Special Proceeding of the 22^{nd} Annual Conference of SSCA held at Savitribai Phule Pune University, Ganeshkhind, Pune, MH, during January 02–04, 2020; pp 181-192

High Frequency Financial Data and Associated Financial Point Process: A Nonparametric Bayesian Perspective

Anuj Mishra and T. V. Ramanathan
Department of Statistics and Centre for Advanced Studies in Statistics
Savitribai Phule Pune University, Pune 411 007, India

Received: 22 August 2020; Revised: 03 September 2020; Accepted: 06 September 2020

Abstract

This paper review the theoretical framework of modeling high frequency financial data (HFD) using a point process approach. We represent the financial event arrival times as the realization of non-homogeneous Poisson process with an intensity function $\lambda(t)$, which is assumed to be periodic. In the case of HFD, this periodic pattern is quite well known, as the intensity of trades is higher in the morning and just before closing the market and lower during the afternoon. We make an attempt to study this intra-day cyclic behaviour with an intensity modelling approach using Bayesian nonparametric method. The posterior consistency of the proposed nonparametric Bayesian procedure is established. The Bayesian estimation of the intensity function is described for a specific case where the conditions of the prior are satisfied. This paper is just a first step towards modelling the HFD using point process approach and the corresponding Bayesian nonparametric analysis. It may be mentioned that lot more computations need to be done to complete this ongoing work.

Key words: Bayesian nonparametrics; Financial point processes; High frequency financial data; Intensity function; Intensity process.

AMS Subject Classifications: 62K05, 05B05

1. Introduction

The empirical studies in finance literature usually concentrate on opening, closing or average prices of stocks from financial markets. However, due to technological advancements, researchers can now work with the high frequency data (HFD), which contains details of all the transactions along with the marks such as price, volume, time of transaction etc. Such data has attracted lot of researchers and this has become a new area of research these days. Known as 'high-frequency finance', it helps to understand the financial markets at a micro level, see Viens *et al.* (2011), Gregoriou (2015) and Florescu *et al.* (2016) for a broad overview. Here, there can be details of hundreds of transactions happening in a micro time interval corresponding to a particular stock from an electronic stock exchange. An important feature of HFD is that the transactions are recorded as and when they occur, hence the observations are irregularly time spaced. This prevents the use of standard time series methods in high frequency finance. The timing of transactions carry substantial amount of information, which can be used in studying the micro structure of a financial market. Therefore, it is very important to model the time interval between transactions (durations) appropriately.

Corresponding Author: T. V. Ramanathan
Email: ram@unipune.ac.in

One way of modeling the durations is to make use of the autoregressive-type conditional duration (ACD) models introduced by Engle and Russell (1998), which attempts to model the time between the events occurring at t_{i-1} and t_i, by defining $x_i = t_i - t_{i-1}$, where $\{t_0, t_1, ..., t_n, ...\}$ is a sequence of arrival times with $0 = t_0 \leq t_1 \leq t_2 \leq ... \leq t_n \leq ...$. The sequence $\{x_1, x_2, ...\}$ of non negative random variables form the durations. Here, the arrival times t_i may not necessarily mean the time corresponding to consecutive trades. There can be other type of events of interest as well. For example, t_i can be time of occurrence of a volume event, which is said to have occurred if the cumulative trade volume since the last volume event is at least a preset amount v. Similarly, price event is said to have occurred if the cumulative price change since last price event is at least of a preset amount p. Thus, x_i is the interval between consecutive events of interest leading to trade, volume or price durations.

Let the conditional expected duration be

$$\psi_i = E(x_i|\mathcal{F}_{i-1}), \tag{1}$$

where, $\mathcal{F}_i$ is the information set at transaction i, that is, $\mathcal{F}_i = \sigma(x_i, x_{i-1}, \cdots, x_1)$. The main assumption of an ACD model is that the durations are of the form

$$x_i = \psi_i \ \epsilon_i, \tag{2}$$

where ϵ_i are independent and identically distributed (i.i.d) random variables with $E(\epsilon_i) = 1$ (In fact, without loss of generality, it is possible to assume that this is true). The above set up is very general and it allows a variety of models which can be obtained by choosing different specifications for the expected duration ψ and different distributions for ϵ. We refer to two interesting review papers Pacurar (2008) and Bhogal and Ramanathan (2019) for detailed discussions on this approach.

Another approach of modeling HFD is using a point process. In this method, one represents the event of arrival times as a realization of non-homogeneous Poisson process with an intensity function $\lambda(t)$ having a specific structure. This method is usually known as *financial point process* method. A further extension of this is modelling based on the intensity function of the process, which leads to more flexible and powerful models. Such an approach is recommended when we deal with multivariate processes, in which case, the conditional duration approach is not very successful (see Russell (1999), Hautsch (2004) and Bauwens and Hautsch (2006)).

Researchers have been using the periodicity adjustment procedure of Engle and Russell (1998), which we have also used to remove intra-day effect in a paper published recently Mishra and Ramanathan (2017). However, WSu (2012) had claimed that this procedure is not very satisfactory. Hence, using a non-homogeneous Poisson process may resolve the issue, as suggested by various researchers in related problems. Belitser *et al.* (2013) proposed an M-estimator to estimate the period of a cyclic non-homogeneous Poisson process, established its consistency and demonstrated the effectiveness by applying it to a call center data. In our case, we already know the period, which is 'daily' and hence we are not interested in estimating the period. Weinberg *et al.* (2007) modelled the day-to-day as well as intraday variations in the same call center data using a normal approximation to Poisson. Specifically, in finance, Andersen *et al.* (2019) develop a procedure to test intra-day periodicity in return volatility. We propose a procedure to adjust the periodicity using a Bayesian approach and prove its consistency.

Bayesian approaches for Hawkes models have received much less attention. The only contributions for the Bayesian inference are due to Gulddahl (2013) and Blundell *et al.* (2012) who explored parametric approaches and used MCMC to approximate the posterior distribution of the parameters. Donnet *et al.* (2018) study the properties of Bayesian nonparametric procedures in the context of multivariate Hawkes processes.

The paper is organized a follows. In Section 2, we discuss the financial point process. Bayesian nonparametric approach for financial point process is described in Section 3. Extensions to Hawke-type processes is described in Section 4. Section 5 concludes with some future directions.

2. Financial Point Processes

Let $\{t_i, i=1,\cdots,n\}$ denote a random sequence of increasing event times $0<t_1<\cdots<t_n$ associated with an orderly (simple) point process. Then,

$$N(t)=\sum_{i\geq 1} I_{\{t_i\leq t\}}$$

define a right-continuous counting function which gives the number of events of some type in the time interval (0,t). The $\mathcal{F}_t$-intensity process $\lambda(t)$ of the counting process $N(t)$ is defined as

$$\lambda(t;\mathcal{F}_t)=lim_{\Delta\to 0}\frac{1}{\Delta}E[N(t+\Delta)-N(t)|\mathcal{F}_t], \tag{3}$$

where $\mathcal{F}_t=\sigma\{N(s);0\leq s\leq t\}$. Therefore, the sequence of event arrival times $\{t_i\}$ can be modeled as a point process by modelling the intensity $\lambda(t)$.

The simplest type of point process is the homogeneous Poisson process defined by

$$\begin{aligned} Pr((N(t+\Delta)-N(t))=1|\mathcal{F}_t) &= \lambda\Delta+o(\Delta), \\ Pr((N(t+\Delta)-N(t))>1|\mathcal{F}_t) &= o(\Delta), \end{aligned} \tag{4}$$

with $o(\Delta)/\Delta\to 0$, as $\Delta\to 0$. Note that in this case, the intensity is constant and it leads to

$$P(t_i>x)=P(N(x)<i)=\sum_{j=0}^{i-1}\frac{e^{-\lambda x}(\lambda x)^j}{j!}.$$

A straight generalisation from here is the case when the intensity function is a deterministic function of time or a non-homogeneous Poisson process with intensity function $\lambda(t)$. This can be particularly useful in modelling the intra-day cyclic behaviour of durations with an appropriate choice of $\lambda(t)$. One another possibility here is to use a marked point process defined with marks such as arrival of buys, sells and certain limit orders, see Bauwens and Hautsch (2006).

3. Bayesian Point Processes

Let $(N_t)_{t\geq 0}$ be a non-homogeneous process on $[0,T]$, that is, the sample paths of $(N_t)_{t\geq 0}$ are right-continuous step functions with $N_0=0$ and with jumps of size 1. Let N_t be the number of jumps in $[0,t]$ and $N_t<\infty$ almost surely. We assume the following about the process $N(t)$ and the intensity function $\lambda(t)$.

- A1 For any disjoint subsets $B_1,B_2,\cdots,B_m\in\mathcal{B}([0,T])$, $\mathcal{B}([0,T])$ the random variables $N(B_1),N(B_2),\cdots,N(B_m)$ are independent random variables denoting the number of jumps in $B_1,B_2,\cdots,B_m$ respectively.

- A2 For any $B\in\mathcal{B}([0,T])$ the random variable $N(B)$ is distributed as Poisson with parameter $\Lambda(B)$; where Λ is a finite measure on $([0,T],\mathcal{B}([0,T]))$, called as the compensator of the process.

- A3 Λ admits a density λ with respect to the Lebesgue measure on $\mathcal{B}([0,T])$. That is,

$$\Lambda_t=\int_0^t \lambda(s)ds,$$

where, $\lambda(t)$ is called the intensity of N_t.

We estimate $\lambda(t)$ using Bayesian procedure and establish the posterior consistency. Consistency results for intensities of Poisson processes can be established by connecting and extending the two main approaches regarding consistency for i.i.d. observations. The first approach, due to Schwartz (1965), Barron *et al.* (1999), and Ghosal *et al.* (1999), requires construction of an increasing sequence of sets, a sieve, and a sequence of uniformly consistent tests. This is the approach followed by Belitser *et al.* (2015) and Donnet *et al.* (2018). An alternative approach, which we follow in this paper, provided by Walker (2004) relies on a martingale sequence to obtain sufficient conditions for posterior consistency in the i.i.d. case. This alternative approach, though equivalent to the use of a suitable sieve, simplifies the verification of necessary conditions for consistency.

Below we state Theorem 1.3 of Kutoyants (1998) as a lemma.

Lemma 1: For any λ, the law P_λ of N under the parameter value λ admits a density p_λ with respect to the measure induced by a standard Poisson point process with intensity 1. This density is given by

$$p(\lambda) = exp(\int_0^T log\lambda(t)dN_t - \int_0^T (\lambda(t)-1)dt).$$

Suppose we observe n independent non-homogeneous Poisson processes $N^{(1)},N^{(2)},\cdots,N^{(n)}$ on $[0,T]$ with a common intensity λ, which is a positive integrable function on $[0,T]$. Then, by Lemma 1, the likelihood is given by

$$L(\lambda) = \prod_{i=1}^{n} exp(\int_0^T log\lambda(t)dN_t^{(i)} - \int_0^T (\lambda(t)-1)dt).$$

We define the parameter space as

$$\mathcal{F} = \left\{ \lambda : [0,T] \to R_+ | \int_0^T \lambda(t)dt < \infty \right\}.$$

Here, we are estimating the intensity function $\lambda(t)$, given $N^{(1)},N^{(2)}\cdots N^{(n)}$, using a Bayesian nonparametric approach. Let λ belong to the class $\mathcal{F}$ of intensities which need not be indexed by a finite dimensional parameter. Let Π be a prior on $\mathcal{F}$, $\Pi : (\mathcal{F},\sigma(\mathcal{F})) \to [0,1]$. Let $\Pi(.|N^{(1)},N^{(2)},\cdots,N^{(n)})$ stand for the posterior distribution of λ given the data. So, if B is a set of intensities, the posterior mass assigned to it is given by

$$\Pi(B|N^{(1)},N^{(2)},\cdots,N^{(n)}) = \frac{\int_B R_n(\lambda)d\Pi(\lambda)}{\int_{\mathcal{F}} R_n(\lambda)d\Pi(\lambda)},$$

where

$$R_n(\lambda) = \prod_{i=1}^{n} \frac{p(N^{(i)},\lambda)}{p(N^{(i)},\lambda_0)}$$

is the likelihood ratio with $\lambda_0 \in \mathcal{F}$ being the true fixed but unknown transition density.

The Bayesian model is consistent if the posterior mass increases around λ_0 as n increases. Suppose that a topology on $\mathcal{F}$ has been specified. Then posterior distribution is said to be consistent at λ_0 if for every neighborhood U of λ_0, we have that,

$$\Pi(U^c|N^{(1)},N^{(2)},\cdots,N^{(n)}) \to 0 \quad a.s.$$

3.1. Posterior consistency

For a continuous function f on $[0,T]$ we define the norms $\|f\|_2$ and $\|f\|_\infty$ as usual by defining

$$\|f\|_2 = \int_0^T f^2(t)dt, \; and \; \|f\|_\infty = sup_{t\in[0,T]}|f(t)|.$$

In the following theorem, we establish the posterior consistency of Bayesian procedure under a couple of conditions on the prior assumed and on the space of intensity functions.

Theorem 1: For any given $\epsilon > 0$, let $A_\epsilon^{'}$ be a set of intensities around the true intensity $\lambda_0 \in \mathcal{F}$ defined as,

$$A_\epsilon^{'} = (\lambda \in \mathcal{F} : \|\sqrt{\lambda} - \sqrt{\lambda_0}\|_2 > \sqrt{2}\epsilon). \tag{5}$$

Let the prior Π be such that

a) $\Pi(\lambda : \|\lambda - \lambda_0)\|_\infty < \epsilon) > 0$ and

b) $\sum_{j=1}^{\infty} \sqrt{\Pi(A_j^{'})} < \infty$, where $\{A_j^{'}\}$ is a countable cover of size $\frac{\delta}{\sqrt{2}}(\delta < \epsilon)$ for $A_\epsilon^{'}$.

Then,

$$\Pi(A_\epsilon^{'} | N^{(1)}, N^{(2)}, \cdots, N^{(n)}) \to 0 \ a.s. \tag{6}$$

Condition (a) is similar to Belitser *et al.* (2015) and Donnet *et al.* (2018). But we have an easily verifiable condition (b), as compared to other conditions provided by the same authors.

Let the square of the Hellinger distance $h(p_\lambda, p_{\lambda'})$ be defined as

$$h^2(p_\lambda, p_{\lambda'}) = 2\left(1 - E_{\lambda'}\left(\sqrt{\frac{p_\lambda(N)}{p_{\lambda'}(N)}}\right)\right),$$

where E_λ is the expectation corresponding to the probability measure under which the process N is a Poisson process with intensity function λ. Let the Kullback-Leibler divergence $K(p_\lambda, p_{\lambda'})$ be defined by

$$K(p_\lambda, p_{\lambda'}) = -E_{\lambda'}\left(log\left(\frac{p_\lambda(N)}{p_{\lambda'}(N)}\right)\right).$$

We state two lemmas before proving Theorem 1. Lemma 2 constitutes a part of Lemma 1 of Belitser *et al.* (2015) and Lemma 3 is nothing but Theorem 4 of Walker (2004), which gives posterior consistency result for density estimation. Proofs of these lemmas are omitted as they are available in the references mentioned.

Lemma 2: For the Hellinger distance $h(p_\lambda, p_{\lambda'})$ and the Kullback-Leibler divergence $K(p_\lambda, p_{\lambda'})$, the following hold good.

(i) $\frac{1}{\sqrt{2}}\|\sqrt{\lambda} - \sqrt{\lambda'}\|_2 \leq h(p_\lambda, p_{\lambda'}) \leq \sqrt{2}(\|\sqrt{\lambda} - \sqrt{\lambda'}\|_2)$

(ii) $\|\lambda - \lambda_0)\|_\infty \leq K(p_\lambda, p_{\lambda'})$

Lemma 3: Let A_ϵ be a set of intensities defined in terms of densities with respect to Poisson measure p_λ h-bounded away from p_{λ_0},

$$A_\epsilon = (\lambda \in \mathcal{F} : h(p_\lambda, p_{\lambda_0}) > \epsilon). \tag{7}$$

Assume that the prior Π has the following properties:

C1. $\Pi(K(p_\lambda, p_{\lambda_0}) < \epsilon) > 0$

C2. $\sum\limits_{j=1}^{\infty} \sqrt{\Pi(A_j)} < \infty$, where $\{A_j\}$ is a countable h-cover of size $\delta(<\epsilon)$ for A_ϵ.

Then,

$$\Pi(A_\epsilon|N^{(1)},N^{(2)},\cdots,N^{(n)})\to 0 \ a.s. \tag{8}$$

Proof: *of Theorem 1:* We can view our problem as a density estimation problem with respect to the Poisson measure, and as a consequence, we have the posterior consistency of Lemma 3. By Lemma 2(i), we have from (7) and (5), $A_\epsilon \subset A_\epsilon^{'}$, which gives

$$\Pi(A_\epsilon)\to 0 \Rightarrow \Pi(A_\epsilon^{'})\to 0. \tag{9}$$

By Lemma 2(ii), we have

$$\Pi(\lambda:\|\lambda-\lambda_0\|_\infty<\epsilon)>0 \Rightarrow \Pi(\lambda:K(p_\lambda,p_{\lambda_0})<\epsilon)>0. \tag{10}$$

Again, by Lemma 2(i), if $\{A_j\}$ is a countable h-cover of size $\delta(<\epsilon)$ for A_ϵ, then $\{A_j^{'}\}$ is a countable cover of size $\frac{\delta}{\sqrt{2}}(\delta<\epsilon)$ for $A_\epsilon^{'}$. Moreover, we also have, again by Lemma 3,

$$\sum_{j=1}^{\infty}\sqrt{\Pi(A_j^{'})}<\infty \Rightarrow \sum_{j=1}^{\infty}\sqrt{\Pi(A_j)}<\infty \tag{11}$$

We have shown in (10) and (11) that condition (a) and (b) of Theorem 1 implies Lemma 3(C1) and 3(C2). Under these conditions, we have the result (8) of Lemma 3, which implies the result (6) of Theorem 1 by (9). □

Thus we have proved a general result. We can use Theorem 1, to obtain the suitability of specific priors. For clarification, we illustrate an example in the subsection below, but would like to emphasize that these results may also be used to verify the posterior consistency of other priors.

3.2. Illustration

There is a huge literature on prior construction for nonparametric models, where space of functions serve as the parameter space, see Chapter 2 of Ghosal and van der Vaart (2017) for details. The simplest and most popular is the random basis expansion, which we discuss briefly.

Given a set of basis functions $\phi_j:[0,T]\to R$, one way of constructing prior on intensity functions $\lambda:[0,T]\to R_+$ is by writing $\lambda=exp(\sum_{j=0}^{\infty}\beta_j\phi_j)$ and putting priors on the coefficients β_j in this representation. There can be many choices of bases, such as, polynomials, trigonometric functions, wavelets, splines, spherical harmonics, etc. See De Boor (1978) for details on splines, Härdle *et al.* (2012) and Donald and Percival (2000) for details on wavelets and Fourier bases. Also see Appendix E of Ghosal and van der Vaart (2017) for their approximation properties. For a given application, the suitability of the prior is determined by the approximation properties of the basis together with the prior on the coefficients. Rivoirard and Rousseau (2012) discussed a very general adaptive priors based on wavelets and Fourier bases. Shen and Ghosal (2015) and Belister *et al.* (2014) carry out a similar study for priors based on spline basis. Shen and Ghosal (2015) assumed the knots as fixed, whereas, Belister *et al.* (2014) considered them as random.

We outline the procedure of estimation without getting into the details of computations.

Given infinite basis functions, the convergence of λ is not guaranteed always. However, it is true if and only if $\sum\beta_j^2<\infty$ *a.s.*. Let $\beta_j\sim N(0,\sigma_j^2)$. To ensure that λ defines a valid intensity function with probability 1, it is sufficient that $\sum\sigma_j<\infty$. Conditions for posterior consistency can be obtained by applying Theorem 1. Walker (2004) studied the same prior with adjustments for density estimation case and obtained a sufficient condition as $\sum(\frac{\sigma_j}{\omega_j})^{2m-\frac{1}{2}}<\infty$ for some sequence ω_j satisfying $\sum\omega_j<\infty$. Basically, if we want to use this basis representation as a prior on intensity functions, then we will put priors on β_j, which will induce a prior on intensity functions. Also we want the prior on β_j to shrink to zero at an appropriate speed. For example, by taking $\sigma_j \propto j^{-1-q}$ for any $q>0$, this will be satisfied (put $\omega_j \propto j^{-1-r}$ for any $r>0$).

What we have provided till now is not a method of estimating the intensity, but, mentioned the conditions that are to be verified, in order to justify the Bayesian procedure.

3.3. Bayesian estimation procedure using HFD

In order to estimate the intensity using the Bayesian procedure, appropriate choice of prior needs to be made along with different bases, and then the posterior has to be computed. The purpose here is not to estimate the intensity in detail, but, to throw some light on the procedure using a particular case. The result will be true for any Bayesian procedure where the conditions on the priors are satisfied.

Let n denote the observed number of days of trading activity and Tn the total time in seconds during which market operates everyday. For example, if market timing is 10 AM to 4 PM, then $T=0$ *sec* corresponds to the time 10:00:00 and $T=21{,}600$ *sec* corresponds to the time 16:00:00. Then the complete event counting process is given by $N=\{N_t : t\in[0,nT]\}$, where N_t denotes the number of events in $[0,t]$. The assumption of periodicity implies that $\lambda(t+T)=\lambda(t)$, $\forall\ t\geq 0$. For $i=1,\cdots,n$, the event arrival counting process during day i may be defined as

$$N_t^{(i)} = N_{(i-1)T+t} - N_{(i-1)T}, \quad t\in[0,T].$$

Since the increments of the process N_t are independent, the processes $N_t^{(i)}$ are independent non-homogeneous Poisson processes with λ as the intensity function, restricted to $[0,T]$.

Our objective is to estimate the intensity function. Let Δ be a small grid of time, say 15 seconds, and $m=T/\Delta$ be the number of grids per day in the data set. Then the number of events in the j^{th} time grid on day i is given by

$$A_{ij} = N_{j\Delta}^{(i)} - N_{(j-1)\Delta}^{(i)},$$

which is assumed to follow a Poisson distribution with mean $\lambda_j = \int_{(j-1)\Delta}^{j\Delta}\lambda(t)$, for every $i=1,\cdots n$ and $j=1,\cdots m$. We denote the available data over grids as $A^n = (A_{ij} : i=1,\cdots,n, j=1,\cdots,m)$. Hence, the likelihood is given by

$$L(\lambda|A^n) = \prod_{i=1}^{n}\prod_{j=1}^{m}\frac{\lambda_j^{A_{ij}}\exp\lambda_j}{A_{ij}!}. \tag{12}$$

Putting a prior of the random basis type on λ, we have

$$\lambda = \sum_{k=1}^{J}\beta_k\phi_k. \tag{13}$$

This finite basis representation may not be truly non-parametric in nature. However, this problem can be addressed by putting another prior on J (see Chapter 2, Ghosal and van der Vaart (2017)). Therefore, a draw from prior Π can be constructed as follows:

1. Draw J from a Poisson distribution with mean μ (around 12).
2. Given $J=j$, draw $\underline{\beta}$ vector of j-independent $N(0,\sigma_j^2)$

Given the data, likelihood (3.8) and the prior (3.9), we can use MCMC to sample from posterior distribution, as it will be difficult to obtain the analytical form of the posterior. As long as the bases considered are fairly easy to evaluate and integrate, we can compute the likelihood and posterior up to a normalising constant, without making any approximation. After that the computations will be a straightforward adoption of existing MCMC methods Polasek (2012). Also, having the data, form of the prior and the likelihood, we can use *Stan* programming language Carpenter *et al.* (2017), which does the posterior computation automatically using the Hamiltonian Monte Carlo (HMC) Betancourt (2017). While using HMC, proper care needs to be taken as a discrete parameter J is introduced into the model, and HMC does not work in discrete parameter case. However, properly augmenting HMC with Gibbs-type update for discrete parameter may circumvent this problem. This demand further investigations in this front.

We need to do a detailed computation considering all the bases and their variations in order to recommend an optimal choice of the prior and the basis for an intra-day periodicity adjustment. To carry out this task with the data, there has to be some additional changes in the structure of the intensity function. For example, here, we have considered the case of a Poisson process where the intensity does not depend on the observations. But in real cases we can incorporate a dynamic behaviour into the structure of the intensity function, so that, it actually affects the intensity of events in the market. This can be achieved by using the Hawkes- type models for the counts. Next we briefly discuss the Hawkes process and some of its generalizations.

4. Extensions to Hakwes-type Processes

A different generalization of the Poisson process is obtained by specifying $\lambda(t)$ as a (linear) self-exciting process given by

$$\lambda(t)=\mu+\int_0^t w(t-u)dN(u)=\mu+\sum_{t_i<t} w(t-t_i) \tag{14}$$

where $\mu(t)$ provides a Poisson base for the process and $w(u)$ is a kernel (exciting). This process is known as Hawkes process and was first proposed by Hawkes (1971) and was applied in seismology. The generalisation capability in Hawkes process over Poisson comes from kernel, which allows contribution by an event that occurs at a previous time $t-k$ to intensity at time t. This kind of a dynamic behaviour is not supported by the Poisson process.

following are some of the kernels that are frequently used in the case of Hawkes process.

1. **Exponential kernel**: The exponential kernel is given by
$$w(u)=\alpha\beta e^{-\beta u},\ u>0. \tag{15}$$
This kernel implies an exponential decay in the effect of an event on future events and β drives the strength of the time decay and α, the overall strength of excitation.

2. **Power-law kernel**: The power-law kernel is given by
$$w(u)=\frac{\alpha\beta}{(1+\beta u)^{1+p}},\ u>0 \tag{16}$$
This kernel implies a hyperbolic decay and capture long range dependence.

Another approach lies in using kernels that take the form of linear combination of exponential and/or power functions with different rate constants, which might help in capturing different short and long range dependence.

The intensity given in (4.1) can also be generalised to accommodate marks. There may be a mark, χ_i associated with the the event t_i, that can affect the intensity. For example, a trade with large volume may excite future trades more than a trade with small volume. Marks may be contained within the kernel function as $w(u,\chi)$. The Hawkes process with such a modification is known as Marked Hawkes process.

Multivariate Hawkes models can be also obtained by a generalization of (4.1). In such a case, $\lambda(t)$ is a $K\times 1$ vector defined by $\lambda(t)=(\lambda^1(t),\cdots,\lambda^K(t))$ with

$$\lambda^k(t)=\mu^k(t)+\sum_{m=1}^{K}\sum_{t_i^k<t}w_{mk}(t-t_i^k). \tag{17}$$

The function w_{mk} is a cross-exciting term with $w_{mk}(t-t_i^k)$ being the contribution to the intensity of type-m events made by a type-k event at t_i^k .

The probabilistic properties of Hawkes processes are discussed in Hawkes (1971), and Brémaud and Massoulié (1996). Hawkes and Oakes (1974) show that every self-exciting Hawkes processes can be represented as a Poisson cluster process. Thinking of each event as a parent, an event occurring at time t_i gives birth to offspring according to a Poisson process with intensity $w(t-t_i)$: these offspring generate their own offspring, and so on. Ogata (1978) discusses the maximum likelihood estimation of Hawkes process, whereas, asymptotic behaviour of such an estimate is investigated by Ozaki (1979).

Despite their usefulness, Hawkes-type models did not find their place in financial econometrics for a long time. Bowsher (2007) applied Hawkes type model in financial econometrics for the first time. He presented a continuous time, bivariate point process model ((4.4) with K=2) of the timing of trades and mid-quote changes for a New York Stock Exchange stock. Estimation was performed using maximum likelihood method as analytic likelihoods were available. Since then, there has been various developments in finance related to applications of Hawkes process. Bacry *et al.* (2012) introduced a non-parametric estimation method for multivariate Hawkes processes based on the spectral factorization of the co-variance matrix and then applied it to tick-by-tick trades data of a futures contract for a total period of 3.5 months. Da Fonseca and Zaatour (2014) proposed an estimation strategy using the method of moments that can be solved almost instantaneously as against the maximum likelihood estimates, and applied to trade arrival times of major stocks for observations of 2 years. Fauth and Tudor (2012) used multivariate marked point processes in order to describe the fluctuation in tick-by-tick data corresponding to trades in currency exchange (EUR,GBP,CHF,JPY). Hawkes (2018), Bacry *et al.* (2015) and Bauwens and Galli (2009) gave excellent reviews of applications of point process in finance. In spite of various applications after Bowsher (2007), there has been no Bayesian study of point processes in finance. Bayesian methods can be advantageous while obtaining the uncertainty about the intensity using the spread of the posterior distribution.

5. Concluding Remarks

In this paper, we have discussed the financial point process associated with the high frequency financial data. With the nonhomogeneous assumption of the count process associated with the durations, it is appropriate to estimate the intensity function $\lambda(t)$ using a nonparametric functional approach. We have addressed the problem using a nonparametric Bayesian method. This review paper is just a first step towards the description of the problem and the associated research. An extensive computational exercise needs to be undertaken by considering different basis combinations for intra-day periodicity adjustment. We also plan to extend this study by investigating various theoretical as well as computational aspects of Bayesian nonparametric approach of modeling the HFD using Hawkes-type models.

Acknowledgements

Authors are thankful to Professor V. K. Gupta, President, Society of Statistics, Computer and Applications for recommending this paper to the special proceedings of the Society's annual conference. T. V. Ramanathan would like to acknowledge the financial support from the Science and Engineering Research Board, Department of Science and Technology, Government of India, Grant reference SR/S4/MS: 866/13.

References

Andersen, T. G., Thyrsgaard, M., and Todorov, V. (2019). Time-varying periodicity in intraday volatility. *Journal of the American Statistical Association*, **114(528)**, 1695-1707.

Bacry, E., Dayri, K., and Muzy, J.-F. (2012). Non-parametric kernel estimation for symmetric hawkes processes. application to high frequency financial data. *The European Physical Journal B*, **85(5)**, 157.

Bacry, E., Mastromatteo, I., and Muzy, J. F. (2015). Hawkes processes in finance. *Market Microstructure and Liquidity*, **1(01)**, 1550005.

Barron, A., Schervish, M. J. and Wasserman, L. (1999). The consistency of posterior distributions in nonparametric problems. *The Annals of Statistics*, **27(2)**, 536-561.

Bauwens, L. and Galli, F. (2009). Efficient importance sampling for ml estimation of scd models. *Computational Statistics and Data Analysis*, **53(6)**, 1974-1992.

Bauwens, L. and Hautsch, N. (2006). Stochastic conditional intensity processes. *Journal of Financial Econometrics*, **4(3)**, 450-493.

Belitser, E. and Serra, P. (2014). Adaptive priors based on splines with random knots. *Bayesian Analysis*, **9(4)**, 859-882.

Belitser, E., Serra, P. and Zanten, H. V. (2013). Estimating the period of a cyclic non-homogeneous poisson process. *Scandinavian Journal of Statistics*, **40(2)**, 204-218.

Belitser, E., Serra, P. and Zanten, H. V. (2015). Rate-optimal Bayesian intensity smoothing for inhomogeneous Poisson processes. *Journal of Statistical Planning and Inference*, **166**, 24-35.

Betancourt, M., Byrne, S., Livingstone, S. and Girolami, M. (2017). The geometric foundations of Hamiltonian Monte Carlo. *Bernoulli*, **23(4A)**, 2257-2298.

Bhogal, S. and Ramanathan, T. (2019). Conditional duration models for high-frequency data: a review on recent developments. *Journal of Economic Surveys*, **33(1)**, 252-273.

Blundell C., Heller, K. A. and Beck, J. M. (2012). Modelling reciprocating relationships with Hawkes processes. In Pereira, F., Burges, C. J. C., Bottou, L., and Weinberger, K. Q., editors, *Advances in Neural Information Processing Systems*, **25**, 2600-2608.

Bowsher, C. G. (2007). Modelling security market events in continuous time: Intensity based, multivariate point process models. *Journal of Econometrics*, **141(2)**, 876-912.

Brémaud, P. and Massoulié, L. (1996). Stability of nonlinear Hawkes processes. *The Annals of Probability*, **24(3)**, 1563-1588.

Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J., Li, P., and Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of Statistical Software*, **76(1)**, 1-32.

Da Fonseca, J. and Zaatour, R. (2014). Hawkes process: Fast calibration, application to trade clustering, and diffusive limit. *Journal of Futures Markets*, **34(6)**, 548-579.

De Boor, C. (1978). *A Practical Guide to Splines*, Springer-Verlag, New York.

Donald B. and Percival, A. T. W. (2000). *Wavelet Methods for Time Series Analysis*. Cambridge University Press, Cambridge.

Donnet, S., Rivoirard, V., Rousseau, J. and Scricciolo, C. (2018). Posterior concentration rates for empirical Bayes procedures with applications to Dirichlet process mixtures. *Bernoulli*, **24(1)**, 231-256.

Engle, R. F. and Russell, J. R. (1998). Autoregressive conditional duration: a new model for irregularly spaced transaction data. *Econometrica*, **66(5)**, 1127-1168.

Fauth, A. and Tudor, C. A. (2012). Modeling first line of an order book with multivariate marked point processes. *arXiv: Trading and Market Microstructure*, https://arxiv.org/pdf/1211.4157.pdf .

Florescu, I., Mariani, M., H.E., S., and Viens, F. (2016). *Handbook of High-Frequency Trading and Modeling in Finance*. Wiley, New York.

Ghosal, S., Ghosh, J. K. and Ramamoorthi, R. (1999). Posterior consistency of Dirichlet mixtures in density estimation. *Annals of Statistics*, **27(1)**, 143-158.

Ghosal, S. and Van der Vaart, A. (2017). *Fundamentals of Nonparametric Bayesian Inference*, Cambridge University Press, Cambridge.

Gregoriou, G. N. (2015). *Handbook of High Frequency Trading*. Academic Press, New York.

Gulddahl, R. J. (2013). Bayesian inference for Hawkes processes. *Methodology and Computing in Applied Probability*, **15(3)**, 623-642.

Härdle, W., Kerkyacharian, G., Picard, D. and Tsybakov, A. (2012). *Wavelets, Aapproximation, and Statistical Applications*, Lecture Notes in Statistics, Springer Science and Business Media, New York.

Hautsch, N. (2004). *Modelling Irregularly Spaced Financial Data: Theory and Practice of Dynamic Duration Models*. Springer, New York.

Hawkes, A. G. (1971). Spectra of some self-exciting and mutually exciting point processes. *Biometrika*, **58(1)**, 83-90.

Hawkes, A. G. (2018). Hawkes processes and their applications to finance: a review. *Quantitative Finance*, **18(2)**, 193-198.

Hawkes, A. G. and Oakes, D. (1974). A cluster process representation of a self-exciting process. *Journal of Applied Probability*, **11(3)**, 493-503.

Kutoyants, Y. A. (1998). *Statistical Inference for Spatial Poisson Processes*. Springer-Verlag, New York.

Mishra, A. and Ramanathan, T. (2017). Nonstationary autoregressive conditional duration models. *Studies in Nonlinear Dynamics and Econometrics*, **21(4)**, https://doi.org/10.1515/snde-2015-0057.

Ogata, Y. (1978). The asymptotic behaviour of maximum likelihood estimators for stationary point processes. *Annals of the Institute of Statistical Mathematics*, **30(1)**, 243-261.

Ozaki, T. (1979). Maximum likelihood estimation of Hawkes' self-exciting point processes. *Annals of the Institute of Statistical Mathematics*, **31(1)**, 145-155.

Pacurar, M. (2008). Autoregressive conditional duration models in finance: a survey of the theoretical and empirical literature. *Journal of Economic Surveys*, **22(4)**, 711-751.

Polasek, W. (2012). Handbook of Markoc Chain Monte Carlo edited by Steve Brooks, Andrew Gelman, Galin Jones, Xiao-Li Meng. *International Statistical Review*, **80(1)**, 184-185.

Rivoirard, V. and Rousseau, J. (2012). Posterior concentration rates for infinite dimensional exponential families. *Bayesian Analysis*, **7(2)**, 311-334.

Russell, J. R. (1999). Econometric modeling of multivariate irregularly-spaced high-frequency data. Manuscript, GSB, University of Chicago.

Schwartz, L. (1965). On Bayes procedures. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, **4(1)**, 10-26.

Shen, W. and Ghosal, S. (2015). Adaptive Bayesian procedures using random series priors. *Scandinavian Journal of Statistics*, **42(4)**, 1194-1213.

Viens, F., Mariani, M., and Florescu, I. (2011). *Handbook of Modeling High-Frequency Data in Finance*. Wiley, New York.

Walker, S. (2004). New approaches to Bayesian consistency. *The Annals of Statistics*, **32(5)**, 2028-2043.

Weinberg, J., Brown, L. D., and Stroud, J. R. (2007). Bayesian forecasting of an inhomogeneous Poisson process with applications to call center data. *Journal of the American Statistical Association*, **102(480)**, 1185-1198.

Wu, Z. (2012). On the intraday periodicity duration adjustment of high-frequency data. *Journal of Empirical Finance*, **19(2)**, 282-291.