Lin_Masters

Frances Lin

June 2021

Major Professor: James Molyneux

Committee Members: Lisa Madsen & Charlotte Wickham

Abstract

1 Introduction

Hawkes process is also known as a self-exciting point process.

Motivation

Applications

Applications of Hawkes (self-exciting) processes can be found in a wide variety of fields such as seismology, criminology, insurance, finance, social network, and neuroscience.

In seismology, an event can be an earthquake occurrence that causes aftershocks. In criminology, an event can be a gang rivalry that triggers retaliations following the gang crime. In insurance, an event can be a standard claim that increases claims. In finance, an event can be a transcation that influences future prices or volumes of transcations or a news that leads to movements in stock prices or trading behaviors. In social network, an event can be a tweet about an event on Twitter that follows a cascade of retweets from other users on the same social networking platform. In neuroscience, an event can be firing of a neuron that triggers spikes (or action potentials) of other neurons.

Objectives

The objectives of this project is to give an overview of various types of point processes, which include

1. defining and discussing properties of counting process, (homogeneous) Poisson process, nonhomogeneous Poisson process, cluster process, Hawkes process, and spatio-temporal Hawkes process,
2. discussing the algorithms and simulating all of the above processes,
3. extending (temporal) Hawkes process to spatio-temporal Hawkes process, and
4. discussing recent and future work of Hawkes process.

2 Definitions and Properties

2.1 Counting Process

SAY SOMETHING HERE

For example, if we were to count $N(t)$ the occurences (or arrivial) of events such as the number of customers arriving at a supermarket or the number of phone calls receiving at the help line up to some time t, we can consider using counting processes.

Definition 2.1.1 (Stochastic Process) A stochastic process is a family of random variables indexed by time t and is defined as

$$\{X(t), t \in T\}$$

Definition 2.1.2 (Counting Process) Let *N(t)* be the total number of events up to some time t such that the values are nonnegative, interger and nondecreasing, a stocastic process is said to be a counting process and is defined as

$$\{N(t), t \geq 0\}$$

Definition 2.1.3 (Point Process) Let $\{T_i, i \in N\}$ be a sequence of non-negative random variables such that $T_i < T_{i+1}\ \forall i \in N$, a point process on R^+ is defined as

$$\{T_i, i \in N\}$$

Definition 2.1.4 (Counting Process) Let $\{T_i, i \in N\}$ be a point process, a counting process associated with $\{T_i, i \in N\}$ is defined as

$$N(t) = \sum_{i \in N} I_{\{T_i \leq t\}}$$

Corollary **2.1.1** A counting process satisfies that

1. $N(t) \geq 0$
2. $N(t)$ is an integer
3. If $s \leq t$, then $N(s) \leq N(t)$
4. If $s < t$, then $N(t) - N(s)$ is the number of events occur in the interval (s, t]

Proposition 2.1.1 A counting process has the following properties

1. (Independence) If. . . , a counting process is said to have independent increments.

e.g. Poisson process, nonhomogeneous Poisson process

2. (Stationarity) If. . . , a counting process is said to have stationary increments.

e.g. Poisson process

3. (Homogeneity)

Stationarity implies homogeneity.

2.2 Poisson Process

One of the simplest and most-widely used counting process is Poisson process.

In the cases of Poisson process, it is assumed that the arrivial rate is constant. For example, if we were to model the number of bus arrivals at a bus stop, the number of car accidents at a site or the requests for documents on a web server, we can consider using Poisson processes.

Definition 2.2.1 (Poisson Process) If the following conditions hold, a counting process $\{N(t), t \geq 0\}$ is said to be a Poisson Process with constant rate (or intensity) $\lambda > 0$

1. $N(0) = 0$
2. $N(t)$ has independent increments
3. $P(N(t+h)) - N(t) = 1)) = \lambda h + o(h)$
4. $P(N(t+h)) - N(t) \geq 2)) = o(h)$

In other words, 1. the process starts at $t = 0$, 2. the increments are independent, 3. λ is the rate (or intensity), and 4. no 2 or more events can occur at the same location.

Proposition 2.2.1 Poisson process has the following properties

1. $N(t)$, the number of events in any interval t, $\sim Pos(\lambda t)$. That is, for all $s, t \geq 0$

$$P(N(t+s) - N(s) = n) = \frac{(\lambda t)^n e^{-\lambda t}}{n!}$$

$n = 0, 1, ...$

2. W, the interarrival times, $\overset{iid}{\sim} exp(\frac{1}{\lambda})$. That is, for rate $\lambda > 0$, the interarrival time W_i $i = 1, 2, ...$

$$P(W_1 > t) = P(N(t) = 0) = e^{-\lambda t}$$

This is because p(1^{st} arrivial arrives after time t) is the same as p(no arrivial in the interval [0,t]),

$$F(t) = P(W_1 \leq t) = 1 - e^{-\lambda t}$$

and

$$f(t) = \lambda e^{-\lambda t}.$$

Similarly,

$$P(W_2 > t | W_1 = s) = P(N(t+s) - N(s)) = P(N(t) = 0) = e^{-\lambda t}.$$

2.3 Nonhomogeneous Poisson Process

In the cases of counting process and Poisson process, we assume that the arrivial rate is constant. If we have reasons to believe that the arrivial rate is not constant, we should consider using nonhomogeneous Poisson processes.

For example, if we were to model the number of customers arriving at a supermarket and we have reasons to believe that the arrivial rate of customers is higher during lunch time as compared to 2pm, we should consider using nonhomogeneous Poisson processes.

Put it more formally, (homogeneous) Poisson process has stationary increments since the distribution of the number of events $N(t)$ that occur in any interval of time t depends only on the length of the time interval but not the location of the interval. In contrast, nonhomogeneous Poisson process does not have stationary increments since the distribution of $N(t)$ can change when shifted in t.

Definition 2.3.1 (Nonhomogeneous Poisson Process) If the following conditions hold, a counting process $\{N(t), t \geq 0\}$ is said to be a nonhomogeneous Poisson Process with intensity function of time $\lambda(t), t > 0$

1. $N(0) = 0$
2. $N(t)$ has independent increments
3. $P(N(t+h)) - N(t) = 1)) = \lambda(t)h + o(h)$
4. $P(N(t+h)) - N(t) \geq 2)) = o(h)$

Proposition 2.2.1 Nonhomogeneous Poisson process has the following properties

1. $N(t)$, the number of events in any interval t, $\sim Pos(\int_0^t \lambda(s)ds)$. That is, for all $s, t \geq 0$

$$P(N(t) = n) = \frac{(\int_0^t \lambda(s)ds)^n e^{\int_0^t \lambda(s)ds}}{n!}$$

or

$$P(N(b) - N(a) = n) = \frac{(\int_a^b \lambda(s)ds)^n e^{\int_a^b \lambda(s)ds}}{n!}$$

$n = 0, 1, ...$

2.4 Cluster Process

In the previous cases, whether the events arrive at a constant rate λ (e.g. Poisson process) or depend on an intensity function $\lambda(t)$ (e.g. nonhomogeneous Poisson Process), they arrive independently. Here, there are cases in which the events do not arrive independently.

For example,

Types of cluster process include Cox process (or doubly stochastic Poisson process), Matérn I process and Matérn II process.

2.5 Hawkes Process

Hawkes process is also known as a self-exciting point process.

In the cases of Hawkes process, the events also do not arrive independently. The arrivial rate of the events depends not only on time t but also past events $\mathcal{H}_t^N$.

Examples that can be modelled using Hawkes process include… In these examples, the occurrence of an event increases the occurrence of future events.

Definition 2.5.1 (Hawkes Process) A counting process $\{N(t), t \geq 0\}$ associated with past events $\{\mathcal{H}_t^N, t > 0\}$ is said to be a Hawkes process with conditional intensity function $\lambda(t|\mathcal{H}_t^N), t > 0$ and takes the form

$$\lambda(t|\mathcal{H}_t^N) = \lambda_0(t) + \sum_{i:T_i<t} \phi(t - T_i)$$

where

- $\lambda_0(t)$ is the base intensity function (or μ the constant background rate)
- $T_i < t$ are the events time occur before current time t
- $\phi(\cdot)$ is the kernel function (or $g(\cdot)$ the triggering function) through which intensity function depends on past events
- $\mathcal{H}_t^N$ is the natural filration (or simply $\mathcal{H}_t$ the past history) which represents the internal history of N up to time t

Corollary **2.5.1** Hawkes process satisfies that

1. $N(t) = 0$
2. $\lambda(t|\mathcal{H}_t^N) = \lambda_0(t) + \int_{-\infty}^{t} \phi(t - T_i)dN(s) = \lambda_0(t) + \sum_{i:T_i<t} \phi(t - T_i)$
3. $P(N(t + h)) - N(t) = 1)|\mathcal{H}_t^N) = \lambda(t)h + o(h)$
4. $P(N(t + h)) - N(t) \geq 2)|\mathcal{H}_t^N) = o(h)$

2.5.1 Choices of $\phi(\cdot)$ include, for example, exponentially decaying function and power-law kernel, and they take the form of

$$\phi(x) = \alpha e^{-\beta x}$$

$$\phi(x) = \frac{\alpha}{(x + \beta)^{\eta+1}}$$

2.5.2 There are two ways to view Hawkes processes

1. Intensity-based Hawkes Process
2. Cluster-based Hawkes Process

2.5 Spatio-Temporal Hawkes Process

Spatio-temporal Hawkes processes is an extention of temporal Hawkes processes. Recall that temporal Hawkes processes take the form of

$$\lambda(t|\mathcal{H}_t) = \mu + \sum_{i:T_i<t} g(t - t_i)$$

Spatio-temporal Hawkes processes take the form of

$$\lambda(t|\mathcal{H}_t) = \mu(s) + \sum_{i:T_i<t} g(s - s_i, t - t_i)$$

where

- $s_i, i = 1, 2, ...$ are the sequence of locations of events
- $t_i, i = 1, 2, ...$ are the times of the events

Next, we

3 Algorithms and Simulations

3.1 Counting Process

3.2 Poisson Process

3.3 Nonhomogeneous Poisson Process

3.4 Cluster Process

3.5 Hawkes Process

4 Conclusions and Discussion

Acknowledgments

Reference

Chen, Y. (2016). Thinning algorithms for simulating point processes. Florida State University, Tallahassee, FL.

Obral, K. (2016). Simulation, estimation and applications of hawkes processes. (Master's thesis, University of Minnesota, Twin Cities, United States).

Rizoiu, M. A., Lee, Y., Mishra, S., & Xie, L. (2017). A tutorial on hawkes processes for events in social media. arXiv preprint arXiv:1708.06401.

Terminology

Counting Processes

Poisson Process

Nonhomogeneous Poisson Processes