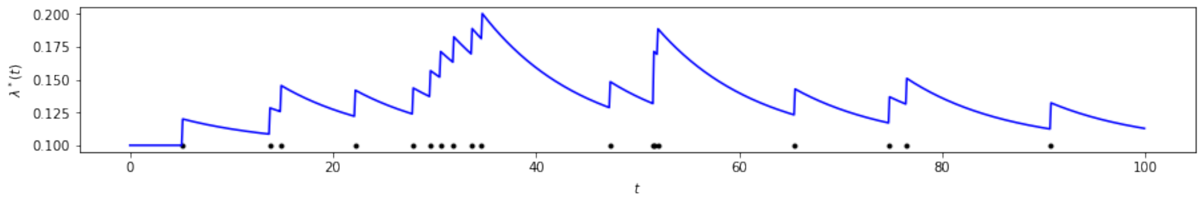
1 Intuition Behind Orderflow Model

A Hawkes process is a member of a family of stochastic processes called point processes. A simple process from this family of processes is something called a Poisson process, which you may be more familiar with. To explain what a Hawkes process is, I'll start by giving an explanation of a Poisson process and then show how it relates to a Hawkes process. An example of a Poisson process from daily experience is cars going past a stop sign. Say during 100 different hours of observation, we notice that on average 20 cars per hours go past a certain stop sign. Then we might model the number of cars that go past the stop sign as a Poisson process with estimated intensity $\hat{\lambda} = 20$. A Hawkes process is just a Poisson process with intensity that increases whenever a car arrives, and then exponentially decays. So a Hawkes process is used to model things that exhibit clustering, including aftershocks from earthquakes, or trade clustering in the stock markets. One realization of the Hawkes Process intensity might look like:



2 Multivariate Hawkes Process

Stock Market Exposition

In the stock market, we know that traders react to the arrivals of both buy and sell market orders. So we expect our buy and sell order arrival-rates to be mutually exciting. We can model this mutual excitation by letting n be the impact a buy order arrival has on the buy order intensity, and by letting v be the impact a sell order impact has on the buy order arrival intensity. I'm going to assume that these impacts are symetrical across buy and sell orders, so v is also the effect that the arrival of a buy order has on sell arrival rates, and n is also the impact that a sell order has on sell arrival rates. I'm also going to assume that both of these processes decay to some background intensity θ at some rate β. This specification can be express with the coupled SDE's:

$$\begin{cases} d\lambda^+ = \beta(\theta - \lambda^+)dt + nd\overline{M_t^+} + vd\overline{M_t^-} \\ d\lambda^- = \beta(\theta - \lambda^-)dt + vd\overline{M_t^+} + nd\overline{M_t^-} \end{cases} \tag{1}$$

Ultimately we want some way to estimate β, θ, n, v from real-world data. To do this, we are going to build a maximum-likelihood estimator for these coefficients. We can start with a heuristic interpretation of the intensity function:

$$\begin{align} \lambda^*(t)dt &= \frac{f^*(t|\mathscr{H}_{t_n})dt}{1 - F^*(t|\mathscr{H}_{t_n})} \tag{2} \\ &= \frac{\mathbb{P}(t_{n+1} \in [t, t+dt]|\mathscr{H}_{t_n})}{\mathbb{P}(t_{n+1} \notin (t_n, t)|\mathscr{H}_{t_n}))} \tag{3} \\ &= \frac{\mathbb{P}(t_{n+1} \in [t, t+dt], t_{n+1} \notin (t_n, t)|\mathscr{H}_{t_n})}{\mathbb{P}(t_{n+1} \notin (t_n, t)|\mathscr{H}_{t_n})} \tag{4} \\ &= \mathbb{P}(t_{n+1} \in [t, t+dt]|t_{n+1} \notin (t_n, t)|\mathscr{H}_{t_n}) \tag{5} \\ &= \mathbb{P}(t_{n+1} \in [t, t+dt]|\mathscr{H}_{t_-}) \tag{6} \\ &= \mathbb{E}[N([t, t+dt))|\mathscr{H}_{t_-}) \tag{7} \end{align}$$

where $N(A)$ is the number of arrivals falling in an interval A and $f^*(t_{n+1}|\mathscr{H}_{t_n})$ is the conditional density function of the next arrival time given the processes history up to time t_n. So already a joint density can be specified as

$$\prod_n f^*(t_n|\mathscr{H}_{t_{n-1}}) \tag{8}$$

Rewriting (1) we have

$$\lambda^*(t) = -\frac{d}{dt}\log(1-F^*(t)) \tag{9}$$

Integrating (7), we see that

$$-\int_{t_k}^{t}\lambda^*(u)du = \log(1-F^*(t)) - \log(1-F^*(t_k)) \tag{10}$$

$$= \log(1-F^*(t)) \tag{11}$$

as $F^*(t_k) = 0$ since a Hawkes process is simple.

$$\implies F^*(t) = 1-\exp(-\int_{t_k}^{t}\lambda^*(u)du), f^*(t) = \lambda^*(t)\exp(-\int_{t_k}^{t}\lambda^*(u)du) \tag{12}$$

Substituting into our expression for the log-likelihood, I get

$$L = \prod_n f^*(t_n) \tag{13}$$

$$= \prod_{i=1}^{k}\lambda^*(t)\exp(-\int_{t_{i-1}}^{t_i}\lambda^*(u)du) \tag{14}$$

$$= (\prod_{i=1}^{k}\lambda^*(t))\exp(-\int_{0}^{t_k}\lambda^*(u)du) \tag{15}$$

Now remember that since the process is observed from 0 to T, we must add a final term $(1-F^*(T))$, the probability of seeing no hits in $(t_k, T]$ to our product, yielding:

$$L = (\prod_{i=1}^{k}\lambda^*(t))\exp(-\int_{0}^{T}\lambda^*(u)du) \tag{16}$$

for a single-variate point process. To arrive at the likelihood we must extend this.

$$\mathscr{L} = (\prod_{i=1}^{k}\lambda^+(t))\exp(-\int_{0}^{T}\lambda^+(u)du)(\prod_{i=1}^{k}\lambda^-(t))\exp(-\int_{0}^{T}\lambda^-(u)du) \tag{17}$$

Taking the log, we have

$$\log(\mathscr{L}) = \sum_{i=1}^{k}\log(\lambda^+_{t_i}) + \sum_{j=1}^{n}\log(\lambda^-_{t_j}) - \int_0^T\lambda^-(u)du - \int_0^T\lambda^+(u)du \tag{18}$$

Starting from the definition of our Hawkes process (4), consider a time interval over which no influential trades occur. Over this interval our dynamics are

$$\begin{cases} d\lambda^+ = \beta(\theta-\lambda^+)dt \\ d\lambda^- = \beta(\theta-\lambda^-)dt \end{cases} \tag{19}$$

Integrating this between trade times, we see that

$$\begin{cases} \lambda^+(t_n) = \theta - e^{-\varepsilon(t_n - t_{(n-1)+})}(\theta - \lambda^+(t_{(n-1)} - \mathscr{B}^+_{n-1}(n,v)) \\ \lambda^-(t_n) = \theta - e^{-\varepsilon(t_n - t_{(n-1)+})}(\theta - \lambda^-(t_{(n-1)} - \mathscr{B}^-_{n-1}(n,v)) \end{cases} \tag{20}$$

Where $\mathscr{B}^+_{n-1}(n,v) = n$ if the influential buy order that arrived was an influential buy and $\mathscr{B}^+_{n-1}(n,v) = v$ if it was an influential sell, and vice-versa for $\mathscr{B}^-_{n-1}(n,v)$. Armed with this, we see that

$$\lambda_t^{\pm} = \theta + \sum_{i=1}^{n} H_i^{\pm} e^{-\beta(t-t_i)} \tag{21}$$

with $H_i^{\pm} = (B_i n + (1-B_i)v, B_i v + (1-B_i)n)$ where B_i is 1 if the order is a buy and 0 if it is a sell by induction on the trades. Looking at

$$\int_0^t \lambda_t^{\pm}(u)du \tag{22}$$

Subsituting our form for lambda, we get

$$\begin{cases} \int_0^T \lambda^+(u)du = \theta T + \sum_{inf BUYS} \frac{-n}{\beta} e^{-\beta(T-t_{buy})} + \sum_{inf SELLS} \frac{-v}{\beta} e^{-\beta(T-t_{sell})} \\ \int_0^T \lambda^-(u)du = \theta T + \sum_{inf BUYS} \frac{-v}{\beta} e^{-\beta(T-t_{buy})} + \sum_{inf SELLS} \frac{-n}{\beta} e^{-\beta(T-t_{sell})} \end{cases} \tag{23}$$

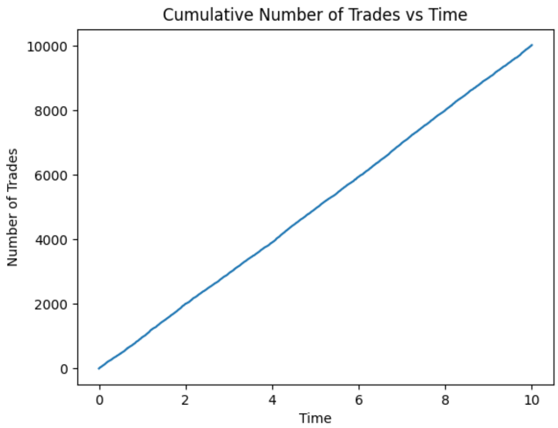
Substituting into our likelihood, we finally get

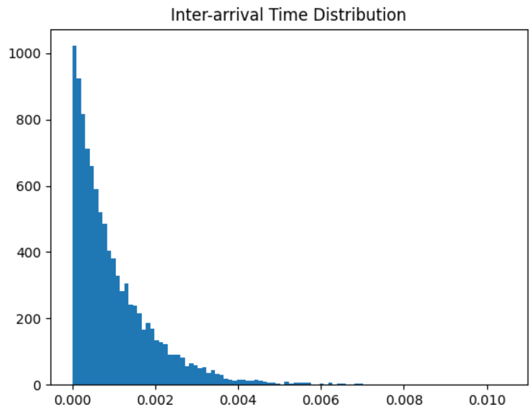
$$\log(\mathscr{L}) = \sum_{i=1}^{k} \log(\lambda_{t_i}^+) + \sum_{j=1}^{n} \log(\lambda_{t_j}^-) - 2\theta T + \sum_{inf TRADES} \frac{n+v}{\beta} e^{-\beta(T-t_i)} \tag{24}$$

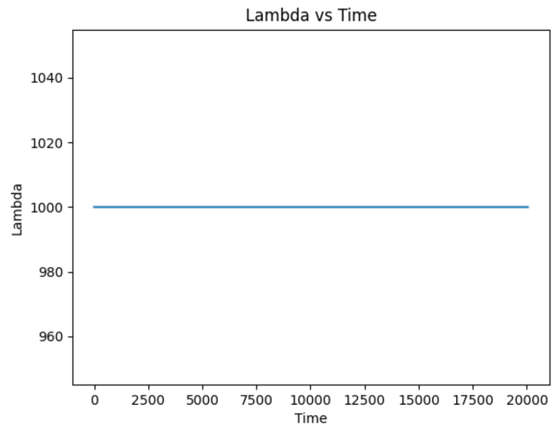
Maximizing this log-likelihood numerically (with initial estimate for theta being the average number of buy/sell trades per hours) will give us the esimtates for β, θ, n, v.

Interesting Things to Note

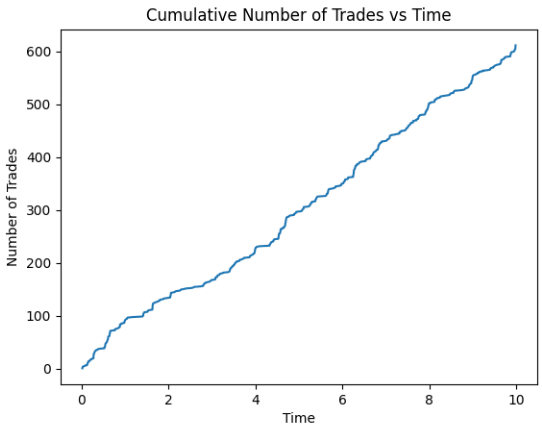
For certain values of β, θ, n, v, the estimator will immeadiately converge and won't update the values of β, n, v. One situation this may happen with is when β sufficiently high compared to n and v. In this case the intensity of the process decays so quickly that its realization still looks statistically like a Poisson process. In general, to determine if a Hawkes process is a good model for a dataset, statistical tests including the Kolmogorov-Smirnov test can be used. As an example, here is data for a Hawkes process with $\beta = 500, \theta = 10000, n = 0, v = 0$:

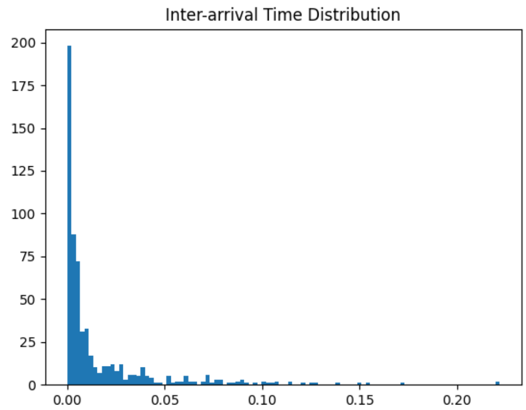


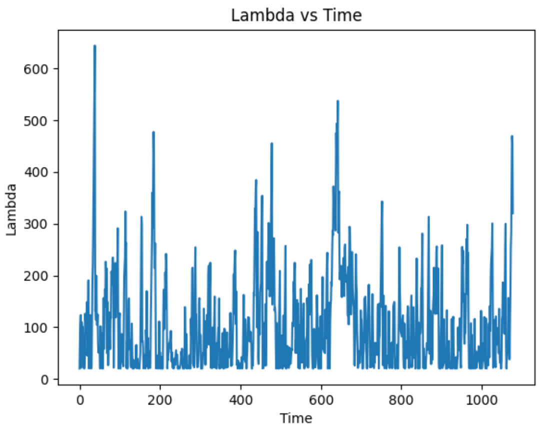




Notice how the data for this Hawkes process exhibits all the features of a Poisson process (exponentially distributed inter-arrival times, constant lambda) - it's a degenerate Hawkes process, and our estimator immeadiately converges. An infinite number of β, n, v could satisfy this data, as long as β is significantly large compared to n and v. Using a KS test would help you evaluate the fit of this model. As an opposite example, here is data for a Hawkes process with $\beta = 200, \theta = 20, n = 100, v = 20$:







This data does not exhibit the features of a Poisson process, and our estimator would effectively search for values of β, n, v. A KS test would help us determine that a Hawkes model would be appropriate.

3 Resources

1. Rasmussen, Jakob Gulddahl. “Lecture Notes: Temporal Point Processes and the Conditional Intensity Function.” arXiv.Org, 1 June 2018, arxiv.org/abs/1806.00221.

2. Laub, Patrick J., et al. “Hawkes Processes.” arXiv.Org, 10 July 2015, arxiv.org/abs/1507.02822.