Real Markets, Microstructure, Clusters with a Branching Process Point of View

Simone Scotti

Université Paris-Diderot

Beijing International Center for Mathematical Research

Summer School on Mathematical Finance and Insurance

Peking University

Beijing, June 2017

Plan

- "Classic" financial framework
- Beyond classic financial framework : news and orders
- Counting processes and Poisson process
- Statistical analysis
- Hawkes process
- Simulations
- Estimation
- How to evaluate the self-exciting effect : Branching point of view

"Classic" Financial Framework

For sake of simplicity, we focus on the price of stock shares.

Shares represent a fraction of ownership in a business.

Evaluation Methods

Entreprise Value: The firm capital is given by the difference between assets (cash, plants, equipments, trademarks, etc.) and liabilities (bonds, salaries, tax, etc.). The value of the stock is given by the ratio between firm capital and number of shares

Discounted Cash-Flow: Let $\{T_i, D_i\}_{i\in\mathbb{N}}$ be the sequence of dividend D_i payed by the firm at time T_i and ρ the discount factor. The actuarial price is

$$S_t = \sum_i D_i e^{-\rho(t-T_i)} \mathbb{I}_{T_i>t}$$

"Classic" Financial Framework

For sake of simplicity, we focus on the price of stock shares.

Shares represent a fraction of ownership in a business.

Evaluation Methods

Entreprise Value The firm capital is given by the difference between assets (cash, plants, equipments, trademarks, etc.) and liabilities (bonds, salaries, tax, etc.). The value of the stock is given by the ratio between firm capital and number of shares

Discounted Cash-Flow Let $\{T_i, D_i\}_{i\in\mathbb{N}}$ be the sequence of dividend D_i payed by the firm at time T_i and ρ the discount factor. The actuarial price is

$$S_t = \sum_i D_i e^{-\rho(t-T_i)} \mathbb{I}_{T_i>t}$$

Market price The price at which it is possible to buy (or to sell). That is "look at the market" or "market has always reason" rules.

Role of Information

Each one of the previous methods is implicitly based on a random framework, since a large part of the terms (for instance the dividend sequence) are random variables. That is, we have to consider a probability space $(\Omega, \mathcal{A}, \mathbb{P})$

Different point of view are possible to describe the uncertainty. I prefer a “Bayesian” approach to highlight the role of information.

Then the price S_t becomes a stochastic process due to the “lack” of information.

It is then important to discuss how the information arrives to the market.

The information is described in stochastic process theory by the filtration $\{\mathcal{F}_t\}_{t\in\mathbb{R}+}$, that is an increasing family of σ-algebra.

Classical Continuous Paradigm

Classical Rules

- The news arrival is smooth, information is revealed progressively. With mathematical work : the filtration is continuous, that is $\mathcal{F}_{t-} = \mathcal{F}_t$.
- The news flow is white, that is the news of today and tomorrow are independent.
- There is no reason for a seasonality on news arrivals and its distribution.

Classical Continuous Paradigm

Classical Rules

- The news arrival is smooth, information is revealed progressively. With mathematical work : the filtration is continuous, that is $\mathcal{F}_{t-} = \mathcal{F}_t$.
- The news flow is white, that is the news of today and tomorrow are independent.
- There is no reason for a seasonality on news arrivals and its distribution.
- The number of shares does not play any particular role. It is possible to split or merge stocks and the evolution is unaffected.

Classical Continuous Paradigm

Classical Rules

- The news arrival is smooth, information is revealed progressively. With mathematical work : the filtration is continuous, that is $\mathcal{F}_{t-} = \mathcal{F}_t$.
- The news flow is white, that is the news of today and tomorrow are independent.
- There is no reason for a seasonality on news arrivals and its distribution.
- The number of shares does not play any particular role. It is possible to split or merge stocks and the evolution is unaffected.

a Mathematical Framework

The three first rules push to consider a Brownian filtration, i.e. $\mathcal{F}_t = \sigma(W_s, s \leq t)$.
The last rule encourage to use a geometric (multiplicative) model

Classical Continuous Paradigm

Classical Rules

- The news arrival is smooth, information is revealed progressively. With mathematical work : the filtration is continuous, that is $\mathcal{F}_{t-} = \mathcal{F}_t$.
- The news flow is white, that is the news of today and tomorrow are independent.
- There is no reason for a seasonality on news arrivals and its distribution.
- The number of shares does not play any particular role. It is possible to split or merge stocks and the evolution is unaffected.

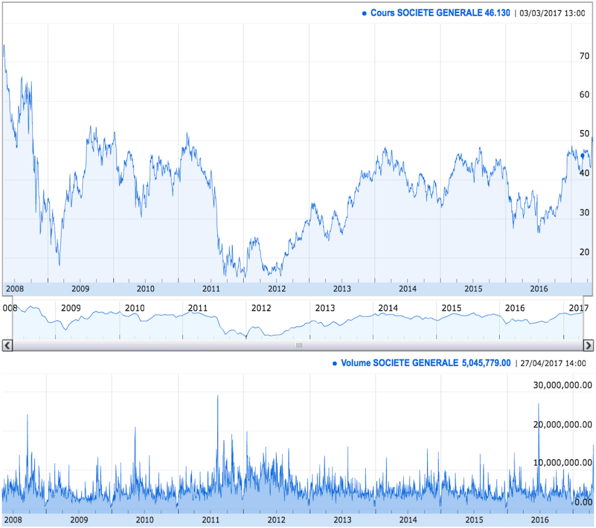
a Mathematical Framework

The three first rules push to consider a Brownian filtration, i.e. $\mathcal{F}_t = \sigma(W_s, s \leq t)$.
The last rule encourage to use a geometric (multiplicative) model

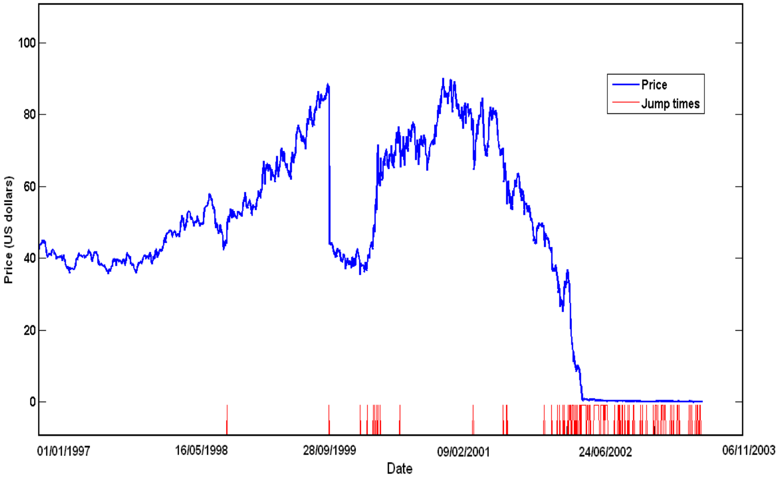
Black Scholes model

$$S_t = S_0 exp\left\{\left(\mu - \frac{1}{2}\sigma^2\right)t + \sigma W_t\right\}$$

Societe Generale



News arrival : ENRON



News arrival : AHOLD

Change of regime

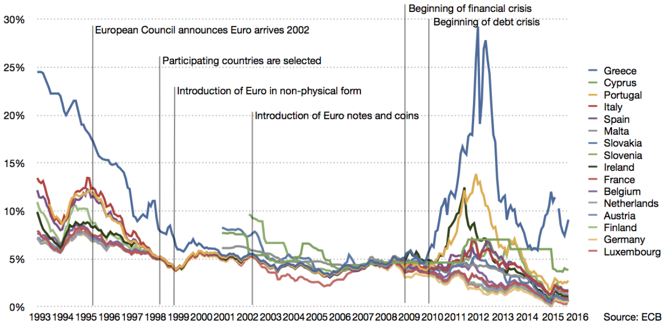


FIGURE: Long term interest rates of Euro area countries.

Why ?

Alice soon came to the conclusion that it was a very difficult game indeed.
The players all played at once without waiting for turns, quarrelling all the while, and fighting for the hedgehogs ; and in a very short time the Queen was in a furious passion, and went stamping about, and shouting "Off with his head !"or "Off with her head !" about once in a minute.

Why ?

Alice soon came to the conclusion that it was a very difficult game indeed. The players all played at once without waiting for turns, quarrelling all the while, and fighting for the hedgehogs ; and in a very short time the Queen was in a furious passion, and went stamping about, and shouting "Off with his head !"or "Off with her head !" about once in a minute.

The Queen's Croquet-Ground
Alice's Adventures in Wonderland
by Lewis Carroll

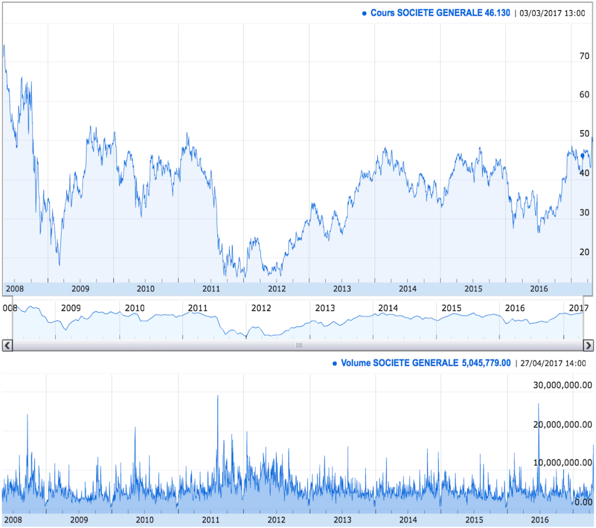
Why ?

Alice soon came to the conclusion that it was a very difficult game indeed.
The players all played at once without waiting for turns, quarrelling all the while, and fighting for the hedgehogs ; and in a very short time the Queen was in a furious passion, and went stamping about, and shouting "Off with his head !"or "Off with her head !" about once in a minute.

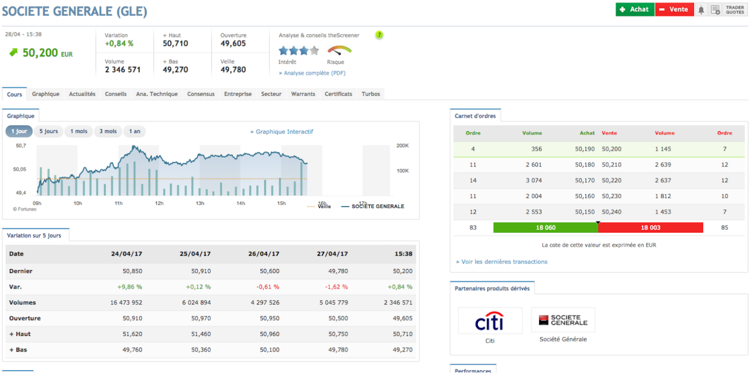
The Queen's Croquet-Ground
Alice's Adventures in Wonderland
by Lewis Carroll

Financial players plays all at once !
Financial markets are, first-of-all, **markets** that is place where people meets to buy and sell.
People or, better, populations have different rules.

Societe Generale



Societe Generale



SOCIETE GENERALE (GLE)

Achat Vente TRADER QUOTES

28/04 - 15:38

50,200 EUR

Variation	+ Haut	Ouverture
+0,84 %	50,710	49,605
Volume	+ Bas	Veille
2 346 571	49,270	49,780

Analyse & conseils theScreener

Intérêt Risque

» Analyse complète (PDF)

Cours Graphique Actualités Conseils Ana. Technique Consensus Entreprise Secteur Warrants Certificats Turbos

Graphique

1 jour 5 jours 1 mois 3 mois 1 an » Graphique Interactif

Variation sur 5 jours

Date	24/04/17	25/04/17	26/04/17	27/04/17	15:38
Dernier	50,850	50,910	50,600	49,780	50,200
Var.	+9,86 %	+0,12 %	-0,61 %	-1,62 %	+0,84 %
Volumes	16 473 952	6 024 894	4 297 526	5 045 779	2 346 571
Ouverture	50,910	50,970	50,950	50,500	49,605
+ Haut	51,620	51,460	50,960	50,750	50,710
+ Bas	49,760	50,360	50,100	49,780	49,270

Carnet d'ordres

Ordre	Volume	Achat	Vente	Volume	Ordre
4	356	50,190	50,200	1 145	7
11	2 601	50,180	50,210	2 639	12
14	3 074	50,170	50,220	2 637	12
11	2 004	50,160	50,230	1 812	10
12	2 553	50,150	50,240	1 453	7
83	18 060			18 003	85

La cote de cette valeur est exprimée en EUR

» Voir les dernières transactions

Partenaires produits dérivés

Citi Société Générale

Performances

Orders

Financial markets are markets. Markets are place where two agents accept to buy/sell an asset. There is then a problem of **counterpart**.

Bargaining

When an agent goes to the market to buy an asset, she has two main strategies

Accept the price: she is sure to buy but paying a very high price.

Bargaining: she propose a price, she is not sure to buy (at the end). This procedure is time-consuming. But the price is lower !

Orders

Financial markets are markets. Markets are place where two agents accept to buy/sell an asset. There is then a problem of **counterpart**.

Bargaining

When an agent goes to the market to buy an asset, she has two main strategies

- **Accept the price** she is sure to buy but paying a very high price.
- **Bargaining** she propose a price, she is not sure to buy (at the end). This procedure is time-consuming. But the price is lower !

Financial Orders

On a market, an agent willing to buy (sell) an asset has two main types of order

- **Market order** she accepts the price, the transaction occurs immediately but she pays an high price (receives a low price).
- **Limit order** she proposes a price and a quantity. Her order is added at the limit bid (ask) book. She has to wait an opposite market order. There is a priority rule. If the transaction occurs, she pays a lower price (receives an higher price).

High frequency trading

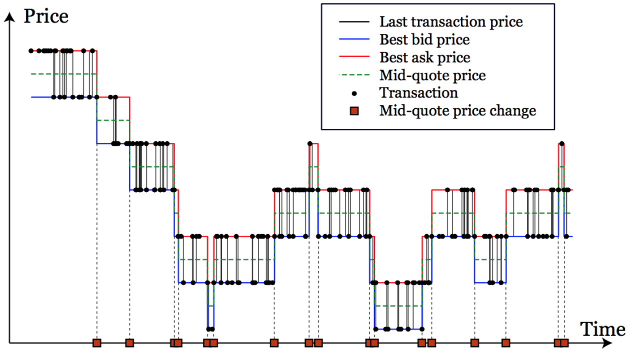


FIGURE: Filimonov et al. *Quantification of the High Level of Endogeneity and of Structural Regime Shifts in Commodity Markets*, 2015.

Financial transactions

Tick effect intrinsically prices evolve by jumps in HFT.

Lack of uniqueness there is at least three prices : last transaction, best bid, best ask.

Clusters the evolution of each of the previous prices exhibits clusters.

Endogenity A large, and fast increasing, part of financial exchange cannot be explained by economic news but only by previous exchanges.

Poisson process and Finance

Result in finance

Big news arrival : arrivals are IID, unexpected (unpredictable) and globally uniformly distributed over wide time windows.

Over/under-reactions Some impacts are highly overestimated or underestimated : jumps gives birth to jumps.

Default times : defaults are unexpected but two defaults never arrives at the same time.

Market microstructure : the financial transactions are globally uniformly distributed over trading days.

Equilibrium issues : unbalance between supply and demand are adjusted immediately.

Change of regime : even consider very regular assets, as bond. The financial data exhibits clusters.

Poisson process and Finance

Result in finance

Big news arrival : arrivals are IID, unexpected (unpredictable) and globally uniformly distributed over wide time windows.

Over/under-reactions Some impacts are highly overestimated or underestimated : jumps gives birth to jumps.

Default times : defaults are unexpected but two defaults never arrives at the same time.

Market microstructure : the financial transactions are globally uniformly distributed over trading days.

Equilibrium issues : unbalance between supply and demand are adjusted immediately.

Change of regime : even considering very regular assets, as bond, the financial data exhibits clusters.

Poisson process and Finance

Result in finance

Big news arrival : arrivals are IID, unexpected (unpredictable) and globally uniformly distributed over wide time windows.

Over/under-reactions Some impacts are highly overestimated or underestimated : jumps gives birth to jumps.

Default times : defaults are unexpected but two defaults never arrives at the same time.

Market microstructure : the financial transactions are globally uniformly distributed over trading days.

Equilibrium issues : unbalance between supply and demand are adjusted immediately.

Change of regime : even considering very regular assets, as bond, the financial data exhibits clusters.

As a matter of fact, big news are as a candy box. Problems look like Pandora's box.

High frequency trading

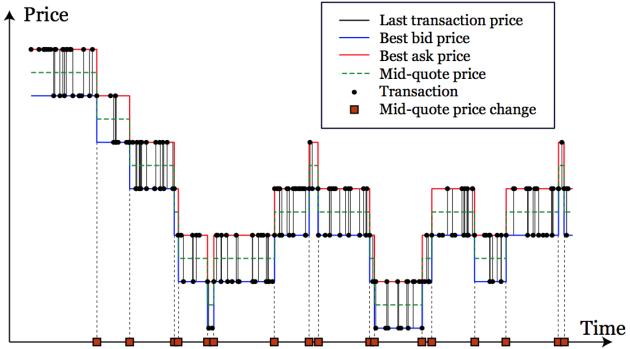


FIGURE: Filimonov et al. *Quantification of the High Level of Endogeneity and of Structural Regime Shifts in Commodity Markets*, 2015.

Equilibrium issues : Sugar

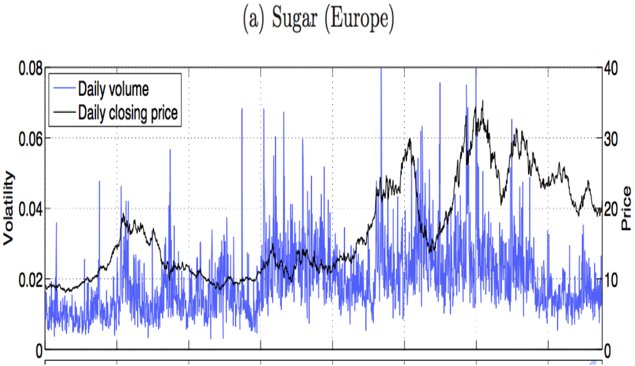


FIGURE: Filimonov et al. *Quantification of the High Level of Endogeneity and of Structural Regime Shifts in Commodity Markets*, 2015.

Equilibrium issues : Energy

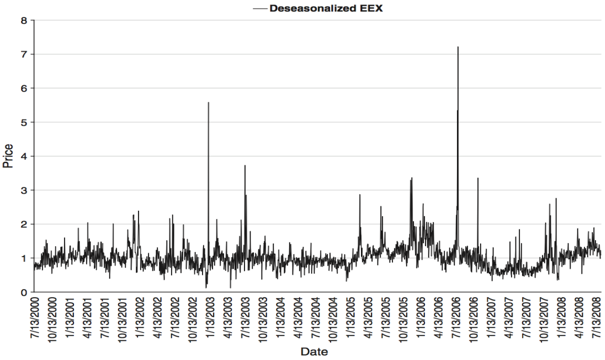


FIGURE: Benth et al. *A critical empirical study of three electricity spot price models*.

Simple point process

We consider a filtered probability space $(\Omega, \mathcal{A}, \{\mathcal{F}_t\}_{t\in\mathbb{R}+}, \mathbb{P})$.
Let $(\tau_i)_{i\in\mathbb{N}}$ be a family of random variables such that

non-negativity $\tau_i \geq 0$;

increasing sequence $\tau_{i+1} > \tau_i, \forall i \in \mathbb{N}$.

The family $\{\tau_i\}_{i\in\mathbb{N}}$ is called a **simple point process**.
An usual convention is to set $\tau_0 = 0$.

Applications in Finance

News arrival : large fluctuations due to unexpected news.

Default times : firms (or countries) stop to reimburse their obligations.

Market microstructure : the financial transactions occur at discrete times.

Equilibrium issues : commodities can exhibit unbalance between supply and demand.

Counting processes and intensity

Let $\mathcal{T} := \{\tau_i\}_{i\in\mathbb{N}}$ be a simple point process. We define the **counting process** N associated with $\mathcal{T}$ as

$$N_t = \sum_{i=1}^{+\infty} \mathbb{I}_{\tau_i \leq t}.$$

Let N be a counting process, we define the related intensity process λ as

$$\lambda(t \,|\, \mathcal{F}_t) := \lim_{h\to 0} \mathbb{E}\left[\left. \frac{N_{t+h} - N_t}{h} \right| \mathcal{F}_t \right].$$

The intensity *depends* on the filtration but, under usual hypotheses, the filtration is the natural one. The indication of the filtration is generally omitted.

Poisson process

Assuming $\lambda(t \,|\, \mathcal{F}_t) = \lambda_0 > 0$. The associated counting process is called a Poisson process.

Properties

- The intensity does not depend on the history (memoryless process), i.e. $N_{t+s} - N_t \perp N_t$.
- The inter-temporal durations $\Delta\tau_i := \tau_i - \tau_{i-1}$ are IID with exponential law $\mathcal{E}(\lambda_0)$.
- A Poisson process never jumps twice at the same time, i.e. $\Delta\tau_i > 0$.

We can easily extend the previous result to inhomogeneous.
An useful result :

Time change Theorem

Let N be a point process with intensity $\lambda(t|\mathcal{F}_t)$. Assuming $\int_0^\infty \lambda(t|\mathcal{F}_t)\,dt = \infty$.
We define the sequence of stopping times $\{t_i\}_{i\in\mathbb{N}}$ by

$$\int_0^{\tau_i} \lambda(t|\mathcal{F}_t)\,dt = t_i.$$

Some statistics

We now focus on the data showed before in a statistical point of view.

The objective is to decide if the data allows us to affirm that the arrival times of the events detailed before do not follow an exponential law.

Two complementary approaches :

1. graphical analysis,
2. statistical tests.

The main conclusion will be that we will **reject** the hypothesis that the arrival times have an exponential law.

Graphical analysis

Graphically there are three main methods to study **the goodness of fit** of the data to a given law.

Density comparaison We plot an histogram of the frequencies comparing with the theoretical density given by the general law (*exponential*) with estimated parameters. A modification of this method is to reconstruct an estimated empirical density coherent with data and compare it with the theoretical one.

CDF comparaison We plot the empirical CDF defined by

$$F_n(x) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{I}_{x_i<x}$$

and we compare with the theoretical one with estimated parameters.

Quantile-Quantile plot Given the empirical and the theoretical CDF, we compute the two inverse functions, i.e. the empirical and the theoretical n-quantile. Then we plot the theoretical quantile (*x-coordinate*) against the empirical one (*y-coordinate*). If the theoretical law fit well then the QQ plot will approximately lie on the line $y = x$.

Exponential law case

1. First estimate the parameter λ of the exponential law, for instance using the reciprocal of the sample mean.

2. Plot the histogram of the true data against the exponential. Using R-language, the code is in two lines :

   ```
   hist(x)
   curve(dexp(x, λ))
   ```

3. Plot the empirical and theoretical CDF :

   ```
   plot(ecdf(x))
   curve(pexp(x, λ))
   ```

4. Plot the QQ plot

   ```
   plot(qexp(ppoints(x), λ), sort(x))
   ```

Kolmogorov Smirnov test

We now focus on statistical test.

The first one is the Kolmogorov Smirnov test, it can reject the goodness of fit of data set to a given law.

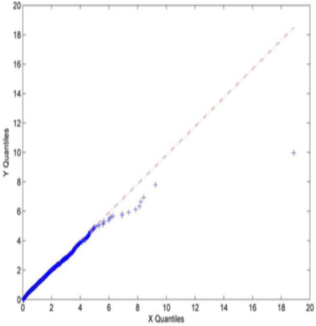
The statistic is given by the supremum of the distance between theoretical and empirical CDF, i.e.

$$d_n = \sqrt{n} \sup_x |F_n(x) - F(x)|.$$

d_n is a deterministic function of the p-value p_n of the test, i.e. the probability of obtaining a result equal or worst to the one actually observed. If $d_n = \phi(p_n) > \phi(\alpha)$ the goodness of fit is rejected with significance α.

ks.test(x, "pexp", λ) $ p.value

High frequency data : EUR/PLN



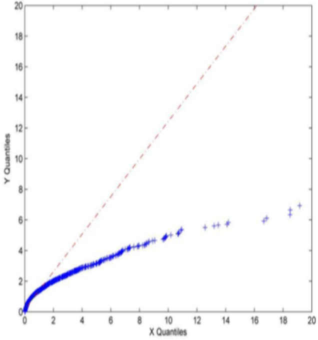


FIGURE: Hewlett 2006, *Clustering of order arrivals, price impact and trade path optimisation*, Q-plot of optimised Hawkes and Poisson Process on EUR/PLN.

Hawkes Process : Framework

Let $(\Omega, \mathcal{A}, \mathbb{P})$ be some probability space and let ν be the counting measure of a non-explosive point process denoted by N.
We denote by $\{\tau_n\}_{n\geq 0}$ the jump times associated to ν, and by $\mathbb{F} := \{\mathcal{F}_t\}_{t\geq 0}$ its natural filtration.
The *compensator* of ν is denoted by λ. We recall that it is the unique predictable measure such that $\tilde{\nu}(dt) := \nu(dt) - \lambda(dt)$ is a martingale measure.

Hypothesis

The compensator μ is absolutely continuous with respect to the Lebesgue measure. Hence, we can write, for short, $\lambda(dt) := \lambda(t)\, dt$, and call the process λ the intensity of the random measure.

Hawkes Process : Definition

Definition

A Hawkes process with exponential kernel is a point process, with intensity satisfying

$$\forall t \geq 0, \quad \lambda(t) = \lambda_0 + \alpha \sum_{\tau_i < t} e^{-\beta\,(t-\tau_i)} \tag{1}$$

where $\lambda_0 > 0$, $\alpha \geq 0$ and $\beta > 0$ are constant parameters.

The Hawkes processes were introduced in Hawkes 1971. For existence of such processes the reader can refer to Hawkes et al. .

Predictable version

It is clear that we have decided to use the predictable version for the intensity. Due to the counting nature of the process, that is it never jumps twice at the same time, our choice has no financial consequences. However, this choice is unusual but is necessary in some of the next results.

Hawkes Process : Nota Bene

Considering a Hawkes process, the crucial aspect is its **intensity**. The process itself

$$N(t) = \int_0^t \nu(ds)$$

is a counting process, that is

$$\Delta N(\tau_i) := N(\tau_i) - N(\tau_i^-) = 1.$$

The knowledge of a counting process is equivalent to the knowledge of the jumps times T_i.

Using mathematical words : the filtration $\mathcal{F}_t$ is equivalent to the sigma-algebra of the sets $\{t \geq \tau_i\}$.

Point Processes beyond Hawkes

Point processes are a very nice subset of stochastic processes. Some results are true only for PP.

For instance, the filtration $\mathcal{F}_t$ is *continuous* in a mild sense, see Dellacherie Meyer. That is $\mathcal{F}_t^-$ coincides with $\mathcal{F}_t$. in particular, let θ be a stopping time, then $\mathcal{F}_{\theta-} = \mathcal{F}_\theta$.

The *main reason* is that the filtration $\mathcal{F}_t$ can be *reduced* to the signature $\mathbb{I}_{t \leq \tau_i}$.

Equivalent forms

$$d\lambda(t) = -\beta\,(\lambda(t) - \lambda_0)\,dt + \alpha\,dN(t)$$

$$\lambda(t) = -\int_0^t \beta\,(\lambda(s) - \lambda_0)\,ds + \alpha \int_0^{t^-} \nu(ds)$$

$$\lambda(t) = \lambda_0 + \alpha \int_0^{t^-} e^{\beta\,(s-t)}\,\nu(ds)$$

How to generate an Hawkes process : Change of Measure

As a direct consequence of Theorem 10.2.6 p. 339 in Last Brant, we can state the following :

Corollary

Let N be a Poisson process with intensity 1. Let λ follows

$$d\lambda(t) = -\beta\,(\lambda(t) - \lambda_0)\,dt + \alpha\,dN(t)$$

and set

$$L[\lambda](t) := \prod_{\tau_n < t} \lambda(T_n) \times \exp\left(-\int_0^t [\lambda(s) - 1]\; ds\right). \tag{2}$$

Then, $L[\lambda](t)$ is the density of a probability $\mathbb{Q}_\lambda$ under which the process N is a Hawkes with intensity λ.

Properties of Hawkes process

Stationarity Average intensity $\mathbb{E}[\lambda(t)] = \mu$, where $\mu = \lambda_0 \dfrac{\beta}{\beta - \alpha}$.
Stationarity iff $\beta > \alpha$.

Infinitesimal covariance $\mathbb{E}\left[\Delta N_t \, \Delta N_s\right] = \left\{\nu_N(t-s) + \mu^2 + \delta_0(t-s)\right\} dt \, ds$,
where ν_N is given by its Laplace transform

$$\widehat{\nu}_N(z) = \frac{\mu}{\left[1 - \widehat{\phi}(-z)\right]\left[1 - \widehat{\phi}(z)\right]} - \mu,$$

and $\widehat{\phi}$ is the Laplace transform
$\widehat{\phi}(z) = \widehat{\alpha \, e^{-\beta \, t}}(z) := \int_0^\infty \alpha \, e^{-\beta \, t} e^{-z \, t} \, dt = \frac{\alpha}{z+\beta}$.
Proof see Bacry, Dayri and Muzy (2012).

How to generate an Hawkes process : simulation via Thinning procedure 1

Rejection sampling

If you want to sample a random variable X with density $f(x)$, but you do not know an explicit way (explicit schemes exist for uniform, exponential, Gaussian r.v., etc).

The idea is then to generate a candidate using a r.v. Y with density $g(x)$ having explicit simulation and then to accept/reject the candidate with probability $f(y)/Mg(y)$.

1. Generate the r.v. $Y \to y$. y is the candidate (sample).
2. Generate a uniform random variable U.
3. Accept the sample if $U < f(y)/Mg(y)$. Otherwise, come back to point 1.

How to generate an Hawkes process : simulation via Thinning procedure 2

Lewis and Shedler 1979 / Ogata 1981

We use a rejection sampling method adapted to the Hawkes case.

1. Initialize the intensity $\lambda(0)$. Fix an auxiliary intensity $\lambda^\star := \lambda(0)$.
2. Generate a sample τ in accord with an exponential random variable with frequency $\lambda^\star$. Generate a uniform r.v. U.
3. Accept the sample if $U < \lambda(\tau)/\lambda^\star$.

   If the sample is accepted : the Point Process increases $\Delta N_\tau = 1$. Actualize the intensity $\lambda(\tau) = \lambda(\tau^-) + \alpha$. Actualize the auxiliary intensity $\lambda^\star := \lambda(\tau)$. Come back to point 2.

   If the sample is rejected : the Point Process does not increase $\Delta N_\tau = 0$. Actualize the auxiliary intensity $\lambda^\star := \lambda(\tau)$.

Questions

- Why it works ?

How to generate an Hawkes process : simulation via Thinning procedure 2

Lewis and Shedler 1979 / Ogata 1981

We use a rejection sampling method adapted to the Hawkes case.

1. Initialize the intensity $\lambda(0)$. Fix an auxiliary intensity $\lambda^\star := \lambda(0)$.
2. Generate a sample τ in accord with an exponential random variable with frequency $\lambda^\star$. Generate a uniform r.v. U.
3. Accept the sample if $U < \lambda(\tau)/\lambda^\star$.

   If the sample is accepted : the Point Process increases $\Delta N_\tau = 1$. Actualize the intensity $\lambda(\tau) = \lambda(\tau^-) + \alpha$. Actualize the auxiliary intensity $\lambda^\star := \lambda(\tau)$. Come back to point 2.

   If the sample is rejected : the Point Process does not increase $\Delta N_\tau = 0$. Actualize the auxiliary intensity $\lambda^\star := \lambda(\tau)$.

Questions

- Why it works ? Memoryless of exponential law. $\lambda(\tau)$ is mean reverting.
- Why we actualize the auxiliary intensity ?

Example of Thinning

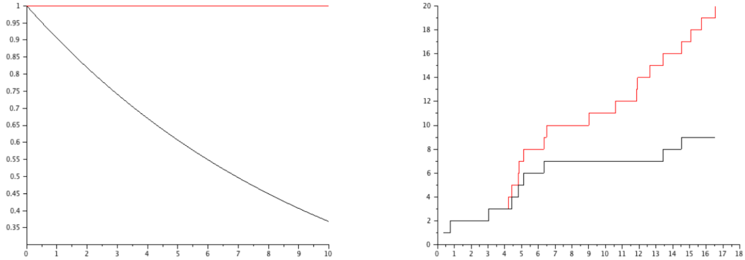


FIGURE: Example of thinning simulation of a Cox process (PhD thesis, Aboumezoued).

Estimation : Likelihood

Ogata 1979

Given a set of occurrence times $\tau_1, \tau_2, \ldots, \tau_n$ of a Point Process over an interval $[0, T]$. Assuming a parametrized intensity function $\lambda(\theta, t|\mathcal{F}_t)$. The associated Likelihood function is

$$L(\theta; \tau_1, \tau_2, \ldots, \tau_n) := \left\{ \prod_{i=1}^{n} \lambda\left(\theta; \tau_i \,|\, \mathcal{F}_{\tau_i}\right) \right\} \exp\left\{ -\int_0^T \lambda(\theta; t \,|\, \mathcal{F}_t)\, dt \right\},$$

$$l(\theta; \tau_1, \tau_2, \ldots, \tau_n) := \sum_{i=1}^{n} \log \lambda\left(\theta; \tau_i \,|\, \mathcal{F}_{\tau_i}\right) - \int_0^T \lambda(\theta; t \,|\, \mathcal{F}_t)\, dt.$$

Hawkes process

$$\lambda(t) = \lambda_0 + \alpha \sum_{\tau_i < t} e^{-\beta\,(t - \tau_i)}$$

$$l(\theta; \tau_1, \tau_2, \ldots, \tau_n) = -\lambda_0\, \tau_n + \sum_{i=1}^{n} \frac{\alpha}{\beta}\left[e^{-\beta\,(\tau_n - \tau_i)} - 1 \right] + \sum_{i=1}^{n} \log\left\{ \lambda_0 + \alpha \sum_{j=1}^{i-1} e^{-\beta\,(\tau_i - \tau_j)} \right\}$$

Estimation : Maximum of Likelihood

$$
\begin{aligned}
\frac{\partial l}{\partial \lambda_0} &= -\tau_n + \sum_{i=1}^{n} \left\{ \lambda_0 + \alpha \sum_{j=1}^{i-1} e^{-\beta\,(\tau_i - \tau_j)} \right\}^{-1} \\
\frac{\partial l}{\partial \alpha} &= \sum_{i=1}^{n} \frac{1}{\beta} \left[e^{-\beta\,(\tau_n - \tau_i)} - 1 \right] + \sum_{i=1}^{n} \frac{\sum_{j=1}^{i-1} e^{-\beta(\tau_i - \tau_j)}}{\lambda_0 + \alpha \sum_{j=1}^{i-1} e^{-\beta\,(\tau_i - \tau_j)}} \\
\frac{\partial l}{\partial \beta} &= \sum_{i=1}^{n} \frac{\alpha}{\beta^2} \left\{ 1 - \left[\beta\,(\tau_n - \tau_i) + 1 \right] e^{-\beta\,(\tau_n - \tau_i)} \right\} + \dots \\
&\quad - \sum_{i=1}^{n} \frac{\alpha \sum_{j=1}^{i-1} (\tau_i - \tau_j)\, e^{-\beta\,(\tau_i - \tau_j)}}{\lambda_0 + \alpha \sum_{j=1}^{i-1} e^{-\beta\,(\tau_i - \tau_j)}}
\end{aligned}
$$

see Jiao, Ma, Scotti, Sgarra (2017) also for extension to marked Hawkes processes. For the Fisher information matrix see Ozaki 1979.

Residual Analysis via Cox process

How to test if the calibrated Hawkes process is a good fit of data ?

Using the calibrated intensity $\lambda(\lambda_0, \alpha, \beta)$, we define the pathwise process

$$\overline{\lambda}(t; \lambda_0, \alpha, \beta; \{\tau_i\}_{i=1,\ldots,n}) := \lambda_0 + \alpha \sum_{i=1}^{n} e^{-\beta(t-\tau_i)}$$

We consider a Cox Point Process, i.e. a Point Process with stochastic intensity, driven by $\overline{\lambda}$.

We can now apply the usual framework since the self-exciting structure is artificially “frozen”.

Residual Analysis

Defining $\overline{\Lambda}_t := \int_0^t \overline{\lambda}(s; \lambda_0, \alpha, \beta; \{\tau_i\}_{i=1,\ldots,n}) ds$.

Compute $t_i = \overline{\Lambda}_{\tau_i}$. The original data set $\{\tau_i\}_{i=1,\ldots,n}$ is transformed into $\{t_i\}_{i=1,\ldots,n}$.

Under the hypothesis that $\{\tau_i\}_{i=1,\ldots,n}$ are the jump times of a Cox PP with intensity $\overline{\lambda}$, the data set $\{t_i\}_{i=1,\ldots,n}$ is the jump times of a Poisson Process with intensity 1.

We can now apply the Kolmogorov Smirnov test.

Extension to multi-dimensional case

We consider a n-variate counting-process $N_t = \{N_t^i\}_{i=1,\dots,n}$
We suppose that each intensity satisfies

$$\lambda_t^i = \lambda_0^i + \sum_{j=1}^{n} \alpha_j^i \int_0^{t^-} e^{\beta_j^i\,(s-t)}\, dN_s^j.$$

Implicit hypotheses

The process is not anticipating.
To obtain an SDE, we need to suppose $\beta_j^i = \beta^i$ for all j.

Extension to multi-dimensional case

We introduce the matrix

$$M = \begin{pmatrix} \frac{\alpha_1^1}{\beta_1^1} & \frac{\alpha_2^1}{\beta_2^1} & \cdots & \frac{\alpha_n^1}{\beta_n^1} \\ \frac{\alpha_1^2}{\beta_1^2} & \frac{\alpha_2^2}{\beta_2^2} & \cdots & \frac{\alpha_n^2}{\beta_n^2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\alpha_1^n}{\beta_1^n} & \frac{\alpha_2^n}{\beta_2^n} & \cdots & \frac{\alpha_n^n}{\beta_n^n} \end{pmatrix}$$

Stability : if the spectral radius (i.e. largest eigenvalue) of M is smaller than 1.

Average intensity : $\Lambda = (\mathbb{I} - M)^{-1}\, \overline{(\lambda_0)}$.

Infinitesimal covariance : $\mathbb{E}\left[\Delta N_t^i\, \Delta N_s^j\right] = \{\nu_{ij}(t-s) + \Lambda_i\, \Lambda_j + \delta_{ij}(t-s)\}\ dt\, ds$, where ν_{ij} is given by its Laplace transform

$$\widehat{\nu}(z) = \left[1 - \widehat{\Phi}(-z)\right]^{-1} \Sigma \left[\left[1 - \widehat{\Phi}(z)\right]^{-1}\right]^T - \Sigma,$$

$\widehat{\Phi}(z)$ is the Laplace transform of the matrix $\alpha_j^i\, e^{-\beta_j^i t}$. And Σ is the diagonal matrix with $\Sigma_{ii} = \Lambda_i$.

Proof see Bacry, Dayri and Muzy (2012).

Motivation : How to evaluate the self-exciting effect

We want to analyse the financial impact of the two terms in

$$\lambda_t = \lambda_0 + \alpha \sum_{i=1}^{n} e^{-\beta\,(t-\tau_i)}$$

The goal is

- to isolate the external influences on the system λ_0
- from the internal feedback $\alpha \sum_{i=1}^{n} e^{-\beta\,(t-\tau_i)}$

Branching language

immigrants: all events "caused" by the external influences ; other words *zero-order events*, *mother*, *exogenous events*, *fundamental events*.

descendants: all events "explained" by the self-exciting ; other words *first- and high-order events*, *daughters*, *endogenous events*.

Financial reasons

- **Behavioural Finance** : Over and under-reaction to news, see Callegaro, Gaigi, Scotti and Sgarra (2016).
- **Margin call** for related futures : automatic liquidation policies
- **Hedging strategies** : Option hedging requires important fluctuations in the same sense of the previous shock.
- **Optimal portfolio execution** : large investors split their orders into smaller ones. Avellaneda and Stoikov (2008), Obizhaeva and Wang (2013), Chevalier, Ly Vath, Roch and Scotti (2016), etc.
- **Technical Analysis and HFT trading** : investors act without changes in economic fundamentals.

Branching ratio

It is impossible to distinguish the nature of an event.

However, it is possible to study the ratio between endogenous and exogenous events. That is the mean number of daughters per mother. This ratio n is called **branching ratio**.

- $n \ll 1$: the system is dominated by external events. Almost no self-exciting.
- $n \lessapprox 1$ the system is dominated by self-exciting. However, the total number of events is finite.
- $n > 1$ the system is dominated by self-exciting and each new events generates an increasing sequence of descendants. The system is said **super-critical** or **explosive**.

Branching ratio

It is impossible to distinguish the nature of an event.

However, it is possible to study the ratio between endogenous and exogenous events. That is the mean number of daughters per mother. This ratio n is called **branching ratio**.

- $n \ll 1$: the system is dominated by external events. Almost no self-exciting.
- $n \lessapprox 1$ the system is dominated by self-exciting. However, the total number of events is finite.
- $n > 1$ the system is dominated by self-exciting and each new events generates an increasing sequence of descendants. The system is said **super-critical** or **explosive**.
- $n = 1$ the system is said **critical**.

Branching ratio

It is impossible to distinguish the nature of an event.

However, it is possible to study the ratio between endogenous and exogenous events. That is the mean number of daughters per mother. This ratio n is called **branching ratio**.

- $n \ll 1$: the system is dominated by external events. Almost no self-exciting.
- $n \lessapprox 1$ the system is dominated by self-exciting. However, the total number of events is finite.
- $n > 1$ the system is dominated by self-exciting and each new events generates an increasing sequence of descendants. The system is said **super-critical** or **explosive**.
- $n = 1$ the system is said **critical**.

why ?

The mean number of events in each cluster is

$$1 + n + n^2 + \ldots = \sum_i n^i = \frac{1}{1-n}.$$

Estimation of branching ratio : stochastic declustering

A *direct* way is to use a **stochastic declustering** method.

$$\lambda(t) = \lambda_0 + \alpha \int_0^{t^-} e^{\beta\,(s-t)}\ \nu(ds).$$

Algorithm : Zhuang, Ogata, Vere-Jones (2012)

Fix the sample $\tau_1, \tau_2, \ldots, \tau_N$

1. Estimate $(\lambda_0, \alpha, \beta)$ using for instance MLE.
2. for each couple of events $\tau_i < \tau_j$, calculate the estimated probability that j-th event is an offspring of i-th one, that is

$$\rho_{i,j} = \frac{\alpha \int_{\tau_i}^{\tau_j} e^{-\beta\,(s-\tau_i)} ds}{\lambda(\tau_j)}$$

3. The total probability that j-th event is an offspring is given by $\rho_j := \sum_{i=1}^{j-1} \rho_{i,j}$
4. Generate a sample U in accord with uniform law. Declare that j-th event is an offspring of i-th one if $U < \rho_{i,j}$.

Estimation of branching ratio : mean value

Looking at the cumulated intensity induced by each event :

$$\alpha \int_0^\infty e^{-\beta\, s} ds = \frac{\alpha}{\beta} = n.$$

Then, it is enough to estimate α and β.

Differences

- Both methods are based on the same estimation of the parameters
- stochastic declustering can help to recreate the *genealogy* of the events.
- both methods require a very high number of events. The main limitation comes from the convergence of MLE.
- second method depends only on the ratio α/β, whereas stochastic declustering works very well if both α and β are large.

Some references : Hawkes processes

- Hawkes 1971. Point Spectra of Some Mutually Exciting Point Processes. Journal of the Royal Statistical Society. Series B (Methodological) 33 (3), 438 ?443.
- Bowsher 2007. Modelling security market events in continuous time : Intensity based, multivariate point process models. Journal of Econometrics 141 (2), 876 ?912.

Some references : Hawkes processes and statistics

- Aït-Sahalia, Y. and Jacod, J. (2009) : Estimating the Degree of Activity of Jumps in High Frequency Data. *The Annals of Statistics*. **37**, 2202-2244.
- Aït-Sahalia, Y., Cacho-Diaz, J., and Laeven, R. J. (2015) : Modeling financial contagion using mutually exciting jump processes. *Journal of Financial Economics*. 117(3), 585-606.
- Jaisson, T., and Rosenbaum, M. (2015) : Limit theorems for nearly unstable Hawkes processes. *Annals of Applied Probability*, 25(2), 600-631.
- Ogata, 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.
- Zhuang, J., Ogata, Y., Vere-Jones, D., 2002. Stochastic declustering of spacetime earthquake occurrences. Journal of the American Statistical Association 97 (458), 369-380.

Some references : Branching

- Dawson, D.A. and Li, Z. (2006) : Skew convolution semigroups and affine Markov processes. *Ann. Probab.* **34**, 1103-1142.
- Dawson, D.A. and Li, Z. (2012) : Stochastic equations, flows and measure-valued processes. *Annals of Probability*. **40** (2), 813-857.
- Duquesne, T. and Le Gall, J-F. (2002) Random Trees, Levy Processes and Spatial Branching Processes Asterisque nr. 281.
- Watson and Galton (1875). On the probability of the extinction of families. The Journal of the Anthropological Institute of Great Britain and Ireland, 4 :138-144.
- Harris, T. E., 2002. The Theory of Branching Processes. Dover Phoenix Editions.
- Lamperti (1967) Continuous state branching processes. Bull. Amer. Math. Soc, 73(3) :382-386.
- Le Gall, J.-F. (1999) Spatial Branching Processes, Random Snakes and Partial Differential Equations. Lectures in Mathematics, ETH Zurich, Birkhauser.
- Pardoux (2008) Continuous branching processes : the discrete hidden in the continuous Dedicated to Claude Lobry. Arima Journal, 9 :211-229.

Some references : micro structure

- Ivanov, Yuen, Podobnik, Lee, 2004. Common scaling patterns in intertrade times of U. S. stocks. Physical Review E 69 (5), 056107.
- Jiang, Chen, Zhou, 2009. Detrended fluctuation analysis of intertrade durations. Physica A : Statistical Mechanics and its Applications 388 (4), 433 ?440.
- Large, 2007. Measuring the resiliency of an electronic limit order book. Journal of Financial Markets 10 (1), 1 ?25.
- Rambaldi, M., Pennesi, P., and Lillo, F. (2015) : Modeling FX market activity around macroeconomic news : a Hawkes process approach. *Physical Review E*, 91(1), 012819.

Some references : order flow

- Bouchaud, Farmer, Lillo, F., 2009. How markets slowly digest changes in supply and demand. In : Handbook of Financial Markets : Dynamics and Evolution. North Holland, Amsterdam, pp. 57 ?160.
- Embrechts, Liniger, Lu, 2011. Multivariate Hawkes Processes : an Application to Financial Data. J. Appl. Probab. 48A, 367 ?378.
- Filimonov, V., Sornette, D., 2012. Quantifying reflexivity in financial markets : Toward a prediction of flash crashes. Physical Review E 85 (5), 056108.
- Hewlett 2006. Clustering of order arrivals, price impact and trade path optimisation. In Workshop on Financial Modeling with Jump processes, Ecole Polytechnique.

Some references : other applications

- Bernis, Salhi, Scotti, (2017). Sensitivity Analysis for Marked Hawkes Processes - Application to CLO Pricing, working paper.
- Dassios, A. and Zhao, H. (2011) : A dynamic contagion process, *Advances in Applied Probability*, 43(3), 814-846.
- Duffie, D., Filipović, D. and Schachermayer, W. (2003) : Affine processes and applications in finance, *Annals of Applied Probability*, 13(3), 984-1053.
- Errais, E., Giesecke, K., Goldberg, L. R., 2010. Affine Point Processes and Portfolio Credit Risk. SIAM Journal on Financial Mathematics 1 (1), 642.
- Filipović, D. (2001) : A general characterization of one factor affine term structure models. *Finance and Stochastics* **5**, 389-412.
- Filimonov, V., Bicchetti, D., Maystre, N., Sornette, D. (2015) :Quantification of the High Level of Endogeneity and Structural Regime Shifts in Commodity Markets, preprint.
- Jiao, Y. ; Ma, C. and Scotti, S. (2017) : Alpha-CIR Model in Sovereign Interest Rate Modelling, *Finance and Stochastics* forthcoming.