Point Process: Theory and Simulation

Lorenzo Mercuri

University of Milan, Italy
CREST Japan Science and Technology Agency

Brixen, 28 June 2019

Outline

A Point Process

Definition (Point Process)

Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a probability space. Let $(T_i)_{i \in \mathbb{N}_0}$ a sequence of non-negative random variables such that $\forall i \in T_i < T_{i+1}$. We say $(T_i)_{i \in \mathbb{N}_0}$ a Point Process on $\mathbb{R}_+$.

In particular, the variable T_i can represent the times of occurrence of events.

Counting Process and Durations

Definition (Counting Process)

Let $(T_i)_{i \in \mathbb{N}_0}$ be a point process. The right-continuous process

$$N_t = \sum_{i \in \mathbb{N}_0} \mathbf{1}_{T_i \leq t}$$

is called the *counting process* associated with $(T_i)_{i \in \mathbb{N}_0}$

Counting Process and Durations

Definition (Counting Process)

Let $(T_i)_{i\in\mathbb{N}_0}$ be a point process. The right-continuous process

$$N_t = \sum_{i\in\mathbb{N}_0} \mathbf{1}_{T_i \le t}$$

is called the *counting process* associated with $(T_i)_{i\in\mathbb{N}_0}$

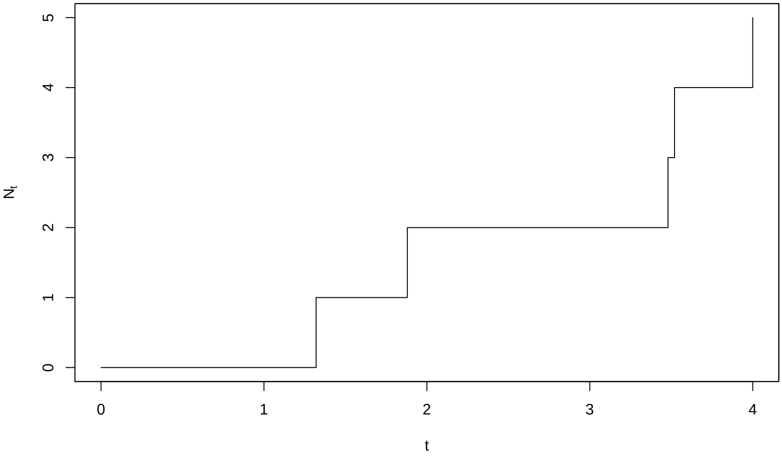
Definition (Duration)

The process ΔT_i defined as:

$$\Delta T_i = T_i - T_{i-1}$$

is called the *duration process* associated with $(T_i)_{i\in\mathbb{N}_0}$

Trajectory of a Point Process 1.



Trajectory of a Point Process 2.

```
##    Duration Point Proc. Count. Proc.
##  0.00000000  0.00000000   0.00000000
##  1.31297803  1.31297803   1.00000000
##  0.55160941  1.86458744   2.00000000
##  1.58521131  3.44979875   3.00000000
##  0.05635054  3.50614928   4.00000000
##  0.45885419  3.96500347   5.00000000
##  1.56435895  5.52936242   6.00000000
```

Intensity 1

Definition (Intensity)

Let N_t be a point process adapted to a filtration $\mathcal{F}_t$. The left-continuous *intensity process* is defined as:

$$\lambda_{t|\mathcal{F}_t} = \lim_{h \to 0} \mathbb{E}\left[\left. \frac{N_{t+h} - N_t}{h} \right| \mathcal{F}_t \right],$$

From hereafter we assume that the filtration is the natural associated with the counting process, denoted $\mathcal{F}_t^N$. We use λ_t instead of $\lambda_{t|\mathcal{F}_t}$

Homogeneous Poisson Process: a short definition

Definition

Poisson Process an homogeneous Poisson Point process is a point process satisfies the following properties:

- $N_0 = 0$
- Stationary and Independent Increments
- $\forall h < t$ the randon variable $N_t - N_h$ is a Poisson with Intensity $\lambda(t-h)$

Homogeneous Poisson Process: some results

Using the previous definition we can see that

$$\mathbf{P}\left[N_{t+h} - N_t = 1 \,|\mathcal{F}_t\right] \quad = \quad \frac{\lambda h e^{-\lambda h}}{1!}$$

Homogeneous Poisson Process: some results

Using the previous definition we can see that

$$\begin{aligned}\mathbf{P}\left[N_{t+h} - N_t = 1 \,|\mathcal{F}_t\right] &= \frac{\lambda h e^{-\lambda h}}{1!} \\ &= \lambda h \left[1 - \lambda h + o\left(h\right)\right]\end{aligned}$$

Homogeneous Poisson Process: some results

Using the previous definition we can see that

$$\begin{aligned}\mathbf{P}\left[N_{t+h} - N_t = 1 \,|\mathcal{F}_t\,\right] &= \frac{\lambda h e^{-\lambda h}}{1!} \\ &= \lambda h\left[1 - \lambda h + o\left(h\right)\right] \\ &= \lambda h + o\left(h\right) \qquad (1)\end{aligned}$$

It means that the probability of a single arrival during a small interval of time h is λh

Homogeneous Poisson Process: some results

Using the previous definition we can see that

$$\begin{aligned} \mathbf{P}\left[N_{t+h} - N_t = 1 \,|\mathcal{F}_t\right] &= \frac{\lambda h e^{-\lambda h}}{1!} \\ &= \lambda h \left[1 - \lambda h + o\left(h\right)\right] \\ &= \lambda h + o\left(h\right) \end{aligned} \tag{1}$$

It means that the probability of a single arrival during a small interval of time h is λh

$$\mathbf{P}\left[N_{t+h} - N_t > 1 \,|\mathcal{F}_t\right] = 1 - \mathbf{P}\left[N_{t+h} - N_t = 0 \,|\mathcal{F}_t\right] - \mathbf{P}\left[N_{t+h} - N_t = 1 \,|\mathcal{F}_t\right]$$

Homogeneous Poisson Process: some results

Using the previous definition we can see that

$$\begin{aligned}\mathbf{P}\left[N_{t+h} - N_t = 1 \,|\mathcal{F}_t\right] &= \frac{\lambda h e^{-\lambda h}}{1!} \\ &= \lambda h \left[1 - \lambda h + o\left(h\right)\right] \\ &= \lambda h + o\left(h\right) \qquad (1)\end{aligned}$$

It means that the probability of a single arrival during a small interval of time h is λh

$$\begin{aligned}\mathbf{P}\left[N_{t+h} - N_t > 1 \,|\mathcal{F}_t\right] &= 1 - \mathbf{P}\left[N_{t+h} - N_t = 0 \,|\mathcal{F}_t\right] - \mathbf{P}\left[N_{t+h} - N_t = 1 \,|\mathcal{F}_t\right] \\ &= 1 - e^{-\lambda h} - \lambda h e^{-\lambda h}\end{aligned}$$

Homogeneous Poisson Process: some results

Using the previous definition we can see that

$$\begin{aligned} \mathbf{P}\left[N_{t+h} - N_t = 1 \,|\mathcal{F}_t\right] &= \frac{\lambda h e^{-\lambda h}}{1!} \\ &= \lambda h \left[1 - \lambda h + o\left(h\right)\right] \\ &= \lambda h + o\left(h\right) \end{aligned} \tag{1}$$

It means that the probability of a single arrival during a small interval of time h is λh

$$\begin{aligned} \mathbf{P}\left[N_{t+h} - N_t > 1 \,|\mathcal{F}_t\right] &= 1 - \mathbf{P}\left[N_{t+h} - N_t = 0 \,|\mathcal{F}_t\right] - \mathbf{P}\left[N_{t+h} - N_t = 1 \,|\mathcal{F}_t\right] \\ &= 1 - e^{-\lambda h} - \lambda h e^{-\lambda h} \\ &= 1 - \left(1 + \lambda h\right) e^{-\lambda h} \end{aligned}$$

Homogeneous Poisson Process: some results

Using the previous definition we can see that

$$\begin{aligned} \mathbf{P}\left[N_{t+h} - N_t = 1 \,|\, \mathcal{F}_t\right] &= \frac{\lambda h e^{-\lambda h}}{1!} \\ &= \lambda h \left[1 - \lambda h + o\left(h\right)\right] \\ &= \lambda h + o\left(h\right) \qquad (1) \end{aligned}$$

It means that the probability of a single arrival during a small interval of time h is λh

$$\begin{aligned} \mathbf{P}\left[N_{t+h} - N_t > 1 \,|\, \mathcal{F}_t\right] &= 1 - \mathbf{P}\left[N_{t+h} - N_t = 0 \,|\, \mathcal{F}_t\right] - \mathbf{P}\left[N_{t+h} - N_t = 1 \,|\, \mathcal{F}_t\right] \\ &= 1 - e^{-\lambda h} - \lambda h e^{-\lambda h} \\ &= 1 - \left(1 + \lambda h\right) e^{-\lambda h} \\ &= 1 - \left(1 + \lambda h\right)\left[1 - \lambda h + o\left(h\right)\right] \end{aligned}$$

Homogeneous Poisson Process: some results

Using the previous definition we can see that

$$\begin{aligned}
\mathbf{P}\left[N_{t+h} - N_t = 1 \,|\, \mathcal{F}_t\right] &= \frac{\lambda h e^{-\lambda h}}{1!} \\
&= \lambda h \left[1 - \lambda h + o\left(h\right)\right] \\
&= \lambda h + o\left(h\right) \qquad (1)
\end{aligned}$$

It means that the probability of a single arrival during a small interval of time h is λh

$$\begin{aligned}
\mathbf{P}\left[N_{t+h} - N_t > 1 \,|\, \mathcal{F}_t\right] &= 1 - \mathbf{P}\left[N_{t+h} - N_t = 0 \,|\, \mathcal{F}_t\right] - \mathbf{P}\left[N_{t+h} - N_t = 1 \,|\, \mathcal{F}_t\right] \\
&= 1 - e^{-\lambda h} - \lambda h e^{-\lambda h} \\
&= 1 - \left(1 + \lambda h\right) e^{-\lambda h} \\
&= 1 - \left(1 + \lambda h\right)\left[1 - \lambda h + o\left(h\right)\right] \\
&= o\left(h\right) \qquad (2)
\end{aligned}$$

The probability of more than a single arrival during a small interval of time is $o\left(h\right)$.

Homogeneous Poisson Process: some results

Using the previous definition we can see that

$$\begin{aligned}
\mathbf{P}\left[N_{t+h} - N_t = 1 \,|\mathcal{F}_t\right] &= \frac{\lambda h e^{-\lambda h}}{1!} \\
&= \lambda h \left[1 - \lambda h + o\left(h\right)\right] \\
&= \lambda h + o\left(h\right) \qquad (1)
\end{aligned}$$

It means that the probability of a single arrival during a small interval of time h is λh

$$\begin{aligned}
\mathbf{P}\left[N_{t+h} - N_t > 1 \,|\mathcal{F}_t\right] &= 1 - \mathbf{P}\left[N_{t+h} - N_t = 0 \,|\mathcal{F}_t\right] - \mathbf{P}\left[N_{t+h} - N_t = 1 \,|\mathcal{F}_t\right] \\
&= 1 - e^{-\lambda h} - \lambda h e^{-\lambda h} \\
&= 1 - \left(1 + \lambda h\right) e^{-\lambda h} \\
&= 1 - \left(1 + \lambda h\right) \left[1 - \lambda h + o\left(h\right)\right] \\
&= o\left(h\right) \qquad (2)
\end{aligned}$$

The probability of more than a single arrival during a small interval of time is $o\left(h\right)$.
The properties (1) and (2) can be used as an alternative definition of a Poisson Process.

Homogeneous Poisson Process: Probability of time arrival

Now we want to establish that in a Poisson Process at least a time arrival T occurs in the interval $(0, t)$. We can write an explicit expression as follows:

$$\begin{aligned} F(t) &= 1 - \mathbf{P}\left(\text{no arrivals before t}\right) \\ &= 1 - \mathbf{P}\left(N_t = 0\right) = 1 - e^{-\lambda t} \end{aligned}$$

Homogeneous Poisson Process: Probability of time arrival

Now we want to establish that in a Poisson Process at least a time arrival T occurs in the interval $(0, t)$. We can write an explicit expression as follows:

$$\begin{aligned} F(t) &= 1 - \mathbf{P}\left(\text{no arrivals before t}\right) \\ &= 1 - \mathbf{P}\left(N_t = 0\right) = 1 - e^{-\lambda t} \end{aligned}$$

$F\left(t\right)$ is the cdf of the time arrivals. Computing the first derivative of $F\left(t\right)$ we get the time arrival density function as:

$$f\left(t\right) = \frac{\partial F\left(t\right)}{\partial t} = \lambda e^{-\lambda t}$$

Homogeneous Poisson Process: Probability of time arrival

Now we want to establish that in a Poisson Process at least a time arrival T occurs in the interval $(0, t)$. We can write an explicit expression as follows:

$$\begin{aligned} F(t) &= 1 - \mathbf{P}\left(\text{no arrivals before t}\right) \\ &= 1 - \mathbf{P}\left(N_t = 0\right) = 1 - e^{-\lambda t} \end{aligned}$$

$F\left(t\right)$ is the cdf of the time arrivals. Computing the first derivative of $F\left(t\right)$ we get the time arrival density function as:

$$f\left(t\right) = \frac{\partial F\left(t\right)}{\partial t} = \lambda e^{-\lambda t}$$

To say that the number of events per time interval follows a Poisson distribution is equivalent to saying that the time between events is exponentially distributed

Homogeneous Poisson Process: Memoryless property.

The Poisson Process has no memory in the sense that the move to a new state depends only upon the current state and is independent of the previous events. In our case:

$$\begin{aligned}
\mathbf{P}\left(T > t_1 + t_2 \,|\, T > t_1\right) &= \frac{\mathbf{P}\left(T > t_1 + t_2 \;\cap\; t > t_1\right)}{\mathbf{P}\left(T > t_1\right)} \\
&= \frac{e^{-\lambda(t_1+t_2)}}{e^{-\lambda t_1}} \\
&= e^{-\lambda t_1} = \mathbf{P}\left(T > t_2\right)
\end{aligned}$$

Homogeneous Poisson Process: Simulation Algorithm.

The simulation algorithm is based on the inversion theorem

Theorem

Let F_X be a strictly increasing CDF. If $u \sim U(0,1)$ and $X = F_X^{-1}(u)$ then X is a random variable with CDF F_X

`Algorithm:`

- Generate $u_i \sim U(0,1)$.
- Set $\Delta t_i = -\frac{-\ln(1-u_i)}{\lambda}$.
- Set $t_i = \sum_{j=1}^{i} \Delta t_i$.

Inhomogeneous Poisson Process: Definition.

Definition (Inhomogeneous Poisson Process)

Let $\lambda_t : \mathbb{R}_+ \to \mathbb{R}_+$ be a positive function, we called N_t a inhomogeneous Poisson process if N_t is a counting process and it satisfies $\forall s < t$ that $N_t - N_s$ is independent of N_s and

$$\mathbf{P}\left(N_{t+h} - N_t = 0 \,|\mathcal{F}_t\right) = 1 - \lambda_t h + o(h)$$

$$\mathbf{P}\left(N_{t+h} - N_t = 1 \,|\mathcal{F}_t\right) = \lambda_t h$$

$$\mathbf{P}\left(N_{t+h} - N_t > 1 \,|\mathcal{F}_t\right) = o(h)$$

Remark. If $\lambda_t = \lambda$ we get the homogeneous Poisson Process as a special case

Point Process Regression Models

In a filtered space $\mathbf{B} := (\Omega, \mathcal{F}, P)$ the d_0-dimensional Point Process Regression Model $Y = (Y_t)_{t \in [t_0, t_1]}$ defined as:

$$Y_t = [X_t, N_t, \lambda_t]^\top \tag{3}$$

where the d_1-dimensional process $X = (X_t)_{t \in [t_0, t_1]}$, denotes covariates, and the $N = (N_t^\alpha)_{t \in [t_0, t_1], \alpha \in \mathcal{I}}, \setminus \mathcal{I} = \{1, \ldots, d\}$, is a d-dimensional counting process with the associated d-dimensional intensity process.

$$\mathrm{d}_0 = \mathrm{d}_1 + 2\mathrm{d}$$

PPR: Covariates

The d-dimensional covariate vector process $X = (X_t)_{t\in[t_0,t_1]}$ satisfies the following system of stochastic differential equations

$$\mathrm{d}X_t = A(t, Y_{t-}, \theta)\,\mathrm{d}t + B(t, Y_{t-}, \theta)\,\mathrm{d}W_t + C(t, Y_{t-}, \theta)\,\mathrm{d}Z_t \quad (4)$$

where $W = (W_T)_{t\in[t_0,t_1]}$ is an s-dimensional standard Wiener process and $Z = (Z_t)_{t\in[t_0,t_1]}$ is an h-dimensional L'evy process of purely discontinuous type. $\theta \in \Theta \subseteq \mathbb{R}^{\mathrm{p}}$ and

$$A : [t_0, t_1] \times \mathbb{R}^{\mathrm{d}_0} \times \Theta \to \mathbb{R}^{\mathrm{d}_1}$$

,

$$B : [t_0, t_1] \times \mathbb{R}^{\mathrm{d}_0} \times \Theta \to \mathbb{R}^{\mathrm{d}_1} \otimes \mathbb{R}^{\mathrm{s}}$$

$$C : [t_0, t_1] \times \mathbb{R}^{\mathrm{d}_0} \times \Theta \to \mathbb{R}^{\mathrm{d}_1} \otimes \mathbb{R}^{\mathrm{h}}$$

PPR: Intensity Process

The d-dimensional vector intensity process λ_t is defined by

$$\lambda_t = g\left(t, Y_{t-}, \theta\right) + \int_{t_0}^{t-} \kappa\left(t-s, Y_{s-}, \theta\right) \mathrm{d} Y_s, \quad t \in \left[t_0, t_1\right], \qquad (5)$$

where $g : [t_0, t_1] \times \mathbb{R}^{\mathsf{d}_0} \times \Theta \to \mathbb{R}_+^{\mathsf{d}}$ and $\kappa : [t_0, t_1] \times \mathbb{R}^{\mathsf{d}_0} \times \Theta \to \mathbb{R}_+^{\mathsf{d} \times \mathsf{d}_0} \subset \mathbb{R}^{\mathsf{d}} \otimes \mathbb{R}^{\mathsf{d}_0}$ are measurable functions. We assume e.g. $\sup_{t \in [t_0, t_1]} |g(t, Y_{t-}, \theta)| < \infty$ and $\sup_{s,t \in [t_0,t_1] : s<t} |\kappa(t-s, Y_{s-}, \theta)| < \infty$ a.s. for each $\theta \in \Theta$, for path-wise integrability of $t \mapsto \lambda_t$.

For each $\theta \in \Theta$, with respect to some filtration, both $g\left(\cdot, Y_{\cdot-}, \theta\right)$ and $\int_{t_0}^{\cdot-} \kappa\left(\cdot - s\cdot, Y_{s-}, \theta\right) dY_s$ are d-dimensional predictable process with non-negative components.

PPR: Counting Processs

The d-dimensional counting process $N = (N_t)_{t \in [t_0, t_1]}$ is characterized by $\lambda = (\lambda_t)_{t \in [t_0, t_1]}$ so that each component of N is a pure jump process with unit jumps and $N - \int_{t_0}^{\cdot} \lambda_s ds$ is a d-dimensional local martingale with respect to a specified filtration.

PPR: Model Classification

There are two different situations based on the way of interaction among the components X, N and its intensity process λ.

- Doubly Stochastic Evolution (PPR-DSE)
- Simultaneous Evolution (PPR-SE)

In the first case there is not the feedback effect of the counting and intensity process in the evolution of the covariates. Therefore Different simulation algorithms are required.

PPR: Doubly Stochastic Evolution

\begin{frame}[fragile]{PPR: Doubly Stochastic Evolution} The d_1-dimensional process $X = (X_t)_{t \in [t_0, t_1]}$ satisfies a stochastic differential equation

$$dX_t = A(t, X_{t-}, \theta)\, dt + B(t, X_{t-}, \theta)\, dW_t + C(t, X_{t-}, \theta)\, dZ_t. \qquad (6)$$

The structure of $\lambda = (\lambda_t)_{t \in [t_0, t_1]}$ to

$$\lambda_t = g(t, Y_{t-}, \theta) + \int_{t_0}^{t-} \kappa(t - s, Y_{s-}, \theta)\, dY_s, \qquad (t \in \mathbb{T}) \qquad (7)$$

The definition (7) admits as a special case the possibility of having a feedback effect of N_t in λ_t. In this case we have a *self-exciting* PPR model meaning that each arrival excites the intensity and increases, for some time period, the probability of subsequent arrivals.

We can simulate separately X_t and N_t. Once the sample path of X_t has been generated, we simulate the time arrivals using the usual scheme used in the Point Process.

Simulation Algorithm: Doubly Stochastic Evolution

Let $N = (N_t)_{t \in [t_0, t_1]}$ be a univariate counting variable in a Point Process Regression model Y_t.
We define the process $\Lambda(t \,|\mathcal{F}_0, T_1, \ldots, T_j)$ as follows

$$\Lambda(t \,|\mathcal{F}_0, T_1, \ldots, T_j) = \int_{T_j}^{t} \lambda_u \mathrm{d}u \tag{8}$$

Using (8), we evaluate the conditional probability of the next random arrival T_{j+1} occurs after $t > T_j$

$$\mathbb{P}\left(T_{j+1} \geq t | \mathcal{F}_0\right) = e^{-\Lambda(t|\mathcal{F}_0, T_1, \ldots, T_j)}.$$

We can use to simulate the arrival T_{j+1} by solving with respect to u the following equation:

$$\ln(\mathcal{U}) + \Lambda(u \,|\mathcal{F}_0, T_1, \ldots, T_j) = 0 \tag{9}$$

where $\mathcal{U} \sim U_{[0,1]}$.

Simulation Algorithm

Notice that the left hand side of (9) is a monotonically increasing differentiable function that starts from the negative value $\ln(\mathcal{U})$ and if the intensity is a strict positive process, we are sure about the existence of the solution u. In general the equation can be solved numerically using Newton-Raphson's algorithm which updates the value of u using the following recursive equation

$$u_{i+1} = u_i - \frac{\ln(\mathcal{U}) + \Lambda(u_i \,|\mathcal{F}_0, T_1, \ldots, T_j)}{\lambda_{u_i}} \tag{10}$$

Simulation Algorithm: Multivariate case

The simulation algorithm based on (9) can be straightforward extended to the multivariate context.
For each component N_t^α of the counting process, we firstly find the T_{k+1}^α as a solution of the equation:

$$\ln(\mathcal{U}_\alpha) + \Lambda^\alpha(u \,|\mathcal{F}_0, T_1, \ldots, T_k) = 0$$

where $\Lambda^\alpha(u \,|\mathcal{F}_0, T_1, \ldots, T_k)$ is the compensator process of the component λ_t^α in the intensity process λ_t. We obtain the next time arrival T_{j+1} as follows:

$$T_{j+1} = \min\left\{T_{j+1}^1, \ldots, T_{j+1}^\alpha, \ldots, T_{j+1}^d\right\}.$$

PPR: Simultaneous Evolution

In this situation we have three different cases:

- Only the Counting process feedbacks the covariates:

$$\mathrm{d}X_t = A\left(t, X_{t-}, N_{t-}, \theta\right)\mathrm{d}t + B\left(t, X_{t-}, N_{t-}, \theta\right)\mathrm{d}W_t + C\left(t, X_{t-}, N_{t-}, \theta\right)\mathrm{d}Z_t. \quad (11)$$

- Only the Intensity feedbacks to covariates:

$$\mathrm{d}X_t = A\left(t, X_{t-}, \lambda_t, \theta\right)\mathrm{d}t + B\left(t, X_{t-}, \lambda_t, \theta\right)\mathrm{d}W_t + C\left(t, X_{t-}, \lambda_t, \theta\right)\mathrm{d}Z_t. \quad (12)$$

- Both counting and intensity processes feedback to covariates:

$$\mathrm{d}X_t = A\left(t, X_{t-}, \lambda_t, N_{t-}, \theta\right)\mathrm{d}t + B\left(t, X_{t-}, \lambda_t, N_{t-}, \theta\right)\mathrm{d}W_t + C\left(t, X_{t-}, \lambda_t, N_{t-}, \theta\right)\mathrm{d}Z_t. \quad (13)$$

In this case it is not possible to simulate separately Counting Process, Intensity and Covariates.

the likelihood function defined as:

$$\mathcal{L}_T(\theta) = \sum_{\alpha=1}^{d} \int_0^T \ln\left(\lambda_s^\alpha\right) \mathrm{d}N_s^\alpha - \sum_{\alpha=1}^{d} \int_0^T \lambda_s^\alpha \mathrm{d}s. \tag{14}$$

For a complete discussion about the optimal properties of the estimates obtained by maimizing the quantity in (14) we refer for the ergodic point process to Ogata (1978), Puri and Tuan (1986) and recently Clinet and Yoshida (2017) while in the context of a non-ergodic point process, Ogihara and Yoshida (2015) derived large sample properties for the maximum likelihood and Bayesian type estimators.

We can also estimate the parameters in the SDE (4) of X with high frequency data of X. If two formulas share some common parameters, we should use the sum of log quasi likelihood functions.