}$ is adapted to its filtration. When working with discrete-time processes, the filtration is defined in terms of the corresponding t-folded σ-algebras for each $t \in \mathcal{T}$. If, in addition, the state-space $(\mathcal{Y}, \mathscr{B}_{\mathcal{Y}})$ happen to be Polish, each element of the filtration of becomes $\mathscr{F}_t^{\mathcal{Y}} = \otimes_{j=1}^{t} \mathscr{B}_{\mathcal{Y}_j}$, the t-folded Borel σ-algebra generated associated to the product space $\mathcal{Y}^t$, for each $t = 1, 2, \ldots$.

According to the Kolmogorov extension theorem, or Daniel's extension theorem in a broad general setting, the probability law of a given stochastic process can be completely characterized in terms of the collection of all its finite dimensional distributions (see, Shiryayev, 1984). This result follows from measure theoretic considerations, as the probability law of a stochastic process is a special case of a probability measure on a specific product space. The collection of finite dimensional distributions are defined as follows.

Definition 2.2.2. (Finite dimensional distributions). Let $\widetilde{\mathcal{T}} = \{(t_1, ..., t_n) : \ 0 \leq t_1 < t_2 < ... < t_n, n = 1, 2, ...\}$ be the collections of all finite increasing sequences in $\mathcal{T}$. Let $\{Y(t) : t \in \mathcal{T}\}$ be a stochastic process. The set of *finite dimensional distributions* as-

sociated with $\{Y(t) : t \in \mathcal{T}\}$ is the collection of functions $\mathcal{P} = \{\mathbb{P}_{(t_1,...,t_n)}(\cdot) : (t_1, ..., t_n) \in \widetilde{\mathcal{T}}\}$ such that

$$\mathbb{P}_{(t_1,...,t_n)}(A) = \mathbb{P}\left((Y(t_1), ..., Y(t_n)) \in A\right) \tag{2.1}$$

any measurable set A in $\mathscr{B}_{\otimes_{j=1}^{n}\mathcal{Y}_{t_j}}$.

When working with stochastic processes in discrete-time, it is common to find the above definition written in terms of finite-consecutive sequences of $\mathcal{T}$.

Let us point out that the probability law of a stochastic process can be alternatively characterized in terms of the collection of regular conditional distribution functions driving the transition dynamics of the process. This result provides us with a more intuitive way to defining stochastic process, which follows from the Ionescu-Tulcea extension theorem (to be summarized in Subsection 2.3.2 for discrete-time stochastic processes).

Concerning finite dimensional distributions of stochastic processes, it is often convenient to study their stability with respect to time shifts. The notion of stability we shall focus on is that of stationarity, which, among different theoretical implications, provides us with the theoretical framework to develop statistical learning for data observed sequentially in time. Let us state the notion of stationarity in the following definition.

Definition 2.2.3. (Strict stationarity). A stochastic process $\{Y(t) : t \in \mathcal{T}\}$ is said to be strictly stationary if for any $s > 0$

$$\{Y(1+s), Y(2+s), \ldots\} \overset{d}{=} \{Y(1), Y(2), \ldots\}. \tag{2.2}$$

The above property is equivalent to saying that for any $s > 0$ and any $(t_1, \ldots, t_n) \in \widetilde{\mathcal{T}}$,

$$\{Y(t_1+s), \ldots, Y(t_n+s)\} \overset{d}{=} \{Y(t_1), \ldots, Y(t_n)\}.$$

Accordingly, under stationary we can understand that the stochastic behavior of the process will remain stable over time. Moreover, it will be actually invariant under shift translations. It is convenient to point out that the notion of stationarity corresponds

to a particular cases of measure probabilities invariant under certain groups of transformations. Hence, the study of certain implication of stationarity can actually be linked to other properties induced by invariance under transformations. In particular, we shall be interested in a particular property related with the long-term behavior of stationary processes, referred to as ergodicity.

In the view of the above observation, let us introduce some further notation. Let $\mathcal{Y}^{\infty} = \otimes_{t=1}^{\infty}\mathcal{Y}_t$, the countably product σ-field corresponding to the stochastic process, and let φ be the one-step time-shift operator on $\mathcal{Y}^{\infty}$, which is defined as the measurable map $\varphi : \mathcal{Y}^{\infty} \to \mathcal{Y}^{\infty}$ such that $\varphi(\{Y(1), Y(2), \ldots\}) = \{Y(2), Y(3), \ldots\}$. Let $\mathcal{I} = \left\{A \in \mathcal{Y}^{\infty} : \varphi^{-1}(A) = A\right\}$ be the σ-field of an invariant subset of $\mathcal{Y}^{\infty}$ with respect to φ. The notion of ergodicity is stated in the following theorem.

Definition 2.2.4. (Ergodicity). A strictly stationary stochastic process $\{Y(t) : t \in \mathcal{T}\}$ is said to be ergodic if the probability of the σ-field $\mathcal{I}$, with respect to φ, is either 0 or 1.

Ergodicity plays a fundamental role in the study of large sample behavior of stochastic processes. The definition given above is connected to the possibility of observing that the process explores all its states after leaving it in free evolution for a significantly long time. But more relevant to that notion is the idea of comparing probabilistic expectations with respect to its sample averages. Such a notion is given in Theorem 2.2.1. It also makes reference to the study of existence of a marginal invariant distribution function to which the transition dynamics of the processes will eventually converge. For the moment, let us state the fundamental ergodic theorem for stochastic processes.

Theorem 2.2.1. Let $\{Y(t)\}_{t=1}^{\infty}$ be a stationary process. The following conditions are equivalent:

1. $\{Y_t\}_{t=1}^{\infty}$ is ergodic;

2. For any shift invariant set A,

$$\mathbb{P}\left((Y_1, Y_2, \ldots) \in A\right) = 0 \text{ or } 1;$$

3. For any measurable set A of $\{Y_t\}_{t=1}^{\infty}$,

$$\lim_{t\to\infty} \frac{1}{t} \sum_{j=1}^{t} \mathbb{1}\{(Y_j, Y_{j+1}, \ldots) \in A\} = \mathbb{P}\left((Y_1, Y_2, \ldots) \in A\right).$$

Two important consequences emerge from the above theorem. On the one hand, it will follow that there is one and only one probability distribution that is invariant under shift, and more generally under φ-operations. And, on the other hand, it will follow that the average time a process spends in a region of its possible states is proportional to the probability measure of that region. Let us mention now that the most contemporaneous notion studied in the ergodic theory deals with the problem of comparing probabilistic means of some characteristics of the stochastic processes with their sample average counterparts and how they coincide after observing a long enough sample paths of the process. Such a property plays a fundamental role for the study or large sample behaviour of statistical procedure involving stochastic processes models.

Ergodicity also plays a fundamental role in Maitra's representation theorem. Let us recall that under the mentioned representation the probability law of a stochastic process who happens to be stationary and ergodic can, conditional on some stochastic element, be decomposed as the product of its corresponding transition distributions. But Maitra's representation theorem is an existence theorem, therefore the existence of such a stochastic component giving rise to the mentioned decomposition is fundamentally subjective to the analyst. See Section 1.2. We shall mention an important result concerning ergodicity for discrete time Markov processes in the following section.

Let us now mention another property that sometimes follows up from stationarity, namely reversibility. This is another type of invariance of the probability law of a given stochastic process, but this time the invariance is defined with respect to inverting direction of the sample paths of the process. Its definition is given below.

Definition 2.2.5. (Reversibility). A stochastic process $\{Y(t) : t \in \mathcal{T}\}$ is said to be reversible if it is stationary and stochastically equivalent under time-reversal transfor-

mations, i.e. if

$$\{Y(t_n), \ldots, Y(t_1)\} \stackrel{d}{=} \{Y(t_1), \ldots, Y(t_n)\}, \tag{2.3}$$

for any $t_1 < \ldots < t_n$ in $\mathcal{T}$.

Refer to Borovkov (1998) for further details.

Let us now proceed with the revision of some theory of discrete-time stochastic process. In particular, the Markov process in discrete time with general state space.

2.3 Discrete-time stochastic processes

2.3.1 Markov processes

Markov processes have one-lagged dependence structures. Their definition encompasses both discrete- and continuous-time stochastic processes. We are going to focus first on the discrete-time version. Before continuing, let us state that the formal definition of a Markov process relies heavily on the notion of a probability transition kernel.

Definition 2.3.1. (Probability transition kernel). Let $(\mathcal{X}, \mathscr{B}_{\mathcal{X}})$ and $(\mathcal{Y}, \mathscr{B}_{\mathcal{Y}})$ be two measurable spaces. A probability (or normalized) transition kernel determining the move from $(\mathcal{X}, \mathscr{B}_{\mathcal{X}})$ to $(\mathcal{Y}, \mathscr{B}_{\mathcal{Y}})$ is a function K from $\mathscr{B}_{\mathcal{Y}} \times \mathcal{X}$ into $[0, 1]$ such that:

1. for any $x \in \mathcal{X}$, $K(\cdot; x)$ is a probability measure on $(\mathcal{Y}, \mathscr{B}_{\mathcal{Y}})$,

2. for any $A \in \mathscr{B}_{\mathcal{Y}}$, the function $K(A; \cdot)$ on $\mathcal{X}$ is measurable.

Probability transition kernels are used in a variety of contexts beyond stochastic processes, as we shall explore in the upcoming chapters.

Restricting our attention to stochastic processes, once more, let us say that a Markov process $\{Y_t\}_{t=1}^{\infty}$, with common state-space $(\mathcal{Y}, \mathscr{B}_{\mathcal{Y}})$ and associated filtration $\{\mathscr{F}_t^{\mathcal{Y}}\}_{t=1}^{\infty}$, is characterised by a system of transition distributions defined for each time-period, as stated in the following definition.

Definition 2.3.2. (Markov process). A stochastic process $\{Y_t\}_{t=1}^{\infty}$ is said to be a Markov process with respect to the filtration $\{\mathscr{F}_t^{\mathcal{Y}}\}_{t=1}^{\infty}$ if for any measurable function g on $\mathcal{Y}$ and any $1 \leq s < t$,

$$\mathbb{E}[g(Y_t)|\,\mathscr{F}_s^{\mathcal{Y}}] = \mathbb{E}[g(Y_t)|\,\mathscr{B}_{\mathcal{Y}_s}]. \tag{2.4}$$

And the probability law of $\{Y_t\}_{t=1}^{\infty}$ is completely characterised by a system of probability transition kernels $\{K_{s,t} : K_{s,t} \text{ maps } \mathscr{B}_{\mathcal{Y}_t} \times \mathcal{Y}_s \text{ into } [0,1], \text{ for } 2 \leq s < t\}$ and an initial distribution P on $\mathscr{B}_{Y_1}$, such that:

$$\mathbb{P}(Y_t \in A|Y_s, \ldots, Y_1) = \mathbb{P}(Y_t \in A|Y_s) = K_{s,t}(Y_t \in A|Y_s). \tag{2.5}$$

We shall notice that the system of probability transition kernels mentioned above can be equivalently specified by the system of one-step probability transition kernels $\{K_t : K_t \text{ maps } \mathscr{B}_{\mathcal{Y}_t} \times \mathcal{Y}_{t-1} \text{ into } [0,1]\}$ and an initial distribution P on $\mathscr{B}_{Y_1}$, as:

$$\begin{aligned} \mathbb{P}(Y_t \in A|y_s) &= K_{s,t}(Y_t \in A|y_s) \\ &= \int \cdots \int K_t(Y_t \in A|y_{t-1}) \prod_{j=s+1}^{t-1} K_j(Y_t \in \mathrm{d}y_j|y_{j-1}), \end{aligned} \tag{2.6}$$

for any $1 \leq s < t$ and any measurable set $A \in \mathscr{B}_{\mathcal{Y}_t}$.

Accordingly, the collection of finite dimensional distributions of a Markov process is composed by elements of the form:

$$\mathbb{P}\left((Y_s, \ldots, Y_t) \in B\right) = \int \cdots \int_B K_t(\mathrm{d}y_t|y_{t-1}) \cdots K_{s+1}(\mathrm{d}y_{s+1}|y_s) P_s(\mathrm{d}y_s), \tag{2.7}$$

for any $1 \leq s < t$ and any set $B \in \mathscr{B}_{\otimes_{j=s}^{t} \mathcal{Y}_j}$, with

$$P(\mathrm{d}y_s) = \int \cdots \int \prod_{j=2}^{s} K_j(\mathrm{d}y_j|y_{j-1}) P(\mathrm{d}y_1).$$

Definition 2.3.3. (Time-homogeneous Markov process). A Markov sequence $\{Y_t\}_{t=1}^{\infty}$ is said to be time-homogeneous if all its transition distributions depend only on $(t-s)$, i.e.

$$K_{s,t}(Y_t \in A|y_s) = K_{t-s}(Y_t \in A|y_s), \tag{2.8}$$

for any $1 \leq s < t$. It is equivalent to say that a Markov sequence is time-homogeneous if its one-step transition kernels are homogeneous in time, i.e. $K_t = K^t$ for some one-step transition kernel K on $\mathscr{B}_{\mathcal{Y}} \times \mathcal{Y}$, where

$$K_t(Y_t \in A|y_0) = K^t(Y_t \in A|y_0) = \int K(Y_t \in A|y_{t-1}) K^{t-1}(\mathrm{d}y_{t-1}|y_0),$$

for any $t = 1, 2, \ldots$.

When a Markov process is time-homogeneous it is characterised by the pair of initial and one-step (or first order) transition distributions. The next definition links the notion of Markov processes with strictly stationary processes.

Definition 2.3.4. (Stationary Markov process). A time-homogeneous Markov process $\{Y_t\}_{t=1}^{\infty}$ with first order transition kernel K is said to have an invariant probability measure, P, if for any $A \in \mathscr{B}_{\mathcal{Y}}$,

$$P(A) = \int K(A|y) P(\mathrm{d}y). \tag{2.9}$$

If such a probability measure P exists, it is said that the pair of marginal invariant distribution, P, and the one-step transition distribution, K, characterizes the probability law of a strictly stationary Markov process (see, Doob, 1953).

It is also said that a time-homogeneous Markov sequence is reversible if for any pair of measurable sets A and B the relation,

$$\iint_{A \times B} K_t(\mathrm{d}y_t|y_{t-1}) P_t(\mathrm{d}y_{t-1}) = \iint_{A \times B} K_t(\mathrm{d}y_{t-1}|y_t) P_t(\mathrm{d}y_t), \tag{2.10}$$

holds for any $t \geq 2$, with K_t and P_t defined as above. As a consequence, any strictly stationary Markov sequence is also reversible.

As we have mentioned earlier in the previous section, ergodicity plays a fundamental role in the study of stochastic process. The study of this property has received considerable attention for Markov processes (see Borovkov, 1998), but attention has been focused on processes with discrete state spaces. However, it can be shown than Markov

processes for which an invariant distribution exists, in terms of the relation (2.9), are ergodic (see Lynch and Sethuraman, 1999). Ergodicity of a Markov process guarantees that in the long run the transition dynamics of the process will coincide with its marginal invariant distribution, which satisfies (2.10). Moreover, such an invariant distribution shall be unique. Refer also to Gray (2009) for further details.

To conclude this revision, let us recall that the probability law of a strictly stationary Markov sequence can be completely specified by the the pair of invariant and first-order transition distributions. Moreover, the invariant distribution function is unique, and all the transition distribution functions will eventually converge to the invariant distribution of the process.

With regard to the existence of the probability law of the whole process induced by the invariant and transition distribution functions follows from the Kolmogorov existence theorem (Shiryayev, 1984). However, it also follows from a more general extension theorem known in measure theory (that we shall state below) which provides us with the framework to construct and extend probability laws for stochastic process with alternative types of dependencies.

2.3.2 An extension theorem

In this section we review an important extension theorem which is used to extend products of transitions probability measures into a probability law for a countable sequence of random variables. This theorem comprises the well known Kolmogorov extension theorem as a particular case. This theorem and its proof can be found in basic books on probability, e.g. Shiryayev (1984) and Ershov (1982).

Theorem 2.3.1. (Ionescu-Tulcea Extension Theorem). Let $\{(\mathcal{Y}_t, \mathscr{B}_{\mathcal{Y}_t})\}_{t=1}^{\infty}$ be a sequence of measurable spaces. Assume that for any $t \geq 2$ there exists a probability kernel P_t on $\mathscr{B}_{\mathcal{Y}_t} \times_{i=1}^{t-1} \mathcal{Y}_i$, with P_1 a probability measure on $(\mathcal{Y}_1, \mathscr{B}_{\mathcal{Y}_1})$. Then there exists

a sequence of random variables $\{Y_t\}_{t=1}^{\infty}$ taking values in $\{\mathcal{Y}_t\}_{t=1}^{\infty}$ such that

$$\mathbb{P}(Y_1 \in A_1, \ldots, Y_t \in A_t) = \int_{A_1} \cdots \int_{A_t} \prod_{i=1}^{t} P_i(\mathrm{d}y_i | y_{i-1}, \ldots, y_1),$$

for $t = 1, 2 \ldots$, where $(A_1, \ldots A_t)$ are measurable sets. This extension is given in such a way that

$$\mathbb{P}(Y_1 \in A_1, \ldots, Y_t \in A_t) = \mathbb{P}(Y_1 \in A_1, \ldots, Y_t \in A_t, Y_{t+1} \in \mathcal{Y}_{t+1}, \ldots),$$

and it is unique.

The Ionescu-Tulcea extension theorem states that if we have a countable collection of transition kernels defined on measurable spaces, then we can use them to define a probability measure on the countable product space defined where the sequence takes values. This theorem naturally applies if the transition kernels are regular conditional distributions. Two examples of this extension theorem are summarized below.

Special cases of the above theorem can be applied to extend products of independent probability measures into a probability law for a countable sequence of random variables, i.e. taking transition densities P_t's as marginal and independent probability measures Q_t each defined on $(\mathcal{Y}_t, \mathscr{B}_{\mathcal{Y}_t})$ extends the product of measures into a probability law $\mathbb{P}$ on $\{(\mathcal{Y}_t, \mathscr{B}_{\mathcal{Y}_t})\}_{t=1}^{\infty}$ such that

$$\mathbb{P}(Y_1 \in A_1, \ldots, Y_t \in A_t, Y_{t+1} \in \mathcal{Y}_{t+1}, \ldots) = \prod_{t=1}^{t} Q_t(A_t),$$

for all $t = 1, 2, \ldots$.

Another special case consists on taking P_t to be Markov transition distributions. Hence, using the above theorem, a probability law for the Markov countable sequence $\{(\mathcal{Y}_t, \mathscr{B}_{\mathcal{Y}_t})\}_{t=1}^{\infty}$ is defined as

$$\mathbb{P}(Y_1 \in A_1, \ldots, Y_t \in A_t, Y_{t+1} \in \mathcal{Y}_{t+1}, \ldots) = \int_{A_1} \cdots \int_{A_t} \prod_{i=1}^{t} P_i(\mathrm{d}y_i | y_{i-1}),$$

with P_1 the marginal probability measure on $(\mathcal{Y}_1, \mathscr{B}_{\mathcal{Y}_1})$.

2.4 Continuous-time stochastic processes

We use continuous-time stochastic process in order to get a better understanding of random probability measures or related Bayesian nonparametric procedures, as we shall describe in Chapter 3. In particular, we shall pay particular attention to summarize some results concerning marked point processes and nondecreasing Lévy processes. To reach that point, we review basic notions of point (Poisson) processes and their connections to Lévy processes. The material presented in this section basically summarizes results developed in Sato (1999) and Daley and Vere-Jones (2002, 2008).

2.4.1 Point processes

Point processes are continuous time processes, $\{Y(t) : t \in \mathcal{T}\}$, whose sample paths represent the occurrence of some phenomenon at time epochs $t_1, t_2, ... \in \mathcal{T}$. These processes can be equivalently described in terms of: i) sequences of points, ii) sequences of intervals (time within occurrences), iii) nondecreasing integer-valued step functions, and iv) counting measures. For the easy of exposition, we shall restrict our attention to point process defined on the (positive) real line. However, it shall be noticed that the description of point processes as a counting measure can be extended to more general spaces (see, Daley and Vere-Jones, 2002, Chapter 9). Let us review some definitions and basic properties.

Definition 2.4.1. A point process on the positive real line is a sequence of random variables $\{Y(t) : t \geq 0\}$ such that

1. $Y(0) = 0$ (a. s.),

2. $Y(s) < \infty$ (a. s.), which implies that $Y(s) \leq Y(t)$ (a. s.), for any $0 \leq s \leq t$.

Their description as a counting measure, in the case the processes are nondecreasing, is given in the next definition, and consists basically in describing the point process as

an integer-valued set function.

Definition 2.4.2. A point process on the positive real line, $\{Y(t_i) : t_i \geq 0, i = 1, 2, \ldots\}$, is described as a counting measure $N(A)$, for any A subset of $\mathbb{R}_+$, which denotes the number of occurrences of the point process in the set A, i.e.

$$N(A) = \#\{i : t_i \in A\}.$$

To see N as a well defined counting measure it is imposed the restriction of A to be an element of the Borel subsets of the positive real line, $\mathscr{B}_{\mathbb{R}_+}$. In addition, N has the property that for any disjoint sets $A_1, ..., A_n$,

$$N\left(\cup_{j=1}^{n} A_j\right) = \sum_{j=1}^{n} N(A_j).$$

The counting measure N is a (positive) integer-valued function defined on the Borel sets $\mathscr{B}_{\mathbb{R}_+}$. Therefore, its description can be contracted to be defined in terms of a real function (a counting function) defined on the positive real line. That is, the point process can be described as $\{N(t) : t \geq 0\}$, such that

$$N(t) = N\left((0, t]\right), \tag{2.11}$$

for any $t \geq 0$. That is, $N(t)$ is an integer-valued function defined on increasing intervals of the positive real line. It follows from the above definition that $\{N(t) : t \geq 0\}$ is a stochastic process with the properties of being:

1. nondecreasing,

2. right-continuous.

This contraction gives rise to the description of a point process as a right-continuous step function.

Two relevant properties for what follows in terms of random probability measures are given by the increments of $N(t)$.

Definition 2.4.3. A counting process $\{N(t) : t \geq 0\}$ is said to have *independent increments* if the number of points occurring at disjoint intervals are independent, i.e. if for any $0 \leq t_1 < t_2 \leq \cdots$, the random variables

$$N(0, t_1], N(t_1, t_2], \ldots$$

are mutually independent.

In addition, the process $\{N(t) : t \geq 0\}$ is said to have *stationary increments* if its distribution is time invariant under the shift-operator, i.e. if

$$N(s, t] \stackrel{d}{=} N(r + s, r + t],$$

for any $r, s, t \geq 0$.

If $N(t)$ satisfies both conditions, it is a counting process with stationary and independent increments.

Among the class of stationary and independent increments point processes, the Poisson process is the most notable example. We shall summarize some of its properties below.

2.4.1.1 Poisson processes

Poisson processes are point process with the property that the number of occurrences in a given interval is Poisson distributed.

Definition 2.4.4. A point (counting) process $\{N(t) : t \geq 0\}$ is a Poisson processes with intensity function $\nu(\cdot) > 0$ if

1. $N(t) - N(s)$ is distributed Poisson with mean $\int_s^t \nu(\mathrm{d}w)$,
2. for any sequence $0 \leq s_1 \leq t_1 \leq \cdots \leq s_n \leq t_n < \infty$ the increments

$$N(t_1) - N(s_1), \ \ldots, \ N(t_n) - N(s_n)$$

are mutually independent.

It is said the $\{N(t) : t \geq 0\}$ is *homogeneous* if the intensity function is constant on t, i.e. $\nu(t) = \nu \cdot t$ for all $t \geq 0$.

Homogeneous Poisson processes, in general point processes, can be used to select points at random over the positive real line. That is, conditioning on having n occurrences up to time t, i.e. $N(t) = n$, there are n ordered times the $(\Xi_1, ..., \Xi_n)$ associated with their times-of-occurrence on the interval $(0, t]$. Moreover, given n, the times-of-occurrence are uniformly distributed over the interval $(0, t]$. Ignoring the order, the n points are independent and distributed $U(0, t)$, conditioning on $N(t) = n$.

2.4.1.2 Marked point process

A marked point process can be understood as a bivariate stochastic process where one of its components is a point process which provides information on the location where the second component arises. The first component is known as the *ground point process* while the second component is called the *marked process* and the space where the second component takes place is referred as the *marked space*. In this section we shall make reference only to marked point processes on the real line with general marked space, i.e. the ground process is defined on the positive real line. But, as with any other point processes, the marked process can be defined having a ground process in a more general state space.

Consistent with the notation we have been using, let $\{T_k\}_{k=1}^{\infty}$ be a point process defined on the positive real line. In addition, we assume that the marked space $(\mathcal{W}, \mathscr{B}_{\mathcal{W}})$ is Polish.

Definition 2.4.5. (Marked point process). A marked point process is a bivariate stochastic process $\{(T_k, W_k)\}_{k=1}^{\infty}$ with state-space $\mathbb{R}_+ \times \mathcal{W}$, endowed with its corresponding Borel σ-field $\mathscr{B}_{\mathbb{R}_+} \otimes \mathscr{B}_{\mathcal{W}}$. Its associated ground process $\{T_k\}_{k=1}^{\infty}$, which is denoted by N_g, is also a point process, such that:

1. for any $A \in \mathscr{B}_{\mathbb{R}_+}$ bounded,

$$N_g(A) = N(A \times \mathcal{W}) < \infty, \tag{2.12}$$

2. $\mathbb{P}(T_k < \infty, W_k \in \mathcal{W}) = \mathbb{P}(T_k < \infty)$.

Accordingly, a marked point process can be seen just as a point process on the product space $\mathbb{R}_+ \times \mathcal{W}$, subject to satisfying the finiteness condition (2.12). Alternatively, a marked point process can be understood as a point process $\{T_k\}$ with an associated sequence of random variables $\{W_k\}$ defined on the marked space $\mathcal{W}$. As discussed in Daley and Vere-Jones (2002), the condition (2.12) is required to make the ground process being a properly defined point process.

The stochastic structure of a marked point process N on the product space $\mathbb{R}_+ \times \mathcal{W}$ is specified in terms of the stochastic structure of the ground point process N_g on $\mathbb{R}_+$, usually determined by its corresponding hazard rate, and a collection of mark kernels $\{K_k\}_{k=1}^{\infty}$ such that K_k is a probability transition kernel on $\mathscr{B}_{\mathcal{W}_k} \times (\mathcal{T}_k \times_{j<k} \mathcal{T}_j \times_{j<k} \mathcal{W}_j)$. Notice that the mark kernel K_k may or may not depend on the past history of the location marks before T_k, but the general case does. Two particular examples of marked point process arise naturally.

Example 2.1. Marked Poisson process with independent marks. Let $N = \{(T_k, W_k)\}_{k=1}^{\infty}$ be a marked point process on the product space $\mathbb{R}_+ \times \mathcal{W}$. It is said that N is a marked Poisson process if:

1. the ground process N_g is a Poisson process, and

2. the mark component has independent mark kernels conditionally on N_g depending only on the current random location, i.e.

$$K_k(W_k \in A | t_k, \ldots, t_1, w_{k-1}, \ldots, w_1) = K_k(W_k \in A | t_k), \tag{2.13}$$

for any $k \geq 1$ and $A \in \mathscr{B}_{\mathcal{W}}$.

Example 2.2. Marked Poisson process with Markovian dependent marks. A marked point process N is said to be a Poisson marked process with Markov marks if

1. the ground process N_g is a Poisson process, and

2. conditionally on the random locations the mark component is Markov, i.e. if

$$K_k(W_k \in A|t_k \ldots, t_1, w_{k-1}, \ldots, w_1) = K_k(W_k \in A|t_k, w_{k-1}), \tag{2.14}$$

for any $k > 1$ and $A \in \mathscr{B}_{\mathcal{W}}$, and a given marginal distribution,

$$K_1(W_1 \in A|t_1). \tag{2.15}$$

Notice that the stability structure of a marked point process is inherited from the stability structure of the associated ground point process, as stated in the following definition.

Definition 2.4.6. A marked point process $\{(T_k, W_k)\}_{k=1}^{\infty}$ with state-space $\mathbb{R}_+ \times \mathcal{W}$ is said to be *stationary* if the probability structure of the process is invariant under shift translations in $\mathbb{R}_+$.

According to the previous definition, it can be said that the marked Poisson processes stated in examples (2.1) and (2.2) are stationary if their corresponding ground process N_g is an homogeneous Poisson process, with intensity $\nu(t) = t \cdot \nu$, for some $\nu > 0$.

For statistical purposes, it is rather convenient to find an expression for the likelihood function associated with a given realization of a marked point process, restricted to a compact set on the positive real line. For that, we are required to find expressions of densities for the locations and marks of a Marked point process. Given that N is a point process, such densities are expressible in terms of its corresponding hazard function, and their existence relies on an additional condition on N to be regular, whose notion is given in the next definition.

Definition 2.4.7. (Regular marked point processes). Let N be a marked point process with state space $\mathcal{X} = \mathbb{R}_+ \times \mathcal{W}$, and let $(\mathcal{W}, \mathscr{B}_{\mathcal{W}})$ be some Euclidean measurable space. It is said that such an N is *regular* if for any bounded Borel set A in $\mathscr{B}_{\mathcal{X}}$, there is a n-fold measure for $\{(T_k, W_k)\}_{k=1}^n$ which is absolutely continuous with respect to the n-fold product measure $\mu \times \mu_{\mathcal{W}}$, where μ denotes the Lebesgue measure on $\mathbb{R}_+$ and $\mu_{\mathcal{W}}$ stands for either the Lebesgue or counting measure on $\mathcal{W}$, depending on whether $\mathcal{W}$ is continuous or discrete.

So, a condition for N to be regular is the ground process N_g to be regular and the conditional distribution of the marks $(W_1, \ldots, W_n)$ given a realization of the locations $(t_1, \ldots, t_n)$, for all $n \geq 1$, to be absolutely continuous with respect to the n-fold product measure $\mu_{\mathcal{W}}$. This condition will be naturally satisfied in all the marked point processes considered in this thesis.

Now, if N is regular it follows that its corresponding densities can be expressed in terms of its intensity function, denoted by $\lambda^*(t, w)$, and represented in piecewise form by

$$\lambda^*(t,w) = \begin{cases} \lambda_g^*(t_1)K_1(w_1|t_1), & 0 < t \leq t_1, \\ \lambda_g^*(t_2|t_1,w_1)K_2(w_2|t_2,t_1,w_1), & t_1 < t \leq t_2, \\ \vdots & \\ \lambda_g^*(t_k|t_{k-1},\ldots,t_1,w_{k-1},\ldots,w_1) & \\ \quad \times K_k(w_k|t_k,,\ldots,t_1,w_{k-1},\ldots,w_1), & t_{k-1} < t \leq t_k, \\ \vdots & \end{cases} \tag{2.16}$$

where λ_g^* is the intensity function of the ground point process N_g and K_k is the density function for the kth-mark given the current location and past history of the process. Now, we can state the likelihood function for a realization of a marked point process.

Proposition 2.4.1. Let N be a regular Marked point process on $\mathbb{R}_+ \times \mathcal{W}$ and let T be a positive and finite number in the real line. Associated with T, let denote a realization of

N over the interval $[0,T]$ by $\{(t_1,w_1),\ldots,(t_{N_g(T)},w_{N_g(T)})\}$. Then, the density function associated with such a realization of the process is expressible in the following form

$$\begin{aligned} f[(t_1,w_1),\ldots,(t_{N_g(T)},w_{N_g(T)})] &= \left[\prod_{k=1}^{N_g(T)}\lambda^*(t_k,w_k)\right]\cdot\exp\left\{-\iint_{[0,T]\times\mathcal{W}}\lambda^*(\mathrm{d}s,\mathrm{d}u)\right\} \\ &= \left[\prod_{k=1}^{N_g(T)}\lambda_g^*(t_k)\right]\cdot\exp\left\{-\int_0^T\lambda_g^*(\mathrm{d}s)\right\} \\ &\quad\times\left[\prod_{k=1}^{N_g(T)}K_k(w_k|t_k,\ldots,t_1,w_{k-1},\ldots,w_1)\right], \end{aligned} \tag{2.17}$$

with λ_g^* and K_k defined as in (2.16).

Marked point processes constitute a natural way to randomize piecewise functions, as it will be required in Chapters 5 and 6. Accordingly, the probability law driving N can be assumed as a probability measure for a given piecewise function, in the spirit of the ideas presented in Chapter 3. For the moment, let us continue with the revision of independent increments processes, which constitute one of the most relevant processes used to randomize probability distribution functions according to the ideas also presented in Chapter 3.

2.4.2 Processes with independent increments

Processes with independent increments constitute one of the richest families of stochastic processes and are frequently used in Financial Mathematics and Bayesian nonparametric statistics. These types of processes make use of a modified version of the strict stationarity assumption presented in previous sections by considering that the process exhibits stationary increments. As we shall discuss, Lévy processes can be decomposed as the sum of three independent components: i) a deterministic drift, ii) a pure jump component, and iii) a diffusion coefficient. Such a decomposition is known as the Lévy-Ito decomposition. A distinctive feature of Lévy processes is that the distributions of their increments belong to the class of infinitely divisible distributions (see, Steutel, 1979).

As such, and due to the assumption of stationary increments, these processes can be completely characterised by a unique infinitely divisible random variable by means of their characteristic function. Such characterization makes evident the aforementioned Lévy-Ito decomposition.

Among the general class of Lévy processes existing in the literature we shall be paying particular attention to the review of those with nonnegative increments. As we shall discuss, these processes are characterised by only two components: i) a deterministic drift, and ii) a pure jump component. The reason of focusing on these processes is that they are used in Bayesian nonparametrics to characterise random probability measures, by means of randomizing cumulative distribution or related functions via suitable transformations. Let us mention that the assumption of nonnegative increments will assist us to characterise σ-finite measures.

Let us notice that processes with independent increments, and Lévy processes in particular, are defined on d-dimensional Euclidean state-spaces. However, for the ease of exposition, and actually their application to us in future chapters, we will restrict our attention to the one-dimensional case with state-space on the real (or positive real) line.

Definition 2.4.8. A real-valued stochastic process $\{Y(t) : t \geq 0\}$ is said to be a *process with independent increments* if for each $n \in \mathbb{N}$ and each $(t_1, t_2, \ldots, t_{n+1}) \in \widetilde{\mathbb{R}}_+$, the random variables

$$Y(t_{j+1}) - Y(t_j), \text{ for } 1 \leq j \leq n, \tag{2.18}$$

are mutually independent. In addition, if the process is such that

$$Y(t+s) - Y(s) \stackrel{d}{=} Y(t) - Y(0), \tag{2.19}$$

for any $t \geq 0$ and $s > 0$, it is said that $\{Y(t) : t \geq 0\}$ is a *process with independent and stationary increments*.

Processes with independent and stationary increments are intrinsically connected to the theory of Lévy processes through the following definition.

Definition 2.4.9. A real-valued stochastic process $\{Y(t) : t \geq 0\}$ is said to be a *Lévy process* if the following conditions are satisfied

1. $Y(0) = 0$, a. s.;
2. $\{Y(t) : t \geq 0\}$ has independent and stationary increments;
3. it is stochastically continuous (or continuous in probability), i.e. if for every $t \geq 0$ and $\epsilon > 0$,
$$\lim_{s \to t} \mathbb{P}\{|Y(t) - Y(s)| > \epsilon\} = 0. \tag{2.20}$$
4. there is $\Omega_0 \in \mathscr{B}_\Omega$ with $\mathbb{P}\{\Omega_0\} = 1$, such that for every $\omega \in \Omega_0$, $Y(t, \omega)$ is right continuous in $t \geq 0$ and has left limits in $t > 0$.

Any process satisfying conditions (1) to (4) is also referred as a Lévy process in law.

Recall that the above definition accommodates $\mathbb{R}^d$-valued Lévy process.

As mentioned before, it is well known that if $\{Y(t) : t \geq 0\}$ is a Lévy process, then $Y(t)$ is infinitely divisible for all $t \geq 0$. Therefore, by virtue of the Lévy-Kintchine representation of infinitely divisible distributions (see, Sato, 1999), the distribution of $Y(t)$ can be characterised by its characteristic function as follows,

$$\mathbb{E}\left[e^{i\theta Y(t)}\right] = \exp\{\psi(\theta, Y(t))\}$$

where $\psi(\theta, Y(t))$ is called the Lévy exponent of $Y(t)$ and it is defined by

$$\psi(\theta, Y(t)) = i\theta b(t) - \frac{\theta^2}{2} a(t) + \int_{\mathbb{R}-\{0\}} \left(e^{i\theta s} - 1 - i\theta s \, \mathbb{1}_{\{|s|<1\}}(s)\right) \nu_t(\mathrm{d}s), \tag{2.21}$$

for any real valued θ, with

- $b(t)$ a real-valued function on $\mathbb{R}_+$,
- $a(t)$ a positive real-valued function on $\mathbb{R}_+$,

- $\nu_t(\mathrm{d}s)$ a measure on $(\mathbb{R}, \mathscr{B}_{\mathbb{R}})$, which must satisfy,

$$\int_{\mathbb{R}-\{0\}} \min\{s^2, 1\}\nu_t(\mathrm{d}s) < \infty,$$

  for all $t \geq 0$.

In the above representation: i) b corresponds to the (deterministic) drift of the process, ii) a corresponds to the scale parameter of the (continuous) Gaussian component of the process, and iii) is the Lévy measure associated to the pure jump component. In that representation, the triplet $\{b, a, \nu\}$ is referred as the *generating triplet* of the Lévy process $\{Y(t) : t \geq 0\}$.

Due to the Lévy-Kintchine representation, it is evident to observe the two canonical examples of Lévy process, named Brownian motion and Poisson processes.

Example 2.3. Brownian motion. A canonical example of Lévy processes is the $\mathbb{R}$-valued Brownian motion $\{B(t) : t \geq 0\}$ with a drift, for which: a) $B(t) \sim \mathrm{N}(b(t), a(t))$, for $t \geq 0$, and b) $B(t)$ has continuous sample paths (a. s.). So, $\{B(t)\}$ defines a Lévy process with Lévy exponent taking the form

$$\psi(\theta, B(t)) = i\theta b(t) - \frac{\theta^2}{2}a(t). \tag{2.22}$$

∘

Example 2.4. Compound Poisson process. The compound Poisson process is another canonical example of Lévy processes. Let $\{Z_n\}_{n=1}^{\infty}$ be a sequence of i. i. d. $\mathbb{R}$-valued random variables with common probability law ν on $(\mathbb{R}, \mathscr{B}_{\mathbb{R}})$, and let N be a Poisson process with intensity $\lambda(t) > 0$ and independent of the Z_n's. The compound process $D(t) = \sum_{j=1}^{N(t)} Z_j$ for $t \geq 0$ defines a Lévy process with Lévy exponent given by

$$\psi(\theta, D(t)) = \int_{\mathbb{R}} (e^{i\theta s} - 1)\nu_t(\mathrm{d}s), \tag{2.23}$$

where $\nu_t(\mathrm{d}s) = \lambda(t)\nu(\mathrm{d}s)$.

Notice that compound Poisson processes have piecewise constant sample paths (a. s.), with jump discontinuities defined at random times $\{T(n)\}_{n=1}^{\infty}$ and random sizes $\{Z_n\}_{n=1}^{\infty}$.

∘

Notice that the general expression of the Lévy exponent (2.21) is the sum of its counterparts for $\{B(t)\}$ and $\{D(t)\}$. However, additional regularity conditions should be added in order to obtain that result. The regularity condition reduces to ensure the inclusion of explosive behaviour in a neighborhood around 0 of the Poisson jump component, which is reflected in the part $\theta s \mathbb{1}_{\{|s|<1\}}(s)$ in the Lévy exponent. The result of having $Y(t) = C(t) + D(t)$ is known as the Lévy-Ito decomposition.

Let us notice that the above representation is also applicable to d-dimensional Lévy processes, for any finite $d > 1$, by means of extending the characterizing triplet. That is, by taking b to be a d-dimensional real valued function, $a(t)$ to be a $d \times d$-valued symmetric positive function, and ν_t to be a measure on $(\mathbb{R}^d, \mathscr{B}_{\mathbb{R}^d})$ such that $\int_{\mathbb{R}^d - \{0\}} \min\{s^2, 1\}\nu_t(\mathrm{d}s) < \infty$. Notice that the Gaussian part of the Lévy exponent should be adapted to have a quadratic form. For what is to come, we shall restrict our attention to one-dimensional Lévy processes. Specifically, on Lévy processes with nondecreasing independent increments.

2.4.2.1 Processes with nondecreasing independent increments

Processes with nondecreasing independent increments play a very important role in Bayesian nonparametric statistics, as many random probability measures can be characterised in terms of suitable transformations. It is worth mentioning here that the random probability measures considered here are expressed as cumulative distribution functions. So, the assumption of having nonnegative increments would guarantee that the process would satisfy the requirements to be considered a proper distribution function. Notice also that as a consequence of assuming nonnegative increments, the Gaussian component of the underlying Lévy process vanishes; hence these processes would be characterized solely by: i) the drift component, and ii) the Lévy measure. Let us state the formal definition.

Definition 2.4.10. Let $\{Y(t) : t \geq 0\}$ be a Lévy process with state-space on $\mathbb{R}_+$. It is said that $\{Y(t)\}$ is a *process with nondecreasing independent increments* if, in addition:

1. $Y(0) = 0$, a. s.;
2. $Y(t) \geq 0$, a. s. for any $t > 0$;
3. $Y(s) \leq Y(t)$, a. s. whenever $s \leq t$;
4. $Y(t)$ is right continuous with respect to t, a. s..

By virtue of the Lévy-Kintchine representation we have that $\{Y(t)\}$ is completely characterised by a drift and a Lévy measure, i.e.

$$\mathbb{E}\left[e^{-\theta \cdot Y(t)}\right] = \exp\left\{-\theta b(t) - \int_0^\infty (1 - e^{-\theta s})\nu_t(\mathrm{d}s)\right\}, \tag{2.24}$$

for every $\theta > 0$ and any $t > 0$; where $b(t)$ is a positive nondecreasing function and $\nu_t(\mathrm{d}s)$ is the Lévy measure on $(\mathbb{R}_+, \mathscr{B}_{\mathbb{R}_+})$ which must satisfy,

$$\int_0^\infty \frac{s}{1+s}\nu_t(\mathrm{d}s) < \infty, \tag{2.25}$$

for any $t > 0$. Notice that in the representation of processes with nondecreasing independent increments the drift component corresponds to a deterministic part of the process. As such, and with no loss of generality, it can be assumed $b(t) = 0$ for all $t > 0$. Therefore, the parameterization of the Lévy measure ν_t will basically determine the Lévy process $\{Y(t) : t \geq 0\}$ in a unique way.

It is worth noticing that processes with nondecreasing independent increments can be defined in terms of a Poisson random measure N on $(\mathbb{R}_+ \times \mathbb{R}_+, \mathscr{B}_{\mathbb{R}_+} \otimes \mathscr{B}_{\mathbb{R}_+})$, by taking N as a compound Poisson process (see, Daley and Vere-Jones, 2008) and defining

$$Y(t) = \int_0^t \int_0^\infty sN(\mathrm{d}u, \mathrm{d}s), \tag{2.26}$$

which has a Laplace transform of the form

$$\mathbb{E}\left[e^{-\theta \cdot Y(t)}\right] = \exp\left\{-\int_0^t \int_0^\infty (1 - e^{-\theta s})\nu(\mathrm{d}u, \mathrm{d}s)\right\}, \tag{2.27}$$

where ν is a Lévy measure on $(\mathbb{R}_+ \times \mathbb{R}_+, \mathscr{B}_{\mathbb{R}_+} \otimes \mathscr{B}_{\mathbb{R}_+})$ such that

1. $\int_0^\infty \int_1^\infty \nu(\mathrm{d}u, \mathrm{d}s) < \infty$,

2. $\int_0^t \int_0^1 s\nu(\mathrm{d}u, \mathrm{d}s) < \infty$,

for any $t > 0$. Notice, as well, that the Lévy measure ν_t used in the conventional notation is expressed by

$$\nu_t(\mathrm{d}s) = \int_0^t \nu(\mathrm{d}u, \mathrm{d}s), \tag{2.28}$$

for any $t > 0$.

Regarding the Lévy measure ν, it is of interest to find the following parameterizations:

- Simple homogeneous:

$$\nu(\mathrm{d}u, \mathrm{d}s) = \nu(\mathrm{d}s)\mathrm{d}u, \tag{2.29}$$

- Generalised homogeneous:

$$\nu(\mathrm{d}u, \mathrm{d}s) = \nu(\mathrm{d}s)\alpha(\mathrm{d}u), \tag{2.30}$$

- Nonhomogeneous:

$$\nu(\mathrm{d}u, \mathrm{d}s) = \nu(\mathrm{d}s|u)\alpha(\mathrm{d}u). \tag{2.31}$$

All are defined on the product space $\mathbb{R}_+ \times \mathbb{R}_+$. For the generalised homogeneous case, it is also assumed that $\alpha : \mathbb{R}_+ \to \mathbb{R}_+$ is nonnegative, continuous and increasing function such that: a) $\alpha(0) = 0$, and b) $\lim_{t\to\infty} \alpha(t) = \infty$. In the nonhomogeneous case, in addition, $\nu(\mathrm{d}s|u)$ is taken to be a probability transition kernel on $\mathscr{B}_{\mathbb{R}_+} \times \mathbb{R}_+$ (see Definition 2.3.1). Accordingly, in the nonhomogeneous case, the distribution of the size of the jumps of the process depends on the locations where such jumps occur. Notice that the simple homogeneous parameterization gives rise to increasing Lévy processes known as *subordinators*, whereas in the generalized homogeneous case they define additive Lévy process. See Sato (1999) for a concise account.

The main contribution of Lévy processes to Bayesian nonparametric statistics is that they serve as a suitable instrument to characterize some specific types of random probability measures. The specification of the random probability measure they characterize

depends on the particular transformation of the underlying process with nondecreasing independent increments taken into consideration. In this respect, we shall consider the transformation inducing neutral-to-the-right processes (see Section 3.6) with nonhomogeneous parametrization of the underlying Lévy measure ν.

It is worth noticing that the compound Poisson process N described above can be extended to a more general product space $\mathbb{R}_+ \times \mathcal{X}$, with $\mathcal{X}$ being a complete and separable measurable space. Therefore, it is also possible to characterise random probability measures on $\mathcal{X}$, under specific normalizing transformations of the underlying Lévy process. We shall briefly discuss more on this matter in Chapter 3.

For the moment, let us present a collection of processes with nondecreasing independent increments frequently used in Bayesian nonparametric settings. We consider examples for the homogeneous and nonhomogeneous parameterizations. All of them characterizes different random probability measures after particular transformations.

Example 2.5. Homogeneous gamma process. The homogeneous gamma process is a $\mathbb{R}_+$-valued increasing Lévy process characterized by the Lévy measure,

$$\nu(\mathrm{d}u, \mathrm{d}s) = \frac{e^{-s\beta}}{s}\mathrm{d}s\ \alpha(\mathrm{d}u), \tag{2.32}$$

which is parameterized by a bounded nondecreasing measure α on $(\mathbb{R}_+, \mathscr{B}_{\mathbb{R}_+})$, with $\alpha(0) = 0$, and a constant parameter $\beta > 0$. The main characteristic of this process is that the process $\{Y(t)\}$ has gamma marginal distribution functions.

∘

Example 2.6. Stable process. Traditionally, stable subordinators are termed as α-stable to make particular emphasis on its index parameter, α. In this case, in order to preserve notational consistence, we denote the index parameter of stable processes by σ, with $0 < \sigma < 1$. This subordinator is characterised by a Lévy measure of the form,

$$\nu(\mathrm{d}u, \mathrm{d}s) = \frac{\sigma}{\Gamma(1-\sigma)}s^{-(1+\sigma)}\mathrm{d}s\ \alpha(\mathrm{d}u), \tag{2.33}$$

where α is a positive measure on $(\mathbb{R}_+, \mathscr{B}_{\mathbb{R}_+})$. Its corresponding Laplace representation is given by

$$\begin{aligned} \mathbb{E}\left[e^{-\theta \cdot Y(t)}\right] &= \exp\left\{-\frac{\sigma}{\Gamma(1-\sigma)}\int_0^t\int_0^\infty (1-e^{-\theta s})s^{-(\sigma+1)}\mathrm{d}\alpha(\mathrm{d}u)\right\} \\ &= \theta^\sigma. \end{aligned}$$

Also notice that the distributions of stable subordinators do not usually have an analytic closed form, with exception of the case $\sigma = 1/2$, which have a marginal distribution with density,

$$f_{Y(t)}(y) = \frac{t}{2\sqrt{\pi}} y^{-3/2} e^{-t^2/4y},$$

for any $t \geq 0$, and Lévy measure given by

$$\nu(\mathrm{d}u, \mathrm{d}s) = \frac{1/2}{\Gamma(1/2)} s^{-3/2}\mathrm{d}s\ \alpha(\mathrm{d}u). \tag{2.34}$$

This subordinator is termed the Lévy subordinator.

∘

Example 2.7. Generalised gamma process. The generalised gamma process were introduced by Brix (1999). This is a particular case of a $\mathbb{R}_+$-valued homogeneous process with nondecreasing independent increments characterized by a Lévy measure of the form

$$\nu(\mathrm{d}u, \mathrm{d}s) = \frac{\sigma}{\Gamma(1-\sigma)} e^{-\tau s} s^{-(1+\sigma)}\mathrm{d}s\ \alpha(\mathrm{d}u), \tag{2.35}$$

with parameters $0 < \sigma < 1$, $\tau > 0$ and $\alpha(\mathrm{d}u)$ a positive measure. Notice that the homogeneous gamma process is a limit case of the generalised gamma processes by taking $\sigma \to 0$. The α-stable process is also a particular case of the generalised gamma process by taking $\sigma = \alpha$ and $\tau = 0$.

∘

Below we present two examples of nondecreasing independent increments process with nonhomogeneous parameterizations of their corresponding Lévy measures.

Example 2.8. Beta process. The Beta processes introduced by Hjort (1990) is a $[0,1]$-valued process with independent increments characterised by the Lévy measure

$$\nu(\mathrm{d}u, \mathrm{d}s) = \frac{c(u)}{s}(1-s)^{c(u)-1}\mathrm{d}s\ \alpha(\mathrm{d}u), \tag{2.36}$$

i.e. $\nu(\mathrm{d}s|u) = \frac{c(u)}{s}(1-s)^{c(u)-1}\mathrm{d}s$, where α is a nondecreasing positive function (typically a cumulative hazard function) and c is a (piecewise) continuous nonnegative function.

∘

Example 2.9. Log-beta process. The log-beta process was first introduced by Walker and Muliere (1997). It is a particular type of $\mathbb{R}_+$-valued nonhomogeneous increasing Lévy procees, which is characterized by a Lévy measure of the form

$$\nu(\mathrm{d}u, \mathrm{d}s) = \frac{e^{-s\beta(u)}}{1-e^{-s}}\mathrm{d}s\ \alpha(\mathrm{d}u). \tag{2.37}$$

where α is a continuous measure on $(\mathbb{R}_+, \mathscr{B}_{\mathbb{R}_+})$, with $\alpha(0) = 0$, and β is a positive function on the positive real line. The nonhomogeneous parameterization of its Lévy measure is given by taking $\nu(\mathrm{d}s|u) = \frac{e^{-s\beta(u)}}{1-e^{-s}}\mathrm{d}s$.

Accordingly, the Lévy-Kintchine representation of $Y(t)$ induced by ν is given by

$$\mathbb{E}[e^{-\theta\cdot Y(t)}] \;=\; \exp\left\{-\int_0^t\int_{\mathbb{R}_+}(1-e^{-\theta s})\frac{\mathrm{d}s}{1-e^{-s}}e^{-s\beta(u)}\alpha(\mathrm{d}u)\right\}, \tag{2.38}$$

for any $\theta > 0$. We can briefly anticipate that the log-beta processes give rise to the beta-Stacy random probability measures (see, Walker and Muliere, 1997). Notice that the Dirichlet process arises as a particular case of the log-beta process by taking the measure α such that $\alpha(\mathbb{R}_+) < \infty$ and $\beta(u) = \alpha(\mathbb{R}_+) - \alpha([0,u))$. Also, the homogeneous gamma process arises as another particular case when β is a positive constant parameter.

∘

Remark 2.1. In the study of processes with nondecreasing independent increments it is possible to relax the assumption of stochastic continuity, giving in (2.20), in order

to accommodate fixed points of discontinuity. This consideration results useful when studying Bayesian nonparametric process, as we shall see in Chapter 3. In particular, these processes admit at most countably many fixed points of discontinuity occurring at 'fixed' times $\{t_k\}_{k=1}^{\infty}$ and with 'random' nonnegative jump-sizes given by $\{S_k\}_{k=1}^{\infty}$, which in addition are assumed to be mutually independent. So, the process we are interested now in studying can be defined as the sum of two independent components: i) a 'fixed' jump component, $Y_d(t) = \sum_{t_k=1}^{\infty} S_k \mathbb{1}\{t_k \leq t\}$, and ii) a (stochastic) continuous component, $Y_c(t)$, i.e.

$$Y(t) = Y_d(t) + Y_c(t), \tag{2.39}$$

such that the continuous component $Y_c(t) = Y(t) - Y_d(t)$ is a process with nondecreasing independent increments with Laplace representation given by (2.27), for a given Lévy measure ν. Hence, the Lévy-Kintchine representation for the process $\{Y(t) : t \geq 0\}$ becomes

$$\mathbb{E}\left[e^{-\theta \cdot Y(t)}\right] = \left\{\prod_{t_k \leq t} \mathbb{E}\left[e^{-\theta \cdot S_k}\right]\right\} \cdot \exp\left\{-\int_0^t \int_0^{\infty} (1 - e^{-\theta s})\nu(\mathrm{d}u, \mathrm{d}s)\right\}. \tag{2.40}$$

The above expression will be useful when working with neutral-to-the-right process in Bayesian nonparametrics (see Section 3.6).

2.5 Discussion

In this chapter we have reviewed some basic notions of the theory of stochastic processes in discrete and continuous time. Continuous type stochastic processes, point and independent increments processes, play a fundamental role to the understanding of Bayesian nonparametric procedures, which shall be briefly described in Chapter 3. Let us anticipate that these processes are used to define randomised versions of probability measures. And the probability law governing those processes will be considered as a probability measure on a suitable space of probability measures. These ideas will be clarified in the next chapter.

On the other hand, the theory we have reviewed concerning discrete time stochastic processes will be relevant to understand the construction and properties of time-series models through the introduction of latent processes. That idea shall be summarised in Chapter 4.