



Study of discrete-time Hawkes process and its compensator

Utpal Jyoti Deba Sarma, Dharmaraja Selvamuthu *

Department of Mathematics, Indian Institute of Technology Delhi, New Delhi, 110016, Delhi, India

ARTICLE INFO

MSC:

60G55

60F05

60F15

Keywords:

Continuous-time Hawkes process

Self-exciting processes

Discrete-time Hawkes process

Compensator

Central limit theorem

Law of large numbers

ABSTRACT

The discrete-time Hawkes process (DTHP) is a sub-class of g -functions that serves as a discrete-time version of the continuous-time Hawkes process (CTHP). Like the CTHP, the DTHP also has the self-exciting property and its intensity depends on the entire history. In this paper, we study the asymptotic behavior of the DTHP and its compensator. We further analyze the moment generating function (MGF) of the DTHP and obtain some bounds and convergence results on the scaled logarithmic MGF of the DTHP.

1. Introduction

The CTHP is a self-exciting point process, and was originally developed by Hawkes in 1971. Hawkes (1971) initially studied the linear CTHP, presenting it through an immigration-birth model. The central limit theorem (CLT), law of large numbers (LLN), large deviations principle (LDP), moderate deviations principle (MDP), etc. have been thoroughly studied by many authors (see, for instance, Daley et al. (2003), Bordenave and Torrisi (2007), Zhu (2013b) and Bacry et al. (2013)).

The nonlinear CTHP, first proposed by Brémaud and Massoulié (1996), are more challenging to study due to the absence of an immigration-birth representation (also called the Galton–Watson theory) and challenges in computational tractability. Nonetheless, there have been attempts to explore these processes as well.

CTHP are utilized across various fields, including finance, DNA modeling, seismology and neuroscience, owing to their clustering and self-exciting characteristics. In finance, these processes are particularly used for evaluating credit derivatives and modeling correlated defaults (see Dassios and Zhao (2011) and Errais et al. (2010)). Despite their broad applicability, CTHP may not be suitable for scenarios where data is captured in discrete time or aggregated formats. In such instances, the DTHP becomes a more relevant tool. Consequently, a discrete-time version of the CTHP has been developed to accommodate these specific data recording methods.

Throughout the paper, we shall define all the stochastic processes on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

Organisation of the paper. Section 2 provides the definitions and literature survey of CTHP, DTHP and the compensator process. Section 3 focuses on deriving the intensity and compensator of the DTHP. In Section 4, we present and prove the strong law of large numbers (SLLN) for DTHP. Section 5 discusses the SLLN, weak law of large numbers (WLLN) and CLT for the compensator of DTHP, along with its proofs. Section 6 examines the MGF of the DTHP and studies the convergence of the scaled logarithmic MGF. The paper concludes in Section 7, summarizing the findings and discussions of the study, and outlines potential directions for future research and extensions of the work.

* Corresponding author.

E-mail address: dharmar@maths.iitd.ac.in (D. Selvamuthu).

2. Preliminaries and literature review

In this section, we present a brief description of CTHP, DTHP, compensator and review their limit theorems.

Continuous-time Hawkes process. Consider a point process $N = \{N_t : t \geq 0\}$ on $(0, \infty)$, where N_t represents the number of arrivals of events in the interval $(0, t]$. Define $\mathcal{F}_t = \sigma(N(D), D \in B((0, \infty)), D \subset (0, t])$, which denotes an increasing family of σ -fields. A non-negative process λ_t that is measurable w.r.t the filtration $\mathcal{F}_t$ and satisfies

$$\mathbb{E}(N(u, v] | \mathcal{F}_u) = \mathbb{E}\left(\int_u^v \lambda_s ds | \mathcal{F}_u\right),$$

for every interval $(u, v]$, almost surely (a.s.), is called an intensity of N . A CTHP is defined as a point process N that possesses an intensity of the form

$$\lambda_t := \psi\left(\int_0^t g(t-s) dN_s\right),$$

where $\psi : (0, \infty) \rightarrow (0, \infty)$ is a left-continuous and locally integrable function, and $g : (0, \infty) \rightarrow (0, \infty)$ is a function with a finite L^1 norm, i.e., $\|g\| = \int_0^\infty g(t) dt < \infty$. The process starts without prior history, denoted by $N(-\infty, 0] = 0$. The functions $\psi(\cdot)$ and $g(\cdot)$ are often called the rate function and exciting function, respectively. Depending on whether $\psi(\cdot)$ is linear or nonlinear, the CTHP is classified as linear or nonlinear respectively.

Discrete-time Hawkes process. Let $(\beta_i)_{i=0}^\infty$ be a sequence of positive real numbers. Consider the following assumptions on this sequence.

1. $\sum_{i=0}^\infty \beta_i < 1$.
2. $\sqrt{n} \sum_{i=n}^\infty \beta_i \rightarrow 0$ as $n \rightarrow \infty$.
3. $\frac{1}{\sqrt{n}} \sum_{i=1}^n i\beta_i \rightarrow 0$ as $n \rightarrow \infty$.
4. $\sum_{i=1}^\infty i\beta_i < \infty$.

Define $\{\xi_n | n = 1, 2, \dots\}$ as a sequence of random variables having image on the set $\{0, 1\}$ as follows.

- $\mathbb{P}(\xi_1 = 1) = \beta_0$; $\mathbb{P}(\xi_1 = 0) = 1 - \beta_0$,
- $\mathbb{P}(\xi_2 = 1 | \xi_1) = \beta_0 + \beta_1 \xi_1$; $\mathbb{P}(\xi_2 = 0 | \xi_1) = 1 - (\beta_0 + \beta_1 \xi_1)$,
- For $n \geq 3$, $\mathbb{P}(\xi_n = 1 | \xi_1, \dots, \xi_{n-1}) = \beta_0 + \sum_{i=1}^{n-1} \beta_{n-i} \xi_i$
 $\mathbb{P}(\xi_n = 0 | \xi_1, \dots, \xi_{n-1}) = 1 - (\beta_0 + \sum_{i=1}^{n-1} \beta_{n-i} \xi_i)$.

Define

$$H_n = \sum_{i=1}^n \xi_i.$$

$\{H_n | n = 1, 2, \dots\}$ is called the DTHP in the literature (see [Seol \(2015\)](#)).

Limit theorems on CTHP. In the literature, the applications of CTHP are mainly focused on the linear case. For the linear case since $\psi(\cdot)$ is linear, specifically $\psi(x) = x + \nu$ for some $\nu > 0$ and $\|g\| < 1$, one can examine the limit theorems using a very beautiful immigration-birth representation. [Daley et al. \(2003\)](#) proved the LLN for linear CTHP. The functional CLT for the linear multi-variate CTHP was proved under specific assumptions by [Bacry et al. \(2013\)](#). This proved the CLT for linear CTHP as well. [Bordenave and Torrisi \(2007\)](#) proved that if $\int_0^\infty g(t)tdt < \infty$ and $\mu \in (0, 1)$, then $\left(\frac{N_t}{t} \in \cdot\right)$ satisfies the LDP, characterized by the rate function $I(x)$, where

$$I(x) = \begin{cases} x \left(\|g\| + \log\left(\frac{x}{x\|g\| + \nu}\right) - 1 \right) + \nu & \text{if } x \geq 0 \\ +\infty & \text{otherwise} \end{cases}.$$

The MDP for linear CTHP is studied by [Zhu \(2013b\)](#).

For nonlinear CTHP, [Brémaud and Massoulié \(1996\)](#) proved that under specific conditions, a stationary version of the nonlinear CTHP exists, which is unique. Furthermore, they established the convergence of a non-stationary version to equilibrium, in terms of both distribution and variation. Afterwards, Zhu conducted extensive studies on both linear and nonlinear CTHP, as seen in his works, [Zhu \(2014, 2013a,b,c\)](#). The CLT for nonlinear CTHP was obtained by [Zhu \(2013a\)](#). [Zhu \(2015\)](#) established the LDP for a particular class of nonlinear CTHP, when the exciting function takes sum of exponentials (including exponential) form. [Zhu \(2014\)](#) also proved a level-3 LDP (also called a process-level LDP) for nonlinear CTHP with a general form of the exciting function, and subsequently derived the level-1 LDP using the contraction principle.

Limit theorems for DTHP. Substantial progress has been made in studying the CTHP, however, DTHP is relatively less studied. Seol (2015) showed a WLLN for the DTHP, i.e.,

$$\frac{H_n}{n} \rightarrow \frac{\beta_0}{1 - \sum_{i=1}^{\infty} \beta_i} \text{ as } n \rightarrow \infty \text{ in probability.} \quad (1)$$

Using the assumptions 2 and 3 of DTHP, Seol (2015) also derived the CLT as follows,

$$\frac{H_n - n\mu}{\sqrt{n}} \rightarrow \mathcal{N}(0, \sigma^2) \text{ as } n \rightarrow \infty \text{ in distribution,} \quad (2)$$

where

$$\mu := \frac{\beta_0}{1 - \sum_{i=1}^{\infty} \beta_i} \text{ and } \sigma^2 := \frac{(1 - \mu)\mu}{\left(1 - \sum_{j=1}^{\infty} \beta_j\right)^2}.$$

(Wang, 2022) studied a marked DTHP introduced in Xu et al. (2021), focusing on a more general exciting function and deduced the LLN and CLT for the model. The MDP and LDP for the model are studied in Wang (2023).

Other related literature on CTHP. Authors have recently started studying different modifications of CTHP. Once such modification is the extended inverse CTHP. The inverse CTHP is characterized by a fixed intensity (in contrast to the CTHP that has stochastic intensity), and jump size, which is stochastic in nature. An extended inverse CTHP is formed through the combination of a CTHP and an inverse CTHP. Selvamuthu et al. (2023) analyzed the limiting properties of this extended inverse CTHP, with a general exciting function, and they derived the CLT the LLN.

Seol (2019) examined the inverse Markovian CTHP, combining features of various self-exciting processes. He specifically conducted an asymptotic analysis for a variant of the inverse Markovian CTHP.

Compensator of a point process. For a point process N , its compensator, denoted by Λ_t , is defined as the unique increasing $\mathcal{F}_t$ -predictable process starting from zero, and satisfying the equation

$$N_t = M_t + \Lambda_t,$$

a.s. for all $t \geq 0$, where M_t is a martingale. The Doob's decomposition theorem assures the presence of such an M_t .

In the literature, authors have studied the compensator of point processes, especially within the context of CTHP. The compensator is crucial in analyzing the dynamical behaviors of point processes because the point process and its compensator is separated only by a martingale. For the importance of compensator, refer to Daley et al. (2003), Guo and Zeng (2008) and Ivanoff et al. (2007). Seol (2017) examined the compensator of linear CTHP, and proved the LLN, CLT and the LDP for the compensator process.

3. Intensity and compensator of DTHP

In this section, we introduce the intensity of the DTHP. Then, we derive the compensator of the DTHP.

Intensity. Let $\mathcal{F}_n = \sigma(\xi_1, \dots, \xi_n)$, i.e., $\mathcal{F}_n$ is the σ -field generated by $\xi_1, \dots, \xi_n$.

From the DTHP described in Section 2, we can conclude that $\lambda_n := \beta_0 + \sum_{i=1}^{n-1} \beta_{n-i} \xi_i$ is the intensity of the DTHP $\{H_n \mid n = 1, 2, \dots\}$. The sequence $(\beta_i)_{i=0}^{\infty}$ plays the role of the exciting function for the DTHP. The arrival of each event at discrete time n increases the intensity, establishing the self-exciting nature of the DTHP. Also, the intensity of the DTHP depends on its entire history. All these characteristics make it a discrete-time equivalent of the CTHP.

Compensator. Now, we show that the DTHP $\{H_n \mid n = 1, 2, \dots\}$ defined in Section 2 is a sub-martingale w.r.t. the filtration $\{\mathcal{F}_n \mid n = 1, 2, \dots\}$. Then we use Doob's decomposition theorem to get the compensator of $\{H_n \mid n = 1, 2, \dots\}$.

Lemma 3.1. *The DTHP is a sub-martingale w.r.t. the filtration $\{\mathcal{F}_n \mid n = 1, 2, \dots\}$.*

Proof. From the definition of $\{H_n \mid n = 1, 2, \dots\}$,

$$|H_n| = \left| \sum_{i=1}^n \xi_i \right| \leq n.$$

Thus, H_n is bounded for each $n = 1, 2, \dots$, which implies H_n is integrable for each $n = 1, 2, \dots$. Also, from the definition, H_n is $\mathcal{F}_n$ -measurable.

Finally

$$\mathbb{E}(H_{n+1} \mid \mathcal{F}_n) = \mathbb{E}(H_n + \xi_{n+1} \mid \mathcal{F}_n) = H_n + \mathbb{E}(\xi_{n+1} \mid \mathcal{F}_n) \geq H_n.$$

Hence, the DTHP is a sub-martingale w.r.t. the filtration $\{\mathcal{F}_n \mid n = 1, 2, \dots\}$. $\square$

Since the DTHP is a sub-martingale, by Doob's decomposition, it can be written as a sum of an increasing predictable process and a martingale, uniquely, i.e.,

$$H_n = M_n + \Lambda_n, \quad (3)$$

where $M_n = \sum_{i=1}^n [\xi_i - \mathbb{E}(\xi_i | \mathcal{F}_{i-1})]$ is a martingale and $\Lambda_n = \sum_{i=1}^n \mathbb{E}(\xi_i | \mathcal{F}_{i-1})$ is an increasing predictable process. The fact that M_n is a martingale and Λ_n is an increasing predictable process is verified as follows.

Claim 3.1. $\{M_n | n = 1, 2, \dots\}$ is a martingale w.r.t. the filtration $\{\mathcal{F}_n | n = 1, 2, \dots\}$.

Proof. From the definition of M_n , M_n is $\mathcal{F}_n$ -measurable.

Also

$$|M_n| = \left| \sum_{i=1}^n (\xi_i - \mathbb{E}(\xi_i | \mathcal{F}_{i-1})) \right| \leq \sum_{i=1}^n |\xi_i - \mathbb{E}(\xi_i | \mathcal{F}_{i-1})| \leq n.$$

Hence, M_n is bounded for each $n = 1, 2, \dots$, which implies that M_n is integrable for each $n = 1, 2, \dots$.

Finally

$$\mathbb{E}(M_{n+1} | \mathcal{F}_n) = \mathbb{E} \left(\sum_{i=1}^{n+1} [\xi_i - \mathbb{E}(\xi_i | \mathcal{F}_{i-1})] | \mathcal{F}_n \right) = \sum_{i=1}^n [\xi_i - \mathbb{E}(\xi_i | \mathcal{F}_{i-1})] + \mathbb{E}(\xi_{n+1} - \mathbb{E}(\xi_{n+1} | \mathcal{F}_n) | \mathcal{F}_n) = M_n.$$

Hence, $\{M_n | n = 1, 2, \dots\}$ is a martingale w.r.t. the filtration $\{\mathcal{F}_n | n = 1, 2, \dots\}$. $\square$

Remark 3.1. We assume for convenience that $M_0 = 0$.

Claim 3.2. $\{\Lambda_n | n = 1, 2, \dots\}$ is an increasing predictable process.

Proof. From the definition of Λ_n , we have

$$\Lambda_{n+1} - \Lambda_n = \sum_{i=1}^{n+1} \mathbb{E}(\xi_i | \mathcal{F}_{i-1}) - \sum_{i=1}^n \mathbb{E}(\xi_i | \mathcal{F}_{i-1}) = \mathbb{E}(\xi_{n+1} | \mathcal{F}_n) \geq 0.$$

Also, from the definition, Λ_n is $\mathcal{F}_{n-1}$ -measurable, which implies it is a predictable process.

Hence, $\{\Lambda_n | n = 1, 2, \dots\}$ is an increasing predictable process. $\square$

This process $\{\Lambda_n | n = 1, 2, \dots\}$ is the compensator of the DTHP $\{H_n | n = 1, 2, \dots\}$.

4. SLLN for DTHP

In this section, we obtain the SLLN for the DTHP $\{H_n | n = 1, 2, \dots\}$ defined in Section 2.

Theorem 4.1. Under assumptions 1, 2, 3 and 4,

$$\frac{H_n}{n} \rightarrow \frac{\beta_0}{1 - \sum_{i=1}^{\infty} \beta_i}, \text{ as } n \rightarrow \infty \text{ a.s.}$$

Proof of Theorem 4.1. From the definition of $\{\xi_n | n = 1, 2, \dots\}$,

$$\mathbb{E}(\xi_i | \mathcal{F}_{i-1}) = 0 \cdot \mathbb{P}(\xi_i = 0 | \mathcal{F}_{i-1}) + 1 \cdot \mathbb{P}(\xi_i = 1 | \mathcal{F}_{i-1}) = \beta_0 + \sum_{i=1}^{n-1} \beta_{n-i} \xi_i.$$

Now

$$M_n = \sum_{i=1}^n [\xi_i - \mathbb{E}(\xi_i | \mathcal{F}_{i-1})] = H_n - \beta_0 n - \sum_{i=1}^n \sum_{j=1}^{i-1} \beta_{i-j} \xi_j.$$

After some computation, the expression $\sum_{i=1}^n \sum_{j=1}^{i-1} \beta_{i-j} \xi_j$ can be rewritten as $\sum_{i=1}^{n-1} \sum_{j=1}^{n-i} \beta_j \xi_i$.

Therefore,

$$M_n = H_n - \beta_0 n - \sum_{i=1}^{n-1} \sum_{j=1}^{n-i} \beta_j \xi_i = \left(1 - \sum_{j=1}^{\infty} \beta_j\right) H_n - \beta_0 n + \left(\sum_{j=1}^{\infty} \beta_j\right) H_n - \sum_{i=1}^{n-1} \sum_{j=1}^{n-i} \beta_j \xi_i.$$

It can be rewritten as

$$M_n = \left(1 - \sum_{j=1}^{\infty} \beta_j\right) H_n - \beta_0 n + \zeta_n, \quad (4)$$

where

$$\zeta_n = \left(\sum_{j=1}^{\infty} \beta_j \right) H_n - \sum_{i=1}^{n-1} \sum_{j=1}^{n-i} \beta_j \xi_i.$$

Lemma 4.1.

$$\frac{\zeta_n}{n} \rightarrow 0, \text{ as } n \rightarrow \infty \text{ a.s.}$$

Proof.

$$\begin{aligned} \zeta_n &= \left(\sum_{j=1}^{\infty} \beta_j \right) H_n - \sum_{i=1}^{n-1} \sum_{j=1}^{n-i} \beta_j \xi_i \\ &= \left(\sum_{j=1}^{\infty} \beta_j \right) H_n - \left[\left(\sum_{i=1}^{n-1} \beta_i \right) \xi_1 + \left(\sum_{i=1}^{n-2} \beta_i \right) \xi_2 + \cdots + \left(\sum_{i=1}^2 \beta_i \right) \xi_{n-2} + \beta_1 \xi_{n-1} \right] \\ &= \beta_1 \xi_n + \beta_2 \left(\sum_{i=0}^1 \xi_{n-i} \right) + \beta_3 \left(\sum_{i=0}^2 \xi_{n-i} \right) + \cdots + \beta_n \left(\sum_{i=0}^{n-1} \xi_{n-i} \right) + \left(\sum_{j=n}^{\infty} \beta_j \right) H_n \\ &\leq \beta_1 + 2\beta_2 + 3\beta_3 + \cdots + n\beta_n + n \left(\sum_{j=n+1}^{\infty} \beta_j \right) \leq \sum_{j=1}^{\infty} j\beta_j. \end{aligned}$$

Hence, we have

$$\frac{\zeta_n}{n} \leq \frac{\sum_{j=1}^{\infty} j\beta_j}{n} \quad \forall n = 1, 2, \dots \quad (5)$$

By assumption 4 of the DTHP described in Section 2, and the fact that $\zeta_n \geq 0 \quad \forall n = 1, 2, \dots$, we get

$$\frac{\zeta_n}{n} \rightarrow 0, \text{ as } n \rightarrow \infty \text{ a.s.} \quad \square$$

Lemma 4.2.

$$\frac{M_n}{n} \rightarrow 0, \text{ as } n \rightarrow \infty \text{ a.s.}$$

Proof. Let

$$A_n = M_n - M_{n-1} = \xi_n - \mathbb{E}(\xi_n \mid \mathcal{F}_{n-1}).$$

This implies

$$|A_n| \leq 1 \text{ and } \mathbb{E}(A_n \mid \mathcal{F}_{n-1}) = 0.$$

Also, for any $n_1 < n_2 < \cdots < n_k \in \mathbb{N}$,

$$\mathbb{E}(A_{n_1} A_{n_2} \cdots A_{n_k}) = \mathbb{E}[\mathbb{E}(A_{n_1} A_{n_2} \cdots A_{n_k} \mid \mathcal{F}_{n_k-1})] = \mathbb{E}[A_{n_1} A_{n_2} \cdots A_{n_{k-1}} \mathbb{E}(A_{n_k} \mid \mathcal{F}_{n_k-1})] = 0.$$

Let

$$B_n = \sum_{i=1}^n \frac{A_i}{i}.$$

By definition, B_n is integrable for each $n = 1, 2, \dots$, and B_n is $\mathcal{F}_n$ -measurable.

Also,

$$\mathbb{E}(B_{n+1} \mid \mathcal{F}_n) = \sum_{i=1}^n \frac{A_i}{i} + \mathbb{E} \left(\frac{A_{n+1}}{n+1} \mid \mathcal{F}_n \right) = B_n.$$

Thus, $\{B_n \mid n = 1, 2, \dots\}$ is a martingale w.r.t the filtration $\{\mathcal{F}_n \mid n = 1, 2, \dots\}$. Now,

$$\mathbb{E}(B_n^2) = \sum_{i=1}^n \frac{\mathbb{E}(A_i^2)}{i^2} + 2 \sum_{i < j} \frac{\mathbb{E}(A_i A_j)}{ij} \leq \sum_{i=1}^n \frac{1}{i^2}.$$

Thus,

$$\sup_{n \in \mathbb{N}} \mathbb{E}(B_n^2) < +\infty.$$

Using Jensen's inequality, $\{B_n^2 \mid n = 1, 2, \dots\}$ is a sub-martingale w.r.t the filtration $\{\mathcal{F}_n \mid n = 1, 2, \dots\}$. Hence, by Doob's Martingale Convergence Theorem, there exists an integrable random variable B such that

$$B_n^2 \rightarrow B \text{ as } n \rightarrow \infty \text{ a.s.}$$

This implies

$$|B_n| \rightarrow \sqrt{B} \text{ as } n \rightarrow \infty \text{ a.s.}$$

Thus,

$$\left| \sum_{i=1}^{\infty} \frac{A_i}{i} \right| < \infty \text{ a.s., which implies } \sum_{i=1}^{\infty} \frac{A_i}{i} < \infty \text{ a.s.}$$

Hence, by Kronecker's lemma,

$$\frac{1}{n} \sum_{i=1}^n i \frac{A_i}{i} \rightarrow 0 \text{ as } n \rightarrow \infty \text{ a.s.,}$$

which implies, using Remark 3.1

$$\frac{M_n}{n} \rightarrow 0 \text{ as } n \rightarrow \infty \text{ a.s.} \quad \square \quad (6)$$

Corollary 4.1. Since convergence a.s. implies convergence in probability, hence, from convergence relation (6) we can conclude that

$$\frac{M_n}{n} \rightarrow 0, \text{ as } n \rightarrow \infty \text{ in probability.} \quad (7)$$

From Eq. (4), we get

$$\frac{M_n}{n} = \left(1 - \sum_{j=1}^{\infty} \beta_j \right) \frac{H_n}{n} - \beta_0 + \frac{\zeta_n}{n}.$$

Using Lemmas 4.1 and 4.2, letting $n \rightarrow \infty$ on both sides we get

$$\frac{H_n}{n} \rightarrow \frac{\beta_0}{1 - \sum_{j=1}^{\infty} \beta_j}, \text{ as } n \rightarrow \infty \text{ a.s.} \quad (8)$$

5. LLN and CLT for the compensator of DTHP

In this section, we obtain the SLLN, WLLN and CLT for the compensator of the DTHP.

Theorem 5.1 (SLLN).

$$\frac{\Lambda_n}{n} \rightarrow \frac{\beta_0}{1 - \sum_{i=1}^{\infty} \beta_i}, \text{ as } n \rightarrow \infty \text{ a.s.}$$

Theorem 5.2 (WLLN).

$$\frac{\Lambda_n}{n} \rightarrow \frac{\beta_0}{1 - \sum_{i=1}^{\infty} \beta_i}, \text{ as } n \rightarrow \infty \text{ in probability.}$$

Theorem 5.3 (CLT).

$$\frac{\Lambda_n - \mu n}{\sqrt{n}} \rightarrow \mathcal{N} \left(0, \left(\sum_{i=1}^{\infty} \beta_i \right)^2 \sigma^2 \right), \text{ as } n \rightarrow \infty \text{ in distribution.}$$

where

$$\mu := \frac{\beta_0}{1 - \sum_{i=1}^{\infty} \beta_i} \text{ and } \sigma^2 := \frac{\mu(1 - \mu)}{(1 - \sum_{i=1}^{\infty} \beta_i)^2}.$$

Proof of Theorem 5.1. From Eq. (3), we have

$$\frac{\Lambda_n}{n} = \frac{H_n}{n} - \frac{M_n}{n}. \quad (9)$$

Hence, using convergence relations (6) and (8), we conclude that

$$\frac{\Lambda_n}{n} \rightarrow \frac{\beta_0}{1 - \sum_{i=1}^{\infty} \beta_i}, \text{ as } n \rightarrow \infty \text{ a.s.}$$

Proof of Theorem 5.2. From Eq. (9), and using WLLN for $\{H_n \mid n = 1, 2, \dots\}$, i.e., convergence relation (1) and convergence relation (7), we conclude that

$$\frac{\Lambda_n}{n} \rightarrow \frac{\beta_0}{1 - \sum_{i=1}^{\infty} \beta_i}, \text{ as } n \rightarrow \infty \text{ in probability.}$$

Proof of Theorem 5.3. From the definition of compensator of DTHP,

$$\Lambda_n = \sum_{i=1}^n \mathbb{E}(\xi_i | \mathcal{F}_{i-1}) = \sum_{i=1}^n \left(\beta_0 + \sum_{j=1}^{i-1} \beta_{i-j} \xi_j \right) = n\beta_0 + \sum_{i=1}^n \sum_{j=1}^{i-1} \beta_{i-j} \xi_j = n\beta_0 + \sum_{i=1}^{n-1} \sum_{j=1}^{n-i} \beta_j \xi_i.$$

Using Eqs. (3) and (4), we get

$$\Lambda_n = \left(\sum_{j=1}^{\infty} \beta_j \right) H_n + \beta_0 n - \zeta_n.$$

From Inequality (5) in the proof of Lemma 4.1, we have the following

$$\frac{\mathbb{E}(\zeta_n)}{\sqrt{n}} \leq \frac{\sum_{j=1}^{\infty} j\beta_j}{\sqrt{n}}. \quad (10)$$

Using the fact that $\zeta_n \geq 0$ and $\sum_{j=1}^{\infty} j\beta_j < \infty$, we get

$$\frac{\mathbb{E}(\zeta_n)}{\sqrt{n}} \rightarrow 0, \text{ as } n \rightarrow \infty \text{ a.s.}$$

Consider the following expression,

$$\begin{aligned} \frac{\Lambda_n - n \frac{\beta_0}{1 - \sum_{i=1}^{\infty} \beta_i}}{\sqrt{n}} &= \frac{\left(\sum_{j=1}^{\infty} \beta_j \right) H_n + \beta_0 n - \zeta_n - n \frac{\beta_0}{1 - \sum_{j=1}^{\infty} \beta_j}}{\sqrt{n}} \\ &= \frac{\left(\sum_{j=1}^{\infty} \beta_j \right) H_n - n \frac{\left(\sum_{j=1}^{\infty} \beta_j \right) \beta_0}{1 - \sum_{j=1}^{\infty} \beta_j} - \zeta_n}{\sqrt{n}} \\ &= \frac{\left(\sum_{j=1}^{\infty} \beta_j \right) \left(H_n - n \frac{\beta_0}{1 - \sum_{j=1}^{\infty} \beta_j} \right) - \zeta_n}{\sqrt{n}}. \end{aligned}$$

Hence, we get

$$\frac{\Lambda_n - n \frac{\beta_0}{1 - \sum_{i=1}^{\infty} \beta_i}}{\sqrt{n}} = \frac{\left(\sum_{j=1}^{\infty} \beta_j \right) \left(H_n - n \frac{\beta_0}{1 - \sum_{j=1}^{\infty} \beta_j} \right)}{\sqrt{n}} - \frac{\zeta_n}{\sqrt{n}}.$$

Using the Slutsky's theorem, CLT of $\{H_n \mid n = 1, 2, \dots\}$, i.e., convergence relation (2), and the fact that $\frac{\mathbb{E}(\zeta_n)}{\sqrt{n}} \rightarrow 0$ as $n \rightarrow \infty$ a.s., we get

$$\frac{\Lambda_n - n\mu}{\sqrt{n}} \rightarrow \mathcal{N} \left(0, \frac{(\sum_{j=1}^{\infty} \beta_j)^2 \mu(1-\mu)}{(1 - \sum_{j=1}^{\infty} \beta_j)^2} \right), \text{ as } n \rightarrow \infty \text{ in distribution.}$$

6. Scaled logarithmic moment generating function of DTHP

In this section, we derive some bounds of the scaled logarithmic MGF of the DTHP, and also prove some important convergence results.

Consider the MGF of H_n .

$$\begin{aligned} \mathbb{E}(e^{tH_n}) &= \mathbb{E}(e^{t(H_{n-1} + \xi_n)}) = \mathbb{E}(\mathbb{E}(e^{tH_{n-1} + t\xi_n} | \mathcal{F}_{n-1})) = \mathbb{E}(e^{tH_{n-1}} \mathbb{E}(e^{t\xi_n} | \mathcal{F}_{n-1})) \\ &= \mathbb{E}(e^{tH_{n-1}} [\mathbb{P}(\xi_n = 0 | \mathcal{F}_{n-1}) + e^t \mathbb{P}(\xi_n = 1 | \mathcal{F}_{n-1})]) \\ &= \mathbb{E}(e^{tH_{n-1}} - e^{tH_{n-1}} \mathbb{P}(\xi_n = 1 | \mathcal{F}_{n-1}) + e^{tH_{n-1}} e^t \mathbb{P}(\xi_n = 1 | \mathcal{F}_{n-1})) \\ &= \mathbb{E}(e^{tH_{n-1}} + \mathbb{P}(\xi_n = 1 | \mathcal{F}_{n-1}) e^{tH_{n-1}} (e^t - 1)) \\ &= \mathbb{E}(e^{tH_{n-1}}) + (e^t - 1) \mathbb{E}(e^{tH_{n-1}} \mathbb{P}(\xi_n = 1 | \mathcal{F}_{n-1})). \end{aligned}$$

Hence, we obtain the relation

$$\mathbb{E}(e^{tH_n}) = \mathbb{E}(e^{tH_{n-1}}) + (e^t - 1) \mathbb{E}(e^{tH_{n-1}} \mathbb{P}(\xi_n = 1 | \mathcal{F}_{n-1})). \quad (11)$$

Define the scaled logarithmic MGF of the DTHP as

$$\Gamma_n(t) = \frac{1}{n} \log \mathbb{E}(e^{tH_n}). \quad (12)$$

We now obtain the bounds for Γ_n .

Theorem 6.1. For each $t \in \mathbb{R}$, $\Gamma_n(t)$ is pointwise bounded, i.e., there exists two functions $\psi(t) : \mathbb{R} \rightarrow \mathbb{R}$ and $\phi(t) : \mathbb{R} \rightarrow \mathbb{R}$ such that

$$\phi(t) \leq \Gamma_n(t) \leq \psi(t).$$

Proof of Theorem 6.1. We divide the proof into four lemmas as follows.

Lemma 6.1. For $t > 0$, $\Gamma_n(t)$ is bounded above.

Proof. From Eq. (11), for $t > 0$

$$\begin{aligned} \mathbb{E}(e^{tH_n}) &= \mathbb{E}(e^{tH_{n-1}}) + (e^t - 1)\mathbb{E}\left(e^{tH_{n-1}}(\beta_0 + \sum_{i=1}^{n-1} \beta_{n-i}\xi_i)\right) \\ &\leq \mathbb{E}(e^{tH_{n-1}}) + (e^t - 1)\mathbb{E}(e^{tH_{n-1}}(\beta_0 + \dots + \beta_{n-1})) \\ &= \mathbb{E}(e^{tH_{n-1}}) + (e^t - 1)(\beta_0 + \dots + \beta_{n-1})\mathbb{E}(e^{tH_{n-1}}) \\ &= \mathbb{E}(e^{tH_{n-1}}) [1 + (e^t - 1)(\beta_0 + \dots + \beta_{n-1})] \\ &= \mathbb{E}(e^{tH_{n-1}}) [e^t(\beta_0 + \dots + \beta_{n-1}) + (1 - (\beta_0 + \dots + \beta_{n-1}))] \\ &\leq e^t \mathbb{E}(e^{tH_{n-1}}). \end{aligned}$$

The last inequality follows from the following. Suppose $g(t) = e^t - [e^t(\beta_0 + \dots + \beta_{n-1})] - [1 - (\beta_0 + \dots + \beta_{n-1})]$. Then, for $t > 0$, $g(t) = (e^t - 1)(1 - (\beta_0 + \dots + \beta_{n-1})) \geq 0$.

Continuing recursively,

$$\mathbb{E}(e^{tH_n}) \leq e^t \mathbb{E}(e^{tH_{n-1}}) \leq e^{2t} \mathbb{E}(e^{tH_{n-2}}) \leq \dots \leq e^{(n-1)t} \mathbb{E}(e^{tH_1}) = e^{(n-1)t}((1 - \beta_0) + e^t \beta_0) \leq e^{nt}.$$

Thus, for $t > 0$

$$\Gamma_n(t) = \frac{\log \mathbb{E}(e^{tH_n})}{n} \leq \frac{\log e^{nt}}{n} = \frac{nt}{n} = t. \quad \square$$

Lemma 6.2. For $t > 0$, $\Gamma_n(t)$ is bounded below.

Proof. For $t > 0$, from Eq. (11) we have,

$$\begin{aligned} \mathbb{E}(e^{tH_n}) &= \mathbb{E}(e^{tH_{n-1}}) + (e^t - 1)\mathbb{E}\left(e^{tH_{n-1}}(\beta_0 + \sum_{i=1}^{n-1} \beta_{n-i}\xi_i)\right) \\ &\geq \mathbb{E}(e^{tH_{n-1}}) + (e^t - 1)\mathbb{E}(e^{tH_{n-1}}\beta_0) \\ &= \mathbb{E}(e^{tH_{n-1}})(1 + \beta_0(e^t - 1)) = \mathbb{E}(e^{tH_{n-1}})\mathbb{E}(e^{tH_1}). \end{aligned}$$

Recursively,

$$\mathbb{E}(e^{tH_n}) \geq \mathbb{E}(e^{tH_{n-1}})\mathbb{E}(e^{tH_1}) \geq \dots \geq \mathbb{E}(e^{tH_1})^{n-1}\mathbb{E}(e^{tH_1}) = (1 - \beta_0 + \beta_0 e^t)^n.$$

Hence, for $t > 0$

$$\Gamma_n(t) = \frac{\log \mathbb{E}(e^{tH_n})}{n} \geq \frac{\log(1 - \beta_0 + \beta_0 e^t)^n}{n} = \frac{n \log(1 - \beta_0 + \beta_0 e^t)}{n} = \log(1 - \beta_0 + \beta_0 e^t). \quad \square$$

Lemma 6.3. For $t < 0$, $\Gamma_n(t)$ is bounded above.

Proof. From Eq. (11), for $t < 0$

$$\mathbb{E}(e^{tH_n}) = \mathbb{E}(e^{tH_{n-1}}) + (e^t - 1)\mathbb{E}\left(e^{tH_{n-1}}(\beta_0 + \sum_{i=1}^{n-1} \beta_{n-i}\xi_i)\right) \leq \mathbb{E}(e^{tH_{n-1}}).$$

Continuing recursively,

$$\mathbb{E}(e^{tH_n}) \leq \mathbb{E}(e^{tH_{n-1}}) \leq \dots \leq \mathbb{E}(e^{tH_1}) = (1 - \beta_0) + \beta_0 e^t.$$

Hence, for $t < 0$,

$$\Gamma_n(t) = \frac{\log \mathbb{E}(e^{tH_n})}{n} \leq \frac{\log((1 - \beta_0) + \beta_0 e^t)}{n} \leq 0. \quad \square$$

Lemma 6.4. For $t < 0$, $\Gamma_n(t)$ is bounded below.

Proof. For any $n = 1, 2, \dots$, the MGF of H_n can also be written as

$$\mathbb{E}(e^{tH_n}) = c_{0,n} + c_{1,n}e^t + \dots + c_{n,n}e^{nt}.$$

where

$$c_{r,n} = \sum_{\substack{a_1 + \dots + a_n = r \\ a_i = 0 \text{ or } 1}} \mathbb{P}(\xi_1 = a_1, \xi_2 = a_2, \dots, \xi_n = a_n), \quad \sum_{r=0}^n c_{r,n} = 1, \quad 0 \leq r \leq n.$$

Using multiplicative rule of probability, we get

$$\begin{aligned} c_{0,n} &= \mathbb{P}(\xi_1 = 0, \xi_2 = 0, \dots, \xi_n = 0) = \mathbb{P}(\xi_1 = 0) \cdot \mathbb{P}(\xi_2 = 0 \mid \xi_1 = 0) \cdot \\ &\quad \dots \cdot \mathbb{P}(\xi_n = 0 \mid \xi_1 = 0, \xi_2 = 0, \dots, \xi_{n-1} = 0) \\ &= (1 - \beta_0)^n. \end{aligned}$$

Similarly,

$$c_{n,n} = \beta_0(\beta_0 + \beta_1) \dots (\beta_0 + \dots + \beta_n).$$

For $t < 0$,

$$\mathbb{E}(e^{tH_n}) = c_{0,n} + c_{1,n}e^t + \dots + c_{n,n}e^{nt} \geq c_{0,n},$$

Hence, for $t < 0$

$$\Gamma_n(t) = \frac{\log \mathbb{E}(e^{tH_n})}{n} \geq \frac{\log c_{0,n}}{n} = \frac{\log(1 - \beta_0)^n}{n} = \log(1 - \beta_0). \quad \square$$

6.1. Convergence of $\Gamma_n(t)$

In this section, we prove that for $t < 0$, $\Gamma_n(t)$ converges to some function of t . First, we will prove the following lemma.

Lemma 6.5. For $t < 0$, $\Gamma_n(t)$ is monotonically decreasing.

Proof. From Eq. (11), for any $n = 2, 3, \dots$ and $t < 0$

$$\mathbb{E}(e^{tH_n}) < \mathbb{E}(e^{tH_{n-1}}).$$

Hence, for $t < 0$,

$$\Gamma_n(t) = \frac{\log \mathbb{E}(e^{tH_n})}{n} < \frac{\log \mathbb{E}(e^{tH_{n-1}})}{n} < \frac{\log \mathbb{E}(e^{tH_{n-1}})}{n-1} = \Gamma_{n-1}(t). \quad \square$$

By Theorem 6.1, $\Gamma_n(t)$ is bounded for all $t \in \mathbb{R}$, and by Lemma 6.5, $\Gamma_n(t)$ is monotonically decreasing for $t < 0$, hence from monotone convergence theorem, we conclude that for $t < 0$, $\Gamma_n(t)$ converges, i.e., for $t < 0$

$$\lim_{n \rightarrow \infty} \Gamma_n(t) = \Gamma(t).$$

7. Conclusion and future work

In this paper, we have extended the work of Seol (2015) to prove the SLLN for DTHP $\{H_n \mid n = 1, 2, \dots\}$. We have identified the compensator of the DTHP and analyzed its asymptotic behavior. We have shown the WLLN, the SLLN and the CLT for the compensator of the DTHP. We have further analyzed the MGF and scaled logarithmic MGF of the DTHP obtaining tight bounds and convergence results. The analysis on the scaled logarithmic MGF is important because these results are necessary conditions to establish the LDP for the DTHP via Gärtner-Ellis Theorem. Therefore, one natural possible direction this work can be extended is to explore the existence of the LDP for the DTHP. Furthermore, one can also study the LDP for the compensator of the DTHP.

Data availability

No data was used for the research described in the article.

References

- Bacry, E., Delattre, S., Hoffmann, M., Muzy, J.-F., 2013. Some limit theorems for Hawkes processes and application to financial statistics. *Stochastic Process. Appl.* 123 (7), 2475–2499.
- Bordenave, C., Torrisi, G.L., 2007. Large deviations of Poisson cluster processes. *Stoch. Models* 23 (4), 593–625.
- Brémaud, P., Massoulié, L., 1996. Stability of nonlinear Hawkes processes. *Ann. Probab.* 1563–1588.
- Daley, D.J., Vere-Jones, D., et al., 2003. *An Introduction to the Theory of Point Processes: Volume I: Elementary Theory and Methods*. Springer.
- Dassios, A., Zhao, H., 2011. A dynamic contagion process. *Adv. Appl. Probab.* 43 (3), 814–846.

- Errais, E., Giesecke, K., Goldberg, L.R., 2010. Affine point processes and portfolio credit risk. *SIAM J. Financial Math.* 1 (1), 642–665.
- Guo, X., Zeng, Y., 2008. Intensity process and compensator: A new filtration expansion approach and the Jeulin–Yor theorem. *Ann. Appl. Probab.* 18 (1), 120–142.
- Hawkes, A.G., 1971. Spectra of some self-exciting and mutually exciting point processes. *Biometrika* 58 (1), 83–90.
- Ivanoff, B.G., Merzbach, E., Plante, M., 2007. A compensator characterization of point processes on topological lattices. *Electron. J. Probab.* 12, no. 2, 47–74.
- Selvamuthu, D., Pandey, S., Tardelli, P., 2023. Limit theorems for an extended inverse Hawkes process with general exciting functions. *Statist. Probab. Lett.* 197, 109817.
- Seol, Y., 2015. Limit theorems for discrete Hawkes processes. *Statist. Probab. Lett.* 99, 223–229.
- Seol, Y., 2017. Limit theorems for the compensator of Hawkes processes. *Statist. Probab. Lett.* 127, 165–172.
- Seol, Y., 2019. Limit theorems for an inverse Markovian Hawkes process. *Statist. Probab. Lett.* 155, 108580.
- Wang, H., 2022. Limit theorems for a discrete-time marked Hawkes process. *Statist. Probab. Lett.* 184, 109368.
- Wang, H., 2023. Large and moderate deviations for a discrete-time marked Hawkes process. *Comm. Statist. Theory Methods* 52 (17), 6037–6062.
- Xu, Y., Zhu, L., Wang, H., 2021. A self-and mutual-exciting model for discrete-time data: Case study on online money market fund. Available at SSRN 3665436.
- Zhu, L., 2013a. Central limit theorem for nonlinear Hawkes processes. *J. Appl. Probab.* 50 (3), 760–771.
- Zhu, L., 2013b. Moderate deviations for Hawkes processes. *Statist. Probab. Lett.* 83 (3), 885–890.
- Zhu, L., 2013c. Ruin probabilities for risk processes with non-stationary arrivals and subexponential claims. *Insurance Math. Econom.* 53 (3), 544–550.
- Zhu, L., 2014. Process-level large deviations for nonlinear Hawkes point processes. *Annales de l’IHP Probabilités et statistiques* 50 (3), 845–871.
- Zhu, L., 2015. Large deviations for Markovian nonlinear Hawkes processes. *Ann. Appl. Probab.* 25 (2), 548–581.