L. DECREUSEFOND

POISSON PROCESSES AND BEYOND

https://perso.telecom-paris.fr/~decreuse

Version, June 2025

Contents

1
Poisson Process

Siméon Denis Poisson (1781-1840)

The modeling of a physical system must comply with two constraints. On the one hand, it must reflect the reality as accurately as possible, and on the other hand, it must have a predictive role, in other words it must provide computational tools for the analysis. Beyond the difficulty to qualitatively and quantitatively determine the pertinent parameters of a physical system, the experience shows that the more one wants an accurate model, the less it will be tractable in practice.

Within the framework of queuing systems, we must, in the first place, model the process of arrivals of the requests. The Poisson process which we study in this Chapter, is the most frequently used model, primarily because it is one of the rare models with which we can make computations. This modeling is found to be highly pertinent for the telephone calls to a commutator. Unfortunately, this is not the same for other types of network, where the traffic is much more versatile. However, as we will see at the end of this chapter, the Poisson process can be modified, so as to reflect this versatility to a certain extent.

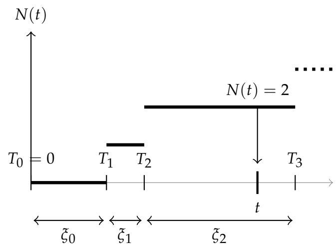


Figure 1.1: Notations for point processes.

A point process is a strictly increasing sequence of positive random variables $(T_1, T_2, \ldots)$ such that $T_n \to \infty$ a.s.. By convention, we adjoin the random variable $T_0 = 0$ a.s. to this sequence. These random variables will represent the arrival times of requests to the system. We can as well describe the sequence by the differences in time which elapses between the successive arrivals: $\xi_n = T_{n+1} - T_n$ is the nth *inter-arrival* time. The sequence $(\xi_n,\, n \in \mathbf{N})$ also characterizes the point process by the relation $T_n = \sum_{i \le n-1} \xi_i$. We will finally denote $N(t)$, the number of points (that is, of arrivals), up to time t.

1.1 Definitions

The Poisson process admits multiple characterizations. As each one of them can be considered as a definition, and the others as properties, we give to all the status of definition and then show that they are equivalent.

The exponential distribution of parameter λ, denoted here by $\mathcal{E}(\lambda)$, has density

$$x \mapsto \lambda e^{-\lambda x} \mathbf{1}_{\mathbf{R}^+}(x).$$

We know

$$\mathbf{E}\left[\mathcal{E}(\lambda)\right] = \frac{1}{\lambda} \text{ et } \operatorname{var}(\mathcal{E}(\lambda)) = \frac{1}{\lambda^2}\cdot$$

It is often useful to remember that

$$\mathbf{P}(\mathcal{E}(\lambda) \ge x) = e^{-\lambda x}.$$

Definition 1.1.– The point process N is a Poisson process of intensity λ if, and only if, the random variables $(\xi_n,\, n \in \mathbf{N})$ are independent and of the same exponential distribution with parameter λ.

DEFINITION 1.2.– The point process N is a Poisson process of intensity λ if, and only if the following two conditions are satisfied:

1. $N(t)$ follows a Poisson distribution with parameter λt;
2. Given $\{N(t) = n\}$, the family $(T_1, \dots, T_n)$ is uniformly distributed over $[0, t]$: $(T_1, \cdots, T_n)$ is distributed as $(U_{(1)}, \cdots, U_{(n)})$ the order statistics (see Lemma 1.3 for the definition) of the vector $(U_1, \cdots, U_n)$ made of independent random variables uniformly distributed over $[0, t]$.

DEFINITION 1.3.– The point process N is a Poisson process of intensity λ if, and only if the following two conditions are satisfied:

1. for any $0 = t_0 < t_1 < \cdots < t_n$, the random variables $(N(t_{i+1}) - N(t_i), 1 \le i \le n-1)$ are independent;
2. for any t, s, the random variables $N(t+s) - N(t)$ follow a Poisson distribution of parameter λs, i.e.

$$\mathbf{P}(N(t+s) - N(t) = k) = \exp(-\lambda s)\frac{(\lambda s)^k}{k!}.$$

DEFINITION 1.4.– The point process N is a Poisson process of intensity λ if and only if, for any function $f : \mathbf{R}^+ \to \mathbf{R}^+$ (or for any function f with compact support on $\mathbf{R}^+$), the following identity holds:

$$\mathbf{E}\left[\exp\left(-\sum_{n\ge 1} f(T_n)\right)\right] = \exp\left(-\int_0^\infty (1 - e^{-f(s)})\,\lambda\,\mathrm{d}s\right). \tag{1.1}$$

In order to show the equivalence between these definitions, we must introduce three technical results.

LEMMA 1.1.– The density of the distribution of $(T_1, \dots, T_n)$ is given by

$$\begin{aligned}\mathrm{d}\mathbf{P}_{(T_1,\cdots,T_n)}&(x_1, \cdots, x_n)\\ &= \lambda^n \exp(-\lambda x_n)\mathbf{1}_{\mathcal{C}}(x_1, \cdots, x_n)\ \mathrm{d}x_1 \dots \mathrm{d}x_n,\end{aligned} \tag{1.2}$$

where

$$\mathcal{C} = \{(y_1, \cdots, y_n) \in (\mathbf{R}^+)^n, 0 \le y_1 \le \dots \le y_n\}.$$

In particular, T_n follows a gamma distribution of parameters n and λ, defined by :

$$\mathrm{d}\mathbf{P}_{T_n}(x) = \lambda^n \exp(-\lambda x)\frac{x^{n-1}}{(n-1)!}\mathbf{1}_{\mathbf{R}^+}(x)\ \mathrm{d}x \tag{1.3}$$

Proof. We proceed by identification. For all bounded measurable f,

$$\begin{aligned}&\mathbf{E}\left[f(T_1, \cdots, T_n)\right]\\ &= \int_{(\mathbf{R}^+)^n} f\left(x_0, x_0 + x_1, \cdots, x_0 + \dots + x_{n-1}\right)\ \mathrm{d}\mathbf{P}_{\xi_0}(x_0)\dots\ \mathrm{d}\mathbf{P}_{\xi_{n-1}}(x_{n-1}).\end{aligned}$$

Perform the change of variable

$$u_1 = x_0,\ u_2 = x_0 + x_1,\ \cdots,\ u_n = x_0 + \ldots + x_{n-1},$$

whose jacobian equals 1. The conditions $x_0 \geq 0, \ldots, x_{n-1} \geq 0$ amounts to $0 \leq u_1 \leq u_2 \ldots \leq u_n$. We then have:

$$\mathbf{E}\left[f(T_1, \cdots, T_n)\right] = \int_{(\mathbf{R}^+)^n} f(u_n)\lambda^n e^{-\lambda u_n} \mathbf{1}_{\mathcal{C}}(u_1, \cdots, u_n)\ \mathrm{d}u_1 \ldots \mathrm{d}u_n.$$

The density of the joint distribution follows from it. If f depends only on T_n, we obtain

$$\begin{aligned}
\mathbf{E}\left[f(T_n)\right] &= \int \ldots \int_{0 \leq u_1 \ldots \leq u_n} f(u_n)\lambda^n \exp(-\lambda u_n)\ \mathrm{d}u_1 \ldots \mathrm{d}u_n \\
&= \int_0^\infty f(u_n)\lambda^n \exp(-\lambda u_n)\left(\int_0^{u_n} \mathrm{d}u_{n-1} \int \ldots \int_0^{u_2} \mathrm{d}u_1\right) \mathrm{d}u_n \\
&= \int_0^\infty f(u_n)\lambda^n \exp(-\lambda u_n)\frac{u_n^{n-1}}{(n-1)!}\ \mathrm{d}u_n.
\end{aligned}$$

The result follows. □

Lemma 1.2.– Let X be a random variable of Poisson distribution with parameter λ. We have

$$\mathbf{E}\left[e^{-sX}\right] = \exp(-\lambda(1 - e^{-s})).$$

Proof. By definition of the Poisson distribution, we have

$$\mathbf{E}\left[e^{-sX}\right] = \sum_{k=0}^{\infty} \mathbf{E}\left[e^{-sk}\right] e^{-\lambda}\frac{\lambda^k}{k!} = \exp(-\lambda + \lambda e^{-s}),$$

hence the result. □

Lemma 1.3.– Let $(U_1, \ldots, U_n)$ be n independent random variables of uniform distribution on $[0, t]$. Let $\bar{U}$ represent the reordering of the n-tuple in increasing order, that is

$$\bar{U}_1(\omega) \leq \bar{U}_2(\omega) \leq \ldots \leq \bar{U}_n(\omega),\ \text{presque sûrement.}$$

The distribution of $\bar{U}$ is given by

$$\mathrm{d}\mathbf{P}_{(\bar{U}_1, \cdots, \bar{U}_n)}(x_1, \cdots, x_n) = \frac{n!}{t^n}\mathbf{1}_{\mathcal{C}}(x_1, \cdots, x_n)\ \mathrm{d}x_1 \ldots \mathrm{d}x_n$$

Proof. Denote σ, the random variable with values in the group of permutations $\mathfrak{S}_n$ of $[\![1, n]\!]$, representing the permutation of indexes necessary to arrange the values of $U_i(\omega)$ in increasing order, e.g. if we have

$$U_2(\omega) \leq U_3(\omega) \leq U_1(\omega),$$

then:

$$\sigma(\omega) = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}.$$

The image of i by $\sigma(\omega)$ is the index of the random variable which is in the ith position for the sample ω. Therefore, by definition we have $\bar{U}_i(\omega) = U_{\sigma(\omega)(i)}(\omega)$. As the random variables U_i, $i \in [\![1, n]\!]$ are independent and of the same distribution, for any $\tau \in \mathfrak{S}_n$, we have

$$\mathrm{d}\mathbf{P}_{\left(U_{\tau(1)},\cdots,U_{\sigma_0(n)}\right)}(u_1, \cdots, u_n) = \otimes_{i=1}^n \frac{1}{t}\mathbf{1}_{[0,t]}(u_i)\ \mathrm{d}u_i.$$

Notice, in particular, that this distribution does not depend on τ. Therefore,

$$\begin{aligned}\mathbf{P}(\sigma = \tau) &= \mathbf{P}\left(U_{\tau(1)} \le \ldots \le U_{\tau(n)}\right) \\ &= \int\ldots\int \mathbf{1}_{\mathcal{C}}(u_1, \cdots, u_n)\ \mathrm{d}\mathbf{P}_{\left(U_{\tau(1)},\cdots,U_{\tau(n)}\right)}(u_1, \cdots, u_n) \\ &= \mathbf{P}(\sigma = \mathrm{Id}).\end{aligned}$$

Thus, σ follows a uniform distribution on $\mathfrak{S}_n$, that is to say $\mathbf{P}(\sigma = \tau) = 1/n!$. To compute the distribution of the n-tuple $\bar{U}$, we partition the probability space in $\cup_{\tau\in\mathfrak{S}_n}(\sigma = \tau)$. For any bounded continuous function f, we have

$$\begin{aligned}\mathbf{E}\left[f\left(\bar{U}_1, \cdots, \bar{U}_n\right)\right] &= \sum_{\tau\in\mathfrak{S}_n} \mathbf{E}\left[f\left(\bar{U}_1, \cdots, \bar{U}_n\right); \sigma = \tau\right] \\ = \sum_{\tau\in\mathfrak{S}_n} &\mathbf{E}\left[f\left(U_{\tau(1)}, \cdots, U_{\tau(n)}\right)\mathbf{1}_{\mathcal{C}}(U_{\tau(1)}, \cdots, U_{\tau(n)})\right] \\ = \sum_{\tau\in\mathfrak{S}_n} &\int\ldots\int f\left(u_1, \cdots, u_n\right)\mathbf{1}_{\mathcal{C}}(u_1, \cdots, u_n)\ \mathrm{d}\mathbf{P}_{(U_{\tau(1)},\cdots,U_{\tau(n)})}(u_1, \cdots, u_n) \\ = \sum_{\tau\in\mathfrak{S}_n} &\int\ldots\int f\left(u_1, \cdots, u_n\right)\mathbf{1}_{\mathcal{C}}(u_1, \cdots, u_n)\ \otimes_{i=1}^n \frac{1}{t}\mathbf{1}_{[0,t]}(u_i)\ \mathrm{d}u_i \\ = \frac{n!}{t^n} &\int\ldots\int f\left(u_1, \cdots, u_n\right)\mathbf{1}_{\mathcal{C}}(u_1, \cdots, u_n)\ \mathrm{d}u_1\ldots\ \mathrm{d}u_n.\end{aligned}$$

Hence the result. □

Proof of equivalence between the definitions. We are going to show the implication chain: 1.1⟹1.2 ⟹ 1.3⟹ 1.4. We save the last implication for further developments.

1.1 ⟹ 1.2.

Let us first show that $N(t)$ follows a Poisson distribution. Since it is clear that the events $\{N(t) = k\}$ and $\{T_k \le t < T_{k+1}\}$ coincide, we have

$$\begin{aligned}\mathbf{P}(N(t) = k) &= \mathbf{P}(T_k \le t < T_k + \xi_{k+1}) \\ &= \iint \mathbf{1}_{\{x\le t\}}\mathbf{1}_{\{x+y>t\}}\ \mathrm{d}\mathbf{P}_{T_k}(x)\ \mathrm{d}\mathbf{P}_{\xi_k}(y) \\ &= \int_0^t \left(\int_{t-x}^\infty \lambda e^{-\lambda y}\, dy\right)\lambda^k \frac{x^{k-1}}{(k-1)!}\exp(-\lambda x)\ \mathrm{d}x \\ &= e^{-\lambda t}\int_0^t \lambda^k \frac{x^{k-1}}{(k-1)!}\ \mathrm{d}x = e^{-\lambda t}\frac{(\lambda x)^k}{k!}.\end{aligned}$$

For the conditional law, we proceed similarly:

$$
\begin{aligned}
&\mathbf{P}\left((T_1, \cdots, T_n) \in A \,|\, N(t) = n\right) \mathbf{P}(N(t) = n) \\
=&\mathbf{P}\left((T_1, \cdots, T_n) \in A, T_n \le t < T_{n+1}\right) \\
=&\underset{0\le u_1 \le \ldots \le u_{n+1}}{\int \ldots \int} \mathbf{1}_A(u_1, \cdots, u_n) \mathbf{1}_{[u_n, u_{n+1}[}(t) \lambda^{n+1} e^{-\lambda u_{n+1}} \, \mathrm{d}u_1 \ldots \, \mathrm{d}u_{n+1} \\
=&\lambda^n \underset{0\le u_1 \le \ldots \le u_{n+1}}{\int \ldots \int} \mathbf{1}_A(u_1, \cdots, u_n) \mathbf{1}_{[u_n, \infty[}(t) \left(\int_t^\infty \lambda e^{-\lambda u_{n+1}} \right) \, \mathrm{d}u_1 \ldots \, \mathrm{d}u_n \\
=&\lambda^n e^{-\lambda t} \underset{0\le u_1 \le \ldots \le u_{n+1}}{\int \ldots \int} \mathbf{1}_A(u_1, \cdots, u_n) \mathbf{1}_{[u_n, \infty[}(t) \, \mathrm{d}u_1 \ldots \, \mathrm{d}u_n.
\end{aligned}
$$

By dividing the term on the right-hand side by $e^{-\lambda t}(\lambda t)^n / n!$, we obtain

$$
\begin{aligned}
&\mathbf{P}\left((T_1, \cdots, T_n) \in A \,|\, N(t) = n\right) \\
&\qquad = \frac{n!}{t^n} \underset{0\le u_1 \le \ldots \le u_{n+1}}{\int \ldots \int} \mathbf{1}_A(u_1, \cdots, u_n) \mathbf{1}_{[u_n, \infty[}(t) \, \mathrm{d}u_1 \ldots \, \mathrm{d}u_n,
\end{aligned}
$$

which, in view of Lemma 1.3, means that $(T_1, \ldots, T_n)$ has conditionally to $\{N(t) = n\}$, the same distribution as the vector $(\bar{U}_1, \ldots, \bar{U}_n)$ defined therein. In other words, conditionally to $\{N(t) = n\}$, the n-tuple $(T_1, \ldots, T_n)$ is uniformly distributed over $[0, t]$.

1.2 $\Longrightarrow$ 1.3.

Let $0 = t_0 < t_1 < \ldots < t_n$ be a family of $n + 1$ real numbers and $i_0, \ldots, i_{n-1}$, a family of n integers. We aim to prove that

$$
\mathbf{P}\left(\bigcap_{l=0}^{n-1} \Big\{ N(t_{l+1}) - N(t_l) = i_l \Big\} \right) = \prod_{l=1}^{n} \mathbf{P}(N(t_{l+1}) - N(t_l) = i_l).
$$

We can always write that

$$
\begin{aligned}
&\mathbf{P}\left(\bigcap_{l=0}^{n-1} \Big\{ N(t_{l+1}) - N(t_l) = i_l \Big\} \right) \\
&= \sum_{k\in\mathbf{N}} \mathbf{P}\left(\bigcap_{l=0}^{n-1} \Big\{ N(t_{l+1}) - N(t_l) = i_l \Big\} \,\Big|\, N(t_n) = k \right) \mathbf{P}(N(t_n) = k).
\end{aligned}
$$

The unique value of k for which the conditional probabilities of the latter quantity are non-zero is $k_0 = \sum_l i_l$. In order to derive the corresponding conditional probability, we know that the points between 0 and t_n are uniformly distributed. This quantity thus equals the probability that k points that are uniformly distributed over an interval, divides into i_1 points in the interval $[0, t_1]$, i_2 points in the interval $]t_1, t_2]$, and so on. Each point belongs to an interval of length x length with probability x/t_n. We thus have

$$
\begin{aligned}
&\mathbf{P}\left(\bigcap_{l=0}^{n-1} \Big\{ N(t_{l+1}) - N(t_l) = i_l \Big\} \,\Big|\, N(t_n) = k_0 \right) \\
&\qquad\qquad = \frac{k_0!}{i_1! \ldots i_n!} \prod_{l=0}^{n-1} \left(\frac{t_{l+1} - t_l}{t_n} \right)^{i_l}.
\end{aligned}
$$

As $N(t_n)$ follows a Poisson distribution with parameter λt_n and $k_0 = \sum_{l=0}^{n-1} i_1$, we deduce that

$$\begin{aligned}\mathbf{P}\left(\bigcap_{l=0}^{n-1}\left\{N(t_{l+1}) - N(t_l) = i_l\right\}\right) \\ &= e^{-\lambda t_n}\frac{(\lambda t_n)^{k_0}}{k_0!}\frac{k_0!}{i_1!\dots i_n!}\prod_{l=0}^{n-1}\left(\frac{t_{l+1}-t_l}{t_n}\right)^{i_l} \\ &= \prod_{l=0}^{n-1} e^{-\lambda(t_{l+1}-t_l)}\frac{\left(\lambda(t_{l+1}-t_l)\right)^{i_l}}{i_l!}.\end{aligned}$$

The probability of the intersection of events thus reads as a product of probabilities, therefore the random variables are independent. By taking $n = 2$, $t_1 = t$, $t_2 = t + s$, $i_0 = i$ and $i_1 = j$, we obtain

$$\mathbf{P}\left(N(t) = i,\, N(t+s) - N(t) = j\right) = e^{-\lambda t}\frac{(\lambda t)^i}{i!}e^{-\lambda s}\frac{(\lambda s)^j}{j!}.$$

Finally, summing over all the values of i yields the desired result.
1.3$\Longrightarrow$ 1.4.
Notice, that taking $f(s) = \mathbf{1}_{[a,b]}(s)$ leads to

$$\sum_n f(T_n) = N(b) - N(a).$$

From Lemma 1.2, we thus deduce that the result is true for the indicator functions and by linearity, for the piece-wise constant functions (that is to say, the linear combinations of indicator functions). By monotone convergence, we deduce that the result holds true for any positive measurable function.

□

1.2 *Properties*

Definition 1.2 might lead to a misinterpretation, and should be clearly understood. The latter stipulates that, conditionally to the number of points on an interval, the points are uniformly distributed over this interval. When we observe a sample path of the process, knowing t and the number of impacts in this interval, we should observe a cloud of point that is uniformly distributed. However, we observe distributions that are similar to that of Figure 1.2. It is the "clusterization" phenomenon: the arrivals give the impression of being grouped. The same observation can be made in actual stores, where after an idle period, many customers may arrive almost at the same time.

Figure 1.2: Four trajectories of a Poisson process. There are no areas of the segment that is not covered by any one of the trajectories, but each of the trajectories present "bursts" of arrivals.

Superposition, thinning

When we have the two point processes N^1 and N^2, the superposition of these processes is the point process, denoted as $N = N^1 + N^2$, whose points are those of N^1 and N^2. With Definition 1.4 in hand, the following result is straightforward.

Theorem 1.4.– The superposition of two independent Poisson processes of respective intensities λ_1 and λ_2 is a Poisson process of intensity $\lambda_1 + \lambda_2$.

n particular, during the superposition of two independent Poisson processes, two clients cannot arrive simultaneously. This result holds true for any two independent point processes.

Let us now assume that a Poisson process N of intensity λ be split into two processes N^1 and N^2 according to a random draw of Bernoulli of probability p, that is to say that at for any point of N, we decide that it belongs to N^1 with probability p and to N^2 with probability $1-p$. This draw is assume as independent of everything else in the model, and in particular of the previous draws (Figure 1.3). It is said that the Poisson process N is *thinned*.

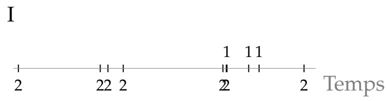


Figure 1.3: Thinning of a Poisson process. The figure above or below each point represents the number of flow to which this point is attributed.

Definition 1.4 ensures that a Poisson process on $\mathbf{R}^+$ is a special case of a spatial Poisson process (see Chapter 5) Therefore, Theorem 5.8 implies the following result. résultat qui suit.

Theorem 1.5.– The processes N^1 and N^2 processes resulting from the thinning of N, are two independent Poisson processes of respective intensities λp and $\lambda(1-p)$.

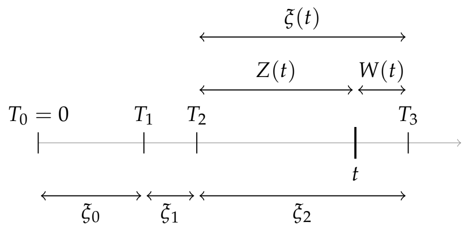


Figure 1.4: Notations

There is a specific result of the one dimensional case, known as the inspection paradox. Let us interpret the points of a Poisson process as the arrival times of buses at a given bus stop. Let

$$\begin{cases} W(t) &= T_{N(t)+1} - t, \\ Z(t) &= t - T_{N(t)}, \\ \xi(t) &= W(t) + Z(t) = T_{N(t)+1} - T_{N(t)}, \end{cases}$$

be the waiting time of the bus when arriving at the stop at time t, the time elapsed since the last bus went by, and the length of the time interval between the bus that we take and that we missed, respectively.

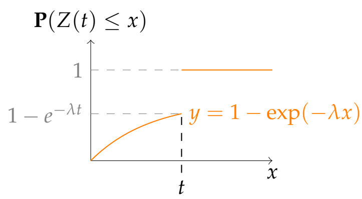


Figure 1.5: Cumulative distribution function of $Z(t)$

Theorem 1.6 (Inspection paradox).– $W(t)$ follows an exponential distribution of parameter λ and is independent of $Z(t)$, whose distribution is given by

$$\mathbf{P}(Z(t) \le x) = \begin{cases} 1 - e^{-\lambda x} & \text{if } 0 \le x < t; \\ 1 & \text{if } x \ge t. \end{cases}$$

Proof. Let us observe that $Z(t)$, by its very definition, cannot be larger than t. For $0 \le x < t$ and $y \ge 0$, by using the independence of T_n and ξ_n for any $n \in \mathbf{N}$, we have

$$\begin{aligned} \mathbf{P}(Z(t) \le x, W(t) > y) &= \sum_{n \in \mathbf{N}} \mathbf{P}(Z(t) \le x, W(t) > y, N(t) = n) \\ &= \sum_{n \in \mathbf{N}} \mathbf{P}(t - T_n \le x, T_n + \xi_n - t > y, T_n \le t < T_n + \xi_n). \end{aligned}$$

For $n = 0$ the latter probability equals

$$\mathbf{P}(t \le x, \xi_0 > t + y) = 0$$

since we have assumed that $x < t$. The previous sum therefore reads

$$\begin{aligned}
\sum_{n>1} \iint \mathbf{1}_{\{t-u\le x\}} \mathbf{1}_{\{u+v-t>y\}} \mathbf{1}_{\{u\le t\}} \, \mathrm{d}\mathbf{P}_{T_n}(u) \, \mathrm{d}\mathbf{P}_{\xi_{n+1}}(v) \\
&= \sum_{n>1} \int_{t-x}^{t} \lambda^n e^{-\lambda u} \frac{u^{n-1}}{(n-1)!} \left(\int_{t+y-u}^{\infty} \lambda e^{\lambda v} \, dv \right) \, \mathrm{d}u \\
&= \sum_{n>1} \lambda^n e^{-\lambda(t+y)} \int_{t-x}^{t} \frac{u^{n-1}}{(n-1)!} \, \mathrm{d}u \\
&= \sum_{n>1} \lambda^n e^{-\lambda(t+y)} \left[\frac{t^n}{n!} - \frac{(t-x)^n}{n!} \right] \\
&= e^{-\lambda(t+y)} \left(\sum_{n>1} \frac{(\lambda t)^n}{n!} - \sum_{n>1} \frac{(\lambda(t-x))^n}{n!} \right) \\
&= e^{-\lambda(t+y)} \left(e^{\lambda t} - 1 - \left(e^{\lambda(t-x)} - 1 \right) \right) \\
&= e^{-\lambda y}(1 - e^{-\lambda x}).
\end{aligned}$$

As $\lim_{x\nearrow t}(1 - e^{-\lambda x}) = 1 - e^{-\lambda t} < 1$, we infer that there is a jump in the distribution function of $Z(t)$ and therefore $\mathbf{P}(Z(t) = t) = e^{-\lambda t} > 0$. □

The mean expectation of $\xi(t)$ is derived in the following manner.

$$\begin{aligned}
\mathbf{E}\left[\xi(t)\right] &= \mathbf{E}\left[W(t)\right] + \mathbf{E}\left[Z(t)\right] \\
&= \frac{1}{\lambda} + t.e^{-\lambda t} + \int_0^t \lambda s e^{-\lambda s} \, \mathrm{d}s \\
&= \frac{1}{\lambda}(2 - e^{-\lambda t}).
\end{aligned}$$

Therefore, as t goes large the expectation of $\xi(t)$ tends to $2/\lambda$ and the average waiting time, which equals $1/\lambda$, represents half of it. This is in accordance with the intuition.

✍ This may seem paradoxical since the average value of W_t, that is to say, the average waiting time, is therefore $1/\lambda$, whereas the average time between two passages of the bus also equals $1/\lambda$. This property is another manifestation of what is commonly called the memoryless property of the exponential distribution, which will be discussed in the next chapter. In fact, everything happens as if, at the moment when we arrive at the bus stop, the counter of time which elapses between two bus arrivals was reset to zero, and if we recounted a time of exponential distribution until the next arrival.

Clearly speaking, this approach is mathematically wrong because if ξ_n follows an exponential distribution of parameter λ for any fixed n, $\xi(t)$ does not have the distribution of a ξ_n. Indeed, the number of the buses that we have missed changes along the samples paths. Conditionally to $\{N(t) = n\}$, $\xi(t)$ indeed has an exponential distribution of parameter λ, but contrary to what we might believe, when we un-condition, the distribution of $\xi(t)$ is no longer exponential.

For the accordance between this model and the reality, you can have a look at the blog *Pythonic perambulations*.

1.3 *Problems*

Exercise 1.3.1: Let N be a Poisson process of intensity λ, we denote T_n as the nth instant of jump. By convention, $T_0 = 0$. Let $(Z_n, n \geq 1)$ be a sequence of random variables of same distribution such that, for any n, T_n and Z_n are independent. Let g be the density of the common distribution of the Z_n's.

1. Show that for any function f,

$$\mathbf{E}\left[\sum_{n\geq 1} f(T_n, Z_n)\right] = \lambda \int_0^{+\infty} \int f(t, z) g(z) \, \mathrm{d}z \, \mathrm{d}t.$$

We assume that the telephonic communications of a subscriber lasts for a random time of exponential distribution of about three minutes in average. These durations are independent of each other. In the last century, the cost of communication was based on its duration t according to the following formula.

$$c(t) = \alpha \text{ si } t \leq t_0 \text{ et } c(t) = \alpha + \beta(t - t_0) \text{ si } t \geq t_0.$$

2. Deduce from the above that the average cost of one complete hour of communication is given by

$$\lambda \int_0^1 c(t) \lambda e^{-\lambda t} \, dt \tag{1.4}$$

with $\lambda = 20$. (*Hint: Consider $Z_n = T_{n+1} - T_n$ and explain why we can apply the previous result*).

Numerical application: for local calls, in 1999, we had the following parameters: $t_0 = 3$ minutes, $\alpha = 0{,}11$ euro and $\beta = 0{,}043$ euro per minute. For national calls, $t_0 = 39$ seconds and $\beta = 0{,}17$ euro per minute. α was the same. For reduced price, divide β by 2. By applying $t_0 = 1$ minute and $\alpha = 0{,}15$ euro, how much is the price of the extra second in mobile telephony in a package, whose amount for 1 hour of communication was 23,62 euros?

Solution on page 15

Exercise 1.3.2: An ATM records the beginning and ending times of queries of the customers, but of course not their arrival times in the queue. A new busy cycle having started at 7:30, we have recorded the following:

Customer number	Beginning of service	End of service
0	7h30	7h34
1	7h34	7h40
2	7h40	7h42
3	7h45	7h50

Let us assume that the arrivals take place according to a Poisson process, what can we say about the arrival time T_1 of customer 1? In particular, give its mean expectation.

Solution on page 15

Exercise 1.3.3: An insurance company must pay for claims at a rate of 5 per day. We assume that the instants of occurrence of disasters follow a Poisson process, and that the total amounts of damages are independent of each other, of exponential distribution with an average of 500 euros. We introduce $(X_i, i \geq 1)$, a sequence of i.i.d. random variables, of exponential distribution of average $1/\mu = 500$ euros, and a Poisson process N of intensity $\lambda = 5$ days^{-1}, independent of the X_i's.

1. What does $Z = \sum_{i=1}^{N(365)} X_i$ represent?

2. Calculate the average total annual amount spent by the insurance company.
3. Calculate $E[e^{-sZ}]$.
4. Infer the variance of Z.

Solution on page 16

1.4 Solution to problems

Solution of Exercise 1.3.1 on page 12:

1. Comme T_n et Z_n sont indépendants,

$$\mathbf{E}\left[f(T_n, Z_n)\right] = \int_0^{+\infty} \int f(t,z)g(z)\lambda e^{-\lambda t} \frac{(\lambda t)^{n-1}}{(n-1)!} \,\mathrm{d}z \,\mathrm{d}t.$$

Par conséquent,

$$\begin{aligned}
\mathbf{E}&\left[\sum_{n\ge 1} f(T_n, Z_n)\right] \\
&= \sum_{n\ge 1} \int_0^{+\infty} \int f(t,z)g(z)\lambda e^{-\lambda t} \frac{(\lambda t)^{n-1}}{(n-1)!} \,\mathrm{d}z \,\mathrm{d}t \\
&= \int_0^{+\infty} \int f(t,z)g(z)\lambda e^{-\lambda t} \left(\sum_{n\ge 1} \frac{(\lambda t)^{n-1}}{(n-1)!}\right) \,\mathrm{d}z \,\mathrm{d}t \\
&= \int_0^{+\infty} \int f(t,z)g(z)\lambda e^{-\lambda t} e^{\lambda t} \,\mathrm{d}z \,\mathrm{d}t
\end{aligned}$$

d'où le résultat.

2. Les durées des appels sont indépendantes les unes des autres et suivent une loi exponentielle donc du point de vue du crédit consommé, on peut mettre bout à bout tous les appels et l'on obtient un processus de Poisson. Le coût d'une heure de communication est donc

$$\sum_{n=1}^{\infty} \mathbf{1}_{T_n \le 1}\, c(T_{n+1} - T_n). \tag{1.7}$$

Comme $Z_n = T_{n+1} - T_n$ est indépendante de T_n, on peut appliquer le résultat précédent, d'où le coût moyen de l'heure de communication est donnée par (1.4). Avec les valeurs numériques proposées, on obtient

$$20 \int_0^{1/60} 20.0{,}15\, e^{-20t} \,\mathrm{d}t + 400.0{,}17.60 \int_{1/60}^{1} (t - \frac{1}{60})e^{-20t} \,\mathrm{d}t = 8{,}16€.$$

En fait, il y a une petite sur-estimation dans l'expression (1.7) parce que l'on compte le coût des minutes après l'heure du dernier appel commencé avant l'heure.

Solution of Exercise 1.3.2 on page 13:
On sait qu'il y a eu exactement deux arrivées entre 7h30 et 7h40 donc T_1 et T_2 suivent la même loi que $(U_{(1)}, U_{(2)})$ où U_1 et U_2 sont deux variables aléatoires indépendantes de loi uniforme sur $[0, 1]$ (en prenant 10 minutes comme unité de

temps). On sait par ailleurs que T_1 est inférieur à $0{,}4 = 2/5$. On doit donc calculer

$$\mathbf{E}\left[T_1 \mid T_1 \leq \frac{2}{5}\right].$$

Comme on a

$$\mathbf{P}(U_{(1)} \geq x) = \mathbf{P}(U_1 \geq x)\mathbf{P}(U_2 \geq x) = (1-x)^2, \tag{1.8}$$

on a

$$\mathbf{P}(T_1 \leq \frac{2}{5}) = \mathbf{P}(U_{(1)} \leq \frac{2}{5}) = 1 - (1 - \frac{2}{5})^2 = \frac{16}{25}.$$

Au vu de (1.8), la densité de la loi de T_1 est

$$2(1-x)\mathbf{1}_{[0,1]}(x)$$

donc

$$\mathbf{E}\left[T_1 \mid T_1 \leq \frac{2}{5}\right] = \frac{25}{16} 2 \int_0^{2/5} x(1-x)\, \mathrm{d}x = \frac{11}{60}.$$

Soit une arrivée en moyenne à $7h30$ plus $11/6$-ième de minute, soit $7h\,31m\ 50s$.

Solution of Exercise 1.3.3 on page 13:

1. Z is the total amount spent by the insurance company during a year.

2. Since N is independent from the X_i's, we can apply the Wald's Formula to get

$$\mathbf{E}[Z] = \mathbf{E}[N(365)]\,\mathbf{E}[X_1] = 5.365.3000 = 5\,475\,000 \text{ euros}$$

3. The difficulty comes from the randomness of the number of elements in the sum:

$$\begin{aligned}
\mathbf{E}\left[e^{-sZ}\right] &= \sum_{k=0}^{\infty} \mathbf{E}\left[e^{-sZ};\, N(365) = k\right] \\
&= \sum_{k=0}^{\infty} \mathbf{E}\left[e^{-s\sum_{i=1}^{k} X_i};\, N(365) = k\right] \\
&= \sum_{k=0}^{\infty} \prod_{i=1}^{k} \mathbf{E}\left[e^{-sX_i}\right] \mathbf{P}(N(365) = k) \\
&= \sum_{k=0}^{\infty} \mathbf{E}\left[e^{-sX_1}\right]^k e^{-365\,\lambda} \frac{(365\,\lambda)^k}{k!} \\
&= \exp\left(-365\,\lambda(1 - \mathbf{E}\left[e^{-sX_1}\right])\right).
\end{aligned}$$

But,

$$\mathbf{E}\left[e^{sX_1}\right] = \int_0^{\infty} e^{-sx} \mu e^{-\mu x}\, \mathrm{d}x = \frac{\mu}{\mu + s}.$$

Hence,

$$\mathbf{E}\left[e^{-sZ}\right] = \exp(-\frac{365\,\lambda\,s}{\mu + s}).$$

4. By differentiating twice the Laplace transform, we get:

$$\mathbf{E}\left[Z\right] = -\frac{d}{ds}\mathbf{E}\left[e^{-sZ}\right]\Big|_{s=0} = 365\frac{\lambda}{\mu}$$

$$\mathbf{E}\left[Z^2\right] = \frac{d^2}{ds^2}\mathbf{E}\left[e^{-sZ}\right]\Big|_{s=0} = 365^2\frac{\lambda^2}{\mu^2} + 730 * \frac{\lambda}{\mu^2}$$

Thus,

$$\mathrm{var}(Z) = 730 * \frac{\lambda}{\mu^2} = 730 * 5 * 3000^2 = 3{,}285.10^{10}.$$

2

Martingales with finite variation

2.1 Finite variation processes

Definition 2.1.– A stochastic process X indexed by $I \subset [0, +\infty)$ on a probability space $(\Omega, \mathcal{A}, \mathbf{P})$ is a map

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

which is measurable from $\mathcal{A} \otimes \mathcal{B}(I)$ to $\mathcal{B}(\mathbf{R})$.

For a fixed $\omega \in \Omega$, the function

$$\begin{aligned} X(\omega, .) \, : \, I &\longrightarrow \mathbf{R} \\ t &\longmapsto X(\omega, t) \end{aligned}$$

is called a sample-path or a trajectory of X. We usually write $X(t)$ instead of $X(\omega, t)$.

When studying stochastic processes, we often have to precise the regularity of its sample-paths. You know continuous and even differentiable functions but there are many more interesting classes of functions. For instance, we have the Hölder continuous functions:

Definition 2.2.– For $\alpha \in (0, 1]$, a function is said to be Hölder continuous of order α on a bounded interval I whenever there exists $c > 0$ such that for $s, t \in I$,

$$|f(t) - f(s)| \le c|t - s|^\alpha.$$

The Hölder norm of such a function is defined by

$$\|f\|_{\mathrm{Hol}(\alpha)} = \|f\|_\infty + \sup_{s \neq t} \frac{|f(t) - f(s)|}{|t - s|^\alpha}.$$

For $\alpha = 1$, the Hölder continuous functions of order 1 are also called Lipschitz functions.

We easily see that for $1 \ge \beta \ge \alpha > 0$,

$$\mathrm{Lip} \subset \mathrm{Hol}(\beta) \subset \mathrm{Hol}(\alpha) \subset \mathcal{C}(I, \mathbf{R}).$$

For functions which are not necessarily continuous, we often use the so-called Skorohod space of *cadlag* functions, i.e. functions which are right-continuous with left-limits (rcll in English but it is often called by its French name because of the importance of the French school in probability in these matters). This space is denoted by $\mathbb{D}([0,T],\mathbf{R})$ and it is equipped with the distance

$$d(f,g)=\inf_{\phi\in\mathrm{Hom}}\left(\|f-g\circ\phi\|_\infty+\|\phi-\mathrm{Id}\,\|_\infty\right)$$

where Hom is the set of homeomorphims from $[0,T]$ onto itself, i.e. time changes. Actually, the main part of the so-called stochastic calculus revolves around the definition of what is an integral with respect to a stochastic process when it is not differential : i.e. can we give a meaning to

$$\int \psi(s)\,\mathrm{d}X(s)?$$

Recall that the Riemann integral is defined as

$$\int_a^b \psi(s)\,\mathrm{d}s=\lim_{n\to\infty}\sum_{i=0}^{n-1}\psi(a+i\delta)\delta$$

where $\delta=(b-a)/n$. Note that

$$\delta=(a+(a+1)\delta)-(a+i\delta).$$

A natural idea to define something which could be denoted as

$$\int_a^b \psi(s)\,\mathrm{d}f(s) \tag{2.1}$$

would then be to replace the increments δ by the increments of the function f between $a+i\delta$ and $a+(i+1)\delta$, i.e. to consider the limit of

$$\sum_{i=0}^{n-1}\psi(a+i\delta)\;\left(f(a+(i+1)\delta)-f(a+i\delta)\right). \tag{2.2}$$

This is called the Young integral and there are many ways to find sufficient condition on ψ and f such that this limit does exist. For instance, if ψ is in $\mathrm{Hol}(\alpha)$ and $f\in\mathrm{Hol}(\beta)$ with $\alpha+\beta>1$, then the sum (2.2) does converge to a quantity denoted as in (2.1). The problem is that the map

$$t\longmapsto\int_0^t\psi(s)\,\mathrm{d}f(s)$$

is no longer Hölder continuous so that we cannot iterate the process and this is rather annoying for applications.

One way to construct such an integral is to assume that f has bounded (or finite) variation.

DEFINITION 2.3.– Let $[a,b]\subset\mathbf{R}$, we call $\mathcal{T}$ a partition of $[a,b]$, a finite set $t_0,\cdots,t_n$ such that:

$$a=t_0<t_1<\ldots<t_n=b.$$

We denote by $|\mathcal{T}|$, the mesh of the partition defined by

$$|\mathcal{T}| = \sup_{t_i \in \mathcal{T}} |t_{i+1} - t_i|.$$

$\mathfrak{T}_{[a,b]}$ is the set of all possible partitions of $[a, b]$.

DEFINITION 2.4.– Let $f : [a, b] \to \mathbf{R}$, f has finite variation whenever

$$\sup_{\mathcal{T} \in \mathfrak{T}_{[a,b]}} \sum_{t_i \in \mathcal{T}} |f(t_{i+1}) - f(t_i)| \text{ is finite.}$$

We denote by $\operatorname{Var}_{[a,b]}(f)$, this upper bound.

It is immediate that non decreasing, non increasing and Lipschitz functions have finite variation.

The proof of the next result is left to the reader.

THEOREM 2.1.– If $[c, d] \subset [a, b]$ and f has finite variation on $[a, b]$ then f has finite variation on $[c, d]$. Moreover,

$$\operatorname{Var}_{[a,c]}(f) + \operatorname{Var}_{[c,b]}(f) = \operatorname{Var}_{[a,b]}(f).$$

THEOREM 2.2.– A right continuous, finite variation function has an at most denumerable set of points of discontinuity.

Proof. We can consider f a non decreasing function on an interval $[a, b]$. Since f is right continuous and non decreasing,

$$f\Big([a, b]\Big) \subset \Big[f(a), f(b)\Big].$$

Let $A_n = \{s, \Delta f(s) \ge 1/n\}$ and $|A_n|$ its cardinality. Since

$$f(a) + |A_n| \frac{1}{n} \le f(b),$$

we see that A_n is finite. The set of points of discontinuity of f is the union of the A_ns hence it is at most denumerable. □

THEOREM 2.3 (Jordan decomposition).– Let f with finite variation on $[a, b]$, there exists a unique pair (g, h) of functions such that

1. $f = g - h + f(a)$;
2. g and h are non-decreasing;
3. $g(a) = h(a) = 0$;
4. $\operatorname{Var}_{[a,b]}(f) = \operatorname{Var}_{[a,b]}(g) + \operatorname{Var}_{[a,b]}(h)$.

Proof. STEP Existence. Set:

$$g(x) = \frac{1}{2}\Big(f(x) - f(a) + \operatorname{Var}_{[a,x]}(f)\Big)$$

and

$$h(x) = \frac{1}{2}\Big(f(a) - f(x) + \operatorname{Var}_{[a,x]}(f)\Big).$$

It remains to verify that these functions satisfy the claimed properties. It is immediate that

$$g(a) = h(a) = 0, \; f = g - h + f(a).$$

Since

$$|f(x) - f(y)| \leq \mathrm{Var}_{[x,y]}(f) = \mathrm{Var}_{[a,y]}(f) - \mathrm{Var}_{[a,x]}(f),$$

g and h are non decreasing. Thus:

$$\mathrm{Var}_{[a,b]}(g) = g(b) - g(a) \text{ et } \mathrm{Var}_{[a,b]}(h) = h(b) - h(a).$$

Hence $\mathrm{Var}_{[a,b]}(f) = \mathrm{Var}_{[a,b]}(g) + \mathrm{Var}_{[a,b]}(h)$.
STEP Uniqueness. Assume there exists another pair (g_1, h_1). Let $x < y$, g_1 is non decreasing hence $g_1(y) - g_1(x) \geq 0$. Furthermore

$$g_1(y) - g_1(x) = f(y) - f(x) + h_1(y) - h_1(x) \geq f(y) - f(x),$$

since h_1 is non decreasing. This implies that

$$\begin{aligned} g_1(y) - g_1(x) &\geq \max(0, f(y) - f(x)) \\ &= \frac{1}{2}(f(y) - f(x) + |f(y) - f(x)|). \end{aligned}$$

For any partition $[x, y]$, we then have

$$\begin{aligned} g_1(y) - g_1(x) &\geq \frac{1}{2}(f(y) - f(x) + \sum_i |f(t_{i+1}) - f(t_i)|) \\ &= \frac{1}{2}(f(y) - f(x) + \mathrm{Var}_{[x,y]}(f)) \\ &= \frac{1}{2}(f(y) - f(x) + \mathrm{Var}_{[a,y]}(f) - \mathrm{Var}_{[a,x]}(f)) \\ &= g(y) - g(x). \end{aligned}$$

The function $\beta \equiv g_1 - g$ is thus non decreasing. Moreover, the relation:

$$f \equiv g - h + f(a) \equiv g_1 - h_1 + f(a)$$

entails that $h_1 \equiv h + \beta$ and $\beta(a) = 0$. Finally, the constraint

$$\mathrm{Var}_{[a,b]}(f) = \mathrm{Var}_{[a,b]}(g_1) + \mathrm{Var}_{[a,b]}(h_1),$$

induces that:

$$g_1(b) - g_1(a) + h_1(b) - h_1(a) = g(b) + h(b).$$

But it holds that $g_1(b) - g_1(a) + h_1(b) - h_1(a) = g(b) + h(b) + 2\beta(b)$, thus $\beta = 0$ which means that $g \equiv g_1$ and $h \equiv h_1$. □

We can then have another characterization of finite variation functions.

THEOREM 2.4.– Let f be a right continuous, finite variation function.

There exists two measures (positive) λ_f^+ and λ_f^- such that

$$f(b) - f(a) = \lambda_f^+((a, b]) - \lambda_f^-((a, b]).$$

This decomposition is unique if we add the constraint for the two measures to have disjoint supports (which is always feasible).

$$|\,\mathrm{d}f(s)| = \,\mathrm{d}\lambda_f^+(s) + \,\mathrm{d}\lambda_f^-(s). \tag{2.3}$$

We then write

$$\int g(s)\,\mathrm{d}f(s) = \int g(s)\,\mathrm{d}\lambda_f^+(s) - \int g(s)\,\mathrm{d}\lambda_f^-(s) \tag{2.4}$$

as soon as

$$\int |g(s)||\,\mathrm{d}f(s)| = \int |g(s)|\,\mathrm{d}(\lambda_f^+ + \lambda_f^-)(s) < \infty. \tag{2.5}$$

Note that

$$\left|\int g(s)\,\mathrm{d}f(s)\right| \le \int |g(s)||\,\mathrm{d}f(s)|. \tag{2.6}$$

When we consider the Riemann integral of a continuous function, the resulting function

$$t \longmapsto \int_0^t f(s)\,\mathrm{d}s$$

is also continuous. There is an analog result for finite variation functions.

LEMMA 2.5 (Variation of integrals).– Let f have a finite variation and u such that

$$\int_0^T |u(s)||\,\mathrm{d}f(s)| < \infty.$$

Then the function

$$I_f\,:\,t \longmapsto \int_0^t u(s)\,\mathrm{d}f(s)$$

has bounded variation and

$$\mathrm{d}I_f(t) = u(t)\,\mathrm{d}f(t). \tag{2.7}$$

Proof. Assume u non negative and f non decreasing then I_t is non decreasing and has bounded variation. Furthermore, by definition

$$\lambda_{I_f}([0,t]) := I_f(t)$$

hence

$$\mathrm{d}\lambda_{I_f}(t) = u(t)\,\mathrm{d}f(t) \tag{2.8}$$

and (2.7) follows.

For the general case, write $u = u^+ - u^-$ and use the Jordan decomposition of f to get four integrals which represents non-decreasing functions and (2.8) follows. □

Before going further, we need to recall the theorem of Lebesgue of decomposition of measures.

Theorem 2.6.– Any measure μ on $\mathbf{R}$ can be uniquely decomposed into three parts:

$$\mu = \mu_{\text{ac}} + \mu_{\text{sc}} + \mu_{\text{d}}$$

where

Absolutely continuous part: μ_{ac} is absolutely continuous with respect to the Lebesgue measure: there exists $\rho \geq 0$ such that

$$\mu_{\text{ac}}(A) = \int_A \rho \, \mathrm{d}\ell$$

where ℓ is the Lebesgue measure on $\mathbf{R}$,

Singular continuous part: μ_{sc} which is singular with respect to the Lebesgue measure, i.e. there exists a set A such that $\ell(A) = 0$ and $\mu_{\text{sc}}(A^c) = 0$, but has no atoms, i.e. $\mu_{\text{sc}}(\{x\}) = 0$ for any $x \in \mathbf{R}$,

Singular atomic part: μ_d is concentrated on an at most denumerable set of $\mathbf{R}$.

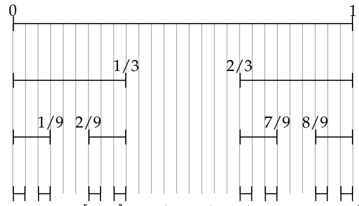
Hence, for f right continuous and non decreasing, we can write

$$f(x) = \int_{-\infty}^{x} h(u) \, \mathrm{d}u + \sum_{s,\, \Delta f(s) \neq 0} \Delta f(s) \mathbf{1}_{]-\infty, x]}(s) + (\lambda_f)_{\text{sc}}(] - \infty, x]).$$

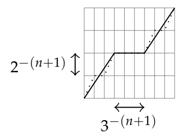
The Cantor triadic set is the set of points in $[0,1]$ that do not contain a 1 in their triadic expansion:

$$x = \sum_{j=1}^{\infty} x_j 3^{-j}$$

with $x_j \in \{0, 2\}$. This is equivalent to dividing the interval $[0,1]$, removing the middle third, and repeating the operation indefinitely on the remaining intervals.



Let $E_0 = [0,1]$ and E_n be the union of intervals remaining at step n. It is easy to see that the length of E_n is $(2/3)^n$ so the function $(3/2)^n \mathbf{1}_{E_n}$ defines a probability density c_n. The cumulative distribution function of c_n is increasing and constant on E_n^c.



Local view of of the cdf of c_n

Using Cauchy's criterion, we show that F_{c_n} converges uniformly to an increasing, continuous, zero function F almost everywhere and yet such that $F(0) = 0$ and $F(1) = 1$. This function defines a singular measure with respect to the Lebesgue measure without presenting any atoms.

Theorem 2.7 (Integration by parts).– Let f and g be two right continuous finite variation functions on $[0, t[$, then:

$$f(t)g(t) - f(0)g(0) = \int_0^t f(s^-) \, \mathrm{d}g(s) + \int_0^t g(s^-) \, \mathrm{d}f(s) + [f, g]_t, \tag{2.9}$$

where

$$[f, g]_t = \sum_{0 \leq s \leq t} \Delta f(s) \Delta g(s) \text{ et } \Delta f(s) = f(s) - f(s^-).$$

Proof. We calculate the integral $\lambda_f \otimes \lambda_g([0,t] \times [0,t])$ of two different ways. On the one hand, by definition of a product measure, we have

$$\lambda_f \otimes \lambda_g([0,t] \times [0,t]) = \lambda_f([0,t])\lambda_g([0,t]) = (f(t) - f(0))(g(t) - g(0)).$$

On the other hand, we break the square into two triangles (lower and upper I and S, respectively) and the diagonal D, see figure 2.1.

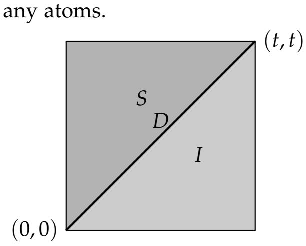


Figure 2.1: Decomposition of $[0,t]^2$

We apply the Fubini's Theorem to the two triangles. For $I \cup D$, we obtain :

$$\begin{aligned}
\lambda_f \otimes \lambda_g(I \cup D) &= \int_0^t \left(\int_0^s \mathrm{d}\lambda_g \right) \mathrm{d}\lambda_f(s) \\
&= \int_0^t (g(s) - g(0)) \, \mathrm{d}\lambda_f(s) \\
&= \int_0^t g(s) \, \mathrm{d}\lambda_f(s) - g(0)(f(t) - f(0)) \qquad (2.10)
\end{aligned}$$

Thus, we get

$$\int_0^t g(s) \, \mathrm{d}\lambda_f(s) = \int_0^t g(s^-) \, \mathrm{d}\lambda_f(s) + \int_0^t (g(s) - g(s^-)) \, \mathrm{d}\lambda_f(s)$$

since the number of discontinuity of g is at most denumerable, only the discrete part of λ_f has to be taken into account, hence

$$\int_0^t g(s)\, \mathrm{d}\lambda_f(s) = \int_0^t g(s^-)\, \mathrm{d}\lambda_f(s) + \sum_{0\le s\le t} \Delta g(s)\Delta f(s) \tag{2.11}$$

For the upper triangle, we have

$$\lambda_f \otimes \lambda_g(S) = \int_0^t f(s^-)\, \mathrm{d}\lambda_g(s) - f(0)(g(t) - g(0)). \tag{2.12}$$

By summing (2.10), (2.11) and (2.12), we get the claimed formula. □

The last theorem is a change of variable theorem. In dimension 1, the usual change of variable theorem states that:

$$F(g(t)) - F(g(0)) = \int_0^t \mathrm{d}(F\circ g)(s) = \int_0^t F'(g(s))g'(s)\, \mathrm{d}s. \tag{2.13}$$

It is in fact the application of the relation $f(t) - f(0) = \int_0^t f'(s)\, \mathrm{d}s$ to the identity $(F\circ g)' = F'\circ g.g'$. In the case of functions with finite variation, we do not have the first relation and the second seems difficult to verify. Nevertheless, we obtain a result similar to (2.13).

THEOREM 2.8 (Change of variable formula or Itô formula).– Let g be right continuous finite variation function and F a $\mathcal{C}^1$ class function. We have:

$$\begin{aligned} F(g(t)) - F(g(0)) &= \int_0^t F'(g(s^-))dg(s) \\ &\quad + \sum_{0\le s\le t} F(g(s)) - F(g(s^-)) - F'(g(s^-))\Delta g(s). \end{aligned} \tag{2.14}$$

REMARK 1.– In particular, if g is continuous with finite variation, we retrieve the usual result 2.13 at the price of replacing the differential element $g'(s)\, \mathrm{d}s$ by the measure $\mathrm{d}g(s)$.

Proof. The Itô formula is proved for $F(x) = x^n$ by induction from the integration by parts formula. The expression being linear in F, (2.14) is therefore true for polynomials. We approximate any function C^1 by a sequence of polynomials in the sense the uniform norm on $[0,t]$, the interversions of limits and integrals are all legitimate, so that we have the result for F of class $\mathcal{C}^1$. □

EXAMPLE 2.1.– Apply the Itô formula to $F = \exp$, we get

$$\begin{aligned} e^{g(t)} &= e^{g(0)} + \int_0^t e^{g(s^-)}\, \mathrm{d}g(s) + \sum_{s\le t} e^{g(s^-)+\Delta g(s)} - e^{g(s^-)} - e^{g(s^-)}\Delta g(s) \\ &= e^{g(0)} + \int_0^t e^{g(s^-)}\, \mathrm{d}g(s) + \sum_{s\le t} e^{g(s^-)} \left(e^{\Delta g(s)} - 1 - \Delta g(s)\right). \end{aligned}$$

2.2 *Martingales*

Definition 2.5.– A filtration $(\mathcal{F}_t,\, t \in \mathbf{R}^+)$ is a non-decreasing family of σ-fields. It is said to be right continuous whenever

$$\bigcap_{s>t} \mathcal{F}_s = \mathcal{F}_t \text{ for all } t \in \mathbf{R}^+.$$

It is assumed in the following that all the filtrations encountered continuous on the right.

Definition 2.6.– Let $(\Omega, \mathcal{F} = (\mathcal{F}_t,\, t \ge 0), \mathbf{P})$ be a filtered space. A stochastic process $M = (M(t),\, t \ge 0)$ is an $\mathcal{F}$-martingale (respectively sub-martingale, super-martingale) when for any $0 \le s \le t$, $M(t) \in L^1(\mathbf{P})$ and:

$$\mathbf{E}\left[M(t) \,|\, \mathcal{F}_s\right] = M(s), \tag{2.15}$$

(respectively $\ge M(s)$ et $\le M(s)$).

We admit that all martingales admit a version with paths which are continuous on the right with limits on the left (cadlag for short). There are two main types of martingales : the continuous ones and the purely discontinuous ones. The continuous martingales whose archetype is the Brownian motion do not have a finite variation. The martingales we are interested in are finite variation martingales and are therefore discontinuous.

Theorem 2.9 (Poisson process).– Let N be a Poisson process of intensity μ. The process $M(t) = N(t) - \mu([0,\, t])$ is a martingale for the filtration generated by the sample-paths of N: $\mathcal{F}_t = \sigma(N(u),\, u \le t)$.

Proof. According to Definition 1.3, $N(t) - N(s)$ is independent of $\mathcal{F}_s$ hence:

$$\mathbf{E}\left[N(t) - N(s) \,|\, \mathcal{F}_s\right] = \mathbf{E}\left[N(t) - N(s)\right] = \mu(]s,\, t]),$$

The proof is thus complete. □

Definition 2.7.– A random variable τ with values in $\mathbf{R}^+ \cup \{+\infty\}$ is a $\mathcal{F}$-stopping time when for for any $t \ge 0$,

$$(\tau \le t) \in \mathcal{F}_t.$$

A set A belongs to the σ-field $\mathcal{F}_\tau$ if and only if

$$A \cap (\tau \le t) \in \mathcal{F}_t.$$

In the particular case where the filtration is generated by a point process, we have an explicit description of the elements of $\mathcal{F}_\tau$.

Theorem 2.10.– Let $\mathcal{F}_t = \sigma\{N(s),\, 0 \le s \le t\}$. A set A belong to $\mathcal{F}_\tau$ if and only if there exist $(\psi_q,\, q \ge 0)$ some $\{0,1\}$-valued functions such that

$$\mathbf{1}_A = \sum_{q=0}^{\infty} \psi_q(T_1(N), \cdots, T_q(N))\, \mathbf{1}_{[T_q(N), T_{q+1}(N))}(\tau)$$

where $T_0(N) = 0$ and $(T_q(N),\, q \geq 1)$ are the jump times of N.

The well known established theorems for discrete time martingales continue to hold without change for continuous time indexed martingales. For instance, if M is a martingale and T is a stopping time, the process

$$M^T = \{M(t \wedge T),\, t \geq 0\},$$

is a martingale.

DEFINITION 2.8.– A $\mathcal{F}$-martingale M is said to be closed if there exists $M_\infty \in L^1$ such that $t \geq 0$, we have:

$$M(t) = \mathbf{E}\left[M_\infty \,|\, \mathcal{F}_t\right].$$

To check that a process is a martingale, it is obviously necessary to show that it satisfies the identity (2.15). But prior to this property, one must be certain of the integrability of M and this is often the problem. When we have a deterministic function whose integrability we want to guarantee, the most common way is to truncate it by considering $(f \wedge M) \vee (-M)$ for example. If we proceed in this way for a process, one will necessarily lose the martingale property. On the other hand, we know that a stopped martingale is a martingale so the idea is to stop the process before its process before its values exceed a certain threshold. This is the idea underlying the notion of local martingale.

DEFINITION 2.9.– An adapted process with cadlag sample paths is local martingale if there exist an increasing sequence of stopping times $(T_n,\, n \geq 1)$, tending to infinity, such that for any $n \geq 1$, M^{T_n} is a closed martingale. The sequence $(T_n,\, n \geq 1)$ is said to reduce M.

THEOREM 2.11.– Let M be a cadlag process which is a local martingale. If there exists $Z \in L^1$ such that $M(t) \leq Z$ for any $t \geq 0$, then M is a martingale.

Proof. Let $(T_n,\, n \geq 1)$ be a sequence of stopping times which reduces M. For $0 \leq s \leq t$, we have:

$$M(s \wedge T_n) = \mathbf{E}\left[M(t \wedge T_n) \,|\, \mathcal{F}_s\right].$$

Since T_n tends to infinity, $s \wedge T_n = s$ from a certain rank onwards (depending on the sample-path) hence by right continuity of the sample-paths of M, $M(s \wedge T_n)$ tends to $M(s)$ almost-surely. By dominated convergence, we obtain $M(s) = \mathbf{E}\left[M(t) \,|\, \mathcal{F}_s\right].$ □

LEMMA 2.12.– Let f be a continuous function on $[0, T]$ and $(\pi_n,\, n \geq 1)$ a sequence of subdivisions of $[0, T]$ such that $|\pi_n| \to 0$. Then,

$$\sup_{i=0}^{l(\pi_n)-1} \left| f(t_{i+1}^{\pi_n}) - f(t_i^{\pi_n}) \right| \xrightarrow{n \to \infty} 0. \tag{2.16}$$

Proof. Since f is continuous on a compact interval, it is uniformly continuous hence for any $\varepsilon > 0$, there exists $\delta > 0$ such that

$$|t-s| < \delta \Longrightarrow |f(t) - f(s)| \le \varepsilon.$$

Fix $\varepsilon > 0$, there exists n_0 such that $n \ge n_0$ implies $|\pi_n| \le \delta$ hence

$$\sup_{n\ge n_0} \sup_{i=0}^{l(\pi_n)-1} \left| f(t_{i+1}^{\pi_n}) - f(t_i^{\pi_n}) \right| \le \varepsilon.$$

That induces (2.16). □

There is a strange link between the martingale and the bounded variation properties.

Theorem 2.13.– Let M be a local martingale with continuous sample-paths and bounded variation. Then $M(t) = M(0)$ for any $t \ge 0$.

Proof. For the sake of simplicity, assume that $M(0) = 0$, otherwise work with $M(t) - M(0)$. Let

$$T_n = \inf\{t \ge 0, \int_0^t |dM(s)| \ge n\}.$$

It is a stopping time and

$$\begin{aligned} |M^{T_n}(t)| &= \left| \int_0^t \, \mathrm{d}M^{T_n}(s) \right| \\ &\le \int_0^t | \, \mathrm{d}M^{T_n}(s)| \\ &\le nT. \end{aligned}$$

Thus according to Theorem 2.11, $N = M^{T_n}$ is a true bounded martingale.

There is a trick for martingales which is that

$$\mathbf{E}\left[(M(t) - M(s))^2 \right] = \mathbf{E}\left[M(t)^2 \right] - \mathbf{E}\left[M(s)^2 \right].$$

Hence, we have

$$\mathbf{E}\left[N(t)^2 \right] = \sum_{t_i \in \tau} \mathbf{E}\left[(N(t_{i+1}) - N(t_i))^2 \right]$$

for any partition τ on $[0, T]$. Thus,

$$\begin{aligned} \mathbf{E}\left[N(t)^2 \right] &\le \mathbf{E}\left[\sup_i |N(t_{i+1}) - N(t_i)| \sum_{t_i \in \tau} |N(t_{i+1}) - N(t_i)| \right] \\ &\le \mathbf{E}\left[\sup_i |N(t_{i+1}) - N(t_i)| \operatorname{var}_{[0,T]} N \right] \\ &\le n\mathbf{E}\left[\sup_i |N(t_{i+1}) - N(t_i)| \right]. \end{aligned}$$

Choose $(\pi_k, k \ge 1)$ a sequence of subdivisions whose mesh tend to 0. Since N has almost surely continuous sample-paths, according to Lemma 2.12

$$\sup_i |N(t_{i+1}^{\pi_k}) - N(t_i^{\pi_k})| \xrightarrow{k\to\infty} 0, \text{ a.e.}$$

Since N is bounded, the dominated convergence theorem entails that $\mathbf{E}\left[N(t)^2\right] = 0$ hence $M^{T_n} = 0$, a.s. By letting n go to infinity, T_n goes to infinity so that $M = 0$ a.e. □

2.3 Compensators

DEFINITION 2.10.– A configuration on $(0, +\infty)$ is a denumerable set $\eta = (t_q, q \ge 1)$ of increasing numbers such that there is a finite number of t_q in any bounded set. We identify the set η and the measure

$$\sum_{q\ge 1} \delta_{t_q}$$

so that

$$\int f \,\mathrm{d}\eta = \sum_{q\ge 1} f(t_q).$$

We denote by $\mathfrak{N}$ the set of such configurations and we equip it with the vague convergence :

$$\eta_n \xrightarrow{\text{vaguely}} \eta \iff \int f \,\mathrm{d}\eta_n \to \int f \,\mathrm{d}\eta$$

for any f continuous with compact support in $(0, +\infty)$.

A point process is a random variable with values in $\mathfrak{N}$. It can be identified with a cadlag function (called a counting process) via the map

$$\begin{aligned} \mathfrak{N} &\longrightarrow \mathbb{D}(\mathbf{R}^+, \mathbf{R}) \\ \eta = (t_q, q \ge 1) &\longmapsto N(t) = \sum_{q\ge 1} \mathbf{1}_{[0,t]}(t_q). \end{aligned}$$

The inverse map is defined on the subspace of $\mathbb{D}$ made by the functions which are constant between their jumps and whose jumps have a height of 1, by

$$\begin{aligned} \mathbb{D}(\mathbf{R}^+, \mathbf{R}) &\longrightarrow \mathfrak{N} \\ N &\longmapsto \{s, \Delta N(s) > 0\}. \end{aligned}$$

The minimal filtration on $\mathbb{D}(\mathbf{R}^+, \mathbf{R})$ is the filtration

$$\mathcal{N}_t = \sigma\{N(s), s \le t\}.$$

DEFINITION 2.11.– Consider that we have chosen a filtration $(\mathcal{F}_t, t \ge 0)$ on Ω. The predictable σ-field on $\Omega \times \mathbf{R}^+$ is generated by the left continuous and adapted processes. It is also generated by the processes of the form

$$u(\omega, t) = \alpha(\omega)\, \mathbf{1}_{(a,b]}(t)$$

where $\alpha \in \mathcal{F}_a$.

When $\mathcal{F}$ is the filtration generated by a point process, nothing happens between two jump times hence we have

THEOREM 2.14.– A predictable process on $(\mathfrak{N}, \mathcal{N})$ can be written as

$$u(N,t) = \sum_{q=0}^{\infty} \varphi_q(T_1(N), \cdots, T_q(N), t)\, \mathbf{1}_{(T_q(N),\, T_{q+1}(N)]}(t)$$

for some φ_0 deterministic and for any $q \ge 1$, φ_q is measurable from $(\mathbf{R}^+)^{q+1}$ into $\mathbf{R}$.

THEOREM 2.15.– For any point process N built on $(\Omega, (\mathcal{F}_t,\, t \ge 0), \mathbf{P})$, there exists a unique non-decreasing, predictable process $(y(N,t),\, t \ge 0)$ such that

$$t \mapsto N(t) - y(N,t)$$

is a local martingale. This process is called the compensator of N under $\mathbf{P}$.

This theorem admits a converse which says that under some hypothesis on the filtration, a probability measure on $\mathfrak{N}$ is fully defined by the compensator.

THEOREM 2.16.– On $(\mathfrak{N}, \mathcal{N})$, consider y a non-decreasing, predictable process, null at time 0. There exists a unique probability measure $\mathbf{P}_y$ for which

$$\eta(t) - y(\eta, t)$$

is a $\mathcal{N}$-local martingale.

REMARK 2.– For the Poisson point process of intensity λ, we already know from Theorem 2.9 that its compensator is the deterministic process $(t \mapsto \lambda t)$.

DEFINITION 2.12.– A point process is said to be a non homogeneous Poisson process whenever its compensator is deterministic.

EXAMPLE 2.2.– Imagine that you are driving along an infinite road with constant speed normalized to be 1. You see the road signs at times described by a Poisson process N of intensity λ. This means that $T_q(N)$ represents the time at which you see the q-th sign.

Now then, another driver has a time varying speed $(\phi(s),\, s \ge 0)$. We denote by Φ the primitive of ϕ:

$$\Phi(t) = \int_0^t \phi(s)\, \mathrm{d}s.$$

If we assume that the other driver never stops, i.e. $\phi > 0$, then Φ is a bijection from $\mathbf{R}^+$ onto itself. The function Φ represents the distance covered during one unit of time. We denote by N' the point process which represents the time at which he sees the road signs. We must have

$$\Phi\big(T_q(N')\big) = T_q(N) \iff T_q(N') = \Phi^{-1}\big(T_q(N)\big).$$

This is equivalent to say that

$$N'(t) = N\big(\Phi(t)\big).$$

Indeed, let

$$R(t) = N\big(\Phi(t)\big).$$

We have

$$\begin{aligned}\Delta R(t) \neq 0 &\iff N\big(\Phi(t)\big) \neq N\big(\Phi(t^-)\big)\\ &\iff \Phi(t) \text{ is a jump time of } N\\ &\iff \Phi\big(T_q(R)\big) = T_q(N).\end{aligned}$$

We clearly have

$$\begin{aligned}\mathcal{R}_t &:= \sigma\{R(s),\, s \le t\}\\ &= \sigma\{N(s),\, s \le \Phi(t)\}\\ &= \mathcal{N}_{\Phi(t)}.\end{aligned}$$

Since Φ is deterministic, the martingale property of N induces that

$$\begin{aligned}\mathbf{E}\left[R(t) - R(s)\,|\,\mathcal{R}_s\right] &= \mathbf{E}\left[N(\Phi(t)) - N(\Phi(s))\,|\,\mathcal{N}_{\Phi(s)}\right]\\ &= \Phi(t) - \Phi(s).\end{aligned}$$

Hence $R - \Phi$ is a martingale, Φ is the compensator of R and ϕ is called the intensity (function) of R.

Pay attention to the filtration used to establish the martingale property.

More generally, given some information on the distributions of the jumps, we can retrieve the compensator.

THEOREM 2.17.– Let N be a point process and denote its jump times by $(T_q(N),\, q \ge 1)$. For any $A \in \mathcal{B}(\mathbf{R}^+)$, define the random measure

$$G_q(N; A) = \mathbf{P}\Big(T_{q+1}(N) \in A \,\Big|\, T_1(N), \cdots, T_q(N)\Big).$$

Then, the compensator of N is equal to

$$y(N,t) = \sum_{q=0}^{\infty} \int_0^t \frac{G_q(N; \mathrm{d}s)}{G_q(N; [s,\infty))}\, \mathbf{1}_{(T_q(N),\, T_{q+1}(N)]}(s)\, \mathrm{d}s. \tag{2.17}$$

DEFINITION 2.13.– A renewal process is a point process where the inter-arrivals are independent and identically distributed.

THEOREM 2.18.– Let N be a renewal process. Let F (respectively f) be the distribution function (respectively the density) of the joint law of inter-arrivals. We then have

$$y(N,t) = \int_0^t \sum_{q=0}^{\infty} \frac{f(s - T_q)}{1 - F(s - T_q)}\, \mathbf{1}_{(T_q,\, T_{q+1}]}(s)\, \mathrm{d}s.$$

Proof. We have

$$\begin{aligned}\mathbf{P}(T_{n+1} \ge s \,|\, T_1, \cdots, T_n) &= \mathbf{P}(T_{n+1} - T_n \ge s - T_n \,|\, T_1, \cdots, T_n)\\ &= \mathbf{P}(\xi_{n+1} \ge s - T_n \,|\, T_n)\\ &= 1 - F(s - T_n).\end{aligned}$$

Hence the conditional density of T_{n+1} given $(T_1, \cdots, T_n)$ is $s \mapsto f(s - T_n)$. We conclude by (2.17). □

REMARK 3.– By the properties of the exponential density, we see that

$$\frac{f(s-T_q)}{1-F(s-T_q)} = \frac{\lambda e^{-\lambda(s-T_q)}}{1-(1-e^{-\lambda(s-T_q)})} = \lambda,$$

so that we retrieve that the compensator of a Poisson process is $t \mapsto \lambda t$.

2.4 *Time change*

We now show that in most of the situations, point processes can be constructed as Poisson processes modified by a time change.

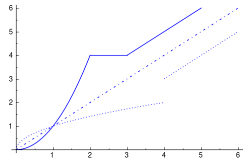


Figure 2.2: Graph of a and of its pseudo-inverse τ(dashed). The plateaux of y are transformed as jumps in y^* and vice-versa.

DEFINITION 2.14.– Let a be positive, non-decreasing, right continuous such that

$$a(0) = 0 \text{ and } a(\infty) = \infty.$$

Consider

$$a^*(t) = \inf\{r,\ a(r) > t\}.$$

LEMMA 2.19.– a^* is right continuous and has a symmetric rôle:

$$a(t) = \inf\{r,\ a^*(r) > t\}.$$

Moreover,

$$\begin{aligned} \left(a^*(s^-) \le r\right) &= \left(s \le a(r)\right) \\ \left(a(s^-) \le r\right) &= \left(s \le a^*(r)\right) \end{aligned}$$

THEOREM 2.20 (Time change).– Let y be a continuous compensator and Y be a point process of compensator y. Denote by y^* the inverse of y. Then,

$$Y\big(y^*(t)\big) - t$$

is a local martingale for the filtration $(\mathcal{F}_{y^*(t)},\ t \ge 0)$.

Stricto sensu, the proof should require to localize everything in order to have well defined expectations. This makes the proof too heavy to be easily understandable.

Proof. Remark that

$$\left(y^*(t) \le s\right) = \left(y(s) \ge t\right) \in \mathcal{F}_s \tag{2.18}$$

hence $y^*(t)$ is stopping time for any $t \ge 0$ and the σ-field is well defined. By definition of a compensator, for $s \le t$,

$$\mathbf{E}\left[Y(t) - y(t) \,|\, \mathcal{F}_s\right] = Y(s) - y(s).$$

If we apply the stopping time theorem to $y^*(t)$ and $y^*(s)$, we get

$$\mathbf{E}\left[Y(y^*(t)) - y(y^*(t)) \,\Big|\, \mathcal{F}_{y^*(s)}\right] = Y(y^*(s)) - y(y^*(s))$$

or equivalently

$$\mathbf{E}\left[Y(y^*(t)) - t \,\Big|\, \mathcal{F}_{y^*(s)}\right] = Y(y^*(s)) - s.$$

Hence the result. $\square$

REMARK 4.– Set
$$Z(t) = Y\Big(y^*(t)\Big). \tag{2.19}$$
A jump occurs for Z at time $T_q(Z)$ if and only if $y^*(T_q(Z))$ is the q-th jump of Y hence we have
$$y^*(T_q(Z) = T_q(Y) \text{ or } T_q(Z) = y\Big(T_q(Y)\Big).$$

REMARK 5.– From (2.18), we see that $y(t)$ is a $\mathcal{F}_{y^*}$-stopping time. Hence, from (2.19), we write
$$Y(t) = Z\Big(y(t)\Big)$$
where Z is a Poisson process of intensity 1. This means that Y is a time changed Poisson process.

2.5 *Problems*

Exercise 2.5.1: Show (2.14) for $F(x) = x^3$. For the sake of simplicity, assume $f(0) = 0$.

Solution on page 36

Exercise 2.5.2: Let N be the point process defined by

- $T_1(N)$ follows an exponential distribution of parameter λ_1 and so do the random variables $T_{2n+1}(N) - T_{2n}(N)$ for any $n \geq 1$.
- The random variables $T_{2n}(N) - T_{2n-1}(N)$ for any $n \geq 1$ follow an exponential distribution of parameter μ.
- The random variables $(T_1(N), T_{n+1}(N) - T_n(N), n \geq 1)$ are independent.

1. Compute the compensator of N.

Solution on page 36

Exercise 2.5.3: For α, γ, τ some positive numbers, for any configuration ω on $\mathbf{R}^+$, let
$$\dot{y}(\omega, s) = \alpha s \mathbf{1}_{[0,\tau]}(s) + \mathbf{1}_{[\tau,\infty[}(s).$$

1. Compute $y(\omega, t)$.
2. How can we simulate a sample-path of Y, the point process of compensator y ?
3. What is the law of the first jump of Y when $\tau > 2/\alpha$?

Solution on page 36

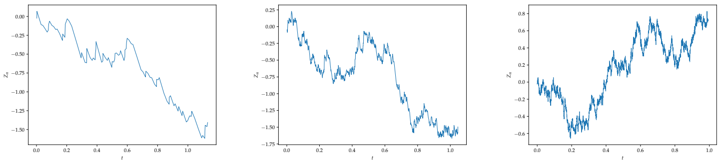


Figure 2.3: Convergence of renormalized Poisson process towards Brownian motion for $n = 100;\ 1,000;\ 10,000$

The next exercise is about the convergence in law of the renormalized Poisson point process towards the Brownian motion when the intensity is large enough as illustrated in Figure 2.3. That is to say that if we denote by

$$Z_n(t) = \frac{N_n(t) - nt}{\sqrt{n}}.$$

where N^n is a Poisson point process of intensity n, then for any continuous function $f : \mathbb{D}([0,T], \mathbf{R}) \to \mathbf{R}$,

$$\mathbf{E}\left[f(Z_n)\right] \xrightarrow{n\to\infty} \mathbf{E}\left[f(B)\right]$$

where B is a standard Brownian motion. Some examples of continuous functions on $\mathbb{D}$ are:

- $f \mapsto \varphi\Big(f(t_1), \cdots, f(t_k)\Big)$ where φ is continuous on $\mathbf{R}^k$,
- $f \mapsto \sup_{s\le T} |f(s)|$,
- $f \mapsto \int_0^T |f(s)|^p \, \mathrm{d}s$, etc.

There is one essential theorem to prove such a convergence:

The reference book for these questions is, for a long time, the book by Billingsley, *Convergence of Probability Measures*. See also Kallenberg, *Foundations of modern probability*.

THEOREM 2.21.– Let $(X_n,\ n \ge 1)$ be a sequence of cadlag processes and ν the law of a cadlag process denoted by X. If

Finite dimensional (fidi) convergence: For any $0 \le t_1 < \ldots < t_k \le T$,

$$\left(X_n(t_1), \cdots, X_n(t_k)\right) \xrightarrow[n\to\infty]{\text{Law}} \left(X(t_1), \cdots, X(t_k)\right),$$

Tightness: there exist $\beta > 0$ and $\alpha > 1/2$ such that for any $0 \le s \le r \le t \le T$,

$$\mathbf{E}\left[|X_n(r) - X_n(s)|^{2\beta}\, |X_n(t) - X_n(r)|^{2\beta}\right] \le (F(t) - F(s))^{2\alpha}$$

for some continuous, non-decreasing function F on $[0,T]$.

Then X_n converges in law in $\mathbb{D}([0,T], \mathbf{R})$ to X.

As a preparation to what follows, there is a classical lemma worth knowing.

LEMMA 2.22.– Assume that for any $n \ge 1$, X_n and Y_n are independent and that they both converge in law to X and Y respectively. Then $\mathbf{P}_{(X_n,Y_n)}$ converges narrowly to $\mathbf{P}_X \otimes \mathbf{P}_Y$: the pair (X_n, Y_n) converges in law to (X', Y') where X' and Y' are independent of respective distribution $\mathbf{P}_X$ and $\mathbf{P}_Y$.

Proof. For any $s, t \in \mathbf{R}$,

$$\begin{aligned}\mathbf{E}\left[e^{i(sX_n+tY_n)}\right] &= \mathbf{E}\left[e^{isX_n}\right]\mathbf{E}\left[e^{itY_n}\right] \\ &\xrightarrow{n\to\infty} \mathbf{E}\left[e^{isX}\right]\mathbf{E}\left[e^{isY}\right].\end{aligned}$$

This means that the characteristic function of the pair (X_n, Y_n) converge to the characteristic function of the measure $\mathbf{P}_X \otimes \mathbf{P}_Y$, otherwise stated, the pair (X_n, Y_n) converges to the product measure whose marginals are $\mathbf{P}_X$ and $\mathbf{P}_Y$. □

We can now prove the convergence we are interested in.

Exercise 2.5.4: Let N_n be a Poisson process of intensity n and

$$Z_n(t) = \frac{N_n(t) - nt}{\sqrt{n}}.$$

1. Compute the characteristic function of a Poisson random variable of parameter λ.
2. Using the Lévy Theorem, show that for $t > 0$ fixed, $Z_n(t)$ converges to a Gaussian law the parameters of which have to be determined.
3. Let $0 < t_1 < \ldots < t_K$ and

$$R_{n,K} = \Big(Z_n(t_1), \cdots, Z_n(t_K) - Z_n(t_{K-1})\Big).$$

   Show that the components of $R_{n,K}$ are independent and find the limit in law of $R_{n,K}$ for K fixed when n goes to infinity.
4. Show

$$\sup_{n\ge1} \mathbf{E}\left[|Z_n(t) - Z_n(s)|^2\right] \le |t - s|$$

5. Show the convergence of Z_n towards B.

Solution on page 37

2.6 Solution to problems

Solution of Exercise 2.5.1 on page 33:
We know that

$$f(t)^2 = 2\int_0^t f(s^-)\,\mathrm{d}f(s) + \sum_{s\le t}(\Delta f(s))^2 \tag{2.22}$$

hence

$$d\left(f(t)^2\right) = 2f(t^-)\,\mathrm{d}f(t) + (\Delta f(t))^2\delta_t. \tag{2.23}$$

It follows that

$$\begin{aligned} f(t)^3 &= f(t)^2 f(t) \\ &= \int_0^t f(s^-)^2\,\mathrm{d}f(s) + \int_0^t f(s^-)\,\mathrm{d}\left(f(s)^2\right) + \sum_{s\le t}\Delta f(s)\,(\Delta f(s))^2. \end{aligned}$$

From (2.23), we have

$$\int_0^t f(s^-)\,\mathrm{d}\left(f(s)^2\right) = 2\int_0^t f(s^-)\,\mathrm{d}f(s) + \sum_{s\le t} f(s^-)(f(s)-f(s^-))^2.$$

If we expand the sum term and then make some rearrangements, we obtain

$$f(t)^3 = 3\int_0^t f(s^-)^2\,\mathrm{d}f(s) + \sum_{s\le t} f(s)^3 - f(s^-)^3 - 3f(s^-)^2\Delta f(s).$$

Solution of Exercise 2.5.2 on page 33:
The law of $T_{2n+1}(N)$ given $T_1(N),\cdots,T_{2n}(N)$ is the law of $T_{2n}(N)+X_\lambda$ where X_λ is an exponentially distributed random variable independent of T_{2n}. It follows from Theorem 2.17 that on the interval $(T_{2q}(N), T_{2q+1}(N)]$, the compensator of N is λt. Similarly on the intervals $(T_{2q-1}(N), T_{2q}(N)]$, it is μt.

Solution of Exercise 2.5.3 on page 33:

1. It is immediate that
$$y(N,t) = \frac{\alpha}{2}(t\wedge\tau)^2 + (t-\tau)^+.$$
2. We see from Remark 5 that in distribution,
$$Y(t) = N(y(t))$$
where N is a Poisson process of intensity 1. Hence
$$y\left(T_q(Y)\right) = T_q(N) \Longleftrightarrow T_q(Y) = y^{-1}(T_q(N)).$$
It remains to compute y^{-1}.
$$v \le v_0 := \frac{\alpha}{2}\tau^2 \Longrightarrow y^{-1}(v) = \sqrt{\frac{2v}{\alpha}}.$$

If $v > v_0$ then

$$\begin{aligned} y(N,t) = v &\iff v_0 + (t-\tau) = v \\ &\iff t = v - v_0 + \tau. \end{aligned}$$

3. We have

$$\begin{aligned} \mathbf{P}(T_1(Y) \leq r) &= \mathbf{P}(y^{-1}(T_1(N)) \leq r) \\ &= \mathbf{P}(T_1(N) \leq y(r)) \\ &= 1 - e^{-y(r)}. \end{aligned}$$

Hence the law of $T_1(Y)$ has density $\left(t \mapsto \dot{y}(t) e^{-y(t)}\right)$.

Solution of Exercise 2.5.4 on page 35:

1. Let R be distributed as a Poisson random variable of parameter λ. We have

$$\begin{aligned} \mathbf{E}\left[e^{itR}\right] &= \sum_{k=0}^{\infty} e^{-\lambda+itk} \frac{\lambda^k}{k!} \\ &= \exp\left(-\lambda + \lambda e^{it}\right) \end{aligned}$$

2. We deduce from the previous question that

$$\begin{aligned} &\mathbf{E}\left[e^{isZ_n(t)}\right] \\ &\quad = \mathbf{E}\left[\exp\left(i\frac{sN_n(t)}{\sqrt{n}}\right)\right] \exp\left(-is\sqrt{n}t\right) \\ &\quad = \exp\left(-nt(1-e^{is/\sqrt{n}}) - is\sqrt{n}\,t\right) \\ &\quad = \exp\left(-nt(-i\frac{is}{\sqrt{n}} + \frac{s^2}{2n} + o(1/n)) - is\sqrt{n}\,t\right) \\ &\quad \xrightarrow{n\to\infty} \exp\left(-\frac{ts^2}{2}\right). \end{aligned}$$

The Lévy Theorem implies that $Z_n(t)$ converges in distribution to a Gaussian distribution $\mathcal{N}(0,t)$.

3. Since the Poisson process has independent increments, the components of $R_{n,K}$ are independent. Since N_n is a stationary process $N_n(t_{i+1}) - N_n(t_i)$ has the same law as $N_n(t_{i+1} - t_i)$. Hence the previous question applied to each component of $R_{n,K}$ says that

$$R_{n,K} \xrightarrow{\text{Law}} \mathcal{N}\left(0, \begin{pmatrix} t_1 & & & \\ & t_2 - t_1 & & \\ & & \ddots & \\ & & & t_K - t_{K-1} \end{pmatrix}\right)$$

4. Recall that $N_n(t) - N_n(s)$ is distributed as a Poisson random variable of parameter $n(t-s)$. For a Poisson random

variables of parameter λ,

$$\mathbf{E}[X] = \mathrm{Var}(X) = \lambda \text{ hence } \mathbf{E}\left[X^2\right] = \lambda + \lambda^2.$$

Expanding the square, it follows that

$$\begin{aligned}
&\mathbf{E}\left[|N_n(t) - N_n(s) - n(t-s)|^2\right] \\
&= \mathbf{E}\left[(N_n(t) - N_s)^2\right] + n^2(t-s)^2 - 2n(t-s)\mathbf{E}\left[N_n(t) - N_n(s)\right] \\
&= n^2(t-s)^2 + n(t-s) + n^2(t-s)^2 - 2n^2(t-s)^2 \\
&= n(t-s).
\end{aligned}$$

Hence

$$\mathbf{E}\left[|Z_n(t) - Z_n(s)|^2\right] = |t-s|$$

5. For $s \le r \le t$, by independence of the increments of the Poisson process, we get

$$\mathbf{E}\left[|Z_n(r) - Z_n(s)|^2 |Z_n(t) - Z_n(r)|^2\right] = |t-r||r-s| \le \frac{1}{4}(t-s)^2.$$

3 Stochastic integrals

3.1 Stochastic integral with respect to a point process

Before going any further, we need to clarify some notations. For N a point process whose jump times are $(T_n, n \geq 1)$. We can see N as a function with finite variation or as a sum of Dirac masses:

$$N(t) = \sum_{n \geq 1} \mathbf{1}_{\{T_n \leq t\}} \longleftrightarrow \sum_{n=1}^{\infty} \delta_{T_n}.$$

With this point of view, we can write

$$\sum_{n=1}^{\infty} f(T_n) = \int f(s) \, \mathrm{d}N(s).$$

If we consider the process

$$N_f \, : \, t \longmapsto \int_0^t f(s) \, \mathrm{d}N(s),$$

it is a process which jumps of height $f(T_n)$ at time T_n. On the one hand, it can be seen as a weighted sum of Dirac masses:

$$\sum_{n \geq 1} f(T_n) \, \delta_{T_n}$$

On the other hand, we can also consider that its trajectories are of finite variation and use the notations defined for these objects. We have then

$$\int g(s) \, \mathrm{d}N_f(s) = \sum_{n \geq 1} g(T_n) f(T_n) = \int g(s) f(s) \, \mathrm{d}N(s).$$

Let us mention a subtle but important point. We are often led to consider integrals of the form

$$\int f(s^-) \, \mathrm{d}N(s) \text{ and } \int f(s^-) \, \mathrm{d}s.$$

These two integrals differ from

$$\int f(s) \, \mathrm{d}N(s) \text{ and } \int f(s) \, \mathrm{d}s.$$

of the respective amounts

$$\int \Delta f(s) \, \mathrm{d}N(s) \text{ and } \int \Delta f(s) \, \mathrm{d}s.$$

If f is finite variation or even better is continuous, the set of points of discontinuity of f is of measure zero (since at most countable) so the function f is Lebesgue-almost certainly zero but not $N(\mathrm{d}s)$ null. This means that we can identify

$$\int f(s^-)\,\mathrm{d}s \text{ and } \int f(s)\,\mathrm{d}s \tag{3.1}$$

but not

$$\int f(s^-)\,\mathrm{d}N(s) \text{ and } \int f(s)\,\mathrm{d}N(s).$$

Note that if X is continuous on the right and adapted then $t \mapsto X(t^-)$ is predictable as a left continuous adapted process.

A stochastic process is said to have a finite variation when almost all its sample-paths have a finite variation. In other words, when we write

$$\int_0^t u(s)\,\mathrm{d}M(s) \tag{3.2}$$

for M with finite variation, we are talking of

$$t \mapsto \int_0^t u(\omega,s)\Big(\,\mathrm{d}\lambda^+_{M(\omega)}(s) - \,\mathrm{d}\lambda^-_{M(\omega)}(s)\Big).$$

For instance,

$$\begin{aligned}\int_0^t \mathbf{1}_{(a,b]}(s)\,\mathrm{d}M(s) &= \lambda^+_{M(\omega)}((a,b]) - \lambda^-_{M(\omega)}((a,b])\\ &= M(\omega,b) - M(\omega,a),\end{aligned}$$

Similarly,

$$\int_0^t u(s)|\,\mathrm{d}M(s)| = \int_0^t u(\omega,s)\Big(\,\mathrm{d}\lambda^+_{M(\omega)}(s) + \,\mathrm{d}\lambda^-_{M(\omega)}(s)\Big).$$

Definition 3.1.– We say that M has integrable variation whenever

$$\mathbf{E}\left[\int_0^1 |\,\mathrm{d}M(s)|\right] < \infty.$$

This means that the measure

$$\nu_M(\,\mathrm{d}\omega, \mathrm{d}s) := |M(\omega, \mathrm{d}s)|\mathbf{P}(\mathrm{d}\omega)$$

is finite over $\Omega \times \mathbf{R}^+$.

Theorem 3.1.– Let M be a martingale with integrable variation and u a predictable process such that :

$$\mathbf{E}\left[\int_0^1 |u(s)|\,|\,\mathrm{d}M(s)|\right] < \infty. \tag{3.3}$$

Then, the process

$$M^u(t) = \int_0^t u(s)\,\mathrm{d}M(s)$$

is a martingale on $[0,1]$.

Proof. According to the hypothesis on u and the properties of the Stieltjes integral, the integrability of $M^u(t)$ is ensured. It remains to prove (2.15). Show it first for u a simple predictable process :

$$u(\omega,\,t) = \mathbf{1}_{(a,\,b]}(t)\mathbf{1}_A(\omega)$$

with $A \in \mathcal{F}_a$. For $0 \le s \le a \le t \le b$, we have:

$$M^u(t) = \mathbf{1}_A \left(M(t) - M(a)\right) \text{ et } M^u(s) = 0.$$

Hence,

$$\begin{aligned}
\mathbf{E}\left[M^u(t) - M^u(s) \,|\, \mathcal{F}_s\right] &= \mathbf{E}\left[\mathbf{1}_A \left(M(t) - M(a)\right) | \mathcal{F}_s\right] \\
&= \mathbf{E}\left[\mathbf{E}\left[\mathbf{1}_A \left(M(t) - M(a)\right) | \mathcal{F}_a\right] | \mathcal{F}_s\right] \\
&= \mathbf{E}\left[\mathbf{1}_A \,\mathbf{E}\left[(M(t) - M(a)) | \mathcal{F}_a\right] | \mathcal{F}_s\right] \\
&= 0,
\end{aligned}$$

since M is a martingale. It is then sufficient, on the same principle, to discuss the other cases according to the relative positions of a, s, t and b.

By a monotone class argument, we can prove that M^u is a martingale for any bounded predictable processes. It remains to pass to non bounded predictable processes satisfying (3.3). □

Definition 3.2.– The space $L^1(\Omega \times [0,1] \to \mathbf{R}, \nu_M)$ is the space of processes $(u(s),\, s \ge 0)$ such that

$$\mathbf{E}\left[\int_0^1 |u(s)| \,|\,\mathrm{d}M(s)|\right] < \infty.$$

The sub-space $L^1_{\mathcal{P}}(\Omega \times [0,1] \to \mathbf{R}, \nu_M)$ of predictable processes in $L^1(\Omega \times [0,1] \to \mathbf{R}, \nu_M)$ is closed and inherits the structure of Banach space from the latter space.

Lemma 3.2.– Let $(M_n(t),\, n \ge 1)$ be a sequence of martingales, if there exists a process M such that

$$M_n(t) \xrightarrow{L^1(\Omega \to \mathbf{R}, \mathbf{P})} M(t),\ \forall t \ge 0$$

then M is a martingale.

Proof. Let $0 \le s \le t \le T$ and $A \in \mathcal{F}_s$, we must prove that

$$\mathbf{E}\left[M(t)\mathbf{1}_A\right] = \mathbf{E}\left[M(s)\mathbf{1}_A\right].$$

For any $n \ge 1$, we have

$$\mathbf{E}\left[M_n(t)\mathbf{1}_A\right] = \mathbf{E}\left[M_n(s)\mathbf{1}_A\right].$$

Since

$$\left|\mathbf{E}\left[(M(t) - M_n(t))\mathbf{1}_A\right]\right| \le \|M_n(t) - M(t)\|_{L^1},$$

we can pass to the limit and obtain the martingale property of M. □

End of the proof of Theorem 3.1. Since ν_M is a finite measure, the space $L^\infty_{\mathcal{P}}(\Omega \times [0,1] \to \mathbf{R}, \nu_M)$ is dense in $L^\infty_{\mathcal{P}}(\Omega \times [0,1] \to \mathbf{R}, \nu_M)$.

Since u belongs to $L^1(\Omega \times [0,T] \to \mathbf{R},\, \nu_M)$, for almost all ω, the function $s \mapsto u(\omega, s)$ satisfies

$$\int_0^T |u(\omega, s)| \,|\,\mathrm{d}M(s)| < \infty$$

hence the process (of bounded variation)

$$t \mapsto \int_0^t u(\omega, s) \, dM(s)$$

is well defined. Now choose $(u_n,\, n \geq 1)$ a sequence of $L^\infty(\Omega \times [0,T] \to \mathbf{R},\, \nu_M)$ which converges in $L^1(\nu_M)$ to u, according to (2.6), we have

$$\begin{aligned}\mathbf{E}\left[\left|\int_0^t u(\omega, s)\, dM(s) - \int_0^t u(\omega, s)\, dM(s)\right|\right] \\ &\leq \mathbf{E}\left[\int_0^T |u(s) - u_n(s)|\,|dM(s)|\right] \\ &= \|u - u_n\|_{L^1(\Omega\times[0,T]\to\mathbf{R},\nu_M)}.\end{aligned}$$

It follows from Lemma 3.2 that M is a martingale. $\square$

Theorem 3.3.– Let N and M be two independent point processes of respective compensators ν_N and ν_M. If ν_N is absolutely continuous, i.e.

$$\nu([0,t]) = \int_0^t h(s)\, ds$$

where h is predictable, then the jumps from N and M have an empty intersection.

Proof. Let $\Delta = \{s,\ \Delta M(s) \neq 0\}$. As M is independent of N, the indicator function of Δ behaves like a deterministic function with respect to N so

$$N_\Delta : t \longmapsto \sum_{T_q(N)\leq t} \mathbf{1}_\Delta(T_q(N))$$

admits as compensator the process

$$t \longmapsto \int_0^t \mathbf{1}_\Delta(s) h(s)\, ds = 0,$$

since Δ is at most countable and therefore of zero Lebesgue measure. Since the null process is the only process with null compensatir, this means that means that N_Δ is the null process so that N and M almost certainly have no jumps in common. $\square$

Lemma 3.4.– Assume that the process N has compensator y. For u a predictable process such that

$$\mathbf{E}\left[\int_0^\infty |1 - u(s)|\,|dy(s)|\right] < \infty,$$

the process

$$M(t) = \exp\left(\int_0^t \ln(u(s))\, dN(s) + \int_0^t (1 - u(s))\, dy(s)\right) \tag{3.4}$$

is a solution of the equation

$$M(t) = 1 + \int_0^t M(s^-)(u(s) - 1)\,(\,dN(s) - \,dy(s)). \tag{3.5}$$

The same restrictions as usual do apply: we should localize every process so that the integrals have a well defined meaning.

Proof. We write

$$m(t) = \int_0^t \ln\big(u(s)\big)\,\mathrm{d}N(s) + \int_0^t (1-u(s))\,\mathrm{d}y(s)$$

and we remark that m has the same discontinuity points as N. Formally, we apply the Itô formula to $F = \exp$ and $g = m$ in (3.4). We get

$$\begin{aligned} M(t) & \\ &= 1 + \int_0^t M(s^-)\,\mathrm{d}m(s) + \sum_{s\le t} M(s^-)\left(e^{\Delta m(s)} - 1 - \Delta m(s)\right). \end{aligned} \tag{3.6}$$

We have

$$\begin{aligned} \Delta m(s) &= \Delta\left(\int_0^t \ln\big(u(s)\big)\,\mathrm{d}N(s)\right) \\ &= \ln\big(u(s)\big)\mathbf{1}_{\{\Delta N(s)=1\}}, \end{aligned}$$

hence

$$e^{\Delta m(s)} - 1 - \Delta m(s) = \Big(u(s) - 1 - \ln\big(u(s)\big)\Big)\mathbf{1}_{\{\Delta N(s)=1\}}. \tag{3.7}$$

It follows that

$$\begin{aligned} &\int_0^t M(s^-)\,\mathrm{d}m(s) \\ &\quad = \int_0^t M(s^-)\ln\big(u(s)\big)\,\mathrm{d}N(s) + \int_0^t M(s^-)(1-u(s))\,\mathrm{d}y(s) \end{aligned} \tag{3.8}$$

and

$$\begin{aligned} &\sum_{s\le t,\ \Delta N(s)=1} M(s^-)\left(e^{\Delta m(s)} - 1 - \Delta m(s)\right) \\ &\qquad = \int_0^t M(s^-)\,(u(s) - 1 - \ln(u(s)))\,\mathrm{d}N(s). \end{aligned} \tag{3.9}$$

Plug (3.8) and (3.9) into (3.6) to obtain (3.5). □

REMARK 6.– As a consequence, M is a local martingale as it is the stochastic integral of a predictable process with respect to a local martingale.

We are now equipped to show how the compensator can yield some information on the law of the point process.

THEOREM 3.5.– Let $\lambda = (\lambda(t),\, t \ge 0)$ be a deterministic, continuous and non-decreasing function such that

$$\lambda(0) = 0,\ \forall t \ge 0,\ \lambda(t) < \infty \text{ and } \lim_{t\to\infty} \lambda(t) = \infty.$$

A point process N is a non homogeneous Poisson point of intensity λ when one the two equivalent properties is satisfied:

1. The Laplace transform is given by

$$\mathbf{E}\left[\exp\left(-\int_0^\infty f(s)\, \mathrm{d}N(s)\right)\right] = \exp\left(-\int_0^\infty (1-e^{-f(s)})\, \mathrm{d}\lambda(s)\right) \quad (3.10)$$

for any $f \geq 0$ or with compact support.

2. The process N admits λ as a compensator.

Proof. Step 1$\Longrightarrow$2. Apply (3.10) to

$$f = \sum_{i=0}^{n-1} \alpha_i \mathbf{1}_{(t_i,\, t_{i+1}]},$$

to obtain:

$$\mathbf{E}\left[\exp\left(-\sum_i \alpha_i (N(t_{i+1}) - N(t_i))\right)\right] = \exp\left(-\lambda \sum_i \int_0^\infty \left(1 - e^{-\alpha_i \mathbf{1}_{(t_i,\, t_{i+1}]}(s)}\right) ds\right).$$

Then remark that the function $s \to \left(1 - e^{-\alpha_i \mathbf{1}_{(t_i,\, t_{i+1}]}(s)}\right)$ vanishes outside $(t_i, t_{i+1}]$ and is equal to $1 - e^{-\alpha_i}$ on this interval. We then have

$$\mathbf{E}\left[\exp\left(-\sum_i \alpha_i (N(t_{i+1}) - N(t_i))\right)\right] = \exp\left(-\sum_i \left(1 - e^{-\alpha_i}\right) \int_{t_i}^{t_{i+1}} \lambda(s)\, \mathrm{d}s\right).$$

We have then obtained that the Laplace transform of the random vector $(N(t_i+1) - N(t_i),\, 1 \leq i \leq n-1)$ is the product of the Laplace tranforms of each of its components. Hence, the random variables $N(t_i+1) - N(t_i),\, 1 \leq i \leq n-1$ are independent. By monotone class theorem, we obtain that $N(t+s) - N(t)$ is independent of $\mathcal{F}_t = \sigma(N(r),\, r \leq t)$. For $n = 2$, $t_1 = a$, $t_2 = b$, we get

$$\mathbf{E}\left[\exp\left(-\alpha\left(N(b) - N(a)\right)\right)\right] = \exp\left(-(1-e^{-\alpha})\right) \int_a^b \lambda(s)\, \mathrm{d}s,$$

which means that $N(b) - N(a)$ follows a Poisson distribution of parameter $\int_a^b \mathrm{d}\lambda(s)$. Hence:

$$\mathbf{E}\left[N(t+s) - N(t) \,|\, \mathcal{F}_t\right] = \mathbf{E}\left[N(t+s) - N(t)\right] = \lambda(b) - \lambda(a)$$

and point 2 follows.

Step 2$\Longrightarrow$1.

Set $u = e^{-f}$, we see that $|1 - u(s)| \leq 1$ and since f is deterministic we have

$$\mathbf{E}\left[\int_0^t |1 - u(s)|\, \mathrm{d}|\lambda|(s)\right] \leq \int_0^t \mathrm{d}\lambda(s) \leq \lambda(t) < \infty.$$

According to Lemma 3.4, this entails that

$$\mathbf{E}\left[\exp\left(-\int_0^t f(s)\,\mathrm{d}N(s)+\int_0^t (1-e^{-f(s)})\,\mathrm{d}\lambda(s)\right)\right]=1.$$

Since λ is deterministic, we obtain

$$\mathbf{E}\left[\exp\left(-\int_0^t f(s)\,\mathrm{d}N(s)\right)\right]=\exp\left(-\int_0^t (1-e^{-f(s)})\,\mathrm{d}\lambda(s)\right).$$

It remains to use the monotone convergence theorem if $f\geq 0$ or to choose t large enough if f has compact support to see that (3.10) holds. □

3.2 *Absolute continuity*

Recall that a probability measure ν is absolutely continuous with respect to a probability measure μ of a σ-field $\mathcal{A}$ when

$$\mu(A)=0\Longrightarrow \nu(A)=0 \text{ for any } A\in\mathcal{A}.$$

The Radon-Nikodym theorem then says that there exists a unique non-negative $f\in L^1(\mu)$ such that

$$\int \psi\,\mathrm{d}\nu=\int \psi\, f\,\mathrm{d}\mu. \tag{3.11}$$

When we have processes, the question of absolute continuity is raised at any time.

Definition 3.3.– A probability measure $\mathbf{Q}$ on $(\Omega,\mathcal{F})$ is locally absolutely continuous with respect to $\mathbf{P}$ when for any $t\geq 0$, for any $A\in\mathcal{F}_t$,

$$\mathbf{Q}(A)=\mathbf{E_P}\left[\mathbf{1}_A L_t\right] \tag{3.12}$$

for some $L_t\geq 0$ which is $\mathcal{F}_t$ measurable. In other words, the restriction of $\mathbf{Q}$ to $\mathcal{F}_t$ is absolutely continuous with respect to the restriction of $\mathbf{P}$ to the same σ-field.

The probability measure $\mathbf{Q}$ is absolutely continuous with respect to $\mathbf{P}$ if (3.12) holds for any $A\in\mathcal{F}_\infty$.

In what follows, we assume that the point process N admits on $(\Omega,\mathcal{F},\mathbf{P})$ a compensator y such that there exists $\dot{y}$ predictable, non-negative such that

$$y(t)=\int_0^t \dot{y}(s)\,\mathrm{d}s.$$

This hypothesis is not necessary to state the forthcoming theorems but it simplifies greatly their expression. Moreover, in most of our applications, we consider the homogeneous Poisson process as reference process for which

$$\dot{y}(s)=\lambda,$$

where the λ is the parameter of the exponential distribution which defines the inter-arrivals (see Definition 1.1.)

THEOREM 3.6 (Girsanov Theorem).– Let $\mathbf{Q}$ be a probability measure on the filtered space $(\Omega, \mathcal{F}, \mathbf{P})$ under which the process N has a compensator denoted by z. If $\mathbf{Q}$ is absolutely continuous with respect to $\mathbf{P}$ on $\mathcal{F}_\infty$, then there exists a process, u, predictable and non negative such that

1. u satisfies the integrability condition:
$$\mathbf{Q}\left(\int_0^\infty \left(1-\sqrt{u(s)}\right)^2 \dot{y}(s)\, \mathrm{d}s < \infty\right) = 1.$$

2. The compensators are absolutely continuous:
$$\mathrm{d}z(s) = u(s)\dot{y}(s)\, \mathrm{d}s.$$
For the sake of notations, we set
$$\dot{z}(s) = u(s)\dot{y}(s).$$

If the filtration $\mathcal{F}$ is the minimal filtration $\mathcal{N}$, then these conditions are sufficient for $\mathbf{Q}$ to be absolutely continuous with respect to $\mathbf{P}$ and furthermore,
$$\left.\frac{\mathrm{d}\mathbf{Q}}{\mathrm{d}\mathbf{P}}\right|_{\mathcal{N}_t} = \exp\left(\int_0^t \ln\big(u(s)\big)\, \mathrm{d}N(s) + \int_0^t (1-u(s))\dot{y}(s)\, \mathrm{d}s\right). \quad (3.13)$$

The minimum filtration associated to a process N is the smallest filtration which makes it adapted:
$$\mathcal{N}_t = \sigma\{N(u),\, u \le t\}.$$

Proof. STEP 1. The proof is a bit abstract, we only show that
$$N(t) - \int_0^t u(s)\dot{y}(s)\, \mathrm{d}s$$
is a $\mathbf{Q}$-local martingale. Set
$$M(t) = \left.\frac{\mathrm{d}\mathbf{Q}}{\mathrm{d}\mathbf{P}}\right|_{\mathcal{N}_t}$$
which we know to be the solution of the equation
$$M(t) = 1 + \int_0^t M(s^-)(u(s)-1)\left(\,\mathrm{d}N(s) - \dot{y}(s)\, \mathrm{d}s\right).$$
We remark that the jumps of M occur at the same times as the jumps of N and according to the expression of M in (3.13),
$$\Delta M(s) = M(s^-)(u(s)-1)\mathbf{1}_{\{\Delta N(s)>0\}}. \quad (3.14)$$
Hence, according to the integration by parts formula
$$\begin{aligned} N(t)M(t) = \int_0^t N(s^-)\, \mathrm{d}M(s) &+ \int_0^t M(s^-)\, \mathrm{d}N(s) \\ &+ \sum_{s, \Delta N(s)>0} \Delta N(s)\, \Delta M(s) \\ &= A_1(t) + A_2(t) + A_3(t). \end{aligned} \quad (3.15)$$
The process A_1 is a $\mathbf{P}$-local martingale according to Theorem 3.1. For A_2, the same theorem says that
$$A_2(t) - \int_0^t M(s^-)\dot{y}(s)\, \mathrm{d}s \quad (3.16)$$

is a **P**-local martingale. As to A_3, according to (3.14), we can write

$$\begin{aligned} A_3(t) &= \int_0^t \Delta M(s)\, \mathrm{d}N(s) \\ &= \int_0^t M(s^-)(u(s)-1)\, \mathrm{d}N(s). \end{aligned}$$

Hence

$$A_3(t) - \int_0^t M(s^-)(u(s)-1)\dot{y}(s)\, \mathrm{d}s \tag{3.17}$$

is a **P**-local martingale. Plug (3.16) and (3.17) into (3.15) to obtain that

$$N(t)M(t) - \int_0^t M(s^-)u(s)\dot{y}(s)\, \mathrm{d}s \tag{3.18}$$

is a **P**-local martingale.

STEP 2. Now then, let $0 \le r < t$,

$$\begin{aligned} &\mathbf{E_Q}\left[N(t) - \int_0^t u(s)\dot{y}(s)\, \mathrm{d}s \,|\, \mathcal{F}_r\right] \\ &\qquad = \mathbf{E_P}\left[M(t)N(t) - M(t)\int_0^t u(s)\dot{y}(s)\, \mathrm{d}s \,|\, \mathcal{F}_r\right] \end{aligned} \tag{3.19}$$

It thus remains to prove that

$$\theta \,:\, t \longmapsto M(t)N(t) - M(t)\int_0^t u(s)\dot{y}(s)\, \mathrm{d}s$$

is a **P**-local martingale. Since the process

$$t \longmapsto \int_0^t u(s)\dot{y}(s)\, \mathrm{d}s$$

is continuous, the integration by parts yields

$$\begin{aligned} M(t)\int_0^t u(s)\dot{y}(s)\, \mathrm{d}s = \int_0^t u(s)\dot{y}(s)\, \mathrm{d}M(s) \\ + \int_0^t M(s^-)u(s)\dot{y}(s)\, \mathrm{d}s. \end{aligned}$$

It follows that

$$\theta(t) = M(t)N(t) - \int_0^t M(s^-)u(s)\dot{y}(s)\, \mathrm{d}s + \int_0^t u(s)\dot{y}(s)\, \mathrm{d}M(s).$$

According to (3.19) and Theorem 3.1, θ is the sum of two local martingales, hence is a local martingale. □

COROLLARY 3.7 (Martingale representation theorem).– We now assume that M is such that

$$\mathbf{E}\left[M^2\right] < \infty.$$

Then there exists a unique predictable process u_M such that

$$M = \mathbf{E}\left[M\right] + \int_0^\infty u(s)\left(\mathrm{d}N(s) - \,\mathrm{d}y(s)\right).$$

Proof. We will show a more precise theorem in the exercises. We just give a hint of the general proof. The process

$$M(t) = \mathbf{E}\left[M \,|\, \mathcal{N}_t\right]$$

is a uniformly integrable martingale. If we assume that $M > 0$ almost-surely. Define the probability measure

$$\mathrm{d}\mathbf{Q} = \frac{M}{\mathbf{E}\left[M\right]}\,\mathrm{d}\mathbf{P}.$$

From the Girsanov Theorem and (3.5), we know that there exists u such that

$$M(t) = \mathbf{E}\left[M\right]\left(1 + \int_0^t M(s^-)u(s)\left(\,\mathrm{d}N(s) - \,\mathrm{d}y(s)\right)\right)$$

hence the result with

$$u_M(s) = \mathbf{E}\left[M\right] M(s^-)u(s).$$

For general M, we apply this first result to M^+ and M^-. □

3.3 Limit theorems

Theorem 3.8.– Let N be a point process of compensator y and U predictable such that

$$\mathbf{E}\left[\int_0^T |U(s)|^2\,\mathrm{d}y(s)\right] < \infty$$

Then,

$$t \longmapsto \left(\int_0^t U(s)\left(\,\mathrm{d}N(s) - \,\mathrm{d}y(s)\right)\right)^2 - \int_0^t U(s)^2\,\mathrm{d}y(s)$$

is a martingale on $[0, T]$. In particular, we have the isometry formula: for any $t \in [0, T]$,

$$\mathbf{E}\left[\left(\int_0^t U(s)\left(\,\mathrm{d}N(s) - \,\mathrm{d}y(s)\right)\right)^2\right] = \mathbf{E}\left[\int_0^t U(s)^2\,\mathrm{d}y(s)\right]. \quad (3.20)$$

Proof. Set

$$M(t) = \int_0^t U(s)\left(\,\mathrm{d}N(s) - \,\mathrm{d}y(s)\right).$$

According to the Itô formula

$$\begin{aligned} M(t)^2 &= 2\int_0^t M(s^-)\,\mathrm{d}M(s) + \sum_{s\le t}(\Delta M(s))^2 \\ &= 2\int_0^t M(s^-)\,\mathrm{d}M(s) + \int_0^t U(s)^2\,\mathrm{d}N(s). \end{aligned}$$

Hence

$$M(t)^2 - \int_0^t U(s)^2\,\mathrm{d}y(s)$$

is the sum of two martingales hence a martingale. □

COROLLARY 3.9.– Take $U(s) = \mathbf{1}$ yields

$$(N(t) - y(t))^2 - y(t)$$

is a local martingale.

The Doob decomposition also holds for continuous time local submartingales:

THEOREM 3.10.– A local submartingale R admits a unique decomposition

$$R(t) = X(t) + A(t)$$

where X is a local martingale and A a predictable, non decreasing process, null at time 0.

When M is a local martingale, we denote $\langle M, M\rangle$, the predictable process which appears in the Doob decomposition of M^2 so that

$$M(t)^2 - \langle M, M\rangle_t$$

is a local martingale.

We then have the following limit theorems:

THEOREM 3.11.– Let M be a local martingale such that

$$\lim_{t\to\infty} \langle M, M\rangle_t = \infty.$$

Then,

$$\frac{M(t)}{\langle M, M\rangle_t} \xrightarrow[t\to\infty]{\text{a.s.}} 0.$$

THEOREM 3.12 (PASTA property).– Let N be a point process of compensator y. Let U be a predictable process for which there exists $c > 0$ such that a.s., for any $t \geq 0$:

$$\int_0^t (1 + U(s))^2 \, \mathrm{d}N(s) \leq cy(t). \tag{3.21}$$

Then,

$$\lim_{t\to\infty} \left(\frac{1}{N(t)} \int_0^t \psi(s) \, \mathrm{d}N(s) - \frac{1}{y(t)} \int_0^t \psi(s) \, \mathrm{d}s \right) = 0.$$

In particular, let N be a Poisson process of intensity λ. Then,

$$\lim_{t\to\infty} \frac{1}{N(t)} \int_0^t U(s) \, \mathrm{d}N(s) = \lim_{t\to\infty} \frac{1}{t} \int_0^t U(s) \, \mathrm{d}s,$$

as soon as one of the limit does exist.

Proof. STEP 1. Remark that Theorem 3.11 entails that

$$\frac{N(t) - y(t)}{y(t)} \xrightarrow{\text{a.s.}} 1$$

thus,

$$\lim_{t\to\infty} \frac{N(t)}{y(t)} = 1. \tag{3.22}$$

We have

$$\begin{aligned}
&\frac{1}{N(t)}\int_0^t \psi(s)\,\mathrm{d}N(s) - \frac{1}{y(t)}\int_0^t \psi(s)\,\mathrm{d}s \\
&= \frac{1}{N(t)}\int_0^t \psi(s)\,\mathrm{d}N(s)\left(\frac{y(t)-N(t)}{y(t)}\right) \\
&\qquad\qquad + \frac{1}{y(t)}\int_0^t \psi(s)\,(\,\mathrm{d}N(s) - \,\mathrm{d}y(s))\,.
\end{aligned}$$

From Theorem 3.11, we know

$$\frac{1}{\int_0^t U(s)^2\,\mathrm{d}y(s)}\int_0^t \psi(s)\,(\,\mathrm{d}N(s) - \,\mathrm{d}y(s)) \xrightarrow{\text{a.s.}} 0$$

and by hypothesis

$$\frac{1}{y(t)}\int_0^t U(s)^2\,\mathrm{d}y(s)$$

is bounded, hence

$$\frac{1}{y(t)}\int_0^t \psi(s)\,(\,\mathrm{d}N(s) - \,\mathrm{d}y(s)) \xrightarrow{\text{a.s.}} 0. \tag{3.23}$$

Let

$$y_U(t) = \int_0^t U(s)\,\mathrm{d}y(s).$$

We have to show that almost-surely,

$$\sup_{t\ge 0}\frac{1}{N(t)}\int_0^t \psi(s)\,\mathrm{d}N(s) < \infty. \tag{3.24}$$

We write

$$\begin{aligned}
&\frac{1}{N(t)}\int_0^t U(s)\,\mathrm{d}N(s) \\
&\quad = \frac{y(t)}{N(t)}\frac{y_{U^2}(t)}{y(t)}\frac{1}{y_{U^2}(t)}\int_0^t U(s)\,(\,\mathrm{d}N(s) - \,\mathrm{d}y(s)) + \frac{y_U(t)}{N(t)}.
\end{aligned}$$

The first term is the product of bounded terms by a term which converges to 0, hence it vanishes at the limit. As to the second term, we have

$$\frac{y_U(t)}{N(t)} = \frac{y(t)}{N(t)}\frac{y_U(t)}{y(t)}. \tag{3.25}$$

The first term is bounded according to (3.22). Since

$$y_U(t) = \int_0^t U(s)\,\mathrm{d}y(s) \le \int_0^t (1+U(s)^2)\,\mathrm{d}y(s),$$

the hypothesis (3.21), the second term of (3.25) is also bounded. We have thus proved (3.24) and the result follows. □

This result is often useful in queueing theory. Consider a system with one server and no buffer. Arrivals and time services are deterministic of respective duration ρ and 1.

The loss probability is the probability that a customer arrives at a time where the system is full. It is here equal to 0 if $\rho < 1$.

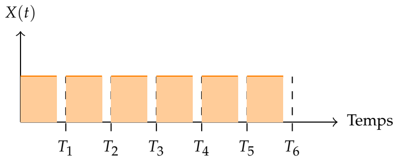


Figure 3.1: M/D/1/1 queue

The blocking probability is actually not a probability but rather the percentage of time during which the system is full. Here, it corresponds to the percentage of the surface colored in orange.

What says the PASTA formula is that if instead of deterministic arrivals, we consider a Poisson point process, the **T**ime **A**verage defined by

$$\lim_{t\to\infty} \int_0^t X(s^-)\, \mathrm{d}s$$

would be equal to

$$\frac{1}{N(t)} \int_0^t X(s^-)\, \mathrm{d}N(s),$$

i.e. the percentage of lost customers as we count one each time an arrival occurs when the system is full which is represented here by the event $(X(s) = 1)$. The meaning of PASTA is then

Poisson **A**rrivals **S**ee **T**ime **A**verage

3.4 Problems

The following exercise illustrates the use of Hilbertian methods which are very common in stochastic calculus. In passing, we precise the process that appears in the representation for martingales (see Corollary 3.7)

Let N be a non-homogeneous Poisson process of intensity $(s \mapsto \lambda(s))$ sur $[0, T]$ et $\tilde{N}$ the compensated process:

$$\tilde{N}(t) = N(t) - \int_0^t \lambda(s)\, \mathrm{d}s.$$

For $f : [0, T] \to \mathbf{R}$, we set

$$U_f(t) = \exp\left(\int_0^t f(s)\, \mathrm{d}N(s) - \int_0^t (e^{f(s)} - 1)\lambda(s)\, \mathrm{d}s \right).$$

We admit that

$$\mathcal{E} = \text{Vect}\left\{ U_f(T),\ f \text{ continuous over } [0, T] \right\}$$

is dense in $L^2(\Omega \to \mathbf{R}, \mathbf{P})$. We introduce the creation and annihilation operators

$$\varepsilon_t^+ F(N) = F(N \backslash \{t\}) + \mathbf{1}_N(t) \text{ et } \varepsilon_t^- F(N) = F(N \backslash \{t\}).$$

For any F, we define

$$D_r F(N) = \varepsilon_r^+ F(N) - \varepsilon_r^- F(N).$$

Exercise 3.4.1: Let N be a Poisson process of intensity $(s \mapsto \lambda(s))$ on $[0, T]$ and $\tilde{N}$ the compensated process:

$$\tilde{N}(t) = N(t) - \int_0^t \lambda(s)\, ds.$$

For $f : [0, T] \to \mathbf{R}$, we set

$$U_f(t) = \exp\left(\int_0^t f(s)\, dN(s) - \int_0^t (e^{f(s)} - 1)\lambda(s)\, ds\right).$$

We admit that

$$\mathcal{E} = \text{Vect}\left\{U_f(T),\ f \text{ continuous on } [0, T]\right\}$$

is dense in $L^2(\Omega \to \mathbf{R}, \mathbf{P})$. We introduce the so-called creation and annihilation operators

$$\varepsilon_t^+ F(N) = F(N \backslash \{t\}) + \mathbf{1}_N(t) \text{ et } \varepsilon_t^- F(N) = F(N \backslash \{t\}).$$

For any F, we set

$$D_r F(N) = \varepsilon_r^+ F(N) - \varepsilon_r^- F(N).$$

1. For f continuous on $[0, T]$, calculate
$$D_x\left(\int_0^t f(s)\, dN(s)\right).$$

2. Deduce that
$$U(t) = \mathbf{E}\left[U(t)\right] + \int_0^t \mathbf{E}\left[D_r U(t) \,|\, \mathcal{F}_{r^-}\right]\, d\tilde{N}(r)$$
where $\mathcal{F}_r = \sigma\{N(u),\, u \le r\}$.

We consider the map

$$\begin{aligned} \partial_N : \mathcal{E} \subset L^2(\Omega, \mathbf{P}) &\longrightarrow H = L_a^2(\Omega \times [0, T] \to \mathbf{R}, \mathbf{P} \otimes\ ds) \\ F &\longmapsto \Big(r \longmapsto \partial_N F(r) = \mathbf{E}\left[D_r F \,|\, \mathcal{F}_{r^-}\right]\Big). \end{aligned}$$

3. Show that for $F \in \mathcal{E}$,
$$\mathbf{E}\left[||\partial_N F||_{L_a^2}^2\right] \le \mathbf{E}\left[F^2\right].$$

4. Show that we can extend ∂_N as a continuous linear map on $L^2(\Omega \to \mathbf{R}, \mathbf{P})$.

We still denote by ∂_N this extension.

5. Show that for $F \in L^2(\Omega \to \mathbf{R}, \mathbf{P})$,
$$F = \mathbf{E}\left[F\right] + \int_0^T \partial_N F(s)\, d\tilde{N}(s),\ \mathbf{P}\text{-a.se.} \tag{3.26}$$

Solution on page 54

Exercise 3.4.2: We work on $(\mathfrak{N}, \mathcal{N}, \pi)$ such that the canonical process N is a Poisson point process of intensity 1. Let $(\dot{y}(\eta, t),\ t \geq 0)$ be a predictable, positive, process on $(\mathfrak{N}, \mathcal{N})$ and

$$y(N, t) = \int_0^t \dot{y}(N, s)\, \mathrm{d}s.$$

We know that there exists a unique probability measure π_y on $(\mathfrak{N}, \mathcal{N})$ such that

$$t \longmapsto N(t) - y(N, t)$$

is a local martingale. We also know that there exists a unique probability measure such that

$$t \longmapsto N(t) - t$$

is a local martingale.

1. Write down the log-likelihood of π_y with respect to π on $\mathcal{N}_t$.

2. Write down its form whenever

$$\dot{y}(\eta, t) = \alpha + \beta \int_0^{t^-} e^{-\gamma(t-s)}\, \mathrm{d}\eta(s) \tag{3.28}$$

for α, β, γ positive real numbers such that $\beta/\gamma < 1$.

Solution on page 56

3.5 *Solution to problems*

Solution of Exercise 3.4.1 on page 52:

1. If $x \in N$ then
$$\varepsilon_x^+ F(N) = \sum_{s \neq x, s \in N} f(s) + f(x) \text{ and } \varepsilon_x^- F(N) = \sum_{s \neq r, s \in N} f(s)$$
hence
$$D_x F(N) = f(x). \tag{3.31}$$
If $x \notin N$,
$$\varepsilon_x^+ F(N) = \sum_{s \in N} f(s) + f(x) \text{ and } \varepsilon_x^- F(N) = \sum_{s \in N} f(s)$$
so that (3.31) holds.

2. We have
$$\begin{aligned} D_r U_f(t) &= U_f(t) \left(\exp\left(D_r \int_0^t f(s) \, \mathrm{d}N(s) \right) - 1 \right) \\ &= U_f(t) \left(e^{f(r)} - 1 \right). \end{aligned}$$
Since f is deterministic and U_f is a martingale, for $s \leq r$,
$$\mathbf{E}\left[U_f(r) \left(e^{f(r)} - 1 \right) \mid \mathcal{F}_s \right] = U_f(s) \left(e^{f(r)} - 1 \right)$$
so that
$$\mathbf{E}\left[D_r U_f(t) \mid \mathcal{F}_r \right] = U_f(r) \left(e^{f(r)} - 1 \right).$$
The process
$$r \mapsto U_f(r) \left(e^{f(r)} - 1 \right)$$
is right-continuous and we consider its predictable projection
$$r \mapsto U_f(r^-) \left(e^{f(r)} - 1 \right)$$
which we also denote as
$$r \mapsto \mathbf{E}\left[U_f(r) \left(e^{f(r)} - 1 \right) \mid \mathcal{F}_{r^-} \right]$$
From Lemma 3.4, we have
$$U_f(t) = 1 + \int_0^t U_f(s^-)(e^{f(s)} - 1) \, \mathrm{d}\tilde{N}(s).$$
It follows that
$$U_f(t) = 1 + \int_0^t \mathbf{E}\left[D_s U_f(t) \mid \mathcal{F}_{s^-} \right] \, \mathrm{d}\tilde{N}(s).$$

3. By linearity, every $F \in \mathcal{E}$ satisfies the identity
$$F = \mathbf{E}[F] + \int_0^T \mathbf{E}[D_s F \mid \mathcal{F}_{s^-}] \, \mathrm{d}\tilde{N}(s).$$

For such F, we then have according to the isometry formula (3.20)

$$\begin{aligned}
\mathbf{E}\left[||\partial_N F||_H^2\right] &= \mathbf{E}\left[\int_0^T \left|\mathbf{E}\left[D_r U_f(t) \,|\, \mathcal{F}_r\right]\right|^2\right] \\
&= \mathbf{E}\left[\left(\int_0^T \mathbf{E}\left[D_r U_f(t) \,|\, \mathcal{F}_r\right] \, \mathrm{d}\tilde{N}(r)\right)^2\right] \\
&= \mathbf{E}\left[|F - \mathbf{E}\left[F\right]|^2\right] \\
&= \mathbf{E}\left[F^2\right] - \mathbf{E}\left[F\right]^2 \\
&\leq \mathbf{E}\left[F^2\right].
\end{aligned}$$

4. We know that $\mathcal{E}$ is dense in H, hence for any $F \in L^2$, there exists $(F_n, n \geq 1)$ a sequence of elements of $\mathcal{E}$ such that F_n converges to F in L^2 The sequence $(F_n, n \geq 1)$ is Cauchy in L^2 and we know that

$$\mathbf{E}\left[\|\partial_N F_p - \partial F_q\|_H^2\right] \leq \mathbf{E}\left[|F_p - F_q|^2\right]$$

so that the sequence $(\partial_N F_n, n \geq 1)$ is Cauchy in H and thus converges to a limit which we denote by $\partial_N F$.

It remains to prove that the limit does not depend on the choice of the approximate sequence. For, consider $(F_n, n \geq 1)$ and $(G_n, n \geq 1)$ two sequences which converge in L^2 to the same limit F. We still have

$$\begin{aligned}
\mathbf{E}\left[\|\partial_N F_n - \partial G_n\|_H^2\right] &\leq \mathbf{E}\left[\|F_n - G_n\|_{L^2}^2\right] \\
&\leq 2\left(\|F_n - F\|_{L^2}^2 + \|G_n - F\|_{L^2}^2\right).
\end{aligned}$$

Hence,

$$\lim_{n\to\infty} \partial_N F_n = \lim_{n\to\infty} \partial_N G_n$$

and the result follows.

5. Let $F \in L^2$ and $(F_n, n \geq 1)$ a sequence of elements of $\mathcal{E}$ which converge to F in L^2. By the Hölder inequality, F_n also converges to F in L^1 hence

$$\mathbf{E}\left[F_n\right] \xrightarrow{n\to\infty} \mathbf{E}\left[F\right].$$

There is a subsequence which converges to F almost-surely hence

$$F_{n_k} \xrightarrow{n\to\infty} F.$$

By the isometry formula,

$$\begin{aligned}\mathbf{E}\left[\left|\int_0^T \partial_N F_n(s)\,\mathrm{d}\tilde{N}(s) - \int_0^T \partial_N F(s)\,\mathrm{d}\tilde{N}(s)\right|^2\right] \\ &= \mathbf{E}\left[\int_0^T |\partial_N F_n(s) - \partial_N F(s)|^2\,\mathrm{d}s\right] \\ &= \mathbf{E}\left[\|\partial_N(F_n - F)\|^2_{L^2_a}\right] \\ &\le \mathbf{E}\left[\|F_n - F\|^2_{L^2}\right].\end{aligned}$$

Hence,

$$\begin{aligned}0 = F_n - \mathbf{E}[F_n] - \int_0^T \partial_N F_n(s)\,\mathrm{d}\tilde{N}(s) \\ &\xrightarrow{\text{in } L^2} F - \mathbf{E}[F] - \int_0^T \partial_N F(s)\,\mathrm{d}\tilde{N}(s)\end{aligned}$$

and (3.26) follows for $F \in L^2$.

Solution of Exercise 3.4.2 on page 53:

1. According to the Girsanov Theorem,

$$\log\left(\left.\frac{\mathrm{d}\pi_y}{\mathrm{d}\pi}\right|_{\mathcal{N}_t}\right) = \int_0^t \log\left(\dot{y}(N,s)\right)\,\mathrm{d}N(s) + t - y(N,t).$$

2. If $\dot{y}$ is given by (3.28), then

$$y(N,t) = \alpha t + \beta \int_0^t \int_0^s e^{-\gamma(s-r)}\,\mathrm{d}N(r)\,\mathrm{d}s.$$

By the Fubini-Tonnelli Theorem

$$\begin{aligned}\int_0^t \int_0^s e^{-\gamma(s-r)}\,\mathrm{d}N(r)\,\mathrm{d}s &= \int_0^t \int_r^t e^{-\gamma(s-r)}\,\mathrm{d}s\,\mathrm{d}N(r) \\ &= \int_0^t \int_0^{t-r} e^{-\gamma s}\,\mathrm{d}s\,\mathrm{d}N(r) \\ &= \frac{1}{\gamma}\int_0^t (1 - e^{-\gamma(t-r)})\,\mathrm{d}N(r) \\ &= \frac{1}{\gamma}\sum_{T_n \le t}\left(1 - e^{-\gamma(t-T_n)}\right).\end{aligned}$$

Hence,

$$y(N,t) = \alpha t + \frac{\beta}{\gamma}\sum_{T_n \le t}\left(1 - e^{-\gamma(t-T_n)}\right). \tag{3.32}$$

Hence,

$$\begin{aligned}\log L(N,t) = \sum_{T_n \le t} \log\left(\alpha + \beta \sum_{j=1}^{n-1} e^{-\gamma(T_n - T_j)}\right) \\ + (1-\alpha)t - \frac{\beta}{\gamma}\sum_{T_n \le t}\left(1 - e^{-\gamma(t-T_n)}\right).\end{aligned} \tag{3.33}$$

4
Multivariate point processes

Definition 4.1.– A multivariate point process R with values in a Polish space E is a sequence $((T_n, Z_n), n \geq 1)$ where $0 < T_n \leq T_{n+1}$ and $Z_n \in E$ for any n. It is said to be integrable whenever $\mathbf{E}\left[\sum_{n=1}^{\infty} \mathbf{1}_{[0,\,t]}(T_n)\right] < \infty$. We write

$$\sum_{n\geq 1} \psi(T_n, Z_n) = \iint_{\mathbf{R}^+\times E} \psi(s, z)\, \mathrm{d}R(s, z),$$

as soon as the left term is a converging series.

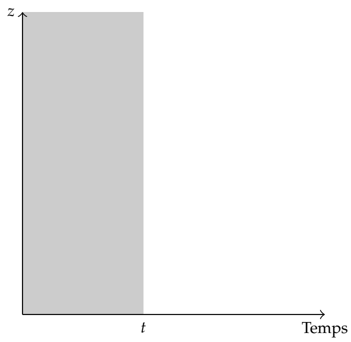
Definition 4.2.– The filtration canonically associated to a multivariate point process is defined by:

$$\mathcal{F}_t = \sigma\Big\{R([0, s] \times B),\, s \leq t,\, B \in \mathcal{B}(E)\Big\}.$$

The predictable σ-field at time t, associated to R, is generated by processes of the form:

$$\psi(\omega, s, z) = \alpha(\omega)\mathbf{1}_{]a,\,b]}(s)g(z),$$

where g is bounded and measurable from $(E, \mathcal{B}(E))$ to $(\mathbf{R}, \mathcal{B}(\mathbf{R}))$ and $\alpha \in \mathcal{F}_a$.



The random variable $U_{t,z}(R)$ is $\mathcal{F}_t$ measurable if and only if it is a function of the points of the rectangle $[0, t] \times \mathbf{R}^+$.

Definition 4.3.– A random measure on $\mathbf{R}^+ \times E$ is said to be predictable if for all $B \in \mathcal{B}(E)$, the process:

$$t \mapsto R([0, t] \times B)$$

is $\mathcal{F}$-predictable.

The following lemma is useful but its proof is far from being trivial.

Lemma 4.1.– Let R be a multivariate point process. With the same notations as above, let $\mathcal{H}_n = \sigma\{(T_j, Z_j), j = 1, \cdots, n\}$. A process X is predictable if and only if there exists a sequence of processes $(Y_n, n \geq 1)$ with Y_n $\mathcal{H}_n$-measurable and

$$X(t) = X(0) + \sum_{n=0}^{\infty} Y_n(s)\, \mathbf{1}_{(T_n, T_{n+1}]}(s)$$

where $X(0)$ is deterministic.

DEFINITION 4.4.– Let R be a multivariate point process. We denote by $\mathbf{Q}_n$ the conditional law of (T_{n+1}, Z_{n+1}) given $\mathcal{H}_n$. We define the following random measure

$$\mathrm{d}\nu(s, z) = \sum_{n\geq 0} \frac{1}{\mathbf{Q}_n([s, \infty[\times E)} \mathbf{1}_{(T_n, T_{n+1}]}(s)\ \mathrm{d}\mathbf{Q}_n(s, z). \tag{4.1}$$

THEOREM 4.2.– Let R be a multivariate point process. For any predicatble process ψ such that:

$$\sup_t \mathbf{E}\left[\int_0^t \int_E |\psi(s, z)|\ \mathrm{d}\nu(s, z)\right] < \infty, \tag{4.2}$$

the process :

$$M^\psi\ :\ t \longmapsto \int_0^t \int_E \psi(s, z)\ \mathrm{d}R(s, z) - \int_0^t \int_E \psi(s, z)\ \mathrm{d}\nu(s, z)$$

is a local martingale.

The measure ν is the intensity (measure) or dual predictable projection of R.

We will retain from this theorem the existence of the intensity measure. Its calculation is rarely feasible except in very special cases such as Poisson processes where by virtue of the theorem 3.5, we already know that $\mathrm{d}\nu(s) = \lambda(s)\ \mathrm{d}s$.

The proof shows that it is sufficient to find the measure to be removed from $R([0, t] \times B)$ to obtain a martingale, for all $B \subset E$.

The nice thing about the dual predictable projection is that it fully characterises the law of a multivariate point process.

THEOREM 4.3.– Let R and S be two multivariate point processes adapted to the minimal filtration on $\mathbb{D}([0, T], \mathbf{R})$. If $\nu_R = \nu_S$ then the two processes have the same law.

Proof of Theorem 4.2. We will not prove the uniqueness of the predictable dual projection which is the official name of ν, only that the given expression of ν satisfies the requirements.

According to Lemma 4.1, the process:

$$t \mapsto \int_0^t \int_E \psi(s, z)\ \mathrm{d}\nu(s, z)$$

is predictable. We now show that for ψ non-negative and predictable, we have $\mathbf{E}\left[M^\psi(t)\right] = 0$. The law of T_{n+1} given $\mathcal{H}_n$ is, by definition, the marginal distribution of $\mathbf{Q}_n$ on $\mathbf{R}^+$ hence

$$\mathbf{E}\left[\mathbf{1}_{[s, \infty)}(T_{n+1}) \,|\, \mathcal{H}_n\right] = \int_s^\infty \int_E \mathrm{d}\mathbf{Q}_n(r, \tau) = \mathbf{Q}_n([s, \infty[\times E).$$

By the very definition of ν,

$$\begin{aligned}\mathbf{E}&\left[\int_0^\infty \int_E \psi(s, z)\ \mathrm{d}\nu(s,z)\right]\\ &= \mathbf{E}\left[\sum_{n\geq 0} \int_0^\infty \int_E \psi(s, z)\ \mathbf{1}_{\{T_n < s \leq T_{n+1}\}} \frac{\mathrm{d}\mathbf{Q}_n(s, z)}{\mathbf{Q}_n([s, \infty[\times E)}\right].\end{aligned} \tag{4.3}$$

On the interval $(T_n, T_{n+1}]$, $\psi(s,z)$ is $\mathcal{H}_n$-measurable, as well as $\mathbf{Q}_n$. It follows that

$$
\begin{aligned}
&\mathbf{E}\left[\int_0^\infty \int_E \psi(s,\,z)\;\mathrm{d}\nu(s,z)\right] \\
&= \sum_{n\geq 0}\mathbf{E}\left[\int_0^\infty \int_E \psi(s,\,z)\;\mathbf{1}_{\{T_n<s\}}\mathbf{E}\left[\mathbf{1}_{\{T_{n+1}\geq s\}}\,|\,\mathcal{H}_n\right]\frac{\mathrm{d}\mathbf{Q}_n(s,\,z)}{\mathbf{Q}_n([s,\,\infty[\times E)}\right] \\
&= \sum_{n\geq 0}\mathbf{E}\left[\int_0^\infty \int_E \psi(s,\,z)\;\mathbf{1}_{\{T_n<s\}}\mathbf{Q}_n([s,\,\infty[\times E)\frac{\mathrm{d}\mathbf{Q}_n(s,\,z)}{\mathbf{Q}_n([s,\,\infty[\times E)}\right] \\
&= \sum_{n\geq 0}\mathbf{E}\left[\int_0^\infty \int_E \psi(s,\,z)\;\mathbf{1}_{\{T_n<s\}}\;\mathrm{d}\mathbf{Q}_n(s,\,z)\right] \\
&= \mathbf{E}\left[\sum_{n\geq 0}\psi(T_{n+1},Z_{n+1})\right] \\
&= \int_0^\infty \int_E \psi(t,z)\;\mathrm{d}R(t,z).
\end{aligned}
$$

Let ψ be a predictable non-negative process and a random variable Y which is $\mathcal{F}_r$ measurable. For $t > r > 0$, the process

$$s \mapsto \psi(s,\,z)Y\,\mathbf{1}_{(r,\,t]}(s)$$

is still predictable. Hence, we have:

$$
\begin{aligned}
&\mathbf{E}\left[M^\psi(t) - M^\psi(r)\,|\,\mathcal{F}_r\right] \\
&\qquad = \mathbf{E}\left[\int_0^t \mathbf{1}_{(r,\,t]}(s)\int_E \psi(s,\,z)(\;\mathrm{d}R(s,\,z) - \;\mathrm{d}\nu(s,\,z))\,|\,\mathcal{F}_r\right]. \qquad (4.4)
\end{aligned}
$$

For $Y \in \mathcal{F}_r$, according to the first part of the proof, we have:

$$\mathbf{E}\left[\int_0^t \int_E Y\,\mathbf{1}_{(r,\,t]}(s)\psi(s,\,z)(\;\mathrm{d}R(s,\,z) - \;\mathrm{d}\nu(s,\,z))\right] = 0,$$

thus $\mathbf{E}\left[M^\psi(t) - M^\psi(r)\,|\,\mathcal{F}_r\right] = 0$ and M^ψ is a (local) martingale. □

All these calculations make sense only if the expectations are finite, which we know nothing about *a priori* for any predictable ψ. In order to make sense of these calculations, we consider the following:

$$\tau_n = \inf\Big\{t,$$
$$\int_0^t \int_E \psi(s,\,z)(\;\mathrm{d}R(s,\,z) + \;\mathrm{d}\nu(s,\,z)) > n\Big\},$$

and we apply the reasoning of (4.4) on $[0,\,t\wedge\tau_k]$ instead of $[0,\,t]$. All the expectations are finite and the calculation is then perfectly rigorous.

REMARK 7 (A word about the vocabulary).– For multivariate point processes, ν is the intensity measure. If we can write

$$\mathrm{d}\nu(s,z) = \rho(s,z)\;\mathrm{d}s \otimes \;\mathrm{d}\mu(z)$$

where μ is a deterministic measure, then ρ is called the intensity. For the point process $t \mapsto R([0,t] \times B)$, the process

$$t \mapsto \nu([0,t] \times B)$$

is its compensator.

DEFINITION 4.5.– A marked Poisson process is a Poisson point process of intensity λ where each point is enriched by a *mark*. The mark of a point located at s is independent of the other marks and has a law, denoted $\rho(s,\;\mathrm{d}z)$ which may depend on the position of the point to which it is associated.

THEOREM 4.4.– A marked Poisson process is a multivariate point process whose compensator is given by

$$\lambda \, \mathrm{d}s \otimes \rho(s, \, \mathrm{d}z).$$

Proof. The law of (T_{n+1}, Z_{n+1}) given $\mathcal{H}_n$ is that of the pair of independent random variables $(T_n + \mathcal{E}(\lambda), Z)$. Since

$$\mathbf{P}(T_n + \mathcal{E}(\lambda) \geq s \,|\, T_n) = e^{-\lambda(s - T_n)},$$

we have

$$\mathrm{d}\mathbf{Q}_n(s,z) = \lambda e^{-\lambda(s-T_n)} \mathbf{1}_{]T_n, +\infty[}(s) \, \mathrm{d}s \otimes \rho(s, \, \mathrm{d}z).$$

Hence,

$$\mathbf{Q}_n([s, +\infty[\times E) = e^{-\lambda(s-T_n)}$$

and

$$\begin{aligned}
\mathrm{d}\nu(s,z) &= \sum_{n \geq 0} \frac{1}{\mathbf{Q}_n([s, \, \infty[\times E)} \mathbf{1}_{(T_n, \, T_{n+1}]}(s) \, \mathrm{d}\mathbf{Q}_n(s, \, z) \\
&= \sum_{n \geq 0} \lambda \, \mathbf{1}_{(T_n, \, T_{n+1}]}(s) \, \mathrm{d}s \otimes \rho(s, \, \mathrm{d}z) \\
&= \lambda \, \mathrm{d}s \otimes \, \mathrm{d}\rho(z).
\end{aligned}$$

The proof is thus complete. □

LEMMA 4.5.– Let R be a marked Poisson process on E with intensity $\lambda \, \mathrm{d}s \otimes \nu(\, \mathrm{d}z)$ where ν is a probability measure on E. For any $B \in \mathcal{B}(\mathbf{R}^+) \otimes \mathcal{B}(E)$, $R(B)$ follows a Poisson distribution of parameter

$$\int_B \lambda \, \mathrm{d}s \otimes \, \mathrm{d}\nu(z). \tag{4.5}$$

Proof. By monotone class theorem, it is sufficient to prove (4.5) for $B = [0, t] \times A$ with $A \in \mathcal{B}(E)$. The process

$$X(t) = \int_0^t \int_A \mathrm{d}N(s,z) - \lambda t \nu(A)$$

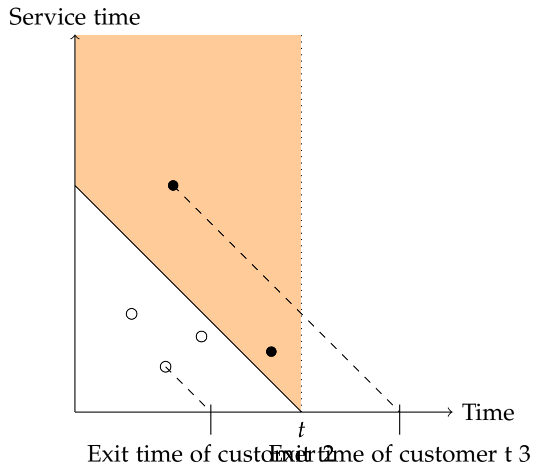
is a local martingale.This means that

$$t \longmapsto \int_0^t \int_A \mathrm{d}N(s,z)$$

is a Poisson process of intensity $\lambda\nu(A)$. Hence the result. □

EXAMPLE 4.1.– The M/M/∞ queue is the queue with Poisson arrivals, independent and identically distributed from exponential distribution service times, and an infinite number of servers (without buffer). It is initially a theoretical object which is particularly simple to analyse and also a model to which we can compare other situations. Due to the independence of the inter-arrivals and service time, according to the second characterisation of Poisson processes, the process:

$$N = \sum_{n \geq 1} \delta_{(T_n, S_n)}$$



where T_n is the instant of nth arrival and S_n the nth service time, is a Poisson process with $d\mu(t,x) = \lambda\, \mathrm{d}t \otimes \mu e^{-\mu x}\, \mathrm{d}x$ intensity in $E = \mathbf{R}^+ \times \mathbf{R}^+$.

The customers who are still in service at the time are those who correspond to the points in the orange trapezium: those for which $S_n \geq t - T_n$.

We deduce that $X(t)$, the number of busy servers at time t follows a Poisson distribution with parameter

$$\int_0^t \left(\int_{t-s}^{\infty} \mu e^{-\mu x}\, \mathrm{d}x \right) \lambda\, \mathrm{d}s = \lambda \int_0^t e^{-\mu(t-s)}\, \mathrm{d}s = \rho(1 - e^{-\mu t}),$$

where $\rho = \lambda/\mu$. If the system is not empty at time 0, we must add $X(t)$ the number of initial customers still in service at time t. If X_0 follows a Poisson distribution with parameter ρ_0, the number of customers in service at time t follows a Poisson distribution with parameter $\rho_0 e^{-\mu t}$ because each and every customer has a probability $e^{-\mu t}$ of being still in service and the total is thus the thinning of a Poisson random variable. In conclusion, $X(t)$ then follows a Poisson distribution with parameter $\rho + (\rho_0 - \rho)e^{-\mu t}$. Irrespective of the value of ρ_0, the stationary probability of X is a Poisson distribution with parameter ρ.

THEOREM 4.6.– Soit R un processus ponctuel multivarié et ψ prévisible tel que

$$\sup_{t \geq 0} \mathbf{E} \left[\int_0^t \int_E \psi^2(s, z)\, \mathrm{d}\nu(s, z) \right] < \infty. \tag{4.6}$$

La martingale locale M^ψ admet comme crochet

$$\langle M^\psi, M^\psi \rangle_t = \int_0^t \int_E \psi^2(s, z)\, \mathrm{d}\nu(s, z).$$

Proof. On note $((T_n, Z_n),\, n \geq 1)$ les sauts de R. D'après le théorème 2.7, on sait que:

$$M^\psi(t)^2 = 2 \int_0^t M^\psi(s^-)\, \mathrm{d}M^\psi(s) + \sum_{s \leq t} (\Delta M^\psi(s))^2 \tag{4.7}$$

Comme

$$\sum_{s\le t}(\Delta M^{\psi}(s))^2 = \sum_{T_n\le t}\psi(T_n, Z_n)^2 = \int_0^t\int_E \psi(s,z)^2\, \mathrm{d}R(s,z).$$

D'après le théorème 4.2,

$$t \longmapsto \int_0^t\int_E \psi(s,z)^2\, \mathrm{d}R(s,z) - \int_0^t\int_E \psi(s,z)^2\, \mathrm{d}\nu(s,z)$$

est une martingale locale. Donc en ôtant la même quantité à $M^{\psi}(t)^2$, (4.7) implique que

$$t \longmapsto M^{\psi}(t)^2 - \int_0^t\int_E \psi(s,z)^2\, \mathrm{d}\nu(s,z)$$

est une martingale locale. □

On tire de l'exemple 2.9 que le compensateur du processus de Poisson d'intensité λ est le processus $A(t) = \lambda([0, t])$.

Dans le cas d'un processus ponctuel, i.e. d'un processus multivarié avec des marques à valeurs dans un singleton, on peut omettre la variable z dans tout ce qui précède donc le compensateur devient juste une mesure aléatoire sur $\mathbf{R}^+$ telle que

$$N(t) - \int_0^t \mathrm{d}\nu(s)$$

soit une martingale locale. Par abus de notation, on appelle aussi

$$A(t) = \int_0^t \mathrm{d}\nu(s)$$

le compensateur de N. Le théorème 4.6 indique alors que

$$(N - A)(t)^2 - A(t)$$

est une martingale, i.e.

$$\langle N - A,\, N - A\rangle_t = A(t).$$

THEOREM 4.7.– Soit M une martingale et $\langle M, M\rangle$ son crochet. Si $\langle M, M\rangle(t)$ tend vers l'infini quand t tend vers l'infini alors:

$$\frac{M(t)}{\langle M, M\rangle(t)} \xrightarrow{t\to\infty} 0.$$

EXAMPLE 4.2.– Si N est un processus de Poisson d'intensité λ. On sait que

$$M(t) = N(t) - \lambda t$$

est une martingale locale de crochet λt. Par conséquent,

$$\frac{N(t)}{t} \xrightarrow{t\to\infty} \lambda.$$

COROLLARY 4.8.– Soit R un processus ponctuel multivarié sur E polonais de compensateur ν. On note N le processus ponctuel associé : $N(t) = R([0, t] \times E)$. Soit $\psi\, :\, \Omega \times \mathbf{R}^+ \times E \to \mathbf{R}$ un processus prévisible. Supposons qu'il existe $c > 0$ tel que presque sûrement, pour tout

$t \geq 0$, on ait:

$$\int_0^t \int_E (1 + \psi^2(s, z)) \, \mathrm{d}\nu(s, z) \leq c \, \nu([0, t] \times E), \tag{4.8}$$

alors, presque sûrement, on a:

$$\lim_{t\to\infty} \Big(\frac{1}{N(t)} \int_0^t \psi(s, z) \, \mathrm{d}R(s, z) - \frac{1}{\nu([0, t] \times E)} \int_0^t \psi(s, z) \, \mathrm{d}\nu(s, z) \Big) = 0. \tag{4.9}$$

Proof. Pour simplifier les notations, on pose:

$$\nu(t) = \nu([0, t] \times E) \text{ et } \nu^{\phi}(t) = \int_0^t \phi(s, z) \, \mathrm{d}\nu(s, z).$$

Remarquons d'abord que d'après le théorème théorème 4.7, on sait que:

$$\frac{1}{\nu(t)}(N(t) - \nu(t)) \xrightarrow{t\to\infty} 0. \tag{4.10}$$

Par conséquent:

$$\frac{N(t)}{\nu(t)} \xrightarrow{t\to\infty} 1. \tag{4.11}$$

On en déduit que:

$$\frac{\nu^{\psi}(t)}{N(t)} = \frac{\nu(t)}{N(t)} \frac{\nu^{\psi}(t)}{\nu(t)} \leq \frac{\nu(t)}{N(t)} \frac{1}{\nu(t)} \int_0^t \int_E (1 + \psi^2)(s, z) \, \mathrm{d}\nu(s, z).$$

D'après (4.8) et (4.11), cette quantité est donc bornée uniformément par rapport au temps. En écrivant:

$$\frac{1}{N(t)} \int_0^t \int_E \psi(s, z) \, \mathrm{d}R(s, z) = \frac{\nu(t)}{N(t)} \frac{\nu^{\psi^2}(t)}{\nu(t)} \times \frac{1}{\nu^{\psi^2}(t)} \int_0^t \int_E \psi(s, z) (\, \mathrm{d}R(s, z) - \, \mathrm{d}\nu(s, z)) + \frac{\nu^{\psi}(t)}{N(t)},$$

on déduit de ce qui précède qu'il existe $r > 0$ tel que:

$$\limsup_{t\to\infty} \frac{1}{N(t)} \int_0^t \int_E \psi(s, z) \, \mathrm{d}R(s, z) \leq r. \tag{4.12}$$

Munis de ces résultats de domination, nous pouvons maintenant procéder au calcul de la limite qui nous intéresse vraiment:

$$\frac{1}{N(t)} \int_0^t \psi \, \mathrm{d}R - \frac{1}{\nu(t)} \int_0^t \psi \, \mathrm{d}\nu$$
$$= \frac{1}{N(t)} (\int_0^t \int_E \psi \, \mathrm{d}R) \left(\frac{\nu(t) - N(t)}{\nu(t)} \right) + \frac{\nu^{\psi^2}(t)}{\nu(t)} \frac{1}{\nu^{\psi^2}(t)} \int \psi(\, \mathrm{d}R - \, \mathrm{d}\nu).$$

On déduit (4.9) du théorème 4.7 et de (4.10), (4.11) et (4.12). □

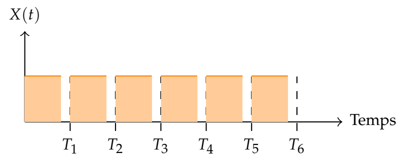
Le théorème suivant est une application directe du théorème 4.2. Il stipule que les moyennes calculées du point de vue des clients sont égales aux moyennes temporelles lorsque le processus d'arrivée est un processus de Poisson.

THEOREM 4.9 (Propriété PASTA).– Si N est un processus de Poisson d'intensité λ alors:

$$\lim_{t\to\infty} \frac{1}{N(t)} \sum_{n:\, T_n \le t} \psi(T_n) = \lim_{t\to\infty} \frac{1}{t} \int_0^t \psi(s)\, \mathrm{d}s,$$

dès que l'une des deux limites existe.

Considérons une situation avec un seul serveur, des arrivées et des temps de service déterministes, de durée respectives $\rho < 1$ et 1.



D'après les définitions, la probabilité de blocage est égale au pourcentage d'aire grisée, soit ρ. La probabilité de perte est nulle puisqu'aucun client n'arrive avec le serveur occupé. En revanche, si les arrivées sont poissonniennes, les probabilités de perte et de blocage coïncident en vertu de la propriété PASTA.

4.1 *Problems*

Exercise 4.1.1: Let X be the number of busy servers in the M/M/∞ queue.

1. Write the law of the difference $X(t+s) - X(t)$ as the difference of two independent Poisson random variables.
2. How is defined $D(t)$ the point process which counts the departures up to time t? What is the law of $D(t)$?
3. Do we have $D(t)$ and $X(t)$ independent?

Solution on page 65

Exercise 4.1.2: Let N be the Poisson marked process of intensity

$$\nu(\,\mathrm{d}s,\, \mathrm{d}z) = \lambda\, \mathrm{d}s \times (p_1\delta_1(z) + p_2\delta_2(z))$$

where $p_1 + p_2 = 1$.

1. What is the nature of the processes

$$N^i\,:\, t \longmapsto \int_0^t \int \mathbf{1}_{\{i\}}(z)\, \mathrm{d}N(s,z)?$$

Are they independent?

We admit that for any $x \in \mathbf{N}$, there exists a unique solution to the equation

$$X^x(t) = x + N^1(t) - \int_0^t \mathbf{1}_{\{X^x(s^-)>0\}}\, \mathrm{d}N^2(s). \tag{4.13}$$

2. Describe the jumps of X.

Fix $x \in \mathbf{R}$ and $\mathbf{E}_x$ denote the expectation given $X(0) = x$.

3. Show that

$$\mathbf{E}_x\left[f(X(t))\right] = f(x) + \mathbf{E}_x\left[\int_0^t (Lf)(X(s))\, \mathrm{d}s\right]$$

where

$$Lf(y) = \lambda\Big(f(y+1) - f(y)\Big) + \mu\Big(f(y-1) - f(y)\Big).$$

4. What is the square bracket of the semi-martingale $(f(X(t)),\, t \ge 0)$?

Solution on page 65

4.2 Solution to problems

Solution of Exercise 4.1.1 on page 64:

1 From the figure, we see that $X(t+s) - X(t)$ is the difference of Y_1 and Y_2 where Y_1 is the number of points of N in the domain

$$\{(r,z), t \le r \le t+s, z \ge t+s-r\}$$

whose measure is

$$\begin{aligned}\lambda \int_t^{t+s} \int_{t+s-r} \mu e^{-\mu z} \, dz \, dr &= e^{-\mu(t+s)} \int_t^{t+s} e^{\mu r} \, dr \\ &= \rho e^{-\mu(t+s)} (e^{\mu(t+s)} - e^{\mu t}) \\ &= \rho(1 - e^{-\mu s}).\end{aligned}$$

Similarly, Y_2 is a Poisson random variable of parameter

$$\begin{aligned}\lambda \int_0^t \int_{t-r}^{t+s-r} \mu e^{-\mu z} \, dz \, dr &= \int_0^t e^{-\mu(t+s-r)} - e^{-\mu(t-r)} \lambda \, dr \\ &= \frac{\lambda}{\mu} \left(e^{-\mu(t+s)} (e^{\mu t} - 1) - e^{-\mu t}(e^{\mu t} - 1) \right) \\ &= \rho \left(e^{-\mu s} - e^{\mu(t+s)} + e^{-\mu t} - 1) \right)\end{aligned}$$

2 We have

$$D(t) = \int_0^t \int_0^{t-s} dN(s,z).$$

By the same principle as above, $D(t)$ follows a Poisson law of parameter

$$\begin{aligned}\int_0^t \int_0^{t-s} \mu e^{-\mu z} \, dz \lambda \, ds &= \int_0^t 1 - e^{-\mu(t-s)} \lambda \, ds \\ &= \lambda t - \rho e^{-\mu t} \int_0^t \mu e^{\mu s} \lambda \, ds \\ &= \lambda t - \rho(1 - e^{-\mu t}).\end{aligned}$$

3 The two domains of integration which give the law of $D(t)$ and $X(t)$ are disjoint, hence according to the properties Poisson processes, $X(t)$ and $D(t)$ are independent.

Solution of Exercise 4.1.2 on page 64:

1 The process N^i has compensator

$$t \mapsto \int_0^t \int \mathbf{1}_{\{i\}}(z) \lambda \, ds \otimes (p_1 \delta_1(z) + p_2 \delta_2(z)) = \lambda p_i t$$

hence it is a Poisson process of intensity λp_i.

2 At each time t, we know that the next arrival occurs in a time which is distributed as an exponential random variable

of parameter λ. If $X(t) > 0$, the part which depends on N^2 is annihilated hence there cannot be a departure in the system.

If $X(t) > 0$, the stochastic integral with respect to N^2 is *active*. Since before a jump of this integral, X can only increase, X remains positive until the next jump of N^2, which occurs in a time with an exponential distribution of parameter μ.

Hence, the next jump occurs in the minimum of these two independent exponential random variables. Its law is an exponential random variable of parameter $\lambda + \mu$. The probability that the exponential of parameter λ is less than the exponential of parameter μ is $\lambda/(\lambda+\mu)$.

3 Apply the Itô formula to (4.13) to obtain

$$
\begin{aligned}
f(X(t)) &= f(x) + \int_0^t f'(X(s^-))\, \mathrm{d}N^1(s) - \int_0^t f'(X(s^-))\mathbf{1}_{X(s^-)>0}\, \mathrm{d}N^2(s) \\
&+ \sum_{s:\Delta N^1(s)>0} f(X(s^-)+1) - f(X(s^-)) - f'(X(s^-)) \\
&+ \sum_{s:\mathbf{1}_{X(s^-)>0}\Delta N^2(s)>0} f(X(s^-)-1) - f(X(s^-)) - f'(X(s^-))(-1).
\end{aligned}
$$

The second and third term can be written as stochastic integral with respect to N^1 and N^2 as

$$
\begin{aligned}
f(X(t)) &= f(x) \\
&+ \int_0^t f'(X(s^-)) + \big(f(X(s^-)+1) - f(X(s^-)) - f'(X(s^-))\big)\, \mathrm{d}N^1(s) \\
&- \int_0^t \big[f'(X(s^-)) + \big(f(X(s^-)-1) - f(X(s^-)) - f'(X(s^-))\big)\big]\, \mathbf{1}_{X(s^-)>0}\, \mathrm{d}N^2(s).
\end{aligned}
$$

After simplification, we get

$$
\begin{aligned}
f(X(t)) = f(x) &+ \int_0^t \big[f(X(s^-)+1) - f(X(s^-))\big]\, \mathrm{d}N^1(s) \\
&- \int_0^t \big[f(X(s^-)-1) - f(X(s^-))\big]\, \mathbf{1}_{X(s^-)>0}\, \mathrm{d}N^2(s).
\end{aligned}
$$

Let

$$\tilde{N}^1(t) = N^1(t) - \lambda t,\ \tilde{N}^2(t) = N^2(t) - \mu t$$

which are local martingales. We can write

$$
\begin{aligned}
f(X(t)) = f(x) &+ \int_0^t \big[f(X(s^- + 1)) - f(X(s^-))\big]\, \mathrm{d}\tilde{N}^1(s) \\
&- \int_0^t \big[f(X(s^-)-1) - f(X(s^-))\big]\, \mathbf{1}_{X(s^-)>0}\, \mathrm{d}\tilde{N}^2(s) \\
&+ \lambda \int_0^t \big[f(X(s^-)+1) - f(X(s^-))\big]\, \mathrm{d}s \\
&- \mu \int_0^t \big[f(X(s^-)-1) - f(X(s^-))\big]\, \mathbf{1}_{X(s^-)>0}\, \mathrm{d}s. \quad (4.15)
\end{aligned}
$$

We can replace s^- by s everywhere in the last two integrals as the number of discontinuity points of X is denumerable at most, hence Lebesgue negligeable. Furthermore, the integrands of the first two compensated integrals are predictable hence these first two terms are (local) martingales and we get

$$\mathbf{E}_x\left[f(X(t))\right] = f(x) + \mathbf{E}_x\left[\int_0^t (Lf)(X(s))\,\mathrm{d}s\right].$$

4 Start from (4.15) to obtain that the square bracket of the semi-martingale $f(X(t))$ is the square bracket of

$$\begin{aligned} t \longmapsto & \int_0^t \left[f(X(s^-+1)) - f(X(s^-))\right]\,\mathrm{d}\tilde{N}^1(s) \\ & - \int_0^t \left[f(X(s^-)-1) - f(X(s^-))\right] \mathbf{1}_{X(s^-)>0}\,\mathrm{d}\tilde{N}^2(s). \end{aligned}$$

According to Theorem 4.6, we get that this is equal to

$$\begin{aligned} t \longmapsto & \int_0^t \left[f(X(s^-+1)) - f(X(s^-))\right]^2 \lambda\,\mathrm{d}s \\ & + \mu \int_0^t \left[f(X(s^-)-1) - f(X(s^-))\right]^2 \mathbf{1}_{X(s^-)>0} \mu\,\mathrm{d}s. \end{aligned}$$

5
Spatial Poisson Process

In this chapter, we assume that E is a metric space for a distance d which is complete (every Cauchy sequence is convergent) and separable (there exists a dense sequence). This ensures that random variables do behave in some sense as ordinary (i.e. real valued) random variables.

A configuration can be seen as a purely atomic measure or as a set

$$(x_n,\, n\ge 1) \simeq \sum_{n\ge 1} \varepsilon_{x_n}$$

Definition 5.1.– A configuration is a set of points of E which is locally finite : there is a finite number of points in any compact set. The set of configurations of E is denoted $\mathfrak{N}_E$.

Definition 5.2.– We define a topology on $\mathfrak{N}_E$ by defining convergent sequences. A sequence $(\xi_n,\, n\ge 1)$ of configurations towards the configuration ξ if for any continuous function with compact support from E into $\mathbf{R}$,

$$\sum_{x\in\xi_n} f(x) \xrightarrow{n\to\infty} \sum_{x\in\xi} f(x).$$

This means that in any compact window, the points of ξ_n converge to those of ξ but we don't know what happens at infinity. For example, the sequence of configurations $(\varepsilon_n,\, n\ge 1)$ tends to zero on $\mathbf{R}$.

Definition 5.3.– A locally finite point process is a random variable N with values in $\mathfrak{N}_E$:

$$\begin{aligned} N,:\ (\Omega,\mathcal{A},\mathbf{P}) &\longrightarrow (\mathfrak{N}_E,\mathcal{B}(\mathfrak{N}_E))\\ \omega &\longmapsto N(\omega) = \text{locally finite set of points of E.}\end{aligned}$$

A point process N is said to be finite when

$$\mathbf{P}\left(N(E)<\infty\right)=1.$$

We note

$$\int_E f\,\mathrm{d}N = \sum_{x\in N} f(x)$$

for $f\ge 0$ or f such that the term on the right is almost certainly finite.

According to the construction of the topology on $\mathfrak{N}_E$, the law of a point process N is characterized by the laws of the variables random variables $(\int f\,\mathrm{d}N,\, f\in\mathcal{C}_K(E;\mathbf{R}))$.

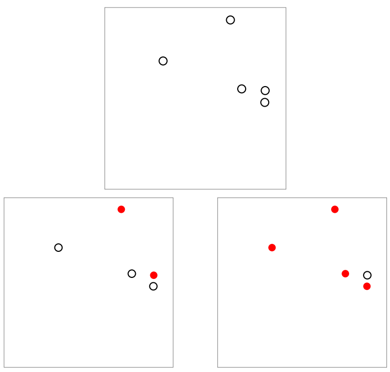


Figure 5.1: Top, the set E. Bottom, two possible configurations. The selected points are filled.

EXAMPLE 5.1 (Bernoulli process).– The Bernoulli point process is a process where E is a finite set: $E = \{x_1, \cdots, x_n\}$. Each of these points is selected independently of the others and with probability p. If we introduce $A_1, \cdots, A_n$ as independent random variables with Bernoulli distribution with parameter p, we can write :

$$N = \sum_{i=1}^{n} A_i \delta_{x_i}.$$

EXAMPLE 5.2 (Binomial point process).– The number of points is fixed at n and we are given μ, a probability measure on $\mathbf{R}^2$. The atoms are drawn at random, independently of each other, according to μ.

We can easily calculate that :

$$\mathbf{P}(N(A) = k) = \binom{n}{k} \mu(A)^k (1 - \mu(A))^{n-k}$$

and for disjoint sets $A_1, \cdots, A_n$, we have:

$$\mathbf{P}(N(A_1) = k_1, \cdots, N(A_n) = k_n) = \frac{(k_1 + \ldots + k_n)!}{k_1! \ldots k_n!} \mu(A_1)^{k_1} \ldots \mu(A_n)^{k_n}. \quad (5.1)$$

The mathematically richest point process is the spatial Poisson process. Poisson process, which is recognized as a generalization of the introduced in Chapter 1.

DEFINITION 5.4.– Let E be a Polish space and μ a finite measure on E. We construct a point process N on E as follows. Let M be a Poisson random variable with parameter $\mu(E)$. We pose

$$N = \sum_{i=1}^{M} \varepsilon_{X_i}$$

where the random variables $(X_n, n \geq 1)$ are independent and have the same distribution given by $\mu(E)^{-1}\mu$. We call N, the Poisson process of control measure μ.

THEOREM 5.1.– For any function $f : E \to \mathbf{R}^+$, we have:

$$\mathbf{E}\left[\exp(-\int_E f \, dN)\right] = \exp\left(-\int_E (1 - e^{-f(s)}) \, d\mu(s)\right). \quad (5.2)$$

Proof. By definition, we have

$$\begin{aligned}\mathbf{E}\left[\exp(-\int_E f \, dN)\right] &= \mathbf{E}\left[\exp(-\sum_{i=1}^{M} f(X_i))\right] \\ &= \sum_{m=0}^{\infty} \mathbf{E}\left[\exp(-\sum_{i=1}^{m} f(X_i))\mathbf{1}_{\{M=m\}}\right]\end{aligned}$$

with the convention $\sum_{i=1}^{0} = 0$. By construction, the X_i sequence is independent of M and the X_i are independent of each other and have same law

$$\mathbf{E}\left[\exp(-\int f \, \mathrm{d}N)\right] = \sum_{m=0}^{\infty} e^{-\mu(E)} \frac{\mu(E)^m}{m!} \left(\frac{1}{\mu(E)} \int_E e^{-f(x)} \, \mathrm{d}\mu(x)\right)^m$$

$$\} = e^{-\mu(E)} \sum_{m=0}^{\infty} \frac{1}{m!} \left(\int_E e^{-f(x)} \, \mathrm{d}\mu(x)\right)^m$$

$$= \exp\left(-\mu(E) + \int_E e^{-f(x)} \, \mathrm{d}\mu(x)\right).$$

Hence the result. □

To indicate that the control measure is μ, we will often index the expectation by μ. By derivation, we immediately deduce Campbell's formula from the definition.

THEOREM 5.2 (Campbell's formula).– Let $f \in L^1(E, \mu)$, we have:

$$\mathbf{E}_\mu\left[\int f \, \mathrm{d}N\right] = \int_E f \, \mathrm{d}\mu.$$

Let $f \in L^2(E \times E, \mu \otimes \mu)$, we have:

$$\mathbf{E}_\mu\left[\sum_{x \neq y \in N} f(x, y)\right] = \iint_{E \times E} f(x, y) \, \mathrm{d}\mu(x) \, \mathrm{d}\mu(y).$$

THEOREM 5.3.– For a Poisson process of control measure μ, we have the following properties:

- for any set $\Lambda \subset E$, $N(\Lambda)$ follows a Poisson distribution with parameter $\mu(\Lambda)$;
- for Λ_1 and Λ_2 two disjoint subsets of $(E, \mathcal{B}(E))$, the random variables $N(\Lambda_1)$ and $N(\Lambda_2)$ are independent.

Proof. Simply calculate the Laplace transform of the pair $(N(\Lambda_1), N(\Lambda_2))$ by taking

$$f = \alpha \mathbf{1}_{\Lambda_1} + \beta \mathbf{1}_{\Lambda_2}$$

in (5.2). □

Up to now, we have only defined the Poisson process for control measures of finite total mass. The construction of a process for a control measurement of infinite total mass is not without difficulty. difficulty.

EXAMPLE 5.3.– For $E = \mathbf{R}^+$, we retrieve the properties of the Poisson process on the real line.

EXAMPLE 5.4.– Recall that the M/M/∞ queue can be represented as a marked point process where the arrival times follow a Poisson process on $\mathbf{R}^+$ of intensity λ and the marks correspond to the service times.

If we restrict the time horizon to $[0,T]$, the $\mathbf{R}^+ \times \mathbf{R}^+$ process

$$\sum_{n\geq 1} \delta_{(T_n,S_n)}$$

is a spatial Poisson process of intensity

$$\mathrm{d}\mu(s,z) = \lambda \mathbf{1}_{[0,T]}(s)\mu e^{-\mu z}\ \mathrm{d}s\ \mathrm{d}z.$$

whose total mass is λT. Actually, let M_T be the number of arrivals between 0 and T. Since, the random variables S_j are independent from the $T_n's$, given $M_T = m$, the random vector

$$\Big((T_1,S_1),\cdots,(T_m,S_m)\Big)$$

has density

$$\rho_m(t_1,\cdots,t_m,s_1,\cdots,s_m) = \prod_{j=1}^{m} \frac{1}{T}\mathbf{1}_{[0,T]}(t_j)g(s_j),$$

where g is the exponential density of parameter μ. Thus, for f : $[0,T]\times\mathbf{R}^+ \to \mathbf{R}^+$,

$$\begin{aligned}
&\mathbf{E}\left[\exp(-\sum_{T_n\leq T} f(T_n,S_n))\right]\\
&\quad = \sum_{m=0}^{\infty} \mathbf{E}\left[\exp(-\sum_{n=0}^{m} f(T_n,S_n)) \mid M_T = m\right] \mathbf{P}(M_T = m)\\
&\quad = \sum_{m=0}^{\infty} \left(\prod_{j=1}^{m} \frac{1}{T}\int_0^T e^{-f(t_j,s_j)} g(s_j)\ \mathrm{d}t_j\ \mathrm{d}s_j\right) \mathbf{P}(M_T = m)\\
&\quad = \sum_{m=0}^{\infty} \left(\int_0^T e^{-f(t,s)}\frac{1}{T}\ \mathrm{d}t\ g(s)\ \mathrm{d}s\right)^m e^{-\lambda T}\frac{(\lambda T)^m}{m!}\\
&\quad = \exp\left(-\int_{[0,T]\times\mathbf{R}^+} (1-e^{-f(t,s)})\lambda\ \mathrm{d}t\ g(s)\ \mathrm{d}s\right).
\end{aligned}$$

Example 5.5.– To evaluate the performance of wireless systems, the radio communication heavily depends on the respective locations of the emitter and of the receptor as the signal power decreases with the distance: the received power P_r is usually modeled as

$$P_r = K\frac{P_e}{d(x,y)^\gamma}$$

K is a constant, P_e is the emitted power and $d(x,y)$ is the distance emitter-receptor. The power loss exponent γ is there to represent the effect of the geometry on the environment: in void, γ should be equal to 2, in real life, because of obstacles (buildings, trees, water, etc.) γ is usually thought to be in $[3.5,5]$ or higher.

We thus need models to represent mobile users and antennas with which the mobile units communicate. The common model for the location of users is that of a Poisson process of intensity $\lambda > 0$.

If we want to go at a finer scale, we can imagine to model streets. A street is represented as a line in the plane. A line itself is parametrized by its distance to the origin and its angle with respect to the abscissa axis. Thus the set of lines is identified with $[0, 2\pi] \times \mathbf{R}^+$. We can then put any distribution we want on this set to define the so-called Poisson Line Tessellation (PLT). Figure 5.5 shows a realization where

$$\mathrm{d}\mu(\theta, r) = \frac{1}{2}\Big(\delta_0(\theta) + \delta_{\pi/2}(\theta)\Big) \otimes e^{-r}\,\mathrm{d}r.$$

Figure 5.5 shows a realization where

$$\mathrm{d}\mu(\theta, r) = \frac{1}{2\pi}\mathbf{1}_{[0,2\pi]}(\theta)\,\mathrm{d}\theta \otimes e^{-r}\,\mathrm{d}r.$$

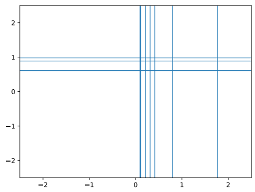


Figure 5.2: A realization with line angle distributed as $1/2(\delta_0 + \delta_{\pi/2})$. Called Manhattan model.

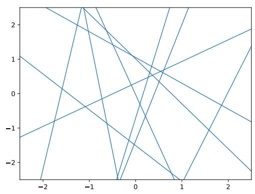


Figure 5.3: A realization with line angle uniformly distributed over $[0, 2\pi]$. Isotropic model.

Definition 5.5.– A Radon measure on Polish E is a finite measure on all compact sets of E.

We admit that for a Polish space E, there always exists an exhaustive sequence of compacts: there exists $(K_n,\, n \geq 1)$ an increasing sequence of compact sets of E such that $E = \cup_n K_n$.

Theorem 5.4.– For μ a Radon measure on E, there exists a point process N called the Poisson control measure μ such that (5.2) is satisfied for any f function of E in $\mathbf{R}$.

Proof. For $m > 0$, we construct N_m the Poisson process of control measure μ restricted to K_m. For f continuous with compact support of E in $\mathbf{R}$, we have

$$\begin{aligned}\mathbf{E}\left[e^{-\int f\,\mathrm{d}N_m}\right] &= \exp\left(-\int_{K_m}(1 - e^{-f(x)})\,\mathrm{d}\mu(x)\right)\\ &\xrightarrow{m\to\infty} \exp\left(-\int_E (1 - e^{-f(x)})\,\mathrm{d}\mu(x)\right)\end{aligned} \tag{5.3}$$

according to the dominated convergence theorem, since f has compact support. We then use Kallenberg's Kallenberg's theorem below, which actually masks the difficulty of the construction. □

Theorem 5.5.– Let $(N_m,\, m \geq 1)$ be a sequence of point processes with values in E. If for any continuous function with compact support f, the sequence of random variables random variables

$$N_m f = \int f(s)\,\mathrm{d}N_m(s),\ k \geq 1$$

converges in law to a random variable N_f then there exists a random variable N with values in $\mathfrak{N}_E$ such that

$$N_m f \xrightarrow[\text{loi}]{m\to\infty} N f.$$

This theorem show how to construct at least formally a Poisson process on the whole plane. Construct a Poisson process N^m on the window $[-m, m]^2$, which has finite Lebesgue measure and then consider the limit in law of the $(N^m,\, m \geq 1)$. It is a theoretical construction as we cannot simulate a Poisson process on an infinite window on a computer.

EXAMPLE 5.6 (Poisson process on the whole line).– On $\mathbf{R}$, there is no "first point" so we cannot use the construction with exponentially distributed inter-arrivals. On $[-m, m]$, we know that the number of points $N([-m, m])$ follows a Poisson distribution of parameter λm. Given $N([-m, m]) = k$, the locations are distributed uniformly on $[-m, m]$. That gives us a first way to simulate such a process.

We can also show that the restriction of N to $\mathbf{R}^+$, denoted by $\hat{N}$, is an ordinary Poisson process on $\mathbf{R}^+$ of intensity λ: for any $f : \mathbf{R}^+ \to \mathbf{R}$ continuous with compact support, f can be viewed as a continuous function with compact support on $\mathbf{R}$ hence

Since $f \in \mathcal{C}_b(\mathbf{R}^+)$, we necessarily have $f(0) = 0$.

$$\begin{aligned}\mathbf{E}\left[\exp\left(-\int_0^\infty f(s)\,\mathrm{d}\hat{N}(s)\right)\right] &= \mathbf{E}\left[\exp\left(-\int_{-\infty}^\infty f(s)\mathbf{1}_{\mathbf{R}^+}(s)\,\mathrm{d}N(s)\right)\right]\\ &= \exp\left(-\int_{\mathbf{R}} 1 - e^{-f\mathbf{1}_{\mathbf{R}^+}}\,\lambda\,\mathrm{d}\ell\right)\\ &= \exp\left(-\int_{\mathbf{R}^+} 1 - e^{-f}\,\lambda\,\mathrm{d}\ell\right).\end{aligned}$$

Consider now $\check{N}$ the process defined for $t \geq 0$, by

$$\check{N}(t) = N([-t, 0]).$$

Hence, for any $f\ \mathbf{R}^+ \to \mathbf{R}$, we have

$$\int f\,\mathrm{d}\check{N} = \int \check{f}\,\mathrm{d}N$$

where $\check{f}(t) = f(-t)$. By the same computations as above, for f continuous with compact support, we get

$$\mathbf{E}\left[\exp\left(-\int_0^\infty f(s)\,\mathrm{d}\check{N}(s)\right)\right] = \exp\left(-\int_{\mathbf{R}^+} 1 - e^{-\check{f}}\,\lambda\,\mathrm{d}\ell\right).$$

This means that $\check{N}$ is also a Poisson point process of intensity λ. We now show that $\hat{N}$ and $\check{N}$ are independent. Choose f and g continuous with compact support in $\mathbf{R}^+$, then f and $\check{g}$ have disjoint supports (only $\{0\}$ may belong to the two supports but this is a Lebesgue-negligeable set) hence

$$\begin{aligned}&\mathbf{E}\left[\exp\left(-\int f\,\mathrm{d}\hat{N}(s) - \int g\,\mathrm{d}\check{N}(s)\right)\right]\\ &\quad= \mathbf{E}\left[\exp\left(-\int (f + \check{g})\,\mathrm{d}N(s)\right)\right]\\ &\quad= \mathbf{E}\left[\exp\left(-\int_0^\infty f(s)\,\mathrm{d}N(s) - \int_{-\infty}^0 \check{g}(s)\,\mathrm{d}N(s)\right)\right]\\ &\quad= \mathbf{E}\left[\exp\left(-\int_0^\infty f(s)\,\mathrm{d}N(s)\right)\right]\mathbf{E}\left[\exp\left(-\int_{-\infty}^0 \check{g}(s)\,\mathrm{d}N(s)\right)\right]\end{aligned}$$

by independence of N restricted to $\mathbf{R}^+$ and N restricted to $(-\infty, 0)$ and the independence of $\hat{N}$ and $\check{N}$ follows.

Another way to see that N is the union of two independent Poisson processes on the half-line is the following. Let T_1 be the point of the point process the closest to 0 on the right:

$$T_1 = \inf\{t \geq 0,\ N([0, t]) > 0\}.$$

The event $(T_1 \geq x)$ is equal to the event that there is no point of N in $[0, x]$, that is to say that a Poisson random variable of parameter λx is null:

$$\begin{aligned}\mathbf{P}(T_1 \geq x) &= \mathbf{P}\big(N([0, x]) = 0\big)\\ &= \exp(-\lambda x).\end{aligned}$$

This means that T_1 follows an exponential random distribution of parameter λ. The same kind of computations shows that

$$T_{-1} = \inf\{t \geq 0,\ N([-t, 0]) > 0\}$$

also follows an exponential distribution of parameter λ. Furthermore, since $(N([0, t]),\ t \geq 0)$ and $(N([-t, 0]),\ t \geq 0)$ are independent (the intervals are disjoint), T_1 and T_{-1} are independent.

By pursuing the reasoning, we can show that we can construct N as two ordinary independent Poisson processes of intensity λ, N^1 and N^2. One, say N^1, is used as is. As to N^2, we apply to it the symmetry with respect to 0 (denoted by s_o) and we have

$$N \stackrel{\text{law}}{=} N^1 + s_o(N^2).$$

Example 5.7 (PLT with users).– We now enrich the PLT model with users on the roads. It can be shown that if E is Polish then $\mathfrak{N}_E$ is Polish. It follows that $E = [0, 2\pi] \times \mathbf{R}^+ \times \mathfrak{N}_\mathbf{R}$ is Polish so we are entitled to consider a Poisson process on this space. Let π_λ the measure on $\mathfrak{N}_\mathbf{R}$ which is the law of a Poisson process of intensity λ on the real line. If we consider the measure

$$\mathrm{d}\mu(\theta, s, \omega) = \frac{1}{2\pi}\mathbf{1}_{[0,2\pi]}(\theta)\ \mathrm{d}\theta \otimes g(r)\ \mathrm{d}r \otimes\ \mathrm{d}\pi_\lambda(\omega),$$

this means that we associate to each line (θ, r), a Poisson process of intensity λ which represents the mobile units on this road.

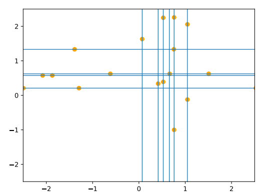


Figure 5.4: The Manhattan model with users distributed as a Poisson process of intensity 4.

Most of the properties of the real Poisson process are transferable to the spatial Poisson process.

Theorem 5.6 (Superposition).– Let N^1 and N^2 be two independent Poisson processes of intensity μ^1 and μ^2 respectively, their superposition N defined by :

$$\int f\ \mathrm{d}N = \int f\ \mathrm{d}N^1 + \int f\ \mathrm{d}N^2$$

is a Poisson process of intensity $\mu^1 + \mu^2$.

Definition 5.6.– Let N be a Poisson process of intensity μ and p a function from E in to $[0, 1]$. The thinned Poisson process with parameter $(\mu,\ p)$ is the process where an atom of the Poisson process N in x is kept with probability $p(x)$.

Theorem 5.7 (Thinning).– A thinned Poisson process of parameters

(μ, p) is a Poisson process of intensity μ_p defined by:

$$\mu_p(A) = \int_A p(x)\, \mathrm{d}\mu(x).$$

Theorem 5.7 is a special case of the displacement theorem.

DEFINITION 5.7.– Let $(\Omega', \mathcal{A}', \mathbf{P}')$ be a probability space and $(F, \mathcal{F})$ a Polish space. A displacement is a measurable application Θ of $\Omega' \times E \longrightarrow F$ such that the random variables $(\Theta(\omega', x), x \in E)$ are independent. For $A \in \mathcal{F}$, introduce:

$$\theta(x, A) = \mathbf{P}'(\omega', : \Theta(\omega', x) \in A).$$

Thus, $\theta(x, A)$ represents the probability that the point x will be moved in A. More mathematically, if we note $\Theta(\omega', .)^*\mu$ the measure image of μ by the application the application $\Theta(\omega', .)$, we have:

$$\begin{aligned}\mathbf{E}_{\mathbf{P}'}\left[\Theta^*\mu(A)\right] &= \mathbf{E}_{\mathbf{P}'}\left[\int \mathbf{1}_{\{\Theta(\omega',x)\in A\}}\, \mathrm{d}\mu(x)\right] \\ &= \int \mathbf{P}'(\Theta(\omega', x) \in A)\, \mathrm{d}\mu(x) = \int \theta(x, A)\, \mathrm{d}\mu(x).\end{aligned}$$

This means that:

$$\mathbf{E}_{\mathbf{P}'}\left[\int \mathbf{1}_A\, \mathrm{d}\Theta^*\mu\right] = \int \int_A \theta(x,\ \mathrm{d}y)\, \mathrm{d}\mu(x),$$

so for a non-negative f function, we get:

$$\mathbf{E}_{\mathbf{P}'}\left[\int f\, \mathrm{d}\Theta^*\mu\right] = \int \int f(y)\theta(x,\ \mathrm{d}y)\, \mathrm{d}\mu(x). \tag{5.4}$$

DEFINITION 5.8.– A displacement is said to be conservative when for any compact $\Lambda \subset E$, the following property is satisfied:

$$\mathbf{E}_{\mathbf{P}'}\left[\Theta^*\mu(\Lambda)\right] = \int_\Lambda \int_F \theta(x,\ \mathrm{d}y)\, \mathrm{d}\mu(x) = \mu(\Lambda).$$

This means that, on average, the total mass of the point process is conserved.

Let Θ be a displacement such that $\int_\Lambda \int_F e^{-f(y)}\theta(x,\ \mathrm{d}y)\, \mathrm{d}\mu(x) = \mu(A)$and N a point process, the point process displaced N^Θ is defined by :

$$N^\Theta(\omega') = \sum_{x\in N} \delta_{\Theta(\omega',x)}.$$

THEOREM 5.8 (Displacement).– Let N be a Poisson process of intensity μ on E and Θ a conservative displacement of E in F. The process N^Θ is a Poisson process of intensity μ^Θ defined by defined by :

$$\mu^\Theta(A) = \int_E \theta(x, A)\, \mathrm{d}\mu(x).$$

Proof. We assume, in a first step, that f is continuous with compact support, denoted by Λ. We known that given $N(\Lambda)$, the points of N

are idependent and identically distributed according to the probability measure $\mu/\mu(\Lambda)$. Hence, we can write:

$$\begin{aligned}
\mathbf{E}&\left[\exp(-\int_\Lambda f \, \mathrm{d}N)\right] \\
&= \sum_{n=0}^{\infty} \mathbf{E}\left[\exp(-\int_\Lambda f \, \mathrm{d}N) \mid N(\Lambda) = n\right] \mathbf{P}(N(\Lambda) = n) \\
&= \frac{e^{-\mu(\Lambda)}\mu(\Lambda)^n}{n!} \int_{E^n} \prod_{j=1}^{n} e^{-f(x_j)} \frac{\mathrm{d}\mu(x_j)}{\mu(\Lambda)}.
\end{aligned}$$

By the construction of N^Θ, the randomness of the displacement is independent of N, thus we have:

$$\begin{aligned}
\mathbf{E}\left[\exp(-\int f \, \mathrm{d}N^\Theta)\right] &= \mathbf{E}_{\mathbf{P}'}\left[\sum_{n=0}^{\infty} \frac{e^{-\mu(\Lambda)}}{n!} \int_{E^n} \prod_{j=1}^{n} e^{-f(\Theta(\omega', x_j))} \, \mathrm{d}\mu(x_j)\right] \\
&= \sum_{n=0}^{\infty} \frac{e^{-\mu(\Lambda)}}{n!} \, \mathbf{E}_{\mathbf{P}'}\left[\int_{E^n} \prod_{j=1}^{n} e^{-f(\Theta(\omega', x_j))} \, \mathrm{d}\mu(x_j)\right].
\end{aligned}$$

By definition of a displacement, the random variables $(\Theta(\omega', x_j), j = 1, \cdots, n)$ are independent. In virtue of (5.4), we obtain:

$$\begin{aligned}
\mathbf{E}\left[\exp(-\int f \, \mathrm{d}N^\Theta)\right] &= \sum_{n=0}^{\infty} \frac{e^{-\mu(\Lambda)}}{n!} \left(\mathbf{E}_{\mathbf{P}'}\left[\int_E e^{-f(\Theta(\omega', x))} \, \mathrm{d}\mu(x)\right]\right)^n \\
&= \sum_{n=0}^{\infty} \frac{e^{-\mu(\Lambda)}}{n!} \left(\int e^{-f} \, \mathrm{d}\mu^\Theta\right)^n \\
&= \exp\left(-\mu(\Lambda) + \int_\Lambda \int_F e^{-f(y)} \theta(x, \, \mathrm{d}y) \, \mathrm{d}\mu(x)\right).
\end{aligned}$$

Since Θ is is conservative, it follows that:

$$\mathbf{E}\left[\exp(-\int f \, \mathrm{d}N^\Theta)\right] = \exp\left(-\int_F (1 - e^{-f(y)}) \int_\Lambda \theta(x, \, \mathrm{d}y) \, \mathrm{d}\mu(x)\right)$$

An increasing sequence of compact sets $(K_n, n \geq 1)$ in E is exhaustive whenever $\cup_n K_n = E$. Think of $[-n, n]^d$ in $\mathbf{R}^d$. More generally, the existence of such a sequence is guaranteed by the Polish property of E.

hence N^Θ is a Poisson point process of control measure μ^Θ.

The general case by applying this result to the sequence $(f\mathbf{1}_{\Lambda_n}, n \geq 1)$ where Λ_n is an exhaustive sequence of compact sets. □

Proof of 5.7. Consider $F = E \cup \Delta$ where Δ is a point external to E. With probability $p(x)$, atom x remains in x, with complementary complementary probability, it is sent to Δ. The restriction of the resulting process to E is the thinning of the initial process. This displacement is indeed conservative, since we keep the same number of atoms. The 5.7 is then a direct consequence of 5.8. □

Applying the 5.8 to the function $(x \in \mathbf{R}^d \mapsto rx)$ where $r \in \mathbf{R}^+$, we obtain a result known as the *scaling*, which is very useful in many applications.

Corollary 5.9.– Let N be a Poisson process of intensity μ on $\mathbf{R}^d$. Let $r > 0$, let $N^{(r)}$ be the dilated process defined by:

$$N^{(r)} = \sum_{x \in N} \delta_{rx}.$$

The process $N^{(r)}$ is a Poisson process of intensity $\mu^{(r)}$ where $\mu^{(r)}(A) = \mu(A/r)$ for all $A \in \mathcal{B}(E)$.

COROLLARY 5.10.– Let N be a Poisson process of control measure $\lambda\, \mathrm{d}x$ on $\mathbf{R}^2$. Let

$$\hat{N} = \sum_{x \in \mathbf{N}} \delta_{\|x\|^2, \mathrm{Arg}(x)}.$$

The $\hat{N}$ process is a multivariate point process where the jump times are those of a Poisson process of intensity $\lambda\pi$ and the marks are independent with a uniform distribution on $[0, 2\pi]$.

Proof. The 5.8 implies that the process:

$$\hat{N} = \sum_{x \in \mathbf{N}} \delta_{\|x\|^2, \mathrm{Arg}(x)}$$

is a Poisson process of control measure given by

$$\begin{aligned}\hat{\mu}([0,r] \times [0,\alpha]) &= \int_{\mathbf{R}^2} \mathbf{1}_{[0,r]}(\|x\|^2)\mathbf{1}_{[0,\alpha]}(\mathrm{Arg}(x))\lambda\, \mathrm{d}x \\ &= \int_0^\infty \int_0^{2\pi} \mathbf{1}_{[0,\sqrt{r}]}(\rho)\mathbf{1}_{[0,\alpha]}(\theta)\lambda\, \mathrm{d}\theta\, \rho\, \mathrm{d}\rho\end{aligned}$$

according to the formula for changing variables in polar coordinates. If we make the change of variables $\tau = \rho^2$, we obtain

$$\hat{\mu}([0,r] \times [0,\alpha]) = \frac{1}{2}\int_0^r \int_0^\alpha \lambda\, \mathrm{d}\theta\, \mathrm{d}\tau. \tag{5.5}$$

Thus the control measure of of $\hat{N}$ is

$$\frac{1}{2}\lambda\, \mathrm{d}\tau \otimes\, \mathrm{d}\theta.$$

If we consider the projection

$$\begin{aligned}\pi_r \;:\; \mathbf{R}^+ \times [0, 2\pi] &\longrightarrow \mathbf{R}^+ \\ (r, \theta) &\longmapsto r^2,\end{aligned}$$

the Theorem 5.8 ensures that the process of modulus is a Poisson process of control measure

$$\begin{aligned}\pi_r^* \hat{\mu}([0,r]) &= \hat{\mu}([0,r] \times [0, 2\pi]) \\ &= \lambda\pi r.\end{aligned}$$

This is therefore a Poisson process of intensity $\lambda\pi$.

Since $\hat{N}$ is a spatial Poisson process, we know that $\hat{N}(]s,t] \times [0,\alpha])$ is independent of the variables $\hat{N}([0,u] \times B)$ for any $u \le s$ and any $B \subset [0, 2\pi]$. Therefore,

$$\mathbf{E}\left[\hat{N}(]s,t] \times [0,\alpha]) \,|\, \mathcal{F}_s\right] = \mathbf{E}\left[\hat{N}(]s,t] \times [0,\alpha])\right].$$

Moreover, according to (5.5), $\hat{N}(]s,t] \times [0,\alpha])$ follows a Poisson distribution with parameter

$$\mathbf{E}\left[\hat{N}(]s,t] \times [0,\alpha])\right] = \frac{1}{2}\lambda\alpha(t-s) = \pi\lambda(t-s) \times \frac{\alpha}{2\pi}.$$

This means that

$$t \longmapsto \hat{N}([0,t] \times [0,\alpha]) - \frac{1}{2}\lambda t\,\alpha$$

is a martingale and therefore $\hat{N}$ is a marked Poisson process where the uniform distribution on $[0,2\pi]$ and the Poisson process of intensity of intensity $\lambda\pi$. □

5.1 Poisson measure

DEFINITION 5.9.– A Poisson measure is a spatial Poisson process on $E = \mathbf{R}^+ \times \mathbf{R}^+$ of control measure $\mathrm{d}s \otimes \mathrm{d}z$.

Since the measure $\mathrm{d}s \otimes \mathrm{d}z$ is diffuse, the probability of having several points of given abscissa corresponds to the probability that a Poisson law of parameter the measure of a half-line is non-zero, and the parameter of such a Poisson law is 0. Consequently, a Poisson measure can be seen as a special case of a multivariate process.

We have already seen how to construct a non-homogeneous Poisson process by time change. This method requires the function $t \mapsto \int_0^t \lambda(s)\,\mathrm{d}s$ to be inverted, which is not necessarily straightforward. We will use Poisson measures to obtain a much simpler algorithm. Let N be a Poisson measure and N_A its restriction to $\mathbf{R}^+ \times [0,A]$. Let $\Big((T_n^A, Z_n^A),\, n \geq 1\Big)$ be its atoms arranged in ascending order of the first component.

LEMMA 5.11.– The point process $M_A = ((T_n^A, Z_n^A), n \geq 1)$ is a marked Poisson process with compensator $A\,\mathrm{d}s \otimes A^{-1}\mathbf{1}_{[0,A]}(z)\,\mathrm{d}z$.

Furthermore, the process

$$M'_A(t) = \int_0^t \int_0^A \mathbf{1}_{z \leq \lambda(s)\}}\,\mathrm{d}M_A(s,z)$$

is a non-homogeneous Poisson process of intensity λ.

Proof. We must show that for any $B \in \mathcal{B}([0,A])$,

$$t \longmapsto M_A\Big([0,t] \times B\Big) - At\int_0^A \mathbf{1}_B(z)\frac{\mathrm{d}z}{A} \tag{5.6}$$

is a martingale. From the definition of a Poisson point process, we know that the variables $M_A(]s,t] \times B)$ and $M_A([0,r] \times C)$ for $r \leq s$ and any $C, B \subset \mathbf{R}^+$ are independent so, by a monotone class argument, $M_A(]s,t] \times B)$ is independent of

$$\mathcal{F}_s = \sigma\big(M_A([0,r] \times C),\, r \leq s,\, C \subset \mathbf{R}^+\big).$$

Hence,

$$\begin{aligned}\mathbf{E}\Big[M_A\Big([0,t] \times B\Big) \mid \mathcal{F}_s\Big] &= M_A\Big([0,s] \times B\Big) + \mathbf{E}\Big[M_A\Big(]s,t] \times B\Big) \mid \mathcal{F}_s\Big] \\ &= M_A\Big([0,s] \times B\Big) + \mathbf{E}\Big[M_A\Big(]s,t] \times B\Big)\Big].\end{aligned}$$

Since $M_A(]s,t] \times B)$ follows a Poisson distribution of parameter

$$\int_s^t \int_B \mathrm{d}r \, \mathrm{d}z = (t-s) \int_{\mathbf{R}^+} \mathbf{1}_B(z) \, \mathrm{d}z,$$

we have (5.6).

We keep only the points of M_A such that

$$Z_n \le \lambda(T_n).$$

That is to say that we consider

$$\begin{aligned} M'_A(t) &= \int_0^t \int_0^A \mathbf{1}_{\{z \le \lambda(s)\}} \, \mathrm{d}M_A(s,z) \\ &= \int_0^t \int_0^\infty \mathbf{1}_{\{z \le \lambda(s)\}} \, \mathrm{d}N(s,z). \end{aligned}$$

We repeat the same reasoning and show that

$$M'_A(t) - \int_0^t \int_0^\infty \mathbf{1}_{\{z \le \lambda(s)\}} \, \mathrm{d}s \, \mathrm{d}z$$

is a local martingale. Now

$$\int_0^t \int_0^\infty \mathbf{1}_{\{z \le \lambda(s)\}} \, \mathrm{d}s \, \mathrm{d}z = \int_0^t \lambda(s) \, \mathrm{d}s,$$

hence the result from the characterisation of Theorem 3.5. □

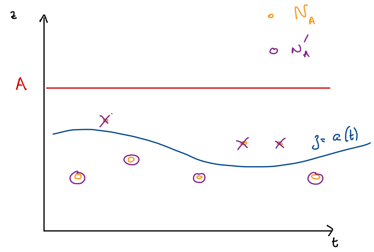


Figure 5.5: Les processus M_A (en orange) et M'_A (en violet)

In practice this result means that to construct a trajectory of a non-homogeneous Poisson process of intensity λ with $\|\lambda\|_\infty \le A$, we construct a Poisson process marked on $\mathbf{R}^+ \times [0, A]$ where the arrivals follow a Poisson of intensity A and the marks are uniformly distributed on $[0, A]$. The point (T_n, Z_n) is kept only if $Z_n \le \lambda(T_n)$.

5.2 *Hawkes processes*

Hawkes processes appear in many fields: in seismology, in biology, in finance, and so on. They are self-exciting processes. Recall the intuitive interpretation of the compensator. We can write

$$\mathbf{E}\left[N(t+h) - N(t) \,|\, \mathcal{F}_t\right] = \int_t^{t+h} \lambda(s) \, \mathrm{d}s$$

which informally translates as: the probability of having an arrival on an infinitesimal interval to the right of t is proportional to $\lambda(t)$.

In the case of an earthquake, an initial tremor generates aftershocks, which in turn generate others, and so on. We assume that earthquakes occur at a rate of λ per unit of time. After a tremor at time T_n, we add to the current intensity a function of the form $\phi(s - T_n)$. In other words, we assume that if N represents the shaking instants,

$$N(t) - \int_0^t \left(\lambda + \int_0^{s^-} \phi(s-u) \, \mathrm{d}N(u) \right) \mathrm{d}s$$

is a local martingale. The first question is to know whether such a process does exist.

As will be apparent in the proof, the filtration which is involved is the filtration generated by a Poisson measure M. This means that y is adapted with respect to this filtration and that it depends on M. Whether there exists a version of N with a compensator predictable with respect to the filtration generated by N itself has been solve recently Coutin and Decreusefond, "Invertibility of functionals of the Poisson process and applications".

Theorem 5.12.– Let $\phi : \mathbf{R}^+ \to \mathbf{R}^+$ such that

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

There exists a unique point process on R^+ of compensator $y(t)$ defined by

$$y(t) = \int_0^t \left(\lambda + \int_0^{s^-} \phi(s-u)\, \mathrm{d}N(u) \right) \mathrm{d}s.$$

Proof. We consider M a Poisson measure and the sequences of intensities and point processes defined by:

$$\begin{aligned} N^0 &= 0 \\ \lambda^0 &= 0 \\ \lambda^{n+1}(t) &= \lambda + \int_0^t \phi(t-u)\, \mathrm{d}N^n(u) && (5.7) \\ N^{n+1}(t) &= \int_0^t \int_0^\infty \mathbf{1}_{\{z \le \lambda^{n+1}(s)\}}\, \mathrm{d}M(s,z). && (5.8) \end{aligned}$$

We first show by induction that

$$\lambda^{n+1}(t) \ge \lambda^n(t)$$

and the jumps of N^n are included in that of N^{n+1}.

It is clearly true for $n = 0$. According to (5.7) and the induction hypothesis, λ^{n+1} differs from λ^n by the quantities $\phi(t - T_k')$ where the T_k' are the jumps of N^n which are not jumps of N^{n-1}. Hence, $\lambda^{n+1}(t) \ge \lambda^{(n)}(t)$ for any $t \ge 0$.

The jumps of N^n (respectively N^{n+1}) are the pairs (T_k, Z_k) of M such that

$$Z_k \le \lambda^n(T_k), \text{ respectively } Z_k \le \lambda^{n+1}(T_k).$$

But $\lambda^n(T_k) \le \lambda^{n+1}(T_k)$, hence the jumps of N^{n+1} contain at least those of N^n. It follows that $N^{n+1}(t) \ge N^n(t)$.

We have two non-decreasing sequences. It is sufficient to bound each of them to ensure their convergence. We deduce from Lemma 5.11 that the intensity of N^n is λ^n hence

$$\begin{aligned} \mathbf{E}\left[N^{n+1}(t)\right] &= \int_0^t \mathbf{E}\left[\lambda^{n+1}(s)\right] \mathrm{d}s \\ &= \int_0^t \left(\lambda + \int_0^s \phi(s-u) \mathbf{E}\left[\lambda^n(u)\right] \mathrm{d}u \right) \mathrm{d}s. \end{aligned}$$

By differentiation, we get

$$\mathbf{E}\left[\lambda^{n+1}(t)\right] = \lambda + \int_0^t \phi(t-u) \mathbf{E}\left[\lambda^n(u)\right] \mathrm{d}u$$

Let

$$v_n(t) = \sup_{s \le t} \mathbf{E}\left[\lambda^n(t)\right].$$

We have

$$\begin{aligned} v_{n+1}(t) &\le \lambda + v_n(t) \int_0^t \phi(t-u)\,\mathrm{d}u \\ &= \lambda + v_n(t) \int_0^t \phi(u)\,\mathrm{d}u \\ &\le \lambda \sum_{k=0}^{n} \left(\int_0^t \phi(u)\,\mathrm{d}u \right)^k. \end{aligned}$$

It follows that

$$\sup_{n\ge 1} \sup_{t\ge 0} v_n(t) \le \frac{\lambda}{1 - \int_0^\infty \phi(u)\,\mathrm{d}u}.$$

Let λ^∞ be the limit of λ^n, which is a priori, a process with values in $\mathbf{R}^+ \cup \{+\infty\}$. According to the Fatou Lemma:

$$\mathbf{E}\left[\int_0^\infty \lambda^\infty(u)\,\mathrm{d}u \right] \le \liminf_{n\to\infty} \mathbf{E}\left[\int_0^\infty \lambda^n(u)\,\mathrm{d}u \right] \le \frac{\lambda}{1 - \int_0^\infty \phi(u)\,\mathrm{d}u}. \tag{5.9}$$

Thus λ^∞ is finite $\mathbf{P} \otimes \ell$ a.e. The monotone limit theorem applied to (5.8) yields that for any $t \ge 0$, $N^n(t)$ converges $\mathbf{P}$-a.e. to the process

$$N^\infty(t) = \int_0^t \int_0^\infty \mathbf{1}_{\{z \le \lambda^\infty(s)\}}\,\mathrm{d}M(s,z),$$

but the set of full measure on which the convergence holds depends on t. It remains to prove that we can choose a full measure set valid for any t. We have

$$0 \le N^\infty(t) - N^n(t) = \int_0^t \int_0^\infty \mathbf{1}_{\{\lambda^n(u) \le z \le \lambda^\infty(u)\}}\,\mathrm{d}M(u,z).$$

Hence,

$$\sup_{t\ge 0} |N^\infty(t) - N^n(t)| \le \int_0^\infty \int_0^\infty \mathbf{1}_{\{\lambda^n(u) \le z \le \lambda^\infty(u)\}}\,\mathrm{d}M(u,z). \tag{5.10}$$

Since

$$\mathbf{E}\left[\int_0^\infty \int_0^\infty \mathbf{1}_{\{z \le \lambda^\infty(u)\}}\,\mathrm{d}M(u,z) \right] = \mathbf{E}\left[\int_0^\infty \lambda^\infty(u)\,\mathrm{d}u \right] < \infty,$$

we have

$$\int_0^\infty \int_0^\infty \mathbf{1}_{\{z \le \lambda^\infty(u)\}}\,\mathrm{d}M(u,z) < \infty, \ \mathbf{P}\text{-a.e.}$$

We can then apply the dominated convergence in the right-hand-side of (5.10) to obtain that

$$\int_0^\infty \int_0^\infty \mathbf{1}_{\{\lambda^n(u) \le z \le \lambda^\infty(u)\}}\,\mathrm{d}M(u,z) \xrightarrow[\mathbf{P}\text{-a.e.}]{n\to\infty} 0.$$

This means that N^n converges uniformly to N^∞ with probability 1. Passing to the monotone limit in (5.7), we see that λ^∞ satisfies

$$\lambda^\infty(t) = \lambda + \int_0^t \phi(t-u)\,\mathrm{d}N^\infty(u).$$

By its very definition, N^∞ is a point process whose compensator is $(t \mapsto \int_0^t \lambda^\infty(s)\,\mathrm{d}s)$. □

Now that we know that there is a point process N, we know that it has jump times $(T_n,\ n \geq 1)$ and that its compensator is written (recall that by convention $T_0 = 0$ and $\sum_1^0 \ldots = 0$):

$$\lambda(t) = \int_0^t \left(\lambda \mathbf{1}_{(0,T_1]}(s) + (\lambda + \phi(s - T_1))\mathbf{1}_{(T_1,T_2]}(s)) + \ldots\right) \text{ d}s$$
$$= \sum_{n \geq 0} \int_0^t \left(\lambda + \sum_{k=1}^n \phi(s - T_k)\right) \mathbf{1}_{(T_n,T_{n+1}]}(s) \text{ d}s$$

In the case where ϕ is decreasing, we derive a simulation algorithm which uses the Poisson measure M.

As before the first jump, the intensity of N is that of a Poisson process of intensity λ, so the first jump takes place in an exponential of parameter λ. From T_1, the intensity becomes non-constant equal to $\lambda + \phi(s - T_1)$. Under the assumption of decreasing ϕ, this intensity is bounded by $\lambda + \phi(0)$. We simulate a marked Poisson process of intensity $(\lambda + \phi(0))ds \otimes (\lambda + \phi(0))^{-1}\mathbf{1}_{[0,\lambda+\phi(0)]}(z)$ dz shifted in time by T_1, its jump instants are therefore $S_1 = T_1 + W_1$, $S_2 = S_1 + W_2, \cdots$ where the $(W_n, n \geq 1)$ are the jump times of a Poisson process of intensity $(\lambda + \phi(0))$ and the corresponding jump heights are uniformly distributed on $[0, \lambda + \phi(0)]$. We have

$$T_2 = \inf\{S_n,\ Z_n \leq \phi(S_n - T_1)\}.$$

Repeat the operation with the intensity

$$\lambda + \phi(s - T_1) + \phi(s - T_2) \leq \lambda + \phi(T_2 - T_1) + \phi(0).$$

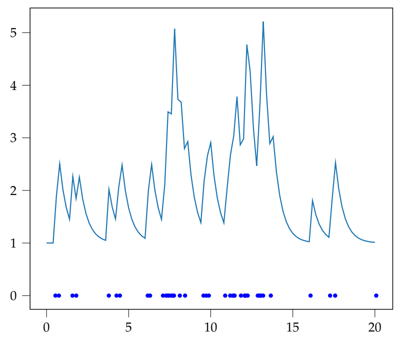


Figure 5.6: A trajectory of a Hawkes process with $\lambda = 1$ and $\phi(t) = \exp(-2t)$

Hawkes processes in finance

After the 2008 crisis, supposedly caused by financial models, mathematicians began trying to model the formation of share prices in more detail. Share prices are formed from an order book.

An order book for a given share looks like this table reftable:ordre1

The table means that 3 people are ready to buy shares up to the price of 118.50 for a total of 57 shares. One person is willing to sell 82 shares at 119.50. If you make a buy offer at 118.50, it will be added to existing offers and will remain pending until a buyer at that price comes forward. If you place a buy order up to the price of 119.60 for 362 shares, they will immediately be yours and the top selling price will be 119.70.

Nombre	Quantité	Prix
3	57	118,50
4	1065	118,20
1	5	118,10

Nombre	Quantité	Prix
1	82	119,50
1	280	119,60
1	500	119,70

Table 5.1: Top picture, buy orders, bottom picture, sell orders

Experiments have shown that after a purchase, the next move is often a sale. One of the models used to represent this is a pair of mutually exciting Hawkes processes : N_1 and N_2 are two point processes with respective intensity

$$\lambda_1(t) = \nu + \int_0^{t^-} \varphi(t - s) \text{ d}N_2(s)$$
$$\lambda_2(t) = \nu + \int_0^{t^-} \varphi(t - s) \text{ d}N_1(s).$$

The price is then given by $S = N_1 - N_2$. As prices are necessarily multiples of 10 cents, we can normalise the amplitude of price variations by 1.

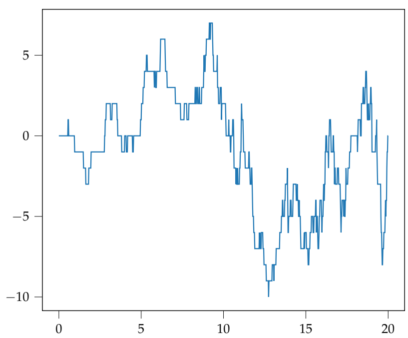


Figure 5.7: A trajectory of an order book simulated by the difference of two Hawkes processes exciting each other

5.3 *Problems*

Exercise 5.3.1: Prove Theorem 5.2.

Solution on page 87

Exercise 5.3.2: Graphe aléatoire géométrique On considère N, un processus de Poisson d'intensité λ dans le plan. On fixe $\varepsilon > 0$ et on construit le graphe $\Gamma_N = (V, E)$ dont les sommets sont les points de N et les arêtes sont construites selon la règle:

$$[xy] \in E(\Gamma_N) \iff d(x,y) \leq \varepsilon.$$

1. Quel est la loi du nombre de voisins noté V_x, d'un point $x \in \Gamma_N$?
2. En déduire le nombre moyen d'arêtes par unité de surface.
3. Soit x, y deux points distincts de N. Montrer qu'il existe trois variables aléatoires A, B et C, indépendantes de loi de Poisson dont on précisera les paramètres telles que

$$V_x = A + B$$
$$V_y = C + B.$$

   We admit that the area of the intersection of two circles of radius ε, at a distance d, is given by

$$2\varepsilon^2 \arccos\left(\frac{d}{2\varepsilon}\right) - d\sqrt{\varepsilon^2 - \left(\frac{d}{2}\right)^2}$$

On considère qu'un virus se transmet le long des arêtes de graphe. On part de l'origine que l'on considère comme un point du processus N et l'on regarde tous les chemins issus de ce point dans le graphe Γ_N.

4. Montrer que si $\lambda\pi\varepsilon^2 < 1$ alors la composante connexe de Γ_N contenant 0 est presque-sûrement finie.

Solution on page 87

Exercise 5.3.3: Pavage de lignes Poisson Une droite dans le plan est paramétrée par sa distance à l'origine et son angle avec l'axe des abscisses. On a donc une bijection entre l'ensemble des droites et l'ensemble $E = [0, 2\pi[\times \mathbf{R}^+$. On considère le processus de Poisson spatial sur E de mesure de contrôle

$$(\frac{1}{2}\delta_0(\theta) + \frac{1}{2}\delta_{\pi/2}(\theta)) \otimes \rho \, \mathrm{d}r.$$

1. Comment sont les droites de chaque réalisation?

2. Quelle est la nature du processus des abscisses des droites verticales?

3. Quelle est la loi de la distance de l'origine à l'ensemble des droites?

Sur chacune des droites, on place des utilisateurs selon un processus de Poisson sur $\mathbf{R}$ d'intensité λ. Les processus de Poisson sont indépendants les uns des autres. On note D la distance de l'origine au plus proche des utilisateurs. Pour $x > 0$, on note N_x le nombre de droites qui intersectent en plus qu'un point $B(0, x)$.

4. Pour une droite verticale qui passe par $(r, 0)$, calculer la longueur de son intersection avec le cercle centré sur l'origine de rayon x.

5. En déduire $\mathbf{P}(D \geq x \mid N_x = k)$.

6. En déduire $\mathbf{P}(D \geq x)$.

Solution on page 88

Exercise 5.3.4: Considérons que les utilisateurs de téléphones mobiles soient répartis dans le plan selon un processus de Poisson d'intensité λ. En un point x du plan, l'interférence créée par les mobiles a pour expression:

$$I(x, N) = \sum_{y \in N} h(y)P(y)l(\|y - x\|),$$

où $P(x)$ est la puissance du signal émis par le mobile en y, l est une fonction de $\mathbf{R}^+$ dans $\mathbf{R}^+$ que l'on prend généralement de la forme :

$$l_0(r) = r^{-\gamma} \text{ ou } l_1(r) = \min(1,\ r^{-\gamma}). \tag{5.11}$$

La deuxième formulation donne des formules moins élégantes mais est plus réaliste (un signal ne vas pas être amplifié sous prétexte que le récepteur est très proche de l'émetteur) et évite les problèmes de divergence d'intégrale. Les variables aléatoires $(h(y), y \in N)$ sont généralement identiquement distribuées (de même loi qu'une variable aléatoire H), indépendantes entre elles et indépendantes de N. Elles représentent le facteur de perte induit par le *fading* (l'atténuation due aux mouvements locaux du récepteur) et le *shadowing* (l'atténuation du signal due aux obstacles entre l'émetteur et le récepteur). Généralement, le *fading* est modélisé par une variable aléatoire de loi exponentielle de paramètre 1. Le *shadowing* est représenté par une loi log-normale, c'est-à-dire l'exponentielle d'une variable gaussienne.

1. Montrer que l'interférence moyenne ne dépend pas de x et la calculer en prenant l_1 comme fonction de path-loss.

Supposons maintenant que les variables aléatoires $(h(x),\, x \in \mathbf{R}^2)$ soient indépendantes de même loi.

2. Montrer que
$$\mathbf{E}\left[e^{-sI(0)}\right] = \exp\left(\int_\Lambda \left(\ln \mathcal{L}_H(sl(\|x\|)) - 1\right) \mathrm{d}\mu(x)\right).$$
où $\mathcal{L}_H$ est la transformée de Laplace de la loi de H.

Solution on page 88

5.4 Solution to problems

Solution of Exercise 5.3.1 on page 84:
Compute

$$\frac{\mathrm{d}}{\mathrm{d}\theta}\mathbf{E}\left[e^{-\theta\int f\,\mathrm{d}N}\right]\Big|_{\theta=0}$$

Solution of Exercise 5.3.2 on page 84:

1 Tous les points à distance inférieure à ε de x sont des voisins de x donc V_x suit une loi de Poisson de paramètre $\lambda\pi\varepsilon^2$.

2 On se restreint à une fenêtre $[-a/2, a/2]^2$ et on compte seulement les arêtes entre points de cette fenêtre. On a la relation

$$2 \text{ nombre d'arêtes } = \text{somme des degrés.}$$

On peut écrire

$$\begin{aligned}\mathbf{E}[A] &= \frac{1}{2}\mathbf{E}\left[\int V_x\,\mathrm{d}N(x)\right]\\ &= \frac{1}{2}\mathbf{E}\left[\int\left(\int_{y\neq x}\mathbf{1}_{\{|y-x|\le\varepsilon\}}\,\mathrm{d}N(y)\right)\,\mathrm{d}N(x)\right]\\ &= \frac{1}{2}\int\left(\int\mathbf{1}_{\{|y-x|\le\varepsilon\}}\lambda\,\mathrm{d}y\right)\lambda\,\mathrm{d}x\end{aligned}$$

d'après la formule de Campbell, Théorème 5.2. On obtient

$$\mathbf{E}[A] = \frac{1}{2}\int_{\mathbb{T}_2}\lambda^2\pi\varepsilon^2\,\mathrm{d}x = \lambda^2\pi\varepsilon^2 a^2.$$

Comme la surface de $\mathbb{T}_2$ est a^2; on voit que le nombre moyen d'arêtes par unité de surface est $\lambda^2\pi\varepsilon^2$.

3 Il y a A qui représente les voisins de x qui ne sont pas voisins de y, C qui représente les voisins de y qui ne sont pas voisins de x et B qui représente les voisins communs. On note B_x respectivement B_y la boule de rayon ε centrée en x, respectivement en y. On note $B_{xy} = B_x \cap B_y$. On a

$$\begin{aligned}A &= N(B_X\backslash B_{xy})\\ B &= N(B_{xy})\\ C &= N(B_y\backslash B_{xy}).\end{aligned}$$

Les trois ensembles sont disjoints donc les variables aléatoires sont indépendantes. On admet que l'aire de l'intersection des deux cercles de rayon ε est donnée par

$$|B_{xy}| = 2\varepsilon^2\arccos\left(\frac{d}{2\varepsilon}\right) - d\sqrt{\varepsilon^2-\left(\frac{d}{2}\right)^2}$$

avec $d = |x - y|$. Au final, A suit une loi de Poisson de paramètre $\lambda|B_x|$, B suit une loi de Poisson de paramètre $\lambda|B_{xy}|$ et C une loi de Poisson de paramètre $\lambda|B_y|$. On remarque que le paramètre de A et C est inférieur à $\lambda\pi\varepsilon^2$.

4 On va construire un processus de branchement sur le graphe géométrique. L'idée est de considérer comme enfants d'un sommet tous ses voisins mais en faisant ça, certains points se retrouveraient dans plusieurs branches, on n'aurait donc plus un arbre mais un graphe. Ce dont on ne veut pas.

On augmente l'espace d'états en attribuant à chaque point une marque de loi uniforme sur $[0,1]$, indépendante du processus et indépendante des autres marques.

A la génération 0, on a $G_0 = \{0\}$. A la génération 1, on a G_1 qui contient tous les voisins de 0. Les choses se compliquent à la génération d'après. On numérote les éléments de $G_1 = \{x_i, i = 1, \cdots, |V_0|\}$ selon l'ordre de leur marque : le premier point est celui qui a la plus petite marque. On pose $G_2 = \emptyset$. On considère que les enfants de x_j sont ses voisins qui ne sont pas dans $G_1 \cup G_2$. On adjoint ces nouveaux enfants à G_2 jusqu'à ce qu'on ait parcouru les $|V_0|$ voisins de l'origine. On continue comme ça jusqu'à ce que l'on n'ait plus de voisins à ajouter. Le nombre d'enfants de chaque point est stochastiquement inférieur à une variable de Poisson de paramètre $\lambda\pi\varepsilon^2$. Cette loi a pour espérance le paramètre dont on a supposé qu'il est strictement inférieur à 1 donc le processus de branchement correspondant s'éteint presque-sûrement. Cela veut dire que l'épidémie s'arrête d'elle même si les mesures de distanciation sociale sont respectées.

Solution of Exercise 5.3.3 on page 84:

Solution of Exercise 5.3.4 on page 85:

1 La formule de Campbell indique que:

$$\mathbf{E}_\mu\left[I(x)\right] = \mathbf{E}\left[H\right] \int P(y) l(\|y - x\|)\, \mathrm{d}\mu(y).$$

On observe immédiatement que la quantité précédente ne dépend pas de x, d'où:

$$\mathbf{E}_\mu\left[I(0)\right] = P\mathbf{E}\left[H\right] \int l(\|y\|)\lambda\, \mathrm{d}y = \lambda\mathbf{E}\left[H\right] \int_0^\infty l(r) r\, \mathrm{d}r.$$

Si l'on prend comme modèle de *path-loss* la fonction l_1 définie dans (5.11), pour une cellule de rayon $R > 1$, on obtient:

$$\mathbf{E}_\lambda\left[I(0)\right] = \mathbf{E}\left[H\right]\lambda\left(\pi + \frac{\pi}{\gamma-2}(1-R^{2-\gamma})\right).$$

2 Pour s réel positif, on a:

$$\begin{aligned}
&\mathbf{E}\left[\exp(-s\int h(x)l(\|x\|)\,\mathrm{d}N(x))\right] \\
&\quad= \mathbf{E}\left[\mathbf{E}\left[\exp(-s\int h(x)l(\|x\|)\,\mathrm{d}N(x))\mid N(\Lambda)\right]\right] \\
&\quad= \mathbf{E}\left[\prod_{x\in N}\int \exp(-sl(\|x\|)y)\,\mathrm{d}\mathbf{P}_H(y)\right],
\end{aligned}$$

puisque conditionnellement au nombre de points dans Λ, les atomes sont indépendants les uns des autres. En notant $\mathcal{L}_H$ la transformée de Laplace de H, on obtient:

$$\begin{aligned}
&\mathbf{E}\left[\exp(-s\int h(x)l(\|x\|)\,\mathrm{d}N(x))\right] \\
&\qquad= \mathbf{E}\left[\prod_{x\in N}\mathcal{L}_H(sl(\|x\|))\right] \\
&\qquad= \mathbf{E}\left[\exp(\int_\Lambda \ln\mathcal{L}_H(sl(\|x\|))\,\mathrm{d}N(x))\right] \\
&\qquad= \exp(-\int_\Lambda 1 - e^{\ln\mathcal{L}_H(sl(\|x\|))}\,\mathrm{d}\mu(x)) \\
&\qquad= \exp(\int_\Lambda (\ln\mathcal{L}_H(sl(\|x\|)) - 1)\,\mathrm{d}\mu(x)).
\end{aligned}$$

Pour le *fading* de Rayleigh, H est la loi exponentielle de paramètre 1. Si l'on suppose que le *path-loss* est donné par l_0, tous les calculs sont faisables et l'on obtient la formule suivante:

$$\mathcal{L}_{I(0)}(s) = \exp(-\pi\lambda s^\delta\frac{\pi\delta}{\sin(\pi\delta)}),$$

où $\delta = 2/\gamma$. On sait alors que cela correspond à une loi stable d'exposant caractéristique δ.

Bibliography

Billingsley, P. *Convergence of Probability Measures*. Second. Wiley Series in Probability and Statistics: Probability and Statistics. New York: John Wiley & Sons Inc., 1999.

Coutin, L. and L. Decreusefond. "Invertibility of functionals of the Poisson process and applications". In: *Annals of probability* (2025), To appear. URL: https://telecom-paris.hal.science/hal-04153437.

Kallenberg, O. *Foundations of modern probability*. Probability and its Applications (New York). Springer-Verlag, 1997, pp. xii+523.

Index