

# HAWKES PROCESSES AND TIME CLUSTERING IN FINANCE

ARIC CUTULI

## 1. INTRODUCTION

On May 6, 2010, U.S. equity markets experienced a then unprecedented intraday price drop and recovery in what would come to be known as the 2010 Flash Crash. Within the five minutes following 2:32 pm EDT, a multi-trillion dollar pullback left the Dow Jones Industrial Average nearly 10% below its opening price for the day. The swift and dramatic drop triggered a half hour of such intense trading volume and market recovery that a 2014 CFTC report would call it “one of the most turbulent periods” in the history of U.S. financial markets[KKST17]. On October 15, 2014, the U.S. Treasury market experienced a similarly drastic excitement[FJKM19], and on October 7, 2016, the sterling took a 9% dive against the dollar in Asian markets before recovering most of the loss within minutes[JCRS17]. These flash events are a byproduct of modernizing electronic markets, and few traditional models of financial processes account for such acute feedback effects. In recent years, however, the Hawkes process has been used to trace these observed phenomena.

$$\lambda(t) = \lambda_0(t) + \int_{-\infty}^t v(t-s) dN(s) \quad (1.1)$$

Classically utilized for the modeling of earthquakes and their aftershocks[Oga88], Hawkes processes are a class of Poisson point processes whose rate is characterized by an exogenous component capturing a latent arrival rate and an endogenous component capturing the realized rate’s dependency on past arrivals, which are recorded in the counting process  $N(s)$ . In particular, through the exogenous term  $\lambda_0(t)$ , any non-stationarities in the underlying process can be accounted for, and the Hawkes kernel  $v$  can be fit to explain the feedback effects that past arrivals have on the present rate. Hence, Hawkes processes offer flexible and interpretable characterizations of many complex phenomena in finance, including, but not limited to, the temporal clustering of trades, mid-price volatility, and credit default events[Bow07, EGG10]. In this article, we outline calibration techniques for Hawkes processes and conclude with practical takeaways.

## 2. EMPIRICAL CALIBRATION

Empirically calibrating the Hawkes process in financial applications consists of pre-processing intra-day data to capture seasonality and then estimating the Hawkes kernel via maximum likelihood estimation. Here, we paraphrase much of section 9.3 of [BBDG18] but outline derivation steps that are omitted from the text.

**2.1. Capture Intra-day Seasonality.** Calendar time can be re-parameterized as business time by measuring the the intra-day activity pattern  $\bar{\lambda}(t)$  as the average activity at time  $t$  over many

days and subsequently defining

$$\hat{t}(t) = \int_0^t \bar{\lambda}(s) ds.$$

Doing so allows for the re-parameterized activity to appear stationary i.e.  $\lambda_0(t) = \lambda_0$ .

**2.2. Estimate the Hawkes Kernel.** Given the process measured in business time, the Hawkes kernel can be estimated using traditional statistical methods such as maximum likelihood estimation, as mentioned earlier.

**2.2.1. Maximum Likelihood Estimation.** Assuming that the kernel function  $v(t)$  can be expressed in a parametric form with a parameter vector  $\boldsymbol{\theta}$ , maximum likelihood estimation (MLE) can be employed to estimate  $\boldsymbol{\theta}$  based on an observed realization of the point process. For  $N$  events occurring in  $[0, T]$ , the interval can be discretized into subintervals so that the likelihood function can be written as

$$\begin{aligned} \mathbf{L}(t_1, t_2, \dots, t_N | \boldsymbol{\theta}) &= \left( \prod_{i=1}^N \lambda(t_i | \boldsymbol{\theta}) e^{-\int_{t_{i-1}}^{t_i} \lambda(s | \boldsymbol{\theta}) ds} \right) e^{-\int_{t_N}^T \lambda(s | \boldsymbol{\theta}) ds} \\ &= \left( \prod_{i=1}^N \lambda(t_i | \boldsymbol{\theta}) \right) e^{-\int_0^T \lambda(s | \boldsymbol{\theta}) ds} \end{aligned}$$

and the log-likelihood as

$$\log \mathbf{L}(t_1, t_2, \dots, t_N | \boldsymbol{\theta}) = \sum_{i=1}^N \log \lambda(t_i | \boldsymbol{\theta}) - \int_0^T \lambda(s | \boldsymbol{\theta}) ds. \quad (2.1)$$

Here,

$$\lambda(t_i | \boldsymbol{\theta}) e^{-\int_{t_{i-1}}^{t_i} \lambda(s | \boldsymbol{\theta}) ds} dt$$

is the probability of an arrival coming in  $[t_i, t_i + dt)$  and not in  $(t_{i-1}, t_i)$ , which is a definition of inhomogeneous Poisson processes. It is important here to note that while a closed-form solution to  $\boldsymbol{\theta}$  would offer nice statistical properties, it is often not attainable in practice unless we assume that the kernel function has a certain parametric form. If this assumption is omitted, it is best to estimate (1.1) via numerical optimization or algorithmically[OC13]. However, imposing, for example, that  $v$  is of exponential form (i.e.  $v(t) = gwe^{-wt}$ ) allows for the direct computation of the maximum likelihood estimator for  $\boldsymbol{\theta}$ .

With this exponential kernel, the parameter vector is  $\boldsymbol{\theta} = (\lambda_0, w, g)$ , where  $w^{-1}$  determines the time at which events stop exciting the process and  $g = \int_0^\infty v(s) ds$  is the norm or branching ratio i.e. the expected number of child events that a parent event produces. A larger  $g$  implies that a large fraction of events come in response to past events i.e. the activity is highly endogeneous. The stochastic differential equation governing the process can be derived by taking the differential on

either side of (1.1):

$$\begin{aligned}
d\lambda(t) &= d\lambda_0 + d \int_{-\infty}^t v(t-s) dN(s) \\
&= \left( \int_{-\infty}^t \frac{dv(t-s)}{dt} dN(s) \right) dt + v(0) dN(t) \\
&= -w \left( \int_{-\infty}^t v(t-s) dN(s) \right) dt + gw dN(t) \\
&= -w(\lambda(t) - \lambda_0) dt + gw dN(t).
\end{aligned}$$

This essentially describes the intensity  $\lambda(t)$  as decaying exponentially toward  $\lambda_0$  but being excited by an amount  $gw$  when an event occurs.

The exponential Hawkes is Markovian in nature, and, given arrival times  $t_1, t_2, \dots, t_N$  in  $(0, T]$ , this enables (1.1) to be discretized to

$$\begin{aligned}
\lambda(t_i) &= \lambda_0 + \sum_{k=1}^{i-1} v(t_i - t_k) \\
&= \lambda_0 + gw \sum_{k=1}^{i-1} e^{-w(t_i - t_k)}.
\end{aligned}$$

Letting  $Z_i = \sum_{k=1}^{i-1} e^{-w(t_i - t_k)}$ , (2.1) can be simplified to

$$\begin{aligned}
\log \mathbf{L}(t_1, t_2, \dots, t_N | \boldsymbol{\theta}) &= \sum_{i=1}^N \log \lambda(t_i | \boldsymbol{\theta}) - \int_0^T \lambda(s | \boldsymbol{\theta}) ds \\
&= \sum_{i=1}^N \log(\lambda_0 + gw Z_i) - \int_0^T \left( \lambda_0 + gw \sum_{k=1}^{N(s)-1} e^{-w(t_{N(s)} - t_k)} \right) dt_{N(s)} \\
&= \sum_{i=1}^N \log(\lambda_0 + gw Z_i) - \lambda_0 T - \int_0^T gw \sum_{k=1}^{N(s)-1} e^{-w(t_{N(s)} - t_k)} dt_{N(s)} \\
&= \sum_{i=1}^N \log(\lambda_0 + gw Z_i) - \lambda_0 T - \sum_{i=1}^N g \int_{t_i}^T w e^{-w(t_{N(s)} - t_i)} dt_{N(s)} \\
&= \sum_{i=1}^N \log(\lambda_0 + gw Z_i) - \lambda_0 T - \sum_{i=1}^N g(e^{-w(T-t_i)} - 1) \\
&= -\lambda_0 T + \sum_{i=1}^N \left[ \log(\lambda_0 + gw Z_i) - g(e^{-w(T-t_i)} - 1) \right].
\end{aligned}$$

Evaluating this expression computationally is an  $O(N^2)$  operation, but  $Z_i$  can be written as a recursion such that the problem is reduced to  $O(N)$ . Namely, recognizing

$$Z_i = e^{-w(t_i - t_{i-1})} \sum_{k=1}^{i-1} e^{-w(t_{i-1} - t_k)} = (1 + Z_{i-1}) e^{-w(t_i - t_{i-1})}$$

allows for the log likelihood and its gradients to be efficiently computed in order to find the MLE for  $\boldsymbol{\theta}$ .

Another popular kernel function is the power-law kernel, which can be approximated as a sum of finite exponential kernels[BC07], and we omit discussing its MLE procedure for brevity. Empirically, the power-law kernel is the more often observed feedback structure[HBB13, BM16].

While MLE is an effective tool, the estimator can be quite biased in small samples and very sensitive to the choice of kernel function over short lags. An alternative method for calibrating the Hawkes process is the method of moments, which consists of matching theoretical and empirically observed first- and second-order moments in order to estimate  $\theta$ . Goodness-of-fit testing can be used to compare the estimated intensity  $\hat{\lambda}(t)$  to the data under the null hypothesis that the data is generated by a Hawkes process, implying the residual process  $\epsilon(t_i) = \int_0^{t_i} \hat{\lambda}(t)dt$  is Poisson with unit intensity[Pap72].

### 3. PRACTICAL TAKEAWAYS

Hawkes processes can be a powerful way to parametrically address the time clustering of events and intermittent arrivals exhibited by many complex phenomena in finance. The key component of a Hawkes model that differentiates it from the more basic Poissonian approach is the self-excitation term that seeks to capture feedback effects from past arrivals.

A brief list of Hawkes representations used in finance includes models of market volatility[FS12], price impact[BM13], and order books[Tok11]. In [CJR14], a high frequency trading strategy is developed that relies on market orders arriving according to a Hawkes process, and [EGG10] use the Hawkes framework to describe portfolio loss due to credit default events.

In practice, such phenomena exhibit high endogeneity (i.e.  $g$  close to 1), suggesting that current markets operate in a close to unstable system (hence the propensity for flash crashes) and/or that Hawkes processes are too simple a framework to model financial markets. The linear Hawkes process that we discussed have been extended by introducing covariates such as price returns[BCT<sup>+</sup>15] and news[RPL14] to improve the ability the explain phenomena. As a financial modeler, it is critical to constantly bear in mind that, as put best in Derman and Wilmott’s manifesto[DW09], “models are not the world.”

### REFERENCES

- [BBDG18] J. Bouchaud, J. Bonart, J. Donier, and M. Gould. *Trades, Quotes and Prices: Financial Markets Under the Microscope*. Cambridge University Press, 2018.
- [BC07] T. Bochud and D. Challet. Optimal approximations of power laws with exponentials: application to volatility models with long memory. *Quantitative Finance*, 7(6):585–589, 2007.
- [BCT<sup>+</sup>15] G. Bormetti, L. M. Calcagnile, M. Treccani, F. Corsi, S. Marmi, and F. Lillo. Modelling systemic price cojumps with hawkes factor models. *Quantitative Finance*, 15(7):1137–1156, 2015.
- [BM13] E. Bacry and J.F. Muzy. Hawkes model for price and trades high-frequency dynamics. *Quantitative Finance*, 14, 2013.
- [BM16] E. Bacry and J.F. Muzy. First- and second-order statistics characterization of hawkes processes and non-parametric estimation. *IEEE Transactions on Information Theory*, 62(4):2184–2202, 2016.
- [Bow07] C. Bowsher. Modelling security market events in continuous time: Intensity based, multivariate point process models. *Journal of Econometrics*, 141(2):876–912, 2007.
- [CJR14] Á. Cartea, S. Jaimungal, and J. Ricci. Buy low, sell high: A high frequency trading perspective. *SIAM Journal on Financial Mathematics*, 5(1):415–444, 2014.

- [DW09] E. Derman and P. Wilmott. The financial modelers’ manifesto. 2009.
- [EGG10] E. Errais, K. Giesecke, and L. R. Goldberg. Affine point processes and portfolio credit risk. *SIAM Journal on Financial Mathematics*, 1(1):642–665, 2010.
- [FJKM19] M. J. Fleming, P. Johansson, F. M. Keane, and J. Meyer. From the vault: A look back at the october 15, 2014, flash rally. *Federal Reserve Bank of New York*, 2019.
- [FS12] V. Filimonov and D. Sornette. Quantifying reflexivity in financial markets: Toward a prediction of flash crashes. *Phys. Rev. E*, 85(5):056108, 2012.
- [HBB13] S. J. Hardiman, N. Bercot, and J. Bouchaud. Critical reflexivity in financial markets: a hawkes process analysis. *The European Physical Journal B*, 86(10):442, 2013.
- [JCRS17] R. Jackson, L. Crowley-Reidy, and A. Schrimpf. The sterling ‘flash event’ of 7 october 2016. *Markets Committee Papers*, (9):1–29, 2017.
- [KKST17] A. Kirilenko, A. S. Kyle, M. Samadi, and T. Tuzun. The flash crash: High-frequency trading in an electronic market. *The Journal of Finance*, 72(3):967–998, 2017.
- [OC13] J. Olson and K. Carley. Exact and approximate em estimation of mutually exciting hawkes processes. *Statistical Inference for Stochastic Processes*, 16:63–80, 2013.
- [Oga88] Y. Ogata. Statistical models for earthquake occurrences and residual analysis for point processes. *Journal of the American Statistical Association*, 83(401):9–27, 1988.
- [Pap72] F. Papangelou. Integrability of expected increments of point processes and a related random change of scale. *Transactions of the American Mathematical Society*, 165:483–506, 1972.
- [RPL14] M. Rambaldi, P. Pennesi, and F. Lillo. Modeling foreign exchange market activity around macroeconomic news: Hawkes-process approach. *Physical review. E, Statistical, nonlinear, and soft matter physics*, 91(1):012819, 2014.
- [Tok11] I. M. Toke. ”market making” in an order book model and its impact on the spread. In *Econophysics of Order-Driven Markets, New Economic Windows*, pages 49–64. Springer Milan, 2011.