# Introduction to Poisson processes

François Chapon

Contents



## 1. Introduction

If you commute by bus to come to the university, you can do a bit of statistics while waiting for the bus: knowing that the buses on your route come on average every 10 minutes, what is your average waiting time? The naive answer you might give is that if buses are coming on average every 10 minutes and you arrive at a random time, your average waiting time will be about 5 minutes. So, it is just that you are so unlucky that, most of the time you wait something like 10 minutes for the bus to arrive?

Not really... The naive answer would be true if the buses arrived exactly every 10 minutes. But things don't go that way, and buses do not arrive exactly on schedule due to random complications, traffic, strike, demonstrations, etc... If you arrive at a random time, the probability that you arrive in a long interval between the arrival of two buses is greater than the probability that you arrive in a shorter interval, and your waiting time will be on average longer than expected.

This phenomenon is known as the waiting time paradox and can be modeled by a counting process such as the Poisson process. This is the basic process for modeling queueing systems.

Let us start with an example of a counting process in discrete time. A heads or tails game is modelized by the Bernoulli process. Let $(\varepsilon_n)_{n\geq 1}$ be i.i.d. random variables with Bernoulli distribution $p\delta_1 + (1-p)\delta_0$ and define $S_0 = 0$, and for all $n \geq 1$,

$$S_n = \varepsilon_1 + \cdots + \varepsilon_n.$$

The Bernoulli process thus counts the number of successes up to time $n$ and of course, $S_n$ has a Binomial distribution with parameters $n$ and $p$. The success times are defined by

$T_0 = 0$ and

$$T_n = \inf\{k > T_{n-1} \,|\, S_k = n\},$$

and correspond to the jump times of $(S_n)_{n\geq 0}$. It is a standard exercice in probability theory to show that the random variables $T_0, T_1 - T_0, \ldots, T_n - T_{n-1}, \ldots$ are independent and identically distributed with geometric distribution with parameter $p$. One recovers the process $S_n$ from its jump times by

$$S_n = \sup\{k \geq 0 \,|\, T_k \leq n\}.$$

Hence, $\{S_n \geq k\} = \{T_k \leq n\}$ so $\{S_n = k\} = \{T_k \leq n, T_{k+1} > n\}$, and a straightforward computation shows that if the increments are i.i.d. with geometric distribution, we have $S_n \sim \mathcal{B}(n,p)$

In applications, modeling by a discrete time process is not always appropiate, and we will see that the Poisson process is a continuous time analogue of the Bernoulli process.

## 2. Definition as a counting process

We need to consider stochastic processes indexed not only by integers but also by continuous time. The generalization is straightforward.

Definition 2.1. *A stochastic process (in continuous time) is a family of random variables $(X_t)_{t\geq 0}$, indexed by $t \in \mathbb{R}_+$, defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$. A filtration on $(\Omega, \mathcal{F}, \mathbb{P})$ is a non-decreasing family of sub-$\sigma$-algebras $(\mathcal{F}_t)_{t\geq 0}$.*

Definition 2.2. *A point process on $[0, +\infty[$ is an increasing sequence of random variables $(T_n)_{n\geq 0}$ with values in $[0, +\infty[$ such that $T_n \to +\infty$ a.s. The counting process $(N_t)_{t\geq 0}$ associated to the sequence $(T_n)_{n\geq 0}$ is the process defined by*

$$N_t = \sum_{n=1}^{+\infty} \mathbb{1}_{[0,t]}(T_n) = \sup\{n \geq 1 \,|\, T_n \leq t\},$$

*for $t \in \mathbb{R}_+$.*

Immediate properties are: almost surely, $t \mapsto N_t$ is non-decreasing, càdlàg ("continue à droite avec limites à gauche", french for right-continuous with left limits), meaning that, for all $t > 0$,

$$\lim_{s \searrow t} N_s = N_t \quad \text{and} \quad \lim_{s \nearrow t} N_s = N_{t^-} \text{ exists.}$$

Moreover, observe that

$$\{T_n \leq t\} = \{N_t \geq n\}$$

and that the point process is recovered from $(N_t)_{t\geq 0}$ by

$$T_n = \inf\{t \geq 0 \,|\, N_t \geq n\}.$$

We now give a first and simple definition of a Poisson process.

Definition 2.3. *Let $(Z_n)_{n\geq 1}$ be a sequence of independent exponentially distributed random variables with parameter $\lambda$. Define, $T_0 = 0$ and for all $n \geq 1$,*

$$T_n = \sum_{k=1}^{n} Z_k.$$

*Then $(T_n)_{n\geq 0}$ is a point process on $[0, +\infty[$ and the associated counting process $(N_t)_{t\geq 0}$ is called a Poisson process with rate $\lambda$.*

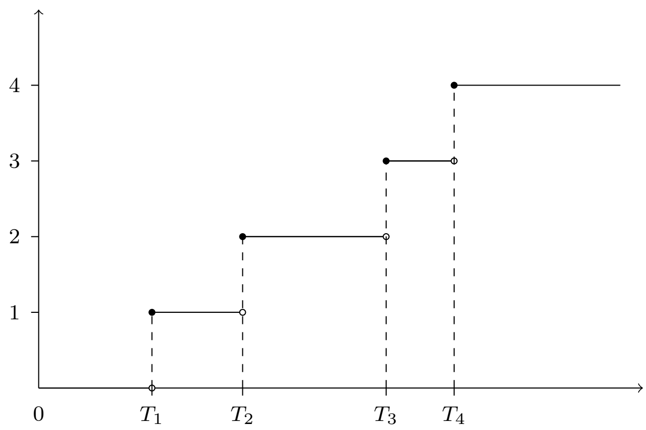


FIGURE 1. A example of a typical trajectory of a Poisson process.

The sequence $(T_n)_{n\geq 0}$ has to be understood as times of occurrences of some events, for instance, arrivals of clients in some queue. The random variables $T_n$ will be called *arrival times* and the increments $T_n - T_{n-1}$ *inter-arrival times*. A typical trajectory of a Poisson process is given by Figure 1.

Recall that a random variable $X$ is exponentially distributed with parameter $\lambda > 0$, if it has density

$$\lambda e^{-\lambda t}\mathbb{1}_{(0,+\infty)}(t)$$

with respect to Lebesgue measure. We denote this distribution by $\mathcal{E}(\lambda)$. Equivalently, the survival probability of $X$ is given by

$$\mathbb{P}(X > t) = e^{-\lambda t}, \quad t \geq 0.$$

The exponential distribution is characterized by the following memoryless property:

PROPOSITION 2.1. *A continuous random variable $X$ with values in $\mathbb{R}_+^*$ is exponentially distributed if and only if for all $t, s \geq 0$,*

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

PROOF. If $X \sim \mathcal{E}(\lambda)$, then

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

Conversely, define $f$ by $f(t) = \log \mathbb{P}(X > t)$. Then $f$ is right continuous, $f(0) = 0$ and satisfies the equation $f(t + s) = f(t) + f(s)$, for all $t, s \geq 0$. One deduces that $f$ is linear, that is there exists $\lambda > 0$ such that $f(t) = -\lambda t$, thus $X$ is exponentially distributed. □

In terms of conditional distribution, it says that if $X$ is exponentially distributed, the conditional distribution of $X - t$ given $\{X > t\}$ is still an exponential distribution with same parameter. This is an essential fact behind the definition of a Poisson process.

Recall that a random variable $X$ has Gamma distribution $\Gamma(\alpha, \beta)$ with parameter shape $\alpha > 0$ and rate $\beta > 0$, if it has the density with respect to Lebesgue measure given by

$$\frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}\mathbb{1}_{(0,+\infty)}(x),$$

where $\Gamma(\alpha) = \int_0^{+\infty} x^{\alpha-1}e^{-x}dx$ is the Gamma function. When $\alpha = n$ is an integer, it is the so-called Erlang's distribution, with denisty

$$\frac{\beta^n}{(n-1)!} x^{n-1} e^{-\beta x}\mathbb{1}_{(0,+\infty)}(x).$$

When $\alpha = 1$, it is the exponential distribution $\mathcal{E}(\beta)$.

PROPOSITION 2.2. *Let $X$ and $Y$ be independent random variables with distribution $\Gamma(\alpha_1, \beta)$ and $\Gamma(\alpha_2, \beta)$ respectively. Then, $X + Y$ has distribution $\Gamma(\alpha_1 + \alpha_2, \beta)$.*

PROOF. Exercise (use the change of variables formula). □

An immediate consequence is:

COROLLARY 2.1. *Let $X_1, \dots, X_n$ be independent random variables with exponential distribution $\mathcal{E}(\alpha)$. Then, $X_1 + \cdots + X_n$ has Gamma distribution $\Gamma(n, \lambda)$.*

We can now state the following properties for a Poisson process:

PROPOSITION 2.3. *Let $(T_n)_{n\geq 0}$ be the point process associated to a Poisson process $(N_t)_{t\geq 0}$ with rate $\lambda$. Let $n \geq 1$. We have:*
*(i) the random variable $T_n$ is distributed according to the $\Gamma(n, \lambda)$ distribution;*
*(ii) for all $t \geq 0$, $N_t$ has a Poisson distribution with parameter $\lambda t$;*
*(iii) given $\{N_t = n\}$, the random vector $(T_1, \dots, T_n)$ is distributed as the order statistics $(U_{(1)}, \dots, U_{(n)})$ of $n$ i.i.d. random variables uniformly distributed on $[0, t]$.*

PROOF. Item (i): This is the above corollary.
Item (ii): Since,

$$\{N_t \geq n\} = \{T_n \leq t\}$$

we have,

$$\{N_t = n\} = \{T_n \leq t < T_{n+1}\}.$$

Hence, using that $T_n \sim \Gamma(n, \lambda)$, $T_{n+1} = T_n + Z_{n+1}$ and the independence of $T_n$ and $Z_{n+1}$, we have, for all $n \geq 0$,

$$\begin{aligned}
\mathbb{P}(N_t = n) &= \mathbb{P}(T_n \leq t, T_n + Z_{n+1} > t) \\
&= \int \mathbb{1}_{\{0<x<t\}} \mathbb{1}_{\{x+z>t\}} \frac{\lambda^n}{(n-1)!} x^{n-1} e^{-\lambda x} \lambda e^{-\lambda z} dx dz \\
&= \int \mathbb{1}_{\{0<x<t\}} \frac{\lambda^n}{(n-1)!} x^{n-1} \left( \int \mathbb{1}_{\{x+z>t\}} \lambda e^{-\lambda(x+z)} dz \right) dx \\
&= \frac{\lambda^n}{(n-1)!} e^{-\lambda t} \int \mathbb{1}_{\{0<x<t\}} x^{n-1} dx \\
&= \frac{(\lambda t)^n}{n!} e^{-\lambda t}.
\end{aligned}$$

Item (iii): Let $U_1, \dots, U_n$ be $n$ i.i.d. random variables, uniformly distributed on $[0, t]$. Recall that the order statistics $(U_{(1)}, \dots, U_{(n)})$ are defined by sorting the values of $U_1, \dots, U_n$ in increasing order. Hence, $U_{(1)}$ is the minimum of $U_1, \dots, U_n$, $U_{(2)}$ is the 2th smallest value, etc...

First we compute the distribution of $(U_{(1)}, \dots, U_{(n)})$. Let $\varphi$ be a bounded measurable function. Then,

$$\begin{aligned}
\mathbb{E}\left[\varphi(U_{(1)}, \dots, U_{(n)})\right] &= \mathbb{E}\left[\varphi(U_{(1)}, \dots, U_{(n)}) \mathbb{1}_{\{\exists\, \sigma \in \mathcal{S}_n,\, U_{\sigma(1)} < \cdots < U_{\sigma(n)}\}}\right] \\
&= \sum_{\sigma \in \mathcal{S}_n} \mathbb{E}\left[\varphi(U_{\sigma(1)}, \dots, U_{\sigma(n)}) \mathbb{1}_{\{U_{\sigma(1)} < \cdots < U_{\sigma(n)}\}}\right] \\
&= n!\, \mathbb{E}\left[\varphi(U_1, \dots, U_n) \mathbb{1}_{\{U_1 < \cdots < U_n\}}\right],
\end{aligned}$$

since $U_1, \dots, U_n$ are i.i.d., hence exchangeable. So we get,

$$\mathbb{E}\left[\varphi(U_{(1)}, \dots, U_{(n)})\right] = \int \mathbb{1}_{\{0<x_1<\cdots<x_n<t\}} \frac{n!}{t^n} dx_1 \cdots dx_n,$$

which gives that the density of $(U_{(1)}, \dots, U_{(n)})$ is

$$(u_1, \dots u_n) \mapsto \frac{n!}{t^n} \mathbb{1}_{\{0<u_1<u_2<\cdots<u_n<t\}},$$

with respect to Lebesgue measure on $\mathbb{R}^n$. Now, we compute the distribution of $(T_1, \dots, T_n)$ conditionally on $\{N_t = n\}$. We have

$$\begin{aligned}
&\mathbb{E}\left[\varphi(T_1, \dots, T_n)\mathbb{1}_{\{N_t=n\}}\right] \\
&= \mathbb{E}\left[\varphi(T_1, \dots, T_n)\mathbb{1}_{\{T_n \le t < T_{n+1}\}}\right] \\
&= \int_{]0,+\infty[^n} \varphi(z_1, \dots, z_1 + \cdots + z_n)\mathbb{1}_{\{z_1+\cdots+z_n<t<z_1+\cdots+z_{n+1}\}}\lambda^{n+1}e^{\lambda(z_1+\cdots+z_{n+1})}\prod_{i=1}^{n+1} dz_i,
\end{aligned}$$

since $T_k$ is the sum $k$ independent and exponentially distributed random variables. Using the change of variables $x_1 = z_1, \dots, x_{n+1} = z_1 + \dots + z_{n+1}$ from $(0, +\infty)^{n+1}$ to $\{0 < x_1 < \cdots < x_{n+1}\}$, with Jacobian equal to one, we get

$$\begin{aligned}
\mathbb{E}\left[\varphi(T_1, \dots, T_n)\mathbb{1}_{\{N_t=n\}}\right] &= \int \varphi(x_1, \dots, x_n)\mathbb{1}_{\{0<x_1<\cdots<x_n<t<x_{n+1}\}}\lambda^{n+1}e^{-\lambda x_{n+1}}\prod_{i=1}^{n+1} dx_i \\
&= \int \varphi(x_1, \dots, x_n)\mathbb{1}_{\{0<x_1<\cdots<x_n<t\}}\lambda^n e^{-\lambda t}\prod_{i=1}^{n} dx_i,
\end{aligned}$$

where we perform the integration with respect to the last variable in the last equality. Since $N_t \sim \mathcal{P}(\lambda t)$, we obtain

$$\mathbb{E}\left[\varphi(T_1, \dots, T_n) \,|\, N_t = n\right] = \int \varphi(x_1, \dots, x_n)\mathbb{1}_{\{0<x_1<\cdots<x_n<t\}}\frac{n!}{t^n}dx_1 \cdots dx_n,$$

hence the result. $\square$

Remark 2.1. The last item of the proposition thus says that if we known that $n$ clients have arrived at time $t$, their arrival times are uniformly distributed on $[0, t]$.

A fundamental result is the following. We will see a converse to the statement in the next section.

Theorem 2.1. *Let $(N_t)_{t\ge0}$ be a Poisson process with rate $\lambda$. Then, $(N_t)_{t\ge0}$ has independent and stationary increments, that is:*

*(i) for all $n \ge 1$, and all $0 = t_0 < t_1 < \cdots < t_n$, the random variables*

$$N_{t_1}, N_{t_2} - N_{t_1}, \dots, N_{t_n} - N_{t_{n-1}}$$

*are independent;*

*(ii) for all $t > s$, $N_t - N_s$ has the same distribution than $N_{t-s}$.*

Proof. Let $(N_t)_{t\ge0}$ be a Poisson process with rate $\lambda$. By definition,

$$N_t = \sum_{n=1}^{+\infty} \mathbb{1}_{[0,t]}(T_n), \quad t \in \mathbb{R}_+,$$

where $(T_n - T_{n-1})_{n\ge1}$ is a sequence of independent and exponentially distributed random variables with parameter $\lambda$.

We have to prove that for all $n \ge 1$, for all $0 = t_0 < t_1 < \cdots < t_n$, the random variables $N_{t_1}, N_{t_2} - N_{t_1}, \dots, N_{t_n} - N_{t_{n-1}}$ are independent. We restrict ourselves to the case $n = 2$, the proof for any $n$ can be done in the same way, but notations become a bit cumbersome. (Exercise: prove the general case).

So let us prove that for all $t, s \geq 0$, for all $k, l \geq 0$, the random variables $N_t - N_s$ and $N_s$ are independent. We can restrict to $l \geq 1$, since the probabilities sum to one. We have, using the relation $\{N_t = n\} = \{T_n \leq t < T_{n+1}\}$,

$$\mathbb{P}(N_s = k, N_t - N_s = l) = \mathbb{P}(N_s = k, N_t = k + l) = \mathbb{P}(T_k \leq s < T_{k+1}, T_{k+l} \leq t < T_{k+l+1}).$$

Hence, since the increments $\Delta T_k = T_k - T_{k-1}$ are independent and exponentially distributed,

$$\begin{aligned}
&\mathbb{P}(N_s = k, N_t - N_s = l) \\
&= \int \mathbb{1}_{\{z_1+\cdots+z_k<s<z_1+\cdots+z_{k+1}, z_1+\cdots+z_{k+l}<t<z_1+\cdots+z_{k+l+1}\}} \lambda^{k+l+1} e^{-\lambda(z_1+\cdots+z_{k+l+1})} \prod_{i=1}^{k+l+1} dz_i \\
&= \int \mathbb{1}_{\{x_k<s<x_{k+1}, x_{k+1}<t<x_{k+l+1}\}} \mathbb{1}_{\{0<x_1<\cdots<x_{k+l+1}\}} \lambda^{k+l+1} e^{-\lambda x_{k+l+1}} \prod_{i=1}^{k+l+1} dx_i \\
&= \int \mathbb{1}_{\{0<x_1<\cdots<x_k<s<x_{k+1}<\cdots<x_{k+l}<t\}} \lambda^{k+l} e^{-\lambda t} \prod_{i=1}^{k+l} dx_i,
\end{aligned}$$

performing the integration with respect to the variable $z_{k+l+1}$. Since,

$$\int \mathbb{1}_{\{0<x_1<\cdots<x_k<s\}} \prod_{i=1}^{k} dx_i = \frac{s^k}{k!}$$

and likewise,

$$\int \mathbb{1}_{\{s<x_{k+1}<\cdots<x_{k+l}<t\}} \prod_{i=k+1}^{k+l} dx_i = \frac{(t-s)^l}{l!},$$

we get

$$\begin{aligned}
\mathbb{P}(N_s = k, N_t - N_s = l) &= \frac{s^k}{k!} \frac{(t-s)^l}{l!} \lambda^{k+l} e^{-\lambda t} \\
&= \frac{(\lambda s)^k}{k!} e^{-\lambda s} \frac{(\lambda(t-s))^l}{l!} e^{-\lambda(t-s)}.
\end{aligned}$$

Hence, $N_s$ and $N_t - N_s$ are independent, with distributions $\mathcal{P}(\lambda s)$ and $\mathcal{P}(\lambda(t-s))$ respectively. Moreover, $N_t - N_s$ has the same distribution than $N_{t-s}$, hence the stationarity of the increments. □

Example 2.1 (Inspection paradox). We return to the example of the introduction concerning the average waiting time between the arrival of two buses. Suppose that buses arrive at your station according to a Poisson process with rate $\lambda$ equal to 0.1 (i.e. on average one bus every 10 minutes). That is, the arrivals of buses are given by the point process $(T_n)_{n\geq 1}$ such that the increments are independent and exponentially distributed with parameter $\lambda$. The counting process $(N_t)_{t\geq 0}$ therefore counts the number of bus stops at the station. A commuter arrives at time $t$ to the station. The waiting time of the commuter is thus $R_t = T_{N_t+1} - t$. Moreover, let $S_t = t - T_{N_t}$ be the amount of time elapsed since the previous arrival. Then, $R_t$ and $S_t$ are independent : let $r > 0$ and $0 < s < t$, then

$$\begin{aligned}
\mathbb{P}(S_t > s, R_t > r) &= \mathbb{P}(T_{N_t} < t - s, T_{N_t+1} > r + t) \\
&= \mathbb{P}(\text{no arrivals between } t - s \text{ and } r + t) \\
&= \mathbb{P}(N_{r+t} - N_{t-s} = 0) \\
&= \mathbb{P}(N_{r+s} = 0)
\end{aligned}$$

by the stationarity of the increments. Hence,

$$\mathbb{P}(S_t > s, R_t > r) = e^{-\lambda(r+s)} = e^{-\lambda s}e^{-\lambda s}.$$

Thus, $R_t$ and $S_t$ are independent, $R_t$ is exponentially distributed with parameter $\lambda$, and $\mathbb{P}(S_t > s) = e^{-\lambda s}$ for $0 < s < t$, that is $S_t$ has distribution $\min(E, t)$ where $E \sim \mathcal{E}(\lambda)$.

Hence, since $S_t + R_t = T_{N_t+1} - T_{N_t}$, we get that

$$\mathbb{E}(T_{N_t+1} - T_{N_t}) = \mathbb{E}(S_t) + \mathbb{E}(R_t) = \frac{1 - e^{-\lambda t}}{\lambda} + \frac{1}{\lambda} > \frac{1}{\lambda}.$$

In average, the interval $[T_{N_t}, T_{N_t+1}]$ of arrival times between a fixed deterministic time is greater than the average time between two arrivals. This is the inspection paradox: the average intertime between the last passage of the bus before $t$ and the next passage of the bus after $t$ is greater than the average passage time of the buses!

## 3. Processes with independent and stationary increments

It may be natural to model arrivals of clients in a queue by a counting process which enjoys the following properties: the number of clients who arrive in the interval $[s, t]$ is independent of the number of clients arrived before time $s$, and such that the distribution of the number of clients in $[s, t]$ is the same than the distribution of the number of clients in $[0, t - s]$. The last theorem shows that the Poisson process enjoys these two properties. We will see than in fact, this is the only counting process with these two properties.

Definition 3.1. *We say that a stochastic process $(X_t)_{t\geq 0}$ is a process with independent and stationary increments if $X_0 = 0$ and*

*(i) the map $t \mapsto X_t$ is right-continuous;*
*(ii) For all $n \geq 1$, for all $0 = t_0 \leq t_1 \leq \cdots \leq t_n$, the random variables $X_{t_1}, X_{t_2} - X_{t_1}, \ldots, X_{t_n} - X_{t_{n-1}}$ are independent;*
*(iii) For all $s, t \geq 0$, $X_{t+s} - X_s$ has the same distribution than $X_t$.*

Remark 3.1. If we replace right-continuous by càdlàg in the definition, the process $(X_t)_{t\geq 0}$ is called a Lévy process.

Remark 3.2. Note that condition that the increments are independent, i.e. item (ii) is equivalent to: for all $s, t \geq 0$,

$$X_{t+s} - X_s \text{ is independent of } \sigma(X_u, 0 \leq u \leq s),$$

which is, by Dynkin's theorem, equivalent to: for all $0 < t_1 < \cdots < t_k$,

$$X_{t_k} - X_{t_{k-1}} \text{ is independent of } \sigma(X_{t_1}, X_{t_1}, \ldots, X_{t_{k-1}}).$$

We first need to generalize the notion of stopping times to the context of continuous time processes.

Definition 3.2. *A stopping time $T$ (relative to a filtration $(\mathcal{F}_t)_{t\geq 0}$) is a random variable $T\colon \Omega \to [0, +\infty]$ such that for all $t \in [0, +\infty[$,*

$$\{T \leq t\} \in \mathcal{F}_t.$$

*The $\sigma$-algebra $\mathcal{F}_T$ of $T$-past is defined as the collection of sets $A \in \mathcal{F}$ such that for all $t \in [0, +\infty[$,*

$$A \cap \{T \leq t\} \in \mathcal{F}_t.$$

Exercise 3.1. Show that $\mathcal{F}_T$ is indeed a $\sigma$-algebra.

REMARK 3.3. Let $(N_t)_{t\geq 0}$ be a counting process associated to the point process $(T_n)_{n\geq 0}$. From the relation

$$\{T_n \leq t\} = \{N_t \geq n\},$$

we see that for all $n \geq 0$, $T_n$ is a stopping time relative to the natural filtration of $(N_t)_{t\geq 0}$.

LEMMA 3.1. *Any stopping time $T$ is a non-increasing limit of a sequence of stopping times $(T_n)_n$, where $T_n$ is valued in $\left\{\frac{k}{2^n}, k \geq 0\right\} \cup \{+\infty\}$.*

PROOF. It suffices to take

$$T_n = \sum_{k\geq 0} \frac{k+1}{2^n} \mathbb{1}_{\{\frac{k}{2^n} \leq T < \frac{k+1}{2^n}\}} + (+\infty)\mathbb{1}_{\{T=+\infty\}},$$

details are left as an exercise. □

Contrary to discrete time, measurability questions are much more difficult in continuous time. But for instance we have the following:

LEMMA 3.2. *Let $(X_t)_{t\geq 0}$ be a stochastic process which is right-continuous. Then, for all a.s. finite stopping time $\tau$, $X_\tau$ is $\mathcal{F}_\tau$-measurable.*

PROOF. By the previous lemma, there exists a non-increasing sequence $\tau_n$ with values in the dyadic numbers such that $\tau_n \searrow \tau$. Writing,

$$X_{t\wedge\tau_n} = \sum_{\substack{k\geq 0 \\ k2^{-n}\leq t}} X_{k2^{-n}} \mathbb{1}_{\{\tau_n = k2^{-n}\}} + X_t \mathbb{1}_{\{\tau_n > t\}},$$

we see that $X_{t\wedge\tau_n}$ is $\mathcal{F}_\tau$-measurable. As $X_{t\wedge\tau} = \lim_{n\to\infty}(\downarrow) X_{t\wedge\tau}$ by right-continuity, we have that $X_{t\wedge\tau}$ is $\mathcal{F}_\tau$-measurable. Eventually, for all measurable set $A$, we have

$$\{X_\tau \in A\} \cap \{\tau \leq t\} = \{X_{t\wedge\tau} \in A\} \cap \{\tau \leq t\},$$

hence $X_\tau$ is $\mathcal{F}_\tau$-measurable. □

DEFINITION 3.3. *We say that two stochastic processes $(X_t)_{t\geq 0}$ and $(Y_t)_{t\geq 0}$ have the same distribution if they have the same finite-dimensional distributions: for all $n \geq 1$ and all $t_1 < \cdots < t_n$, we have*

$$(X_{t_1}, X_{t_2}, \ldots, X_{t_n}) \overset{(d)}{=} (Y_{t_1}, Y_{t_2}, \ldots, Y_{t_n}),$$

*where $\overset{(d)}{=}$ denotes equality in distribution.*

Processes with independent and stationary increments enjoy the following Markov property:

THEOREM 3.1. *Let $(X_t)_{t\geq 0}$ be a process with independent and stationary increments, then, for all $s \geq 0$, the process $(X_{t+s} - X_s)_{t\geq 0}$ is independent of $\mathcal{F}_s$ and has the same distribution than $(X_t)_{t\geq 0}$.*

This is more or less obvious from the definition, but we provide the proof as it offers useful insight for the proof of the strong Markov property, which we will see next.

PROOF. Denote $Y_t = X_{t+s} - X_s$, for all $s$. Let $0 = t_0 < t_1 < \cdots < t_n$, and $f_1, \ldots, f_n$ bounded measurable functions. Then, since the increments of $(X_t)_{t\geq 0}$ are independent, we have

$$\begin{aligned}
&\mathbb{E}\left[f_1(Y_{t_1})f_2(Y_{t_2}-Y_{t_1})\cdots f_n(Y_{t_n}-Y_{t_{n-1}})\right]\\
&=\mathbb{E}\left[f_1(X_{t_1+s}-X_s)f_2(X_{t_2+s}-X_{t_1+s})\cdots f_n(X_{t_n+s}-X_{t_{n-1}+s})\right]\\
&=\prod_{i=1}^n \mathbb{E}\left[f_i(X_{t_i+s}-X_{t_{i-1}+s})\right]\\
&=\prod_{i=1}^n \mathbb{E}\left[f_i(Y_{t_i}-Y_{t_{i-1}})\right],
\end{aligned}$$

proving that the increments of $(Y_t)_{t\geq 0}$ are also independent. Moreover,

$$Y_{t_i}-Y_{t_{i-1}}=X_{t_i+s}-X_{t_{i-1}+s}\overset{(\mathrm{d})}{=}X_{t_i-t_{i-1}}$$

by the stationarity of the increments of $(X_t)_{t\geq 0}$. Hence, the increments of $(Y_t)_{t\geq 0}$ are independent and stationary, and the processes $(Y_t)_{t\geq 0}$ and $(X_t)_{t\geq 0}$ have the same distribution. $\square$

Now we extend the Markov property to stopping times.

Theorem 3.2. *Let $(X_t)_{t\geq 0}$ be a process with independent and stationary increments and let $T$ be finite stopping time. Then, the process $(X_{t+T}-X_T)_{t\geq 0}$ is independent of $\mathcal{F}_T$ and has the same distribution than $(X_t)_{t\geq 0}$.*

Proof. Let $Y_T = X_{t+T}-X_T$. By proposition 3.1, $T$ can be approximate by a non-increasing sequence $(\tau_n)_{n\geq 1}$ of stopping times with values in the set of dyadic numbers. Then, using the fact that the trajectories are right-continuous, we can write, for $f_1,\dots,f_n$ bounded measurable functions,

$$\begin{aligned}
&\mathbb{E}\left[f_1(Y_{t_1})f_2(Y_{t_2}-Y_{t_1})\cdots f_n(Y_{t_n}-Y_{t_{n-1}})\right]\\
&=\lim_{n\to\infty}\mathbb{E}\left[f_1(X_{t_1+\tau_n}-X_{\tau_n})f_2(X_{t_2+\tau_n}-X_{t_1+\tau_n})\cdots f_n(X_{t_n+\tau_n}-X_{t_{n-1}+\tau_n})\right]\\
&=\lim_{n\to\infty}\sum_{k\geq 0}\mathbb{E}\Big[\mathbb{1}_{\{\tau_n=k2^{-n}\}}f_1(X_{t_1+k2^{-n}}-X_{k2^{-n}})f_2(X_{t_2+k2^{-n}}-X_{t_1+k2^{-n}})\\
&\qquad\cdots f_n(X_{t_n+k2^{-n}}-X_{t_{n-1}+k2^{-n}})\Big],
\end{aligned}$$

where we use the dominated convergence theorem to exchange expectation and limit. Since $\tau_n$ is a stopping time, we have $\{\tau_n=k2^{-n}\}\in\mathcal{F}_{k2^{-n}}$, and by the previous theorem, the process $(X_{t+k2^{-n}}-X_{k2^{-n}})_{t\geq 0}$ is independent of $\mathcal{F}_{k2^{-n}}$ and has the same distribution than $(X_t)_{t\geq 0}$. Hence, we get

$$\begin{aligned}
&\mathbb{E}\left[f_1(Y_{t_1})f_2(Y_{t_2}-Y_{t_1})\cdots f_n(Y_{t_n}-Y_{t_{n-1}})\right]\\
&=\lim_{n\to\infty}\sum_{k\geq 0}\mathbb{E}\left[\mathbb{1}_{\{\tau_n=k2^{-n}\}}\right]E\Big[f_1(X_{t_1+k2^{-n}}-X_{k2^{-n}})f_2(X_{t_2+k2^{-n}}-X_{t_1+k2^{-n}})\\
&\qquad\cdots f_n(X_{t_n+k2^{-n}}-X_{t_{n-1}+k2^{-n}})\Big]\\
&=\lim_{n\to\infty}\sum_{k\geq 0}\mathbb{E}\left[\mathbb{1}_{\{\tau_n=k2^{-n}\}}\right]\mathbb{E}\left[f_1(X_{t_1})f_2(X_{t_2}-X_{t_1})\cdots f_n(X_{t_n}-X_{t_{n-1}})\right]\\
&=\prod_{i=1}^n\mathbb{E}\left[f_i(X_{t_i+s}-X_{t_{i-1}+s})\right],
\end{aligned}$$

since $T<\infty$ a.s. The theorem follows. $\square$

We now state the fundamental result:

THEOREM 3.3. *Let $(T_n)_{n\geq 0}$ be a point process on $[0,+\infty[$ and $(N_t)_{t\geq 0}$ the counting process associated to $(T_n)_{n\geq 0}$. Suppose that $(N_t)_{t\geq 0}$ is a process with independent and stationary increments. Then $(N_t)_{t\geq 0}$ is a Poisson process (with some parameter $\lambda$).*

PROOF. We have to prove that the increments $(T_k - T_{k-1})_{k\geq 1}$ are i.i.d. and exponentially distributed with parameter some $\lambda$. First we show that $T_1 \sim \mathcal{E}(\lambda)$. We have,

$$\mathbb{P}(T_1 > t+s) = \mathbb{P}(N_{t+s} = 0) = \mathbb{P}(N_{t+s} - N_s = 0, N_s = 0).$$

Since the increments of $(N_t)_{t\geq 0}$ are independent and stationary, we have

$$\mathbb{P}(T_1 > t+s) = \mathbb{P}(N_{t+s} - N_s = 0)\,\mathbb{P}(N_s = 0) = \mathbb{P}(N_t = 0)\,\mathbb{P}(N_s = 0) = \mathbb{P}(T_1 > t)\,\mathbb{P}(T_1 > s).$$

Hence, the distribution of $T_1$ has the memoryless property, and thus is the exponential distribution with parameter $\lambda = -\log \mathbb{P}(T_1 > 1)$. Now we use the strong Markov property: since $T_k$ is a stopping time, the process $(Y_t)_{t\geq 0}$ defined by $Y_t = N_{T_k + t} - N_{T_k}$, has the same distribution than $(N_t)_{t\geq 0}$ and is independent of $\mathcal{F}_{T_k}$. Hence, define

$$T_1' = \inf\{t \geq 0 \,|\, Y_t = 1\}.$$

Then, $T_1'$ has the same distribution than $T_1$, thus is exponentially distributed with parameter $\lambda$, and is independent of $\mathcal{F}_{T_k}$. But since $N_{T_k} = k$ by definition of $T_k$, we have

$$\begin{aligned} T_1' &= \inf\{t \geq 0 \,|\, N_{T_k+t} - N_{T_k} = 1\} \\ &= \inf\{t \geq 0 \,|\, N_{T_k+t} = k+1\} \\ &= \inf\{s \geq 0 \,|\, N_s = k+1\} - T_k \\ &= T_{k+1} - T_k. \end{aligned}$$

Hence, $T_{k+1} - T_k$ is exponentially distributed with parameter $\lambda$ and is independent of $\mathcal{F}_{T_k}$, hence independent of the increments $T_k - T_{k-1}, \ldots, T_2 - T_1, T_1$. $\square$