Hawkes Processes in Finance

An overview

D. Astaburuaga[1,2]

[1]Departamento de Matemática
Universidad Técnica Federico Santa María, Casa Central.

[2]Departamento de Industrias
Universidad Técnica Federico Santa María, Casa Central.

Keep Learing Proyect, May 2025

Content

Background of Point Processes

Definition

A **point process** is a stochastic process that consists of random collections of points (or events) occurring in time or space. In one dimension (time), it can be seen as a random sequence of arrival times $\{t_1, t_2, \ldots\}$.

- It generalizes the Poisson process, which has exponentially distributed interarrival times.
- It can be described by its **conditional intensity function** $\lambda(t)$, which measures the instantaneous likelihood of an event at time t, given the history up to time t.

Example

The times of arrival of buy and sell orders in a financial market form a point process on the real line.

Why Hawkes Processes?

The Poisson process has limitations as:

- It assumes that events are independent of each other.
- It cannot capture the clustering behavior observed in real-world data (e.g., trades or market orders arriving in bursts).

In this sense, the **Hawkes Process** extends the Poisson process by including a self-exciting behavior, meaning that the occurrence of an event increases the likelihood of future events in the short term.
This property makes Hawkes processes particularly useful for:

- Modeling aftershocks in seismology.
- Capturing volatility clustering and order arrivals in financial time series.

Overview of Financial Applications

- **High-Frequency Trading (HFT):** Model arrival times of trades and orders to capture market microstructure effects.
- **Volatility Clustering:** Explain periods of intense activity and calm in asset price volatility.
- **Systemic Risk and Credit Contagion:** Model how the default of one firm can trigger subsequent defaults (domino effect).
- **Cross-Asset Excitation:** Use multivariate Hawkes processes to understand dependencies across multiple assets or markets.
- **Market Impact Modeling:** Assess how large trades influence future order flows and prices.

Counting Processes

Definition

A **counting process** $N(t)$, defined for $t \geq 0$, is a stochastic process that models the cumulative number of events occurring up to time t. Formally:

- $N(0) = 0$, and
- $N(t) = \sum_{t_i \leq t} 1$,

where $\{t_i\}$ denotes the set of event times.

- **Interpretation:** $N(t)$ counts how many events have occurred by time t.
- **Examples:** Number of trades in a financial market or number of order arrivals up to time t.

Conditional Intensity

Definition

The **conditional intensity process** of N_t is defined by:

$$\lambda(t) = \lim_{\Delta \downarrow 0} \frac{\mathbb{E}[N_{t+\Delta} - N_t \mid \mathcal{H}_t]}{\Delta}, \quad t \geq 0,$$

if this limit exists.
Here, $\mathcal{H}_t$ (the *history* of N_t) denotes the filtration generated by $\{N_s\}_{s<t}$.
We use λ_t and $\lambda(t)$ interchangeably.

- **Interpretation:** It measures the instantaneous rate at which events are expected to occur at time t, given the past up to t.
- **Key Concept:** Conditional intensity characterizes the dependence of the process on its own history.

Hawkes Process

Definition

A **Hawkes process** is a counting process N_t whose conditional intensity process, for $t \geq 0$, is given by:

$$\lambda(t) = \mu(t) + \sum_{t_i < t} \alpha(t - t_i),$$

where:

- $\mu(t) > 0$ is the **background (exogenous) intensity**, and
- $\alpha : \mathbb{R}^+ \to \mathbb{R}^+$ is the **excitation (kernel) function**, which models how past events influence the future intensity.

- **Self-exciting property:** Each event increases the likelihood of future events, capturing clustering effects in financial data.

Kernel Functions

Common choices for the excitation (kernel) function $\alpha(\cdot)$ in Hawkes processes include:

- **Exponential kernel:**

$$\alpha(t) = \alpha_0 e^{-\beta t}, \quad \alpha_0, \beta > 0.$$

  It captures short-term memory and leads to Markovian dynamics.

- **Power-law kernel:**

$$\alpha(t) = \frac{c}{(t + c_0)^{1+\gamma}}, \quad c, c_0, \gamma > 0.$$

  It captures long-range dependence and heavy tails in financial time series.

- The choice of kernel determines how past events influence the process over time.

Stationarity and Stability Conditions

- **Stationarity:** A Hawkes process is stationary if its statistical properties do not depend on time.
- **Branching ratio:** The integral of the kernel function:

$$\eta = \int_0^\infty \alpha(s)\, ds.$$

- **Stability condition:** For the process to be stationary and not explode, it is required that:

$$\eta < 1.$$

- In financial applications, η represents the average number of offspring events triggered by a single event (degree of endogeneity).

Ogata's Thinning Algorithm

- A popular algorithm for simulating non-homogeneous point processes[1].
- Steps:
  1. Initialize $t = 0$, $N = 0$.
  2. At each iteration, simulate a candidate event time using an upper bound λ^*.
  3. Accept or reject the candidate using a uniform random variable and the ratio $\lambda(t)/\lambda^*$.
  4. Update the intensity $\lambda(t)$ with the newly accepted events.
  5. Repeat until desired time horizon T is reached.

[1] This algorithm is included in the library `tick`.

Maximum Likelihood Estimation (MLE)

- The parameters of a Hawkes process (e.g., μ, α, β) can be estimated by maximizing the likelihood function.
- The log-likelihood function for a Hawkes process up to time T is:

$$\ell(\theta) = \sum_{t_i \leq T} \log \lambda(t_i) - \int_0^T \lambda(s)ds,$$

  where:
  - $\{t_i\}$ are observed event times.
  - θ are the parameters to estimate.
  - $\lambda(t)$ is the conditional intensity at time t.
- The MLE involves balancing the fit to the observed intensities and the cumulative intensity (compensator).

Numerical Optimization

- The log-likelihood function is generally non-linear and requires numerical optimization.
- Common algorithms:
  - **Gradient-based methods:** e.g., BFGS, L-BFGS.
  - **Expectation-Maximization (EM) algorithm:** for exponential kernels.
  - **Adaptive methods:** e.g., Adam for large-scale problems.
- **Software tools:** Libraries such as `tick` (Python) and `PtProcess` (R) provide built-in optimization routines for Hawkes processes.

Goodness-of-Fit

- Assessing the adequacy of the Hawkes process model is crucial.
- **Time-rescaling theorem:** Transform event times $\{t_i\}$ into residuals $\{\tau_i\}$:

$$\tau_i = \int_{t_{i-1}}^{t_i} \lambda(s)\, ds.$$

  If the model is correct, $\{\tau_i\}$ should be i.i.d. exponential(1).
- **Diagnostic tools:**
  - QQ-plots or KS-plots of $\{\tau_i\}$ against the exponential(1) distribution.
  - Residual-based tests for over-dispersion or clustering.
- **Visual checks:** Plotting conditional intensities vs. realized events.

Modeling High-Frequency Financial Data

- **Hawkes processes** have been extensively used to model the arrival times of market orders, trades, and limit order book updates.
- They capture the **self-exciting and clustering behavior** typical of high-frequency trading (HFT) data.
- **Example:** Bacry et al. (2015) [**bacry2015**] use Hawkes processes to describe the dynamics of order arrivals in equity markets.
- In practice, high-frequency event data (millisecond resolution) from exchanges like NASDAQ or Euronext are modeled with Hawkes processes.
- **Real Data:** Limit order book (LOB) data with timestamped trades can be analyzed to calibrate Hawkes models.

Systemic Risk and Credit Contagion

- In credit risk, Hawkes processes help model **default contagion** and systemic risk.
- **Jarrow and Yu (2001):** Use self-exciting point processes to capture the impact of one firm's default on the default probabilities of others.
- **Aït-Sahalia et al. (2014):** Develop a contagion framework for credit markets using multivariate Hawkes processes.
- These models can incorporate **default intensities** that jump upon contagion events, linking default clustering to financial crises.
- **Application:** Stress-testing and scenario analysis of default cascades in large credit portfolios.

Volatility Clustering

- Hawkes processes can be used to model **volatility bursts** in asset returns, connecting microstructure events to macroscopic volatility.
- **Example:** Bacry et al. (2013) show that trade arrival processes modeled by Hawkes processes explain volatility clustering observed in financial time series.
- **Marked Hawkes models:** Incorporate trade sizes (marks) to directly link order flow to price volatility.
- **Real Data:** Empirical studies on equity and FX markets show that the intensity of trade arrival processes strongly correlates with volatility spikes.

Recent Research Directions

- **Non-parametric estimation:** Methods to estimate the kernel shape from data directly (e.g., Bacry and Muzy, 2014).
- **Marked Hawkes processes:** Including trade volumes, bid-ask spreads, or asset classes to capture richer dependencies.
- **Deep learning approaches:** Neural networks for estimating complex, non-linear dependencies in Hawkes processes (Zhu, 2020).
- **Rough Hawkes processes:** Linking rough volatility models and Hawkes processes to explain the persistence of volatility (El Euch et al., 2020).

- Active area of research combining high-frequency data, statistical learning, and market microstructure.

Marked Hawkes Processes

- **Motivation:** In financial markets, events have additional attributes (marks) such as trade size, price impact, or spread.
- **Definition:** A **marked Hawkes process** associates a random mark m_i to each event time t_i.
- **Conditional intensity:**

$$\lambda(t) = \mu + \sum_{t_i < t} \alpha(t - t_i, m_i),$$

  where the kernel depends on both the time since past events and their marks.
- **Applications:**
  - Modeling how large trades affect future order flow and market volatility.
  - Capturing the impact of order book events on subsequent market activity.

Multivariate Hawkes Processes

- **Motivation:** Many financial systems involve interactions across multiple assets, markets, or event types.
- **Definition:** A d-dimensional Hawkes process $\mathbf{N}(t) = (N_1(t), \ldots, N_d(t))$ with conditional intensities:

$$\lambda_i(t) = \mu_i + \sum_{j=1}^{d} \int_{-\infty}^{t} \phi_{ij}(t-s)\, dN_j(s).$$

- **Cross-excitation:** Events in process j increase the intensity of process i, capturing mutual excitation (e.g., between stocks in the same sector).
- **Applications:**
  - Modeling joint arrival of orders in correlated assets.
  - Spillover effects across asset classes or markets.

Branching Structure and Endogeneity

- **Branching representation:** Hawkes processes can be interpreted as a **branching process**:
  - Immigrants (exogenous events) arrive with intensity μ.
  - Each event triggers offspring (endogenous events) according to the kernel $\alpha(\cdot)$.
- **Branching ratio:**

$$\eta = \int_0^\infty \alpha(s)\, ds.$$

- **Financial interpretation:** η measures the degree of endogeneity:
  - $\eta < 1$ for stationarity (sub-critical regime).
  - $\eta \approx 1$ implies a large fraction of events are self-excited (high endogeneity in markets).

Power-Law Kernels and Long-Range Dependence

- **Limitation of exponential kernels:** They decay too fast to capture long memory effects in financial markets.
- **Power-law kernels:**

$$\alpha(t) = \frac{c}{(t + c_0)^{1+\gamma}}, \quad \gamma > 0,$$

  which decay more slowly and capture long-range dependence.
- **Applications:**
  - Explaining long memory in order flow (e.g., persistence in trade signs).
  - Modeling rough volatility phenomena in high-frequency data.
- **Example:** Bacry et al. (2013) link power-law Hawkes kernels to volatility clustering and rough volatility models.

Data Requirements and Challenges

- **Data precision:** High-resolution timestamps (microseconds or milliseconds) are essential to capture true dynamics.
- **Order book data:** Level-1 and Level-2 data provide insights into order arrivals and cancellations.
- **Market fragmentation:** Data from multiple trading venues can be complex to consolidate.
- **Challenges:**
  - Handling asynchronous data feeds.
  - Correctly identifying trade initiators.
  - Adjusting for market microstructure noise.

Example Case Study

- **Context:** Analysis of trade arrivals in NASDAQ LOB data for a large-cap stock (e.g., AAPL).
- **Steps:**
  1. Preprocess trade data: filter for single-stock events, merge trades at the same timestamp.
  2. Fit an exponential Hawkes model to the trade arrival times.
  3. Estimate parameters using MLE (e.g., with `tick`).
  4. Evaluate branching ratio to assess endogeneity of order flow.
- **Findings:** High branching ratio ($\eta \approx 0{,}7$) suggests substantial endogenous dynamics in trade arrivals.

Limitations and Model Misspecification

- **Model assumptions:**
  - Exponential kernels may be too simplistic for real data.
  - Stationarity assumption may not hold in turbulent market regimes.
- **Overfitting risk:** Fitting too many parameters can lead to poor out-of-sample performance.
- **Exogenous events:** Hawkes models may miss external drivers like macroeconomic news or regulatory changes.
- **Robustness:** Model misspecification can bias parameter estimates and risk measures.

Key Takeaways

- Hawkes processes are powerful tools to model self-exciting and clustering behaviors in financial events.
- Widely used in high-frequency trading, systemic risk analysis, and volatility modeling.
- Calibration involves MLE, numerical optimization, and residual-based model validation.
- Multivariate and marked extensions enrich the modeling capabilities for real financial data.

Current Research and Open Questions

- **Non-parametric estimation:** How to flexibly learn kernel shapes from data.
- **Rough volatility:** Connections between power-law kernels and rough volatility dynamics.
- **High-dimensional applications:** Scaling Hawkes models to many assets and event types.
- **Incorporating exogenous drivers:** Integrating macroeconomic or fundamental signals into Hawkes-type models.
- **Deep learning approaches:** Using neural networks to approximate complex, non-linear intensities.

References and Disclaimer

- **References:**
  - Hawkes Models And Their Applications. Patrick J. Laub, Young Lee, Philip K. Pollett, Thomas Taimre.
  - Modelling financial high frequency data using point processes. Luc Bauwens, Nikolaus Hautsch.
- **Disclaimer:** This Beamer presentation was created using a ChatGPT customized model to accelerate content development. Double-check all information, equations, and methods before applying them in practice.