



A stochastic process represents a system, usually evolving along time, which incorporates an element of randomness, as opposed to a deterministic process.

The independent sequences of random variables, the Markov chains and the martingales have been presented above. In the present chapter, we investigate more general real stochastic processes, indexed by $\mathbb{T} \subset \mathbb{R}$, with $t \in \mathbb{T}$ representing the time, in a wide sense. Depending on the context, when $\mathbb{T} = \mathbb{N}$ or \mathbb{R}_+ , they are called sequences of random variables—in short random sequences, stochastic processes with continuous or discrete time, signals, times series, etc.

First, we generalize to stochastic processes the notions of distributions, distribution functions, etc., already encountered in the previous chapters for random sequences. Then, we study some typical families of stochastic processes, with a stress on their asymptotic behavior: stationary or ergodic processes, ARMA processes, processes with independent increments such as the Brownian motion, point processes—especially renewal and Poisson processes, jump Markov processes and semi-Markov processes will be especially investigated in the final chapter.

4.1 General Notions

We will first define general stochastic elements, and then extend to stochastic processes notions such as distributions, stopping times, and results such as the law of large numbers, central limit theorem.

Let (Ω, \mathcal{F}) and $(\mathbf{E}, \mathcal{E})$ be two measurable spaces. In probability theory, a measurable function $X : (\Omega, \mathcal{F}) \rightarrow (\mathbf{E}, \mathcal{E})$, that is such that $X^{-1}(\mathcal{E}) \subset \mathcal{F}$, is called a stochastic element.

The stochastic elements may take values in any measurable space. If $\mathbf{E} = \mathbb{R}$ (\mathbb{R}^d , $d > 1$, $\mathbb{R}^{\mathbb{N}} = \mathbb{R} \times \mathbb{R} \times \dots, \mathbb{R}^{\mathbb{R}}$), then X is a real random variable (random vector, random sequence or stochastic process with discrete time, stochastic process with continuous time).

It is worth noticing that even in the general theory of stochastic processes, \mathbb{T} is typically a subset of \mathbb{R} , as we will assume thereafter. For any fixed $\omega \in \Omega$, the function $t \rightarrow X_t(\omega)$ is called a trajectory or realization of the process. The value of this function is determined by the result of a random phenomenon at the time t of its observation. For example, one may observe tossing of a coin, the fluctuations of the generations of a population, the day temperature in a given place, etc. If the trajectories of the process are continuous functions (on the left or on the right), the process itself is said to be continuous (on the left or on the right).

In general, if $(\mathbf{E}, \mathcal{E}) = (\mathbb{R}^{\mathbb{T}}, \mathcal{B}(\mathbb{R}^{\mathbb{T}}))$, where \mathbb{T} is any convenient set of indices, a stochastic element can be represented as a family of real random variables $\mathbf{X} = (X_t)_{t \in \mathbb{T}}$. It can also be regarded as a function of two variables

$$\begin{aligned} \mathbf{X} : \Omega \times \mathbb{T} &\longrightarrow \mathbb{R} \\ (\omega, t) &\longrightarrow X_t(\omega) \end{aligned}$$

such that X_t is a real random variable defined on (Ω, \mathcal{F}) for any fixed t . If \mathbb{T} is a totally ordered enumerable set (as \mathbb{N} or \mathbb{R}_+), then \mathbf{X} is called a time series. If $\mathbb{T} = \mathbb{R}^d$, with $d > 1$ or one of its subsets, the process is a multidimensional process. If $\mathbf{E} = (\mathbb{R}^d)^{\mathbb{T}}$, with $d > 1$, it is a multivariate or vector process.

The canonical space of a real process $\mathbf{X} = (X_t)_{t \in \mathbb{T}}$ with distribution $\mathbb{P}_{\mathbf{X}}$ is the triple $(\mathbb{R}^{\mathbb{T}}, \mathcal{B}(\mathbb{R}^{\mathbb{T}}), \mathbb{P}_{\mathbf{X}})$. Generally, we will consider that the real stochastic processes are defined on their canonical spaces.

Note that a function $\theta_s : \mathbb{R}^{\mathbb{T}} \rightarrow \mathbb{R}^{\mathbb{T}}$ is called a translation (or shift) operator on $\mathbb{R}^{\mathbb{T}}$ if and only if $\theta_s(x_t) = x_{t+s}$ for all $s, t \in \mathbb{T}$ and all $x = (x_t) \in \mathbb{R}^{\mathbb{T}}$. When $\mathbb{T} = \mathbb{N}$, θ_n is the n -th iterate of the one step translation operator θ_1 (also denoted by θ , see above Definition 3.21).

A stochastic process \mathbf{X} defined on $(\Omega, \mathcal{F}, \mathbb{P})$ is said to be an L^p process if $X_t \in L^p(\Omega, \mathcal{F}, \mathbb{P})$ for all $t \in \mathbb{T}$, for $p \in \mathbb{N}^*$. For $p = 1$ it is also said to be integrable, and for $p = 2$ to be a second order process.

Here are some classical examples of stochastic processes, the sinusoidal signals, the Gaussian processes, and the ARMA processes.

▷ *Example 4.1 (Sinusoidal Signal)* For the process defined by

$$X_t = A \cos(2\pi \nu t + \varphi), \quad t \in \mathbb{R},$$

φ is called the phase, A the amplitude and ν the frequency. These parameters can be either constant or random. For instance, if A and ν are real constants and φ is a random variable, the signal is a monochromatic wave with random phase. ◁

▷ *Example 4.2 (Gaussian White Noise)* Let $(\varepsilon_t)_{t \in \mathbb{Z}}$ be such that $\varepsilon_t \sim \mathcal{N}(0, 1)$ for all $t \in \mathbb{T}$ and $\mathbb{E}(\varepsilon_{t_1} \varepsilon_{t_2}) = 0$ for all $t_1 \neq t_2$. The process (ε_t) is called a Gaussian white noise. ◁

▷ *Example 4.3 (Gaussian Process)* Let $\mathbf{X} = (X_t)_{t \in \mathbb{T}}$ be such that the vector $(X_{t_1}, \dots, X_{t_n})$ is a Gaussian vector for all integers n and all t_1, \dots, t_n . The process \mathbf{X} is said to be Gaussian. ◁

Definition 4.4 A sequence of random variables $X = (X_n)_{n \in \mathbb{Z}}$ is called an auto-regressive moving-average process with orders p and q , or ARMA(p, q), if

$$X_n + \sum_{i=1}^p a_i X_{n-i} = \varepsilon_n + \sum_{j=1}^q b_j \varepsilon_{n-j}, \quad n \in \mathbb{Z},$$

where ε is a Gaussian white noise, p and q are integers and a_1, \dots, a_p and b_1, \dots, b_q are all real numbers, with $a_p \neq 0$ and $b_q \neq 0$.

This type of processes is often used for modeling signals. If $q = 0$, the process is said to be an auto-regressive process, or AR(p). This latter models for example economical data depending linearly of the p past values, up to a fluctuation—the noise. If $a = -1$ and $p = 1$, it is a random walk. If $p = 0$, the process is said to be a moving-average process, or MA(q).

The ARMA processes can also be indexed by \mathbb{N} , by fixing the value of X_0, \dots, X_p , as shown in Exercise 4.1.

▷ *Example 4.5 (MA(1) Process)* Let ε be a Gaussian white noise with variance σ^2 and let (X_n) be a sequence of random variables such that $X_n = b\varepsilon_{n-1} + \varepsilon_n$, for $n > 0$, with $X_0 = \varepsilon_0$. The vector (X_0, \dots, X_n) is the linear transform of the Gaussian vector $(\varepsilon_0, \dots, \varepsilon_n)$, and hence is a Gaussian vector too, that is centered. For $n \geq 1$ and $m \geq 1$,

$$\begin{aligned} \text{Cov}(X_n, X_m) &= b^2 \mathbb{E} \varepsilon_{n-1} \varepsilon_{m-1} + b(\mathbb{E} \varepsilon_{n-1} \varepsilon_m + \mathbb{E} \varepsilon_n \varepsilon_{m-1}) + \mathbb{E} \varepsilon_n \varepsilon_m \\ &= \begin{cases} 0 & \text{if } |n - m| \geq 2, \\ b\sigma^2 & \text{if } |n - m| = 1, \\ (b^2 + 1)\sigma^2 & \text{if } n = m. \end{cases} \end{aligned}$$

Finally, we compute $\text{Var } X_0 = \sigma^2$, $\text{Cov}(X_0, X_1) = b\sigma^2$ and $\text{Cov}(X_0, X_m) = 0$ for all $m > 1$. ◁

Definition 4.6 Let (Ω, \mathcal{F}) be a measurable set.

1. If \mathcal{F}_t is a σ -algebra included in \mathcal{F} for all $t \in \mathbb{T}$, and if $\mathcal{F}_s \subset \mathcal{F}_t$ for all $s < t$, then $\mathbf{F} = (\mathcal{F}_t)_{t \in \mathbb{T}}$ is called a filtration of (Ω, \mathcal{F}) . Especially, if $\mathbf{X} = (X_t)_{t \in \mathbb{T}}$ is a stochastic process, the filtration $\mathbf{F} = (\sigma(X_s; s \leq t))_{t \in \mathbb{T}}$ is called the natural filtration of \mathbf{X} .

2. A stochastic process \mathbf{X} is said to be adapted to a filtration $(\mathcal{F}_t)_{t \in \mathbb{T}}$ if for all $t \geq 0$, the random variable X_t is \mathcal{F}_t -measurable.

Obviously, every process is adapted to its natural filtration. For investigating stochastic processes, it is necessary to extend the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ by including a filtration $\mathbf{F} = (\mathcal{F}_t)_{t \in \mathbb{T}}$; then $(\Omega, \mathcal{F}, \mathbf{F}, \mathbb{P})$ is called a stochastic basis. Unless otherwise stated, the filtration will be the natural filtration of the studied process.

Theorem-Definition 4.7 *Let \mathbf{F} be a filtration.*

A random variable $T : (\Omega, \mathcal{F}) \rightarrow \overline{\mathbb{R}}_+$ such that $(T \leq t) \in \mathcal{F}_t$, for all $t \in \mathbb{R}_+$, is called an \mathbf{F} -stopping time. Then, the family of events

$$\mathcal{F}_T = \{A \in \mathcal{F} : A \cap (T \leq t) \in \mathcal{F}_t, t \in \mathbb{R}_+\}$$

is a σ -algebra, called the σ -algebra of events previous to T .

A stopping time adapted to the natural filtration of some stochastic process is said to be adapted to this process. The properties of the stopping times taking values in \mathbb{R}_+ derive directly from the properties of stopping times taking values in \mathbb{N} studied in Chap. 2. Note that for any stopping time T , and for the translation operator θ_s for $s \in \mathbb{R}_+$, we have $(T \circ \theta_s = t + s) = (T = t)$.

Definition 4.8 Let $\mathbf{X} = (X_t)_{t \in \mathbb{T}}$ be a stochastic process defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, where $\mathbb{T} \subset \mathbb{R}$.

1. The probability $\mathbb{P}_{\mathbf{X}} = \mathbb{P} \circ \mathbf{X}^{-1}$ defined on $(\mathbb{R}^{\mathbb{T}}, \mathcal{B}(\mathbb{R}^{\mathbb{T}}))$ by

$$\mathbb{P}_{\mathbf{X}}(B) = \mathbb{P}(\mathbf{X} \in B), \quad B \in \mathcal{B}(\mathbb{R}^{\mathbb{T}}),$$

is called the distribution of \mathbf{X} .

2. The probabilities $\mathbb{P}_{t_1, \dots, t_n}$ defined by

$$\mathbb{P}_{t_1, \dots, t_n}(B_1 \times \dots \times B_n) = \mathbb{P}(X_{t_1} \in B_1, \dots, X_{t_n} \in B_n),$$

for $t_1 < \dots < t_n$, $t_i \in \mathbb{T}$, are called the finite dimensional distributions of \mathbf{X} . More generally, the restriction of $\mathbb{P}_{\mathbf{X}}$ to $\mathbb{R}^{\mathbb{S}}$ for any $\mathbb{S} \subset \mathbb{T}$ is called the marginal distribution of \mathbf{X} on \mathbb{S} .

3. The functions F_{t_1, \dots, t_n} defined by

$$F_{t_1, \dots, t_n}(x_1, \dots, x_n) = \mathbb{P}(X_{t_1} \leq x_1, \dots, X_{t_n} \leq x_n),$$

for $t_1 < \dots < t_n$, $t_i \in \mathbb{T}$ and $x_i \in \mathbb{R}$, are called the finite dimensional distribution functions of \mathbf{X} .

▷ *Example 4.9 (Finite Dimensional Distribution Functions)* Let X be a positive random variable with distribution function F . Let \mathbf{X} be the stochastic process defined by $X_t = X - (t \wedge X)$ for $t \geq 0$. Its one-dimensional distribution functions are given by

$$\begin{aligned} F_t(x) &= P(X_t \leq x) = P(X_t \leq x, X \leq t) + P(X_t \leq x, X > t) \\ &= \mathbb{P}(0 \leq x, X \leq t) + \mathbb{P}(X - t \leq x, X > t) \\ &= F(t) + F(t+x) - F(t) = F(t+x), \quad x \geq 0, \end{aligned}$$

for $t \geq 0$, and its two-dimensional distribution functions are

$$\begin{aligned} F_{t_1, t_2}(x_1, x_2) &= F((x_1 + t_1) \wedge (x_2 + t_2)) + F((x_1 + t_1) \wedge t_2) \\ &\quad + F(t_1 \wedge (x_2 + t_2)) - 2F(t_1 \vee t_2), \quad x_1 \geq 0, x_2 \geq 0, \end{aligned}$$

for $t_1 \geq 0$ et $t_2 \geq 0$. ◁

For a given family of distribution functions, does a stochastic process on some probability space exist with these functions as distribution functions? The answer may be positive, for instance under the conditions given by the following theorem, which we state without proof.

Theorem 4.10 (Kolmogorov) *Let (F_{t_1, \dots, t_n}) , for $t_1 < \dots < t_n$ in \mathbb{T} , be a family of distribution functions satisfying for any $(x_1, \dots, x_n) \in \mathbb{R}^n$ the two following coherence conditions:*

1. *for any permutation (i_1, \dots, i_n) of $(1, \dots, n)$,*

$$F_{t_1, \dots, t_n}(x_1, \dots, x_n) = F_{t_{i_1}, \dots, t_{i_n}}(x_{i_1}, \dots, x_{i_n});$$

2. *for all $k \in \llbracket 1, n-1 \rrbracket$,*

$$F_{t_1, \dots, t_k, \dots, t_n}(x_1, \dots, x_k, +\infty, \dots, +\infty) = F_{t_1, \dots, t_k}(x_1, \dots, x_k).$$

Then, there exists a stochastic process $\mathbf{X} = (X_t)_{t \in \mathbb{T}}$ defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ such that

$$\mathbb{P}(X_{t_1} \leq x_1, \dots, X_{t_n} \leq x_n) = F_{t_1, \dots, t_n}(x_1, \dots, x_n).$$

For stochastic processes, the notion of equivalence takes the following form.

Definition 4.11 Let $\mathbf{X} = (X_t)_{t \in \mathbb{T}}$ and $\mathbf{Y} = (Y_t)_{t \in \mathbb{T}}$ be two stochastic processes both defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and taking values in $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$. They are said to be:

1. weakly stochastically equivalent if for all $(t_1, \dots, t_n) \in \mathbb{T}^n$, all $(B_1, \dots, B_n) \in \mathcal{B}(\mathbb{R}^d)^n$ and all $n \in \mathbb{N}^*$,

$$\mathbb{P}(X_{t_1} \in B_1, \dots, X_{t_n} \in B_n) = \mathbb{P}(Y_{t_1} \in B_1, \dots, Y_{t_n} \in B_n). \quad (4.1)$$

2. stochastically equivalent if

$$\mathbb{P}(X_t = Y_t) = 1, \quad t \in \mathbb{T}. \quad (4.2)$$

Then, \mathbf{Y} is called a version of \mathbf{X} .

3. indistinguishable if

$$\mathbb{P}(X_t = Y_t, \forall t \in \mathbb{T}) = 1. \quad (4.3)$$

The trajectories of two indistinguishable processes are a.s. equal. This latter property is stronger than the two former ones; precisely

$$(4.3) \implies (4.2) \implies (4.1).$$

The converse implications do not hold, as the following example shows.

▷ *Example 4.12* Let Z be a continuous positive random variable. Let \mathbf{X} and \mathbf{Y} be two processes indexed by \mathbb{R}_+ , defined by $X_t = 0$ for all t and $Y_t = \mathbb{1}_{(Z=t)}$. These two processes are stochastically equivalent, but not indistinguishable. Indeed, $\mathbb{P}(X_t \neq Y_t) = \mathbb{P}(Z = t) = 0$, for all $t \geq 0$, but $\mathbb{P}(X_t = Y_t, \forall t \geq 0) = 0$. ◁

Definition 4.13 A stochastic process $\mathbf{X} = (X_t)_{t \in \mathbb{R}}$ is said to be stochastically continuous at $s \in \mathbb{R}$ if, for all $\varepsilon > 0$,

$$\lim_{t \rightarrow s} \mathbb{P}(|X_t - X_s| > \varepsilon) = 0.$$

Note that stochastic continuity does not imply continuity of the trajectories of the process.

Most of the probabilistic notions defined for sequences of random variables indexed by \mathbb{N} in Chap. 1 extend naturally to stochastic processes indexed by \mathbb{R}_+ .

Definition 4.14 Let \mathbf{X} be a stochastic process. The quantities

$$\mathbb{E}(X_{t_1}^{n_1} \dots X_{t_d}^{n_d}), \quad d \in \mathbb{N}^*, n_i \in \mathbb{N}^*, \sum_{i=1}^d n_i = n,$$

are called the order n moments of \mathbf{X} , when they are finite.

Especially, $\mathbb{E} X_t$ is called the mean value at time t and $\text{Var} X_t$ the instantaneous variance. The latter is also called power by analogy to the electric tension X_t in a resistance, for which $\mathbb{E} (X_t^2)$ is the instantaneous power.

The natural extension of the covariance matrix for finite families is the covariance function.

Definition 4.15 Let \mathbf{X} be a stochastic process. The function $R_{\mathbf{X}}: \mathbb{T} \times \mathbb{T} \longrightarrow \overline{\mathbb{R}}$ defined by

$$R_{\mathbf{X}}(t_1, t_2) = \text{Cov}(X_1, X_2) = \mathbb{E}(X_{t_1} X_{t_2}) - (\mathbb{E} X_{t_1})(\mathbb{E} X_{t_2}),$$

is called the covariance function of the process.

4.1.1 Properties of Covariance Functions

1. If $R_{\mathbf{X}}$ takes values in \mathbb{R} , then \mathbf{X} is a second order process, and $\text{Var} X_t = R_{\mathbf{X}}(t, t)$.
2. A covariance function is a positive semi-definite function. Indeed, for all $(c_1, \dots, c_n) \in \mathbb{R}^n$, since covariance is bilinear,

$$\begin{aligned} \sum_{i=1}^n \sum_{j=1}^n c_i c_j R_{\mathbf{X}}(t_i, t_j) &= \text{Cov}\left(\sum_{i=1}^n c_i X_{t_i}, \sum_{j=1}^n c_j X_{t_j}\right) \\ &= \text{Var}\left(\sum_{i=1}^n c_i X_{t_i}\right) \geq 0. \end{aligned}$$

3. For a centered process,

$$R_{\mathbf{X}}(t_1, t_2)^2 \leq R_{\mathbf{X}}(t_1, t_1) R_{\mathbf{X}}(t_2, t_2),$$

by Cauchy-Schwarz inequality.

The moments of a stochastic process are obtained by averaging over Ω . They are called space averages. Another notion of average exists for processes, on the set of indices \mathbb{T} ; here we take $\mathbb{T} = \mathbb{N}$ for simplification.

Definition 4.16 Let $\mathbf{X} = (X_n)_{n \in \mathbb{N}}$ be a second order random sequence. The random variable

$$\bar{X} = \lim_{N \rightarrow +\infty} \frac{1}{N} \sum_{n=1}^N X_n,$$

is called the time mean of \mathbf{X} and the random variable

$$\overline{\mathbf{X}(m)} = \lim_{N \rightarrow +\infty} \frac{1}{N} \sum_{n=1}^N X_n X_{n+m}$$

is called the (mean) time power of \mathbf{X} on $m \in \mathbb{N}^*$. Limit is taken in the sense of convergence in mean square.

The time mean is linear as is the expectation. If $\mathbf{Y} = a\mathbf{X} + b$, then $\overline{\mathbf{Y}(m)} = a^2 \overline{\mathbf{X}(m)}$ as for the variance. In the next section, we will make precise the links between space and time averages.

The different notions of convergence defined in Chap. 1 for random sequences indexed by \mathbb{N} extend to stochastic processes indexed by \mathbb{R}_+ in a natural way. The next extension of the large numbers law has been proven for randomly indexed sequences in Theorem 1.93 in Chap. 1; the next proof is an interesting alternative. It is then completed by a central limit theorem.

Theorem 4.17 *Let (X_n) be an integrable i.i.d. random sequence. If $(N_t)_{t \in \mathbb{R}_+}$ is a stochastic process taking values in \mathbb{N}^* , a.s. finite for all t , independent of (X_n) and converging to infinity a.s. when t tends to infinity, then*

$$\frac{1}{N_t}(X_1 + \cdots + X_{N_t}) \xrightarrow{\text{a.s.}} \mathbb{E} X_1, \quad t \rightarrow +\infty.$$

Proof Set $\overline{X}_n = (X_1 + \cdots + X_n)/n$ for $n \geq 1$. By the strong law of large numbers, \overline{X}_n converges to $\mathbb{E} X_1$ almost surely, that is on $\Omega \setminus A$, where $A = \{\omega : Y_n(\omega) \not\rightarrow \mathbb{E} X_1\}$. Note that $(\overline{X}_{N_t(\omega)}(\omega))$ is a sub-sequence of $(\overline{X}_n(\omega))$, and set $B = \{\omega : N_t(\omega) \not\rightarrow \infty\}$ and $C = \{\omega : \overline{X}_{N_t(\omega)}(\omega) \not\rightarrow \mathbb{E} X_1\}$. Then $C \subset A \cap B$, and the proof is completed. \square

Theorem 4.18 (Anscombe) *Let $(X_n)_{n \in \mathbb{N}^*}$ be an i.i.d. random sequence with centered distribution with finite variance σ^2 . If $(N_t)_{t \in \mathbb{R}_+}$ is a stochastic process taking values in \mathbb{N}^* , a.s. finite for all t , independent of (X_n) and converging to infinity a.s. when t tends to infinity, then*

$$\frac{1}{\sigma \sqrt{N_t}}(X_1 + \cdots + X_{N_t}) \xrightarrow{\mathcal{D}} \mathcal{N}(0, 1), \quad t \rightarrow +\infty.$$

Proof Set $S_n = X_1 + \cdots + X_n$. We have

$$\mathbb{E}(e^{itS_{N_t}/\sigma\sqrt{N_t}}) = \mathbb{E}\left(\sum_{n \geq 1} e^{itS_n/\sigma\sqrt{n}} \mathbf{1}_{(N_t=n)}\right) = \sum_{n \geq 1} \mathbb{P}(N_t = n) \mathbb{E}(e^{itS_n/\sigma\sqrt{n}}),$$

so

$$|\mathbb{E}(e^{itS_{N_t}/\sigma\sqrt{N_t}}) - e^{-\sigma^2 t^2/2}| \leq \sum_{n \geq 1} \mathbb{P}(N_t = n) |\mathbb{E}(e^{itS_n/\sigma\sqrt{n}}) - e^{-\sigma^2 t^2/2}|.$$

On the one hand, due to the central limit theorem, $S_n/\sigma\sqrt{n}$ converges in distribution to a standard Gaussian variable, so, for all $\varepsilon > 0$, we obtain $|\mathbb{E}(e^{itS_n/\sigma\sqrt{n}}) - e^{-\sigma^2 t^2/2}| < \varepsilon$ for $n > n_\varepsilon$. Therefore,

$$\sum_{n > n_\varepsilon} \mathbb{P}(N_t = n) |\mathbb{E}(e^{itS_n/\sigma\sqrt{n}}) - e^{-\sigma^2 t^2/2}| \leq \varepsilon \mathbb{P}(N_t > n_\varepsilon) \leq \varepsilon.$$

On the other hand,

$$\sum_{n=1}^{n_\varepsilon} \mathbb{P}(N_t = n) |\mathbb{E}(e^{itS_n/\sigma\sqrt{n}}) - e^{-\sigma^2 t^2/2}| \leq 2\mathbb{P}(N_t \leq n_\varepsilon),$$

and the result follows because N_t converges to infinity and $\mathbb{P}(N_t \leq n_\varepsilon)$ tends to zero when t tends to infinity. \square

Results of the same type can be stated for more general functionals, as for example the ergodic theorems.

4.2 Stationarity and Ergodicity

We will study here classical properties of processes taking values in \mathbb{R} . First, a stochastic process is stationary if it is invariant by translation of time, that is to say if it has no interne clock.

Definition 4.19 Let $\mathbf{X} = (X_t)_{t \in \mathbb{T}}$ be a stochastic process. It is said to be strictly stationary if

$$(X_{t_1}, \dots, X_{t_n}) \sim (X_{t_1+s}, \dots, X_{t_n+s}), \quad (s, t_1, \dots, t_n) \in \mathbb{T}^{n+1}.$$

\triangleright *Example 4.20 (Some Stationary Sequences)* An i.i.d. random sequence is stationary. An ergodic Markov chain whose initial distribution is the stationary distribution is stationary. \triangleleft

Ergodicity, a notion of invariance on the space Ω , has been introduced for Markov chains in Chap. 3. It extends to general random sequences as follows.

Definition 4.21 A random sequence (X_n) is said to be strictly ergodic:

1. at order 1 if for any real number x the random variable

$$\lim_{N \rightarrow +\infty} \frac{1}{N} \sum_{n=0}^{N-1} \mathbb{1}_{(X_n \leq x)}$$

is constant;

2. at order 2 if for all real numbers x_1 and x_2 the random variable

$$\lim_{N \rightarrow +\infty} \frac{1}{N} \sum_{n=0}^{N-1} \mathbb{1}_{(X_n \leq x_1, X_{n+m} \leq x_2)}$$

is constant for all $m \in \mathbb{N}$.

Since $\mathbb{P}(X_n \leq x) = \mathbb{E}[\mathbb{1}_{(X_n \leq x)}]$, a random sequence is ergodic at order 1 (order 2) if the distributions of the random variables X_n (pairs (X_n, X_{n+m})) can be obtained by time average.

▷ *Example 4.22 (Markov Chains)* Any aperiodic positive recurrent Markov chain is ergodic, as seen in Chap. 3. ◁

▷ *Example 4.23* The stochastic process modeling the sound coming out of a tape recorder, with null ageing but parasite noise, is ergodic. Indeed, the observation of one long-time recording is obviously similar to the observation of several short-time recordings. This process is also obviously stationary. ◁

The following convergence result is of paramount importance in statistics. We state it without proof.

Theorem 4.24 (Ergodic) *If (X_n) is a strictly stationary and ergodic random sequence and if $g : \mathbb{R} \rightarrow \mathbb{R}$ is a Borel function, then*

$$\frac{1}{N} \sum_{n=0}^{N-1} g(X_n) \rightarrow \mathbb{E} g(X_0), \quad N \rightarrow +\infty.$$

In ergodic theory, the strict stationarity and ergodicity of a stochastic process are expressed using set transformations.

Let (E, \mathcal{E}, μ) be a measured set. A measurable function $S : E \rightarrow E$ is called a transformation of E ; we denote by Sx the image of an element $x \in E$ and we set $S^{-1}B = \{x \in E : Sx \in B\}$, for $B \in \mathcal{E}$.

A set B of \mathcal{E} is said to be S -invariant if $\mu(S^{-1}(B)) = \mu(B)$. The collection of S -invariant sets constitutes a σ -field, denoted by \mathcal{J} . The function S is said to be

μ -invariant if the elements of \mathcal{E} are all S -invariant. The transformation is said to be strictly ergodic if for all $B \in \mathcal{E}$,

$$S^{-1}(B) = B \implies \mu(B) = 0 \text{ or } 1.$$

The following convergence result, stated without proof, is also called the pointwise ergodic theorem.

Theorem 4.25 (Birkhoff) *Let μ be a finite measure on a measurable set (E, \mathcal{E}) . If μ is S -invariant for some transformation $S : E \rightarrow E$, then, for all μ -integrable functions $f : (E, \mathcal{E}, \mu) \rightarrow \mathbb{R}$,*

$$\frac{1}{n} \sum_{i=0}^{n-1} f(S^i x) \longrightarrow \bar{f}(x) \quad \mu - a.e., \quad x \in E,$$

where $\bar{f} : E \rightarrow \mathbb{R}$ is a μ -integrable function such that $\bar{f}(x) = \bar{f}(Sx)$ μ -a.e., and

$$\int_E f(t) d\mu(t) = \frac{1}{\mu(E)} \int_E \bar{f}(t) d\mu(t).$$

If, moreover, S is ergodic, then

$$\bar{f}(x) = \int_E f(t) d\mu(t), \quad a.e..$$

Note that if μ is a probability, then by definition, \bar{f} is the conditional expectation of f given \mathcal{J} , or $\bar{f} = \mathbb{E}(f | \mathcal{J})$. If S is ergodic, then $\bar{f} = \mathbb{E} f$ a.s.

▷ *Example 4.26 (Interpretation in Physics)* Let us observe at discrete times a system evolving in an Euclidean space E .

Let x_0, x_1, x_2, \dots be the points successively occupied by the system. Let S transform the elements of E through $x_n = Sx_{n-1}$, for $n \geq 1$. Let S^n denote the n -th iterate of S , that is $x_n = S^n x_0$. Thus, the sequence $(x_0, x_1, x_2, \dots, x_n, \dots)$ can be written $(x_0, Sx_0, S^2x_0, \dots, S^n x_0, \dots)$ and is called an orbit of the point x_0 (or trajectory).

Let $f : E \rightarrow \mathbb{R}$ be a function whose values $f(x)$ express a physical measure (speed, temperature, pressure, etc.) at $x \in E$. Thanks to former experiments, the measure $f(x)$ is known to be subject to error and the empirical mean $[f(x_0) + f(x_1) + \dots + f(x_{n-1})]/n$ to give a better approximation of the true measured quantity. For n large enough, this mean is close to the limit

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{i=0}^{n-1} f(S^i x),$$

which should be equal to the true physical quantity. ◁

Any stochastic process $\mathbf{X} = (X_t)_{t \in \mathbb{T}}$ on $(\Omega, \mathcal{F}, \mathbb{P})$, taking values in E , can be considered as defined by

$$X_t(\omega) = X(S_t\omega), \quad \omega \in \Omega, \quad t \in \mathbb{T},$$

where $X : \Omega \rightarrow E$ is a random variable and $\{S_t : \Omega \rightarrow \Omega : t \in \mathbb{T}\}$ is a group of transformations. If $\mathbb{T} = \mathbb{N}$, then S_n is the n -th iterate of some transformation $S : \Omega \rightarrow \Omega$. Clearly, the process is strictly stationary if

$$\mathbb{P}((S_t)^{-1}(A)) = \mathbb{P}(A), \quad A \in \mathcal{F}, \quad t \in \mathbb{T},$$

and is strictly ergodic at order 1 if

$$\forall A \in \mathcal{F}, \quad (S_t)^{-1}(A) = A, \quad t \in \mathbb{T} \implies \mathbb{P}(A) = 0 \text{ or } 1.$$

Any real valued process $\mathbf{X} = (X_t)_{t \in \mathbb{T}}$ can be defined on its canonic space $(\mathbb{R}^{\mathbb{T}}, \mathcal{B}(\mathbb{R}^{\mathbb{T}}), \mathbb{P}_{\mathbf{X}})$ by setting

$$X_t(\omega) = \omega(t), \quad \omega \in \mathbb{R}^{\mathbb{T}}, \quad t \in \mathbb{T}.$$

Strict stationarity thus appears as a property of invariance of the probability $\mathbb{P}_{\mathbf{X}}$ with respect to translation operators, that is $\mathbb{P}_{\mathbf{X}} \circ \theta_t^{-1} = \mathbb{P}_{\mathbf{X}}$ for all $t \in \mathbb{T}$, or

$$\mathbb{P}_{\mathbf{X}}(\theta_t^{-1}(B)) = \mathbb{P}_{\mathbf{X}}(B), \quad B \in \mathcal{B}(\mathbb{R}^{\mathbb{T}}), \quad t \in \mathbb{T}.$$

Strict ergodicity can be expressed using θ_t -invariant sets. Precisely, the process is ergodic if

$$\forall B \in \mathcal{B}(\mathbb{R}^{\mathbb{T}}), \quad \theta_t^{-1}(B) = B, \quad t \in \mathbb{T} \implies \mathbb{P}_{\mathbf{X}}(B) = 0 \text{ or } 1.$$

For a real stationary and integrable random sequence (X_n) , Birkhoff's ergodic theorem yields

$$\frac{1}{n} \sum_{i=0}^{n-1} X_i \xrightarrow{a.s.} \bar{X} = \mathbb{E}(X_0 | \mathcal{J}),$$

where \mathcal{J} is the σ -field of θ_1 -invariant sets. For an ergodic sequence—such that θ_1 -invariant events have probability 0 or 1, the limit is $\bar{X} = \mathbb{E} X_0$.

Note that, according to Kolmogorov 0–1 law, any i.i.d. sequence is stationary and ergodic.

Further, the entropy rate of a random sequence has been defined in Chap. 1. The entropy rate of a continuous time stochastic process indexed by \mathbb{R}_+ is defined similarly.

Definition 4.27 Let $\mathbf{X} = (X_t)_{t \in \mathbb{R}_+}$ be a stochastic process such that the marginal distribution of $\mathbf{X} = (X_t)_{t \in [0, T]}$ has a density $p_T^{\mathbf{X}}$ with respect either to the Lebesgue or to the counting measure, for all T . The entropy up to time T of \mathbf{X} is the entropy of its marginal distribution, that is

$$\mathbb{H}_T(\mathbf{X}) = -\mathbb{E} \log p_T^{\mathbf{X}}(X).$$

If $\mathbb{H}_T(\mathbf{X})/T$ has a limit when T tends to infinity, this limit is called the entropy rate of the process and is denoted by $\mathbb{H}(\mathbf{X})$.

The next result, stated here for stationary and ergodic sequences is of paramount importance in information theory. Note that its proof is based on Birkhoff's ergodic theorem.

Theorem 4.28 (Ergodic Theorem of Information Theory) Let $\mathbf{X} = (X_t)_{t \in \mathbb{T}}$, with $\mathbb{T} = \mathbb{R}_+$ or \mathbb{N} , be a stationary and ergodic stochastic process. Suppose that the marginal distribution of $\mathbf{X} = (X_t)_{t \in [0, T]}$ has a density $p_T^{\mathbf{X}}$ with respect to the Lebesgue or counting measure, for all T . If the entropy rate $\mathbb{H}(\mathbf{X})$ is finite, then $-\log p_T^{\mathbf{X}}(X)/T$ converges in mean and a.s. to $\mathbb{H}(\mathbf{X})$.

Explicit expressions of $\mathbb{H}(\mathbf{X})$ for jump Markov and semi-Markov processes will be developed in the next chapter.

Other notions of stationarity exist, less strict and more convenient for applications; we define here only first and second order stationarity.

Definition 4.29 A second order process \mathbf{X} is said to be weakly stationary:

1. to the order 1 if $\mathbb{E} X_{t_1} = \mathbb{E} X_{t_2} = m_X$ for all $(t_1, t_2) \in \mathbb{T}^2$.
2. to the order 2 if, moreover, $R_{\mathbf{X}}(t_1, t_2) = R_{\mathbf{X}}(t_1 + \tau, t_2 + \tau)$ for all $\tau \in \mathbb{T}$ and $(t_1, t_2) \in \mathbb{T}^2$.

The mean and the variance of X_t are then constant in t , but the X_t do not necessarily have the same distribution.

▷ *Example 4.30 (Sinusoidal Signal (Continuation of Example 4.1))* Let \mathbf{X} be as defined in Example 4.1. Suppose that ν is constant, and that $\varphi \sim \mathcal{U}[-\pi, \pi]$ and A are independent variables. Then $\mathbb{E} X_t = 0$ and $R_{\mathbf{X}}(t_1, t_2) = \mathbb{E}(A^2) \cos[2\pi \nu(t_1 - t_2)]/2$; hence, this sinusoidal signal is second order stationary. Clearly, it is not strictly stationary. ◁

For a second order process, stationarity induces weak stationarity. If \mathbf{X} is a Gaussian process, the converse holds true. Indeed, then $\mathbb{E}(X_{t_1}, \dots, X_{t_n}) = M$ is constant and the covariance matrix $\Gamma_{X_{t_1}, \dots, X_{t_n}}$ is a matrix valued function of the differences $t_j - t_i$. The distribution of $(X_{t_1}, \dots, X_{t_n})$ depends only on M and Γ .

Further, the covariance function of a weakly stationary process can be written $R_{\mathbf{X}}(t_1, t_2) = r_{\mathbf{X}}(t_2 - t_1)$, where $r_{\mathbf{X}}$ is an even function of only one variable, called the auto-correlation function of the process. The faster \mathbf{X} fluctuates, the faster $r_{\mathbf{X}}$ decreases.

If \mathbf{X} is a second order process, the Cauchy-Schwarz inequality yields $|r_{\mathbf{X}}(t_2 - t_1)| \leq r_{\mathbf{X}}(0)$. The function $r_{\mathbf{X}}$ is positive definite; if, moreover, it is continuous, then it has a spectral representation—by Bochner’s theorem, meaning that it is the Fourier transform of a bounded positive measure. For example, if $\mathbb{T} = \mathbb{R}(\mathbb{Z})$,

$$r_{\mathbf{X}}(t) = \int_{\Lambda} e^{i\lambda t} \mu(d\lambda)$$

where $\Lambda = \mathbb{R}([0, 2\pi])$ and μ is called the spectral measure of the process. If μ is absolutely continuous with respect to the Lebesgue measure, its density is called the spectral density of the process. The term “spectral” designs what is connected to frequencies.

Then the integral representation of the process follows—by Karhunen’s theorem,

$$X_t = \int_{\Lambda} e^{i\lambda t} Z(d\lambda),$$

where \mathbf{Z} is a second order process with independent increments—see next section, indexed by Λ .

▷ *Example 4.31 (White Noise)* Let $\boldsymbol{\varepsilon} = (\varepsilon_n)_{n \in \mathbb{Z}}$ be a stationary to the order 2 centered process, with $\mathbb{V}\text{ar } \varepsilon_n = 1$ and

$$\mathbb{E}(\varepsilon_n \varepsilon_m) = 0 = \int_0^{2\pi} \frac{1}{2\pi} e^{i\lambda(n-m)} d\lambda, \quad m \neq n.$$

Therefore, the spectral density of $\boldsymbol{\varepsilon}$ is constant. All the frequencies are represented with the same power in the signal, hence its name: white noise, by analogy to white light from an incandescent body. ◁

▷ *Example 4.32 (ARMA Processes)* An $\text{MA}(q)$ process \mathbf{X} is obtained from a white noise $\boldsymbol{\varepsilon}$ by composition with the function $f(x_1, \dots, x_n) = \sum_{k=1}^q b_k x_{n-k}$, called linear filter. We compute $\mathbb{E}(X_n X_{n+m}) = \sum_{k=1}^q b_k b_{m+k}$. The spectral density of \mathbf{X} is obtained from that of $\boldsymbol{\varepsilon}$, that is

$$h_{\mathbf{X}}(\lambda) = \frac{1}{2\pi} \left| \sum_{k=1}^q b_k e^{-ik\lambda} \right|^2.$$

In the same way, an AR(p) process \mathbf{Y} is obtained by recursive filtering of order p from $\boldsymbol{\varepsilon}$, and

$$h_{\mathbf{Y}}(\lambda) = \frac{1}{2\pi} \left| \sum_{k=1}^q a_k e^{-ik\lambda} \right|^{-2}.$$

Finally, the spectral density of an ARMA(p, q) process \mathbf{Z} is

$$h_{\mathbf{Z}}(\lambda) = \frac{1}{2\pi} \frac{|\sum_{k=1}^q b_k e^{-ik\lambda}|^2}{|\sum_{k=1}^p a_k e^{-ik\lambda}|^2},$$

for $\lambda \in [0, 2\pi]$. ◁

Just as strict stationarity, strict ergodicity is rarely satisfied in applications, and hence weaker notions of ergodicity are considered.

Definition 4.33 A random sequence (X_n) is said to be:

1. first order (or mean) ergodic if the random variable $\overline{\mathbf{X}}$ is a.s. constant.
2. second order ergodic if, moreover, for all $\tau \in \mathbb{R}$, the random variable $\overline{\mathbf{X}(m)}$ is a.s. constant.

Mean ergodicity is satisfied under the following simple condition.

Proposition 4.34 A random sequence (X_n) is first order ergodic if and only if its covariance function $R_{\mathbf{X}}$ is summable.

Proof For a centered sequence \mathbf{X} .

We compute

$$\mathbb{E} \left(\frac{1}{N} \sum_{n=0}^{N-1} X_n \right)^2 = \frac{1}{N^2} \sum_{n=1}^N \sum_{m=1}^N R_{\mathbf{X}}(n, m)$$

If the sequence is first order ergodic, $\frac{1}{N} \sum_{n=1}^N X_n$ converges in L^2 to a constant, so

$$\sum_{n=0}^{N-1} \sum_{m=0}^{N-1} R_{\mathbf{X}}(n, m) < +\infty.$$

Conversely, if $R_{\mathbf{X}}$ is summable,

$$\frac{1}{N^2} \sum_{n=0}^{N-1} \sum_{m=0}^{N-1} R_{\mathbf{X}}(n, m) \xrightarrow{L^2} 0, \quad N \rightarrow +\infty.$$

and the sequence is first order ergodic. ◻

Stationarity and ergodicity are not equivalent, as shown by the following examples.

▷ *Example 4.35* A process \mathbf{X} constant with respect to t , equal to some random variable V , is obviously stationary. Since $\overline{\mathbf{X}} = V$, it is first order ergodic only if V is a.s. constant.

By linearity of integration, the sum of two ergodic processes is ergodic too. The sum of two independent second order stationary processes is second order stationary too. On the contrary, the sum of two dependent second order stationary processes is not stationary in general.

Finally, let \mathbf{X} be a process and let V be a random variable independent of \mathbf{X} . If \mathbf{X} is stationary, the processes $(V X_t)$ and $(X_t + V)$ are stationary too; on the contrary, if \mathbf{X} is ergodic, $(V X_t)$ and $(X_t + V)$ are ergodic too only if V is a.s. constant. ◁

If the process is only weakly stationary and ergodic, the ergodic theorem does not hold in general. Still, the following result holds true. In this case, the mean power is equal to the instantaneous power.

Proposition 4.36 *If (X_n) is stationary and ergodic to the order 2, then $\mathbb{E} X_n = \overline{\mathbf{X}}$ for $n \in \mathbb{N}$ and $\overline{\mathbf{X}(m)} = r_{\mathbf{X}}(m) + \overline{\mathbf{X}}^2$ a.s. for $m \in \mathbb{N}$.*

Proof The random variable $\overline{\mathbf{X}}$ is a.s. constant, and all the X_n have the same distribution, so $\mathbb{E} X_n = \overline{\mathbf{X}}$ by linearity. In the same way, since $\mathbb{E} X_n = \mathbb{E} X_{n+m} = \overline{\mathbf{X}}$ and $r_{\mathbf{X}}(m) = \mathbb{E}(X_n X_{n+m}) - \mathbb{E} X_n \mathbb{E} X_{n+m}$ does not depend on n , we can write

$$r_{\mathbf{X}}(m) = \frac{1}{N} \sum_{n=0}^{N-1} \mathbb{E}(X_n X_{n+m}) - \left(\frac{1}{N} \sum_{n=0}^{N-1} \overline{\mathbf{X}} \right)^2,$$

and $r_{\mathbf{X}}(m) = \overline{\mathbf{X}(m)} - \overline{\mathbf{X}}^2$ a.s. ◻

Thus, the moments of the marginal distributions of weakly stationary and ergodic sequences appear to be characterized by time averaging, that is by the knowledge of one—and only one—full trajectory. This explains the importance of ergodicity in the statistical study of marginal distributions.

4.3 Processes with Independent Increments

Numerous types of processes have independent increments, among which we will present the Brownian motion, Poisson processes, compound Poisson processes, etc.

Definition 4.37 Let $\mathbf{X} = (X_t)_{t \in \mathbb{T}}$ be a stochastic process adapted to a filtration (\mathcal{F}_t) , where $\mathbb{T} = \mathbb{R}$ or $\mathbb{T} = \mathbb{N}$.

1. The process is said have independent increments if for all $s < t$ in \mathbb{T} , the random variable $X_t - X_s$ is independent of the σ -algebra \mathcal{F}_s .
2. The process is said to have stationary increments if the distribution of $X_t - X_s$, for $s < t$ in \mathbb{T} , depends only on $t - s$.
3. A process with independent and stationary increments is said to be homogeneous (with respect to time).

When $\mathbb{T} = \mathbb{N}$ and \mathbf{X} is a process with independent increments, the random variables $X_{t_0}, X_{t_1} - X_{t_0}, \dots, X_{t_n} - X_{t_{n-1}}$ are independent for all $n \in \mathbb{N}$ and $t_0 < t_1 < \dots < t_n$. Since $X_n = \sum_{i=1}^n (X_i - X_{i-1})$, knowing a process with independent increments is equivalent to knowing its increments. If \mathbf{X} is a homogeneous process, then necessarily, $X_0 = 0$ a.s.

Let us now present a typical process with independent increments, the Brownian motion. It is an example of Gaussian, ergodic, non stationary process, with independent increments. A trajectory of a Brownian motion is shown in Fig. 4.1.

Definition 4.38 The process $\mathbf{W} = (W_t)_{t \in \mathbb{R}_+}$ taking values in \mathbb{R} , with independent increments and such that $W_0 = 0$, and for all $0 \leq t_1 < t_2$,

$$W_{t_2} - W_{t_1} \sim \mathcal{N}(0, t_2 - t_1),$$

is called a standard Brownian motion (or Wiener process) with parameters ν and σ^2 .

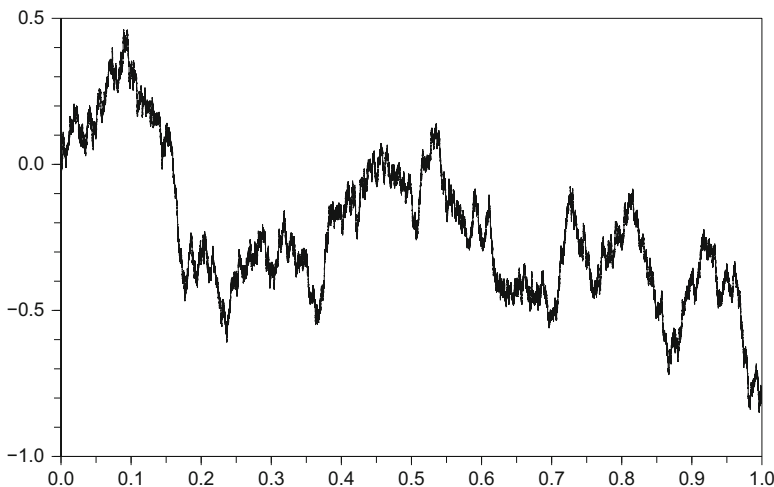


Fig. 4.1 A trajectory of a Brownian motion

The process $\mathbf{W} = (W_t)_{t \in \mathbb{R}_+}$ taking values in \mathbb{R} , with independent increments and such that $W_0 = 0$, and for all $0 \leq t_1 < t_2$,

$$W_{t_2} - W_{t_1} \sim \mathcal{N}(v(t_2 - t_1), \sigma^2(t_2 - t_1)),$$

is called a Brownian motion with drift parameter v and diffusion parameter σ^2 .

Considering the random walk (S_n) of Example 1.77, with $p = 1/2$, yields an elementary construction of a Brownian motion. Indeed, setting

$$X_t = S_{[nT]}, \quad t \in \mathbb{R}_+$$

defines a continuous time process. If $t = nT$, then $\mathbb{E} X_t = 0$ and $\text{Var} X_t = ts^2/T$. Let now t be fixed. If both s and T tend to 0, then the variance of X_t remains fixed and non null if and only if $s \approx \sqrt{T}$.

Let us set $s^2 = \sigma^2 T$ where $\sigma^2 \in \mathbb{R}_+^*$ and define a process \mathbf{W} by

$$W_t(\omega) = \lim_{T \rightarrow 0} X_t(\omega), \quad t \in \mathbb{R}_+, \omega \in \Omega.$$

Taking the limit yields $\mathbb{E} W_t = 0$ and $\text{Var} W_t = \sigma^2 t$.

Let us show that $W_t \sim \mathcal{N}(0, \sigma^2 t)$ by determining its distribution function at any $w \in \mathbb{R}_+$. Set $r = w/s$ and $T = t/n$. If w and t are fixed and T tends to 0, since $s \approx \sqrt{T}$, we get $r \approx \sqrt{n}$. Since $\mathbb{P}[X_{nT} = (2k - n)s] = \binom{n}{k} p^k (1-p)^{n-k}$, we obtain by applying de Moivre-Laplace theorem

$$\mathbb{P}(S_n \leq rs) \approx \int_{-\infty}^{r/\sqrt{n}} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt,$$

or, since $r/\sqrt{n} = w/\sqrt{\sigma^2 t}$,

$$\mathbb{P}(W_t \leq w) \approx \int_{-\infty}^{w/\sqrt{\sigma^2 t}} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt.$$

Let us show that if $0 \leq t_1 < t_2 < t_3$, then $W_{t_2} - W_{t_1}$ and $W_{t_3} - W_{t_2}$ are independent. If $0 < n_1 < n_2 < n_3$, then the number of “heads” obtained between the n_1 -th and the n_2 -th tossings is independent of the number of “heads” obtained between the n_2 -th and n_3 -th tossings. Hence $S_{n_2} - S_{n_1}$ and $S_{n_3} - S_{n_2}$ are independent, and taking the limit yields the result.

Finally, let us compute the covariance function of \mathbf{W} . If $t_1 < t_2$, then $W_{t_2} - W_{t_1}$ and $W_{t_1} - W_0$ are independent. But $W_0 = 0$, so

$$\mathbb{E}[(W_{t_2} - W_{t_1})W_{t_1}] = [\mathbb{E}(W_{t_2} - W_{t_1})]\mathbb{E}W_{t_1} = 0.$$

Since we also have $\mathbb{E}[(W_{t_2} - W_{t_1})W_{t_1}] = \mathbb{E}(W_{t_2}W_{t_1}) - \mathbb{E}(W_{t_1}^2)$, we obtain $\mathbb{E}(W_{t_2}W_{t_1}) = \mathbb{E}W_{t_1}^2 = \sigma^2 t_1$, or $R_W(t_1, t_2) = \sigma^2(t_1 \wedge t_2)$.

Note that $X_i/\sqrt{n} = S_{[ni]}/\sqrt{n}$ gives an approximation of W_t that can be simulated as a sum of Bernoulli variables.

4.4 Point Processes on the Line

A point process is a stochastic process consisting of a finite or enumerable family of points set at random in an arbitrary space, for example gravel on a road, stars in a part of the sky, times of occurrences of failures of a given system, positions of one of the bases A, C, G, T of a DNA sequence, etc.

In a mathematical sense, a point can be multiple. Even if the space to which the considered points belong can be any topological space, we will here consider \mathbb{R}^d , for $d \geq 1$. After some general notions on point processes we will present the renewal process and the Poisson process on \mathbb{R}_+ .

4.4.1 Basics on General Point Processes

The point processes are naturally defined through random point measures.

Definition 4.39 Let μ be a measure on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ and let $(x_i)_{i \in I}$ be a sequence of points in \mathbb{R}^d , with $I \subset \mathbb{N}$. If

$$\mu = \sum_{i \in I} \delta_{x_i},$$

and if $\mu(K) < +\infty$ for all compact subsets K of \mathbb{R}^d , then μ is called a point measure on \mathbb{R}^d .

A point measure is a discrete measure. The multiplicity of $x \in \mathbb{R}^d$ is $\mu(\{x\})$. When $\mu(\{x\}) = 0$ or 1 for all $x \in \mathbb{R}^d$, then μ is said to be simple. If $\mu(\{x\}) = 1$ for all $x \in \mathbb{R}^d$, the measure $\mu(A)$ is equal to the number of points belonging to A , for all $A \in \mathcal{B}(\mathbb{R}^d)$.

Definition 4.40 A function $\mu : \Omega \times \mathcal{B}(\mathbb{R}^d) \rightarrow \overline{\mathbb{R}}$ such that $\mu(\omega, \cdot)$ is a point measure on \mathbb{R}^d for all ω is called a random point measure on \mathbb{R}^d .

When $\mu(\omega, \{x\}) = 0$ or 1 for all ω and x , the random measure μ is also said to be simple.

▷ *Example 4.41* Let n points in \mathbb{R}^d be set at random positions X_1, \dots, X_n . A random point measure on \mathbb{R}^d is defined by setting for all $A \in \mathcal{B}(\mathbb{R}^d)$,

$$\mu(A) = \sum_{i=1}^n \delta_{X_i}(A),$$

the random number of points belonging to A . ◁

▷ *Example 4.42* Let $(T_n)_{n \in \mathbb{N}^*}$ be the sequence of times in \mathbb{R}_+ of failures of a system. Then setting for all $s < t$ in \mathbb{R}_+ ,

$$\mu([s, t]) = \sum_{i \geq 1} \delta_{T_i}([s, t]),$$

the random number of failures observed in the time interval $[s, t]$, defines a random point measure on \mathbb{R}_+ . ◁

Let $M_p(\mathbb{R}^d)$ denote the set of all point measures on \mathbb{R}^d .

Definition 4.43 A function $\mathbf{N} : \Omega \times \mathcal{B}(\mathbb{R}^d) \rightarrow M_p(\mathbb{R}^d)$ such that $N(\cdot, A)$ are random variables for all $A \in \mathcal{B}(\mathbb{R}^d)$ is called a point process.

The variables $N(\cdot, A)$, taking values in $\overline{\mathbb{N}}$, are called the counting random variables of the point process \mathbf{N} , and the measure m defined by

$$m(A) = \mathbb{E}[N(\cdot, A)], \quad A \in \mathcal{B}(\mathbb{R}^d),$$

is its mean measure.

Each point process \mathbf{N} is associated with the random point measure μ defined by

$$\mu(\cdot, A) = \int \mathbb{1}_A d\mu = N(\cdot, A).$$

It is said to be simple if m is simple.

Thus, $N(\omega, A)$ counts the number of points of the process belonging to A for the outcome $\omega \in \Omega$. Note that $m(A)$ can be infinite even when $N(\cdot, A)$ is a.s. finite.

If m has a density $\lambda : \mathbb{R}^d \rightarrow \mathbb{R}_+$, that is $m(dx) = \lambda(x)dx$ or $\mathbb{P}[N(\cdot, dx) = 1] = \lambda(x)dx$, or

$$m(A) = \int_A \lambda(x)dx, \quad A \in \mathcal{B}(\mathbb{R}^d),$$

then the function λ , called the intensity of the process \mathbf{N} , is locally summable—meaning that λ is integrable over all bounded rectangles of \mathbb{R}^d .

Let us present some properties of integration with respect to a point measure.

Let $f : \mathbb{R}^d \rightarrow \mathbb{R}$ be a Borel function and μ a point measure on \mathbb{R}^d . The integral of f with respect to μ is

$$\mu(f) = \int_{\mathbb{R}^d} f d\mu = \sum_{i \in I} \int_{\mathbb{R}^d} f d\delta_{x_i} = \sum_{i \in I} \delta_{x_i}(f) = \sum_{i \in I} f(x_i).$$

The mean is

$$\mathbb{E} N(f) = \int_{\mathbb{R}^d} f dm,$$

and the Laplace functional associated with \mathbf{N} is defined as

$$L_N(f) = \mathbb{E} \left(\exp \left[- \int_{\mathbb{R}^d} f(x) N(\cdot, dx) \right] \right).$$

Especially, if f is positive and if $N = \sum_{n \geq 1} \delta_{X_n}$, then

$$L_N(f) = \mathbb{E} \left(\exp \left[- \sum_{n \geq 1} f(X_n) \right] \right).$$

The Laplace functional characterizes the distribution of a point process. Indeed, the distribution of a point process \mathbf{N} is given by its finite-dimensional distributions, that is by the distributions of all the random vectors $(N(\cdot, A_1), \dots, N(\cdot, A_n))$ for $n \geq 1$ and $A_1, \dots, A_n \in \mathcal{B}(\mathbb{R}^d)$. The function

$$L_N \left(\sum_{i=1}^n s_i \mathbb{1}_{A_i} \right) = \mathbb{E} \left(\exp \left[- \sum_{i=1}^n s_i N(\cdot, A_i) \right] \right)$$

is precisely the Laplace transform of the distribution of $(N(\cdot, A_1), \dots, N(\cdot, A_n))$ that characterizes its distribution.

4.4.2 Renewal Processes

Renewal processes are punctual processes defined on \mathbb{R}_+ , modeling many experiments in applied probability—in reliability, queues, insurance, risk theory... They are also valuable theoretical tools for investigating more complex processes, such as regenerative, Markov or semi-Markov processes. A renewal process can be regarded as a random walk with positive increments; the times between occurring events are i.i.d. It is not a Markovian process but a semi-Markov process, that is studied by specific methods.

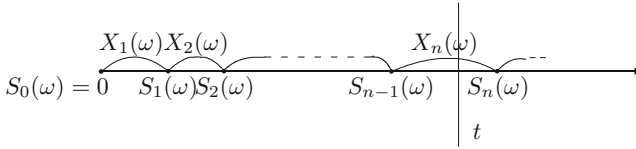


Fig. 4.2 A trajectory of a renewal process, on $(N_t = n)$

Definition 4.44 Let (X_n) be an i.i.d. sequence of positive variables. Set

$$S_0 = 0 \quad \text{and} \quad S_n = X_1 + \dots + X_n, \quad n \geq 1. \tag{4.4}$$

The random sequence (S_n) is called a renewal process. The random times S_n are called renewal times. The associated counting process is defined by $\mathbf{N} = (N_t)_{t \in \mathbb{R}_+}$, where

$$N_t = \sum_{n \geq 0} \mathbb{1}_{[0, t]}(S_n) = \inf\{n \geq 0 : S_n \leq t\}, \quad t \in \mathbb{R}_+. \tag{4.5}$$

Generally, (X_n) is the sequence of inter-arrival times of some sequence of events, and N_t counts the number of events in the time interval $[0, t]$. Note that $(N_t = n) = (S_{n-1} \leq t < S_n)$ and that $N_0 = 1$. A trajectory of a renewal process is shown in Fig. 4.2.

When each of the variables X_n has an exponential distribution with parameter λ , the counting process \mathbf{N} is called a Poisson process with parameter λ , in which case it is usual to set $N_0 = 0$; the Poisson processes will be especially investigated in the next section.

A counting process can also be regarded as a simple point process. Indeed,

$$N(\cdot, A) = \sum_{n \geq 0} \delta_{S_n}(A), \quad A \in \mathcal{B}(\mathbb{R}_+), \tag{4.6}$$

defines a point process, and we obtain for $A = [0, t]$,

$$N_t = N(\cdot, [0, t]), \quad t \geq 0.$$

▷ *Example 4.45 (A Renewal Process in Reliability)* A new component begins operating at time $S_0 = 0$. Let X_1 denote its lifetime. When it fails, it is automatically and instantly replaced by a new identical component. When the latter fails after a time X_2 , it is renewed, and so on.

If (X_n) is supposed to be i.i.d., then (4.4) defines a renewal process (S_n) whose distribution is that of the sum of the life durations of the components. The counting process $(N_t)_{t \in \mathbb{R}_+}$ gives the number of components used in $[0, t]$, of which the last component still works at time t . ◁

The expectation of the counting process at time t (or expected number of renewals), is

$$m(t) = \mathbb{E} N_t = \sum_{n \geq 0} \mathbb{P}(S_n \leq t),$$

and m is called the renewal function. If the variables X_n are not degenerated, this function is well-defined. Therefore, we will suppose thereafter that $F(0) < 1$, where F is the distribution function of X_n . The distribution function of S_n is the n -th Lebesgue-Stieltjes convolution of F , that is

$$F^{*(n)}(t) = \int_{\mathbb{R}_+} F^{*(n-1)}(t-x) dF(x),$$

with $F^{*(0)}(t) = \mathbb{1}_{\mathbb{R}_+}(t)$ and $F^{*(1)}(t) = F(t)$, so that

$$m(t) = \sum_{n \geq 0} F^{*(n)}(t), \quad t \in \mathbb{R}_+. \quad (4.7)$$

The mean measure m of the point process (4.6) is given by

$$m(A) = \mathbb{E} N(\cdot, A), \quad A \in \mathcal{B}(\mathbb{R}_+),$$

and we have $m(t) = m([0, t])$ pour $t \in \mathbb{R}_+$. Note that we use the same notation for both the renewal function and the mean measure.

When F is absolutely continuous with respect to the Lebesgue measure, the derivative of the renewal function $\lambda(t) = m'(t)$ is called the renewal density (or renewal rate) of the process.

Proposition 4.46 *The renewal function is increasing and finite.*

Proof For all $s > 0$, we have $N_{t+s} \geq N_t$, so m is increasing.

Assume that $F(t) < 1$ for $t > 0$. Then $F^{*(n)}(t) \leq [F(t)]^n$ for all n , and hence

$$m(t) \leq 1 + F(t) + [F(t)]^2 + \cdots \leq \frac{1}{1 - F(t)},$$

so $m(t)$ is finite. The general case is omitted. □

Relation (4.7) implies straightforwardly that $m(t)$ is a solution of

$$m(t) = 1 + \int_0^t m(t-x) dF(x), \quad t \in \mathbb{R}_+.$$

This equation is a particular case of the scalar renewal equation

$$h = g + F * h, \quad (4.8)$$

where h and g are functions bounded on the finite intervals of \mathbb{R}_+ . The solution of this equation is determined by use of the renewal function m , as follows.

Proposition 4.47 *If $g: \mathbb{R}_+ \rightarrow \mathbb{R}$ is bounded on the finite intervals of \mathbb{R} , then (4.8) has a unique solution $h: \mathbb{R}_+ \rightarrow \mathbb{R}$ that is bounded on the finite intervals of \mathbb{R} , given by*

$$h(t) = m * g(t) = \int_0^t g(t-x) dm(x),$$

where m is defined by (4.7), and extended to \mathbb{R}_- by 0.

Proof We deduce from (4.7) that

$$F * (m * g) = F * g + F^{*(2)} * g + \dots = m * g - g.$$

Thus, $m * g$ is a solution of (4.8).

Suppose that $k: \mathbb{R}_+ \rightarrow \mathbb{R}$ is another solution of (4.8) bounded on the finite intervals of \mathbb{R} . Then $h - k = F * (h - k)$. Since $F^{*(n)}$ tends to zero when n tends to infinity and $h - k$ is bounded, it follows that $h - k = F^{*(n)} * (h - k)$ and hence $k = h$. \square

Many extensions of renewal processes exist.

If S_0 is a nonnegative variable not identically zero, independent of (X_n) and with distribution function F_0 different from the renewal distribution function F , then the process (S_n) is said to be delayed or modified. When

$$F_0(x) = \frac{1}{\mu} \int_0^x [1 - F(u)] du, \quad x > 0,$$

the delayed renewal process is said to be stationary.

Definition 4.48 Let (Y_n) and (Z_n) be two i.i.d. nonnegative independent random sequences, with respective distribution functions G and H . The associated process defined by

$$S_0 = 0 \quad \text{and} \quad S_n = S_{n-1} + Y_n + Z_n, \quad n \geq 1,$$

is called an alternated renewal process.

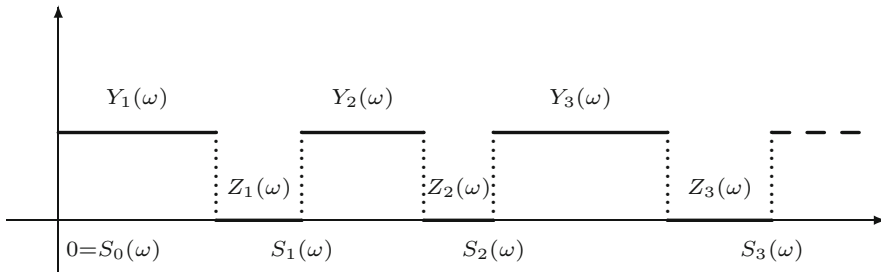


Fig. 4.3 A trajectory of an alternated renewal process

Such an alternated renewal process (S_n) is shown to be a renewal process in the sense of Definition 4.44 by setting $X_n = Y_n + Z_n$ and $F = G * H$. A trajectory of an alternated renewal process is shown in Fig. 4.3.

Still another extension is the stopped renewal process, also called transient renewal process.

Definition 4.49 Let (X_n) be an i.i.d. random sequence taking values in $\overline{\mathbb{R}}_+$, with defective distribution function F . The associated renewal process defined by (4.4) is called a stopped renewal process.

The life duration of this process is $T = \sum_{n=1}^N X_n$, where N is the number of events at the stopping time of the process, defined by

$$(N = n) = (X_1 < +\infty, \dots, X_n < +\infty, X_{n+1} = +\infty).$$

▷ *Example 4.50* If N has a geometric distribution on \mathbb{N}^* with parameter q , then the distribution function of T is

$$F_T(t) = \sum_{n \geq 1} F^{*(n)}(t)(1 - q)^{n-1}q.$$

Indeed, $\mathbb{P}(N = n) = (1 - q)^{n-1}q$ for $n \geq 1$, and the distribution function of T follows by Point 2. of Proposition 1.73. ◁

▷ *Example 4.51 (Risk Process in Insurance)* Let $u > 0$ be the initial capital of an insurance company. Let (S_n) be the sequence of times at which accidents occur, and let $\mathbf{N} = (N_t)$ denote the associated counting process. Let (Y_n) be the sequence of compensations paid at each accident. The capital of the company at time t is $U_t = u + ct - \sum_{n=1}^{N_t} Y_n$, where c is the rate of subscriptions in $[0, t]$. A trajectory of such a process is shown in Fig. 4.4. The time until ruin is T .

The linked problems are the ruin in a given time interval, that is $\mathbb{P}(U_t \leq 0)$ with limit $\mathbb{P}(\lim_{t \rightarrow +\infty} U_t \leq 0)$, and the mean viability of the company, that is

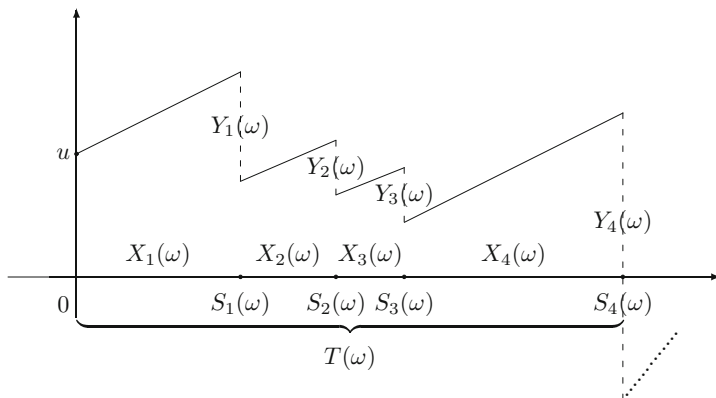


Fig. 4.4 A trajectory of a risk process

$\mathbb{E} U_t \geq 0$. Different approaches exist for solving these issues, involving either renewal theory or martingale theory. Note that a particular case of risk process, the Cramér-Lundberg process, will be presented in the next section. \triangleleft

Renewal process can also be considered in the vector case. Another extension will also be studied below, the Markov renewal process.

4.4.3 Poisson Processes

The renewal process whose counting process is a Poisson process is the most used in modeling real experiments. We keep the notation of the preceding section.

Definition 4.52 If $X_n \sim \mathcal{E}(\lambda)$ for $n \in \mathbb{N}^*$, the counting process $\mathbf{N} = (N_t)_{t \in \mathbb{R}_+}$ defined by (4.5) p. 195 for $t \in \mathbb{R}_+$, with $N_0 = 0$, is called a homogeneous Poisson process with intensity (or parameter) λ .

Thanks to the absence of memory of the exponential distribution, the probability that an event occurs for the first time after time $s + t$ given that it did not occur before time t is equal to the probability that it occurs after time s . More generally, the following result holds true.

Theorem 4.53 A Poisson process \mathbf{N} with intensity λ is homogeneous—with independent and stationary increments—and satisfies $N_t \sim \mathcal{P}(\lambda t)$ for $t \in \mathbb{R}_+$.

Proof Let us show first that $N_t \sim \mathcal{P}(\lambda t)$ for all $t > 0$. We have

$$\mathbb{P}(N_t = k) = \mathbb{P}(S_k \leq t) - \mathbb{P}(S_{k+1} \leq t).$$

According to Example 1.66, $S_n \sim \gamma(n, \lambda)$, so

$$\mathbb{P}(S_n \leq x) = \int_0^x \frac{1}{(n-1)!} \lambda^n e^{-\lambda t} t^{n-1} dt = 1 - e^{-\lambda x} \sum_{k=1}^{n-1} \frac{(\lambda x)^k}{k!},$$

and hence $\mathbb{P}(N_t = k) = e^{-\lambda t} (\lambda t)^k / k!$.

Let us now show that the increments of the process are independent. We begin by determining the joint distribution of N_t and N_s , for $0 \leq s \leq t$.

$$\begin{aligned} \mathbb{P}(N_s = k, N_t = n) &= \\ &= \mathbb{P}(S_k \leq s < S_{k+1}, S_n \leq t < S_{n+1}) \\ &= \int_{E_{n+1}} \mathbb{1}_{[x_k, x_{k+1}[}(s) \lambda^{n+1} e^{-\lambda x_{n+1}} \mathbb{1}_{]x_n, x_{n+1}[}(t) dx_1 \dots dx_{n+1} \end{aligned}$$

where $E_n = \{0 < x_1 < \dots < x_n\}$. Thus, by Fubini's theorem,

$$\begin{aligned} \mathbb{P}(N_s = k, N_t = n) &= \\ &= \int_{E_n} \mathbb{1}_{[x_k, x_{k+1}[}(s) \lambda^n \mathbb{1}_{]x_n, +\infty[}(t) dx_1 \dots dx_n \int_t^{+\infty} \lambda e^{-\lambda x_{n+1}} dx_{n+1} \\ &= \lambda^n e^{-\lambda t} \int_{E_n} \mathbb{1}_{]x_k, x_{k+1}[}(s) \mathbb{1}_{]x_n, +\infty[}(t) dx_1 \dots dx_n \\ &= \lambda^n e^{-\lambda t} \int_E dx_1 \dots dx_k \int_F dx_{k+1} \dots dx_n = \lambda^n e^{-\lambda t} \frac{s^k}{k!} \frac{(t-s)^{n-k}}{(n-k)!}, \end{aligned}$$

where $E = \{0 < x_1 < \dots < x_k < s\}$ and $F = \{s < x_{k+1} < \dots < x_n < t\}$. Therefore,

$$\mathbb{P}(N_s = k, N_t - N_s = l) = e^{-\lambda s} \frac{(\lambda s)^k}{k!} \cdot e^{-\lambda(t-s)} \frac{[\lambda(t-s)]^l}{l!},$$

and the result follows. \square

Note that Theorem 4.53 can also be taken as an alternative definition of the Poisson process, under the following form.

Definition 4.54 A stochastic process $\mathbf{N} = (N_t)_{t \in \mathbb{R}_+}$ is a homogeneous Poisson process if it is a process with independent and stationary increments such that $N_t - N_s \sim \mathcal{P}(\lambda(t-s))$ for all $t \geq s > 0$.

Indeed, the associated renewal process can be defined by $S_0 = 0$ and then S_n recursively through the relation $(N_t = n) = (S_n \leq t < S_{n+1})$. Setting $X_n = S_n - S_{n-1}$, we get $\mathbb{P}(X_1 > t) = \mathbb{P}(N_t = 0) = e^{-\lambda t}$, and hence $X_1 \sim \mathcal{E}(\lambda)$.

The Poisson process can also be defined more qualitatively, as follows.

Definition 4.55 A stochastic process $\mathbf{N} = (N_t)_{t \in \mathbb{R}_+}$ is a Poisson process if it is a process with independent increments such that $t \rightarrow N_t(\omega)$ is for almost all ω an increasing step function with jumps of size 1.

Let us now define the compound Poisson process.

Definition 4.56 Let \mathbf{N} be a (homogeneous) Poisson process. Let (Y_n) be an i.i.d. random sequence with finite mean and variance, and independent of \mathbf{N} . The stochastic process ξ defined on \mathbb{R}_+ by

$$\xi_t = \sum_{n=1}^{N_t} Y_n, \quad t \geq 0, \quad (4.9)$$

with $\xi_t = 0$ if $N_t = 0$, is called a (homogeneous) compound Poisson process.

The compound Poisson process has independent increments.

▷ *Example 4.57 (Cramér-Lundberg Process)* With the notation of Example 4.51, we suppose here that (Y_n) is i.i.d. with distribution function G with mean μ and that \mathbf{N} is a homogeneous Poisson process with intensity λ independent of (Y_n) . Thus, the process ξ defined by (4.9) is a homogeneous compound Poisson process, and (U_t) is called a Cramér-Lundberg process.

We compute $\mathbb{E} U_t = u + ct - \mathbb{E} N_t \mathbb{E} Y_1 = u + ct - \lambda \mu t$. This gives a condition of viability of the company, namely $c - \lambda \mu > 0$. The probability of ruin before time t is $r(t) = \mathbb{P}(U_t \leq 0) = \mathbb{P}(\xi_t \geq u + ct)$. Since the distribution function of ξ_t is

$$\begin{aligned} \mathbb{P}(\xi_t \leq x) &= \sum_{n \geq 0} \mathbb{P}\left(\sum_{i=1}^n Y_i \leq x, N_t = n\right) = \sum_{n \geq 0} \mathbb{P}\left(\sum_{i=1}^n Y_i \leq x\right) \mathbb{P}(N_t = n) \\ &= \sum_{n \geq 0} e^{-\lambda t} \frac{(\lambda t)^n}{n!} G^{*(n)}(x), \end{aligned}$$

we get

$$r(t) = \sum_{n \geq 0} e^{-\lambda t} \frac{(\lambda t)^n}{n!} G^{*(n)}(u + ct).$$

The probability of ruin of the company—in other words the probability that the life duration T of the process is finite—is

$$\mathbb{P}(\lim_{t \rightarrow +\infty} U_t \leq 0) = \mathbb{P}(T < +\infty) = \lim_{t \rightarrow +\infty} r(t),$$

but this quantity remains difficult to compute under this form. ◁

4.4.4 Asymptotic Results for Renewal Processes

All throughout this section, (S_n) will be a renewal process such that $0 < \mu = \mathbb{E} X_1 < +\infty$, with counting process $(N_t)_{t \in \mathbb{R}_+}$. Recall that $m(t) = \mathbb{E} N_t$ defines the renewal function of the process, and that we suppose $F(0) < 1$, where F is the distribution function of X_n .

Proposition 4.58 *The following statements hold true:*

1. S_n tends a.s. to infinity when n tends to infinity.
2. S_n and N_t are a.s. finite for all $n \in \mathbb{N}$ and $t \in \mathbb{R}_+^*$.
3. N_t tends a.s. to infinity and $m(t)$ tends to infinity when t tends to infinity.

Proof

1. The law of large numbers implies that S_n/n tends a.s. to μ , so S_n tends a.s. to infinity.
2. Since $(S_n \leq t) = (N_t \geq n + 1)$, it derives from Point 1. that N_t is a.s. finite.
 Moreover, $(S_n < +\infty) = \cap_{i=1}^n (X_i < +\infty)$ for $n \in \mathbb{N}^*$, and X_n is \mathbb{P} -a.s. finite, so S_n is a.s. finite for all $n \in \mathbb{N}^*$.
3. We compute

$$\mathbb{P}(\lim_{t \rightarrow +\infty} N_t < +\infty) = \mathbb{P}[\cup_{n \geq 1} (X_n = +\infty)] \leq \sum_{n \geq 1} \mathbb{P}(X_n = +\infty) = 0.$$

Therefore, N_t tends a.s. to infinity, from which it follows that $m(t)$ tends to infinity. ◻

Proposition 4.59 *The following convergence holds true,*

$$\frac{1}{t} N_t \xrightarrow{a.s.} \frac{1}{\mu}, \quad t \rightarrow +\infty.$$

The induced convergence of $m(t)/t$ to $1/\mu$ is known as the elementary renewal theorem.

Proof Thanks to the law of large numbers, S_n/n tends a.s. to μ . Moreover, N_t is a.s. finite and tends a.s. to infinity when t tends to infinity. Thanks to Theorem 4.17, S_{N_t}/N_t tends a.s. to μ , and the inequality

$$\frac{S_{N_t}}{N_t} \leq \frac{t}{N_t} < \frac{S_{N_t+1}}{N_t+1} \frac{N_t+1}{N_t}$$

yields the result. \square

\triangleright *Example 4.60 (Cramér-Lundberg Process (Continuation of Example 4.57))* We have

$$\frac{1}{t}U_t = \frac{u}{t} + c - \frac{N_t}{t} \frac{1}{N_t} \sum_{n=1}^{N_t} X_n.$$

According to Proposition 4.59, N_t/t tends to $1/\lambda$. Hence, thanks to Theorem 4.17, U_t/t tends a.s. to $c - \mu/\lambda$ when t tends to infinity. \triangleleft

The next result is an extension of the central limit theorem to renewal processes.

Theorem 4.61 *If $0 < \sigma^2 = \mathbb{V}\text{ar } X_1 < +\infty$, then*

$$\frac{N_t - t/\mu}{\sqrt{t\sigma^2/\mu^3}} \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1), \quad t \rightarrow +\infty.$$

Proof Set

$$Z_t = \sqrt{t} \sqrt{\frac{\mu^3}{\sigma^2}} \left(\frac{N_t}{t} - \frac{1}{\mu} \right).$$

We compute $\mathbb{P}(Z_t \leq x) = \mathbb{P}(N_t \leq t/\mu + x\sqrt{t\sigma^2/\mu^3})$. If n_t denotes the integer part of $t/\mu + x\sqrt{t\sigma^2/\mu^3}$, then

$$\mathbb{P}(Z_t \leq x) = \mathbb{P}(S_{n_t} \geq t) = \mathbb{P}\left(\frac{S_{n_t} - n_t\mu}{\sigma\sqrt{n_t}} \geq \frac{t - n_t\mu}{\sigma\sqrt{n_t}} \right).$$

By the central limit theorem, $(S_{n_t} - n_t\mu)/\sigma\sqrt{n_t}$ tends in distribution to a standard Gaussian variable. Moreover, $n_t \approx x\sqrt{t\sigma^2/\mu^3} + t/\mu$ when t tends to infinity. Hence $t - n_t\mu \approx -x\sqrt{t\sigma^2/\mu}$ and $\sigma\sqrt{n_t} \approx \sigma\sqrt{t/\mu}$, so $(t - n_t\mu)/\sigma\sqrt{n_t} \approx -x$, and the conclusion follows. \square

We state without proof the following two renewal theorems. A distribution function F is said to be arithmetic with period δ if the distribution is concentrated on $\{x_0 + n\delta : n \in \mathbb{N}\}$.

Theorem 4.62 (Blackwell’s Renewal) *If F is a non arithmetic distribution function on \mathbb{R}_+ , then, for all $h > 0$,*

$$m(t) - m(t - h) \longrightarrow \frac{h}{\mu}, \quad t \rightarrow +\infty. \tag{4.10}$$

If F is arithmetic with period δ , the above result remains valid provided that h is a multiple of δ .

▷ *Example 4.63 (Poisson Process)* In this case, $F(t) = (1 - e^{-\lambda t})\mathbb{1}_{\mathbb{R}_+}(t)$, whose Laplace transform is $\tilde{F}(s) = \lambda/(\lambda + s)$. Moreover, $N_0 = 0$, so the Laplace transform of m is

$$\tilde{m}(t) = \sum_{n \geq 1} (\tilde{F}(s))^n = \frac{1}{1 - \tilde{F}(s)} - 1 = \frac{\lambda}{s}.$$

Inverting Laplace transform yields $m(t) = \lambda t$. Thus, the relation $m(t) - m(t - h) = \lambda h$ holds for all $t > h$. ◁

In order to state the key renewal theorem, let us introduce the direct Riemann integrable functions.

Definition 4.64 Let g be a function defined on $\overline{\mathbb{R}_+}$. Let $\overline{m}_n(a)$ denote the maximum and $\underline{m}_n(a)$ the minimum of g on $[(n - 1)a, na]$, for all $a > 0$. Then g is said to be direct Riemann integrable if $\sum_{n \geq 1} \underline{m}_n(a)$ and $\sum_{n \geq 1} \overline{m}_n(a)$ are finite for all $a > 0$ and if $\lim_{a \rightarrow 0} \sum_{n=1}^{\infty} \underline{m}_n(a) = \lim_{a \rightarrow 0} \sum_{n=1}^{\infty} \overline{m}_n(a)$.

▷ *Example 4.65* Any nonnegative, decreasing function integrable over \mathbb{R}_+ is direct Riemann integrable. Indeed, we get

$$\sum_{n \geq 1} (\overline{m}_n(a) - \underline{m}_n(a)) \leq g(0).$$

Any nonnegative function integrable over \mathbb{R}_+ and with a compact support is also direct Riemann integrable. ◁

Theorem 4.66 (Key Renewal) *If F is a non arithmetic distribution function on \mathbb{R}_+ and if $g : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ is direct Riemann integrable, then*

$$\int_0^t g(t - x) dm(x) \longrightarrow \frac{1}{\mu} \int_0^{+\infty} g(x) dx, \quad t \rightarrow +\infty. \tag{4.11}$$

If F is arithmetic with period δ and if $\sum_{k \geq 0} g(x + k\delta) < +\infty$, then

$$m * g(x + n\delta) \longrightarrow \frac{\delta}{\mu} \sum_{k \geq 0} g(x + k\delta), \quad n \rightarrow +\infty.$$

Note that Blackwell's renewal theorem and the key renewal theorem are equivalent in the sense that (4.10) and (4.11) are equivalent.

For stopping renewal processes, the key renewal theorem takes the following form.

Proposition 4.67 *Let F be a defective distribution function on \mathbb{R}_+ . If $g : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ is direct Riemann integrable and such that $g(+\infty) = \lim_{t \rightarrow +\infty} g(t)$ exists, then the solution of the renewal equation (4.8) p. 198 satisfies*

$$h(t) = m * g(t) \longrightarrow \frac{g(+\infty)}{q}, \quad t \rightarrow +\infty,$$

where $q = 1 - F(+\infty)$.

Proof According to Proposition 4.47, $h(t) = m * g(t)$. According to relation (4.7) p. 197, the limit of $m(t)$ when t tends to infinity is

$$1 + F(+\infty) + F(+\infty)^2 + \dots = \frac{1}{1 - F(+\infty)} = \frac{1}{q},$$

and the result follows. □

▷ *Example 4.68 (Cramér-Lundberg Process (Continuation or Example 4.57))* Let us determine the probability of ruin of the Cramér-Lundberg process. Set $\zeta(u) = \mathbb{P}(T = +\infty) = 1 - \mathbb{P}(T < +\infty)$, where T is the time of ruin of the process. We compute

$$\begin{aligned} \zeta(u) &= \int_0^{+\infty} \int_0^{u+cs} \mathbb{P}(S_1 \in ds, Y_1 \in dy, T \circ \theta_s = +\infty) \\ &= \int_0^{+\infty} \int_0^{u+cs} \mathbb{P}(S_1 \in ds) \mathbb{P}(Y_1 \in dy, T \circ \theta_s = +\infty \mid S_1 = s) \\ &= \int_0^{+\infty} \mathbb{P}(S_1 \in ds) \int_0^{u+cs} \mathbb{P}(T \circ \theta_s = +\infty \mid S_1 = s, Y_1 = y) \mathbb{P}(Y_1 \in dy) \\ &= \int_0^{+\infty} \lambda e^{-\lambda s} ds \int_0^{u+cs} \zeta(u + cs - y) dG(y). \end{aligned}$$

By the change of variable $v = u + cs$, we get

$$\zeta(u) = \lambda_0 \int_u^{+\infty} e^{-\lambda_0(v-u)} dv \int_0^v \zeta(v-y) dG(y) = \lambda_0 e^{\lambda_0 u} g(u),$$

where

$$g(u) = \int_u^{+\infty} \int_0^v \lambda_0 e^{-\lambda_0(v-u)} dv \zeta(v-y) dG(y),$$

and $\lambda_0 = \lambda/c$. Differentiating this relation gives

$$\zeta'(u) = \lambda_0 e^{\lambda_0 u} [\lambda_0 g(u) + g'(u)] = \lambda_0 \zeta(u) - \lambda_0 G * \zeta(u) = \lambda_0 [1 - G] * \zeta(u).$$

Integrating the above differential equation on $[0, u]$ yields

$$\zeta(u) = \zeta(0) + \lambda_0 \int_0^u \zeta(u-y) [1 - G(y)] dy, \quad u \geq 0. \quad (4.12)$$

This equation is a renewal equation with defective distribution function with density $L(y) = \lambda_0 [1 - G(y)]$, where $L(+\infty) = \lambda_0 \mu < 1$. The case $L(+\infty) = 1$ is excluded because then $\zeta(u) = 0$ for all $u \in \mathbb{R}_+$. Thanks to the key renewal theorem for stopped renewal processes,

$$\zeta(+\infty) = \frac{\zeta(0)}{1 - L(+\infty)},$$

or, finally, since $\zeta(+\infty) = 1$,

$$\zeta(0) = 1 - \frac{\lambda \mu}{c}.$$

This allows the computation of $\zeta(u)$ for all $u \in \mathbb{R}_+$ through (4.12).

We have considered only nonnegative Y_1 . The result remains valid for any variable Y_1 : it is sufficient to take $-\infty$ and $u + cs$ instead of 0 and $u + cs$ as bounds of the second integral in the above computation of $\zeta(u)$. \triangleleft

4.5 Exercises

∇ **Exercise 4.1 (The AR(1) Process on \mathbb{N})** Let $(\varepsilon_n)_{n \in \mathbb{N}}$ be a white noise with variance 1. Let $a \in]-1, 1[$. An AR(1) process on \mathbb{N} is defined by setting $X_n = aX_{n-1} + \varepsilon_n$ for $n > 0$ and $X_0 = \varepsilon_0$.

1. a. Write X_n as a function of $\varepsilon_0, \dots, \varepsilon_n$ and determine its distribution.
- b. Determine the characteristic function of X_n .
- c. Give the distribution of (X_0, \dots, X_n) .

2. Show that:

- a. (X_n) converges in distribution and give the limit.
- b. $(X_0 + \cdots + X_n)/n$ converges in probability to 0;
- c. $(X_0 + \cdots + X_n)/\sqrt{n}$ converges in distribution; give the limit.

Solution

1. a. For $n \geq 1$, we have $X_n = \sum_{p=1}^n a^{n-p} \varepsilon_p$, and hence $\mathbb{E} X_n = 0$ and

$$\text{Var } X_n = a^{2n} + \cdots + a^2 + 1 = \frac{1 - a^{2(n+1)}}{1 - a^2}.$$

- b. Since $(\varepsilon_1, \dots, \varepsilon_n)$ is a Gaussian vector, X_n is a Gaussian variable, so the characteristic function of X_n is

$$\phi_{X_n}(t) = \exp\left(-\frac{1 - a^{2(n+1)}}{1 - a^2} t^2\right).$$

- c. The vector (X_0, \dots, X_n) is a linear transform of the standard Gaussian vector $(\varepsilon_1, \dots, \varepsilon_n)$, so is a Gaussian vector too, with mean $(0, \dots, 0)$ and covariance matrix given by

$$\begin{aligned} \text{Cov}(X_j, X_{j+k}) &= \mathbb{E}(X_j X_{j+k}) = \mathbb{E}\left[\left(\sum_{p=0}^j a^{j-p} \varepsilon_p\right)\left(\sum_{q=0}^{j+k} a^{j+k-q} \varepsilon_q\right)\right] \\ &= a^k \sum_{l=0}^j a^{2l} = \frac{1 - a^{2(j+1)}}{1 - a^2} a^k. \end{aligned}$$

2. a. Clearly, $\phi_{X_n}(t)$ tends to $\exp[-t^2/(1 - a^2)]$, so X_n converges in distribution to a random variable with distribution $\mathcal{N}(0, 1/(1 - a^2))$.
- b. We compute

$$\begin{aligned} \frac{1}{n}(X_0 + \cdots + X_n) &= \frac{X_0}{n} + \frac{1}{n} \sum_{i=1}^n (a X_{i-1} + \varepsilon_i) = \frac{a}{n} \sum_{i=0}^{n-1} X_i + \frac{1}{n} \sum_{i=0}^n \varepsilon_i \\ &= \frac{a}{n}(X_0 + \cdots + X_n) - \frac{a}{n} X_n + \frac{1}{n} \sum_{i=0}^n \varepsilon_i. \end{aligned}$$

We know that $X_n \sim \mathcal{N}(0, (1 - a^{2(n+1)})/(1 - a^2))$, so X_n/n converges to 0 in probability. Moreover, by the strong law of large numbers, $\sum_{i=0}^n \varepsilon_i/n$ converges a.s. to 0. Therefore, $(1 - a)(X_0 + \cdots + X_n)/n$ converges to 0 in probability, and hence $(X_0 + \cdots + X_n)/n$ too.

- c. In the same way, by the central limit theorem, $\sum_{i=0}^n \varepsilon_i / \sqrt{n}$ converges in distribution to $\mathcal{N}(0, 1)$, so $(1-a)(X_0 + \cdots + X_n) / \sqrt{n}$ too, from which it follows that $(X_0 + \cdots + X_n) / \sqrt{n}$ converges in distribution to $\mathcal{N}(0, 1/(1-a)^2)$. \triangle

▽ Exercise 4.2 (Generalization of AR and MA Processes) Let (γ_n) be a sequence of i.i.d. standard random variables. Let $\theta \in]-1, 1[$.

1. Set $V_n = \gamma_1 + \theta\gamma_2 + \cdots + \theta^{n-1}\gamma_n$, for $n \in \mathbb{N}^*$.
 - a. Show that V_n converges in square mean—use the Cauchy criterion.
 - b. Set $V = \sum_{i \geq 1} \theta^{i-1}\gamma_i$. Show that V_n tends a.s. to V .
2. Let X_0 be a random variable independent of (γ_n) . Set $X_n = \theta X_{n-1} + \gamma_n$ for $n \geq 1$.
 - a. Show that V_n and $X_n - \theta^n X_0$ have the same distribution for $n \geq 1$.
 - b. Let ρ denote the distribution of V . Compute the mean and the variance of ρ . Show that X_n tends in distribution to ρ .
 - c. Assume that $X_0 \sim \rho$. Show that $X_n \sim \rho$ for all n .

Solution

1. a. We have

$$\begin{aligned} \mathbb{E}(V_m - V_{n+m})^2 &= \mathbb{E}(\theta^m \gamma_{m+1} + \cdots + \theta^{n+m-1} \gamma_{n+m})^2 \\ &= \sum_{i=0}^{n-1} \theta^{2(m+i)} \mathbb{E}(\gamma_{m+i}^2) + 2 \sum_{i < j} \theta^{m+i} \theta^{m+j} \mathbb{E}(\gamma_{m+i} \gamma_{m+j}), \end{aligned}$$

so that $\mathbb{E}(V_m - V_{n+m})^2 = \sum_{i=0}^{n-1} \theta^{2(m+i)}$, which converges to 0.

- b. According to Proposition 1.80, it is sufficient to show that $\mathbb{P}(\overline{\lim} |V_n - V| > \varepsilon) = 0$, or, using Borel-Cantelli lemma, that $\sum_{n \geq 0} \mathbb{P}(|V_n - V| > \varepsilon)$ is finite for all $\varepsilon > 0$.

Chebyshev's inequality gives $\mathbb{P}(|V_n - V| > \varepsilon) \leq \mathbb{E}[(V_n - V)^2] / \varepsilon^2$. Moreover,

$$\mathbb{E}[(V_n - V)^2] = \mathbb{E}[(\theta^{n+1}\gamma_{n+2} + \theta^{n+2}\gamma_{n+3} + \cdots)^2] = \frac{\theta^{2n}}{1 - \theta^2},$$

so $\mathbb{P}(|V_n - V| > \varepsilon) \leq \theta^{2n} / (1 - \theta^2) \varepsilon^2$, and the sum of the series is finite.

2. a. We can write $X_n - \theta^n X_0 = \theta^{n-1}\gamma_1 + \cdots + \theta\gamma_{n-1} + \gamma_n$, from which the result follows, because all the γ_i have the same distribution.

- b. The sequence (V_n) converges to V in square mean—so also in mean, hence $\mathbb{E} V = \lim_{n \rightarrow +\infty} \mathbb{E} V_n = 0$ and

$$\begin{aligned} \text{Var } V &= \mathbb{E}(V^2) = \lim_{n \rightarrow +\infty} \mathbb{E}(V_n^2) = \lim_{n \rightarrow +\infty} \sum_{i=0}^n \theta^{2i} \mathbb{E}(\gamma_{i+1}^2) \\ &= \lim_{n \rightarrow +\infty} \frac{1 - \theta^{2n}}{1 - \theta^2} = \frac{1}{1 - \theta^2}. \end{aligned}$$

Since $X_n - \theta^n X_0 \sim V_n$, we can write $X_n = \theta^n X_0 + U_n$, where U_n is a variable with the same distribution as V_n . Therefore, (U_n) converges in distribution to ρ too, and since $(\theta^n X_0)$ converges a.s. to 0, (X_n) converges in distribution to ρ .

- c. We have $X_0 \sim V$ so $\theta X \sim \sum_{n \geq 0} \theta^{n+1} \gamma_{n+1}$ or $\theta X \sim \sum_{n \geq 1} \theta^n \gamma_n$. Since $\gamma_1 \sim \gamma_0$, it follows that $X_1 \sim \sum_{n \geq 0} \theta^n \gamma_n$; since $\gamma_n \sim \gamma_{n+1}$ for all n , we obtain $X_1 \sim \sum_{n \geq 0} \theta^n \gamma_{n+1}$, that is $X_1 \sim V \sim \rho$. The result follows by induction.

Note that if $\rho = \mathcal{N}(0, 1)$ and $\gamma_n \sim \mathcal{N}(0, 1)$, then X_n is an AR(1) process and (V_n) is an MA($n - 1$) process on \mathbb{N} . \triangle

Exercise 4.3 (Sinusoidal Signals and Stationarity) Let \mathbf{X} be the stochastic process defined in Example 4.1. The variables ν , A and ϕ are not supposed to be constant, unless specifically stated. The variable ϕ takes values in $[0, 2\pi[$.

1. Assume that $A = 1$ and $\phi \sim \mathcal{U}(0, 2\pi)$.
 - a. Show that if ν is a.s. constant, then \mathbf{X} is strictly stationary to the order 1.
 - b. Show that if ν is continuous with density f and is independent of ϕ , then \mathbf{X} is weakly stationary to the order 2; determine its spectral density h .
2. Suppose that ν and ϕ are constant.
 - a. Give a necessary and sufficient condition on A for \mathbf{X} to be weakly stationary to the order 1.
 - b. Can \mathbf{X} be weakly stationary to the order 2?
 - c. Let $\mathbf{S} = \mathbf{X} + \mathbf{Y}$. Give a necessary and sufficient condition on A and B for \mathbf{S} to be weakly stationary to the order 1, and then to the order 2.
3. Suppose that ν is a.s. constant, that A is nonnegative and that A and ϕ are independent.
 - a. Give a necessary and sufficient condition on ϕ for \mathbf{X} to be weakly stationary to the order 1 and then 2.
 - b. Give a necessary and sufficient condition on ϕ for \mathbf{X} to be strictly stationary to the order 1.
 - c. Let \mathbf{Z} be the stochastic process defined by

$$Z_t = A \cos(\nu t + \phi) + B \sin(\nu t + \phi),$$

where $\varphi \sim \mathcal{U}(0, 2\pi)$ is independent of A and B . Show that \mathbf{Z} is weakly stationary to the order 2.

Solution

1. a. Setting $\psi = \nu\tau + \varphi$, we obtain $X_{t+\tau} = \cos(\nu t + \psi)$. Since $\varphi \sim \mathcal{U}[0, 2\pi]$ and $\nu\tau$ is constant for a fixed τ , we have $\psi \sim \mathcal{U}[\nu\tau, \nu\tau + 2\pi]$, and hence $X_t \sim X_{t+\tau}$.
- b. We have $\mathbb{E} X_t = 0$ and $2\mathbb{E}(X_t X_{t+\tau}) = \mathbb{E}(\cos[\nu(2t + \tau) + 2\varphi]) + \mathbb{E}(\cos(\nu\tau))$. We compute

$$\begin{aligned} \mathbb{E}(\cos[\nu(2t + \tau) + 2\varphi]) &= \\ &= \mathbb{E}(\cos[\nu(2t + \tau)])\mathbb{E}(\cos(2\varphi)) - \mathbb{E}(\sin[\nu(2t + \tau)])\mathbb{E}(\sin(2\varphi)) = 0, \end{aligned}$$

and $\mathbb{E}(\cos(2\varphi)) = \mathbb{E}(\sin(2\varphi)) = 0$, so \mathbf{X} is weakly stationary to the order 2, with $r(\tau) = \int_{\mathbb{R}} \cos(\lambda\tau) f(\lambda) d\lambda$. Since r is even and real,

$$r(\tau) = \int_{\mathbb{R}} e^{i\lambda\tau} h(\lambda) d\lambda = \int_{\mathbb{R}_+} 2 \cos(\lambda\tau) h(\lambda) d\lambda,$$

and hence $h(\lambda) = [f(-\lambda) + f(\lambda)]/4$.

2. a. We have $\mathbb{E} X_t = [\cos(\nu t)]\mathbb{E} A$, so \mathbf{X} is weakly stationary to the order 1 if and only if A is centered.
- b. We compute $2\mathbb{E}(X_t X_{t+\tau}) = [\cos(2\nu t + \nu\tau) + \cos(\nu\tau)]\mathbb{E}(A^2)$. This is a function of τ only if $\mathbb{E}(A^2) = 0$, that is to say if A is null. Then the signal itself is null.
- c. Therefore, $\mathbb{E} S_t = [\cos(\nu t)]\mathbb{E} A + [\sin(\nu t)]\mathbb{E} B$, which is constant in t if and only if A and B are centered. Moreover,

$$\begin{aligned} \mathbb{E}(S_t S_{t+\tau}) &= (\text{Var } A + \text{Var } B) \cos(\nu t) + (\text{Var } A - \text{Var } B) \cos(2\nu t + \nu\tau) \\ &\quad + \text{Cov}(A, B) \sin(2\nu t + \nu\tau), \end{aligned}$$

so S is weakly stationary to the order 2 if A and B are uncorrelated and have the same variance.

3. a. We have $\mathbb{E} X_t = [\cos(\nu t)\mathbb{E}(\cos \varphi) - \sin(\nu t)\mathbb{E}(\sin \varphi)]\mathbb{E} A$, which is constant in t if $\mathbb{E}(\cos \varphi) = \mathbb{E}(\sin \varphi) = 0$. Similarly,

$$\begin{aligned} 2\mathbb{E}(X_t X_{t+\tau}) &= \\ &= [\cos(\nu\tau) + \cos(2\nu t + \nu\tau)]\mathbb{E}(\cos 2\varphi) - \sin(2\nu t + \nu\tau)\mathbb{E}(\sin 2\varphi)]\mathbb{E}(A^2) \end{aligned}$$

so \mathbf{X} is weakly stationary to the order 2 if $\mathbb{E}(\cos \varphi) = \mathbb{E}(\sin \varphi) = 0$ and $\mathbb{E}(\cos 2\varphi) = \mathbb{E}(\sin 2\varphi) = 0$.

- b. The variables $X_{t+\tau} = A \cos(\nu t + \varphi + \nu\tau)$ and $X_t = A \cos(\nu t + \varphi)$ have the same distribution if $(A, \varphi + \nu\tau) \sim (A, \varphi)$ for all τ . This condition is fulfilled if and only if $\varphi \sim \mathcal{U}(0, 2\pi)$ and is independent of A .
- c. Since $\mathbb{E}[\cos(\nu t + \varphi)] = \mathbb{E}[\sin(\nu t + \varphi)] = 0$, the process is centered. Set $X_t = A \cos(\nu t + \varphi)$ and $Y_t = B \sin(\nu t + \varphi)$. We have

$$\mathbb{E}(Z_t Z_{t+\tau}) = \mathbb{E}(X_t X_{t+\tau}) + \mathbb{E}(Y_t Y_{t+\tau}) + \mathbb{E}(X_t Y_{t+\tau}) + \mathbb{E}(Y_t X_{t+\tau}).$$

We compute $2\mathbb{E}(X_t X_{t+\tau}) = \cos(\nu\tau)\mathbb{E}(A^2)$ and $2\mathbb{E}(Y_t Y_{t+\tau}) = \cos(\nu\tau)\mathbb{E}(B^2)$, and also $2\mathbb{E}(X_t Y_{t+\tau}) = \sin(\nu\tau)\mathbb{E}(AB) = -2\mathbb{E}(Y_t X_{t+\tau})$, so \mathbf{Z} is indeed weakly stationary to the order 2. \triangle

▽ Exercise 4.4 (Alternated Renewal Process and Availability) Consider a component starting operating at time $S_0 = 0$. When it fails, it is renewed. When the second fails, it is renewed and so on. Suppose the sequence of time durations (X_n) of the successive components is i.i.d. and that of the replacing times (Y_n) is also i.i.d. and is independent of (X_n) .

- Show that (S_n) , defined by $S_0 = 0$ and $S_n = S_{n-1} + X_n + Y_n$ for $n \geq 1$, is a renewal process.
- Write the event E_t ="the system is in good shape at time t " as a function of (X_n) and (S_n) .
 - Infer from a. the instantaneous availability $A(t) = \mathbb{P}(E_t)$.
 - If $\mathbb{E}X_1 + \mathbb{E}Y_1 < +\infty$, compute the limit availability $A = \lim_{t \rightarrow +\infty} A(t)$.

Solution

- The sequence $(T_n) = (X_n + Y_n)$ is i.i.d., and, according to Definition 4.44, (S_n) is indeed a renewal process.
- a. We can write

$$E_t = (X_1 > t) \cup \left[\bigcup_{n \geq 1} (S_n \leq t) \cap (X_{n+1} > t - S_n) \right].$$

- b. Let F and G denote the respective distribution functions of X_1 and Y_1 . The distribution function of T_n is

$$H(t) = F * G(t) = \int_0^t F(t-x)dG(x), \quad t \geq 0.$$

Therefore, setting $R(t) = 1 - F(t)$,

$$A(t) = \mathbb{P}(E_t) = \mathbb{P}(X_1 > t) + \sum_{n \geq 1} \mathbb{P}(S_n \leq t, X_{n+1} > t - S_n).$$

We compute

$$\begin{aligned}\mathbb{P}(S_n \leq t, X_{n+1} > t - S_n) &= \mathbb{E}[\mathbb{1}_{(S_n \leq t)} \mathbb{P}(X_{n+1} > t - S_n \mid S_n)] \\ &= \int_0^t R(t - y) dH^{*(n)}(y).\end{aligned}$$

Thus

$$A(t) = \int_0^t R(t - y) dM(y) = R * m(t),$$

where $M(t) = \sum_{n \geq 1} H^{*(n)}(t)$ and $m(t) = \mathbb{1}_{\mathbb{R}_+}(t) + M(t)$. Finally, the key renewal theorem yields

$$A = \lim_{t \rightarrow +\infty} A(t) = \frac{\mathbb{E} X}{\mathbb{E} X + \mathbb{E} Y},$$

Note that the same problem will be modelled by a semi-Markov process in Exercise 5.5. △

▽ **Exercise 4.5 (Poisson Process)** The notation is that of Sect. 4.4.3.

1. Set $U_n = S_n/n$, for $n \in \mathbb{N}^*$. Show that (U_n) and $(n^{1/2}(\lambda U_n - 1))$ converge in distribution and give the limits.
2. Set $V_n = N_n/n$, for $n \in \mathbb{N}^*$. Determine the characteristic function of V_n . Show that (V_n) and $((n/\lambda)^{1/2}(V_n - \lambda))$ converge in distribution; give the limits.

Solution

1. Thanks to the law of large numbers, U_n converges a.s. to $1/\lambda$, and, thanks to the central limit theorem, $n^{1/2}(\lambda U_n - 1)$ converges in distribution to the standard normal distribution.
2. We compute

$$\mathbb{E}(e^{iuN_t}) = \sum_{k \geq 0} e^{iuk} \mathbb{P}(N_t = k) = e^{-\lambda t} \sum_{k \geq 0} \frac{(\lambda t e^{iu})^k}{k!} = e^{-\lambda t(1 - e^{iu})}.$$

Hence $\phi_{V_n}(u) = e^{-\lambda n(1 - e^{iu})}$ converges to $e^{iu\lambda}$, and V_n converges in distribution to λ . Moreover, $Z_n = \sqrt{n}(V_n - \lambda)/\sqrt{\lambda} = N_n/\sqrt{n\lambda} - \sqrt{n\lambda}$, so $\phi_{Z_n}(u) = \exp[-\lambda n(1 - e^{iu/\sqrt{n\lambda}}) - iu\sqrt{n\lambda}]$, which converges to $e^{-u^2/2}$, and Z_n converges in distribution to the standard normal distribution. △

∇ Exercise 4.6 (Superposition and Decomposition of Poisson Processes)

1. Let \mathbf{N} and $\tilde{\mathbf{N}}$ be two independent Poisson processes with respective intensities $\lambda > 0$ and $\mu > 0$. Show that the process \mathbf{K} , defined by $K_t = N_t + \tilde{N}_t$, is a Poisson process, called superposition. Give its intensity.
2. Let \mathbf{N} be a Poisson process with intensity $\lambda > 0$, whose arrivals are of two different types, A and B , with respective probabilities p and $1 - p$ independent of the arrivals times. Show that the counting process $\mathbf{M} = (M_t)_{t \in \mathbb{R}_+}$ of the arrivals of type A is a Poisson process. Give its intensity.

Solution

1. We can write $K_t - K_s = (N_t - N_s) + (\tilde{N}_t - \tilde{N}_s)$, so \mathbf{K} is indeed a process with independent increments. Moreover,

$$\begin{aligned} \mathbb{P}(K_t = n) &= \sum_{m=0}^n \mathbb{P}(N_t = m) \mathbb{P}(\tilde{N}_t = n - m) \\ &= \sum_{m=0}^n e^{-\lambda t} \frac{(\lambda t)^m}{m!} e^{-\mu t} \frac{(\mu t)^{n-m}}{(n-m)!} = e^{-(\lambda+\mu)t} \frac{[(\lambda + \mu)t]^n}{n!}. \end{aligned}$$

One can show similarly that $K_{t+s} - K_t \sim \mathcal{P}((\lambda + \mu)s)$. Proposition 4.54 yields that \mathbf{K} is a Poisson process, with intensity $\lambda + \mu$.

2. By definition, the process \mathbf{N} is the counting process of the renewal process associated with the sequence of arrivals times of A and B . These times are i.i.d., hence in particular the arrivals times of A are i.i.d., and the associated counting process is \mathbf{M} . Moreover,

$$\begin{aligned} \mathbb{P}(M_t = k) &= \sum_{n \geq k} \mathbb{P}(M_t = k \mid N_t = n) \mathbb{P}(N_t = n) \\ &= \sum_{n \geq k} \binom{n}{k} p^k (1-p)^{n-k} e^{-\lambda t} \frac{(\lambda t)^n}{n!} = e^{-p\lambda t} \frac{(p\lambda t)^k}{k!}. \end{aligned}$$

One can show similarly that $M_{t+s} - M_t \sim \mathcal{P}(p\lambda s)$ so that \mathbf{M} is a Poisson process with intensity $p\lambda$. \triangle