An expansion formula for Hawkes processes and application to cyber-insurance derivatives.

Caroline Hillairet, ENSAE Paris, CREST

with the support of the **Research Initiative** "Cyber Risk : actuarial modeling", funded by Axa Research Funds.

Joint work with A. Réveillac (Institut Mathématiques de Toulouse) and M. Rosenbaum (Ecole polytechnique).

Conference "New Trends on Hawkes Processes"

June 1-2, 2022, Toulouse

Outline

Ruin Theory framework

Cumulative Loss process :

$$L_t := \sum_{i=1}^{N_t} X_i, \quad t \in [0, T],$$

- Frequency: Claims arrival modeled by a jump process $N := (N_t)_{t\in[0,T]}$, jumping at time $(\tau_i)_{i\in\mathbb{N}^*}$,
- Severity: claims sizes $(X_i)_{i\in\mathbb{N}^*}$

Classical **Cramer-Lundberg** model

- N is a Poisson process (inter-arrivals $(\tau_i - \tau_{i-1})$ are iid)
- N is independent of the claims sizes (X_i),
- (X_i) iid random variables.
- ... but the independence assumptions are in practice often too restrictive

Relaxing independence assumptions

- **Dependencies between claims arrival N and claims sizes (X_i).**
  - see H., Jiao and Réveillac (2018)
  - we do not assume a Markovian framework
  - extend the mixing approach of Albrecher et al. (2011) by allowing of non-exchangeable family of random variables for the claim size.
- **Self-exciting arrival of claims** and clustering effects
  - Clustering and contagion of cyber-events: Baldwin et al. (2017), Bessy-Roland, Boumezoued, H. (2020)
  - Modeling through Hawkes process H.
    Papers dedicated to Hawkes processes (mainly with exponential kernel) in insurance: Dassios and Zhao (2012), Magnusson Thesis (2015), Gao and Zhu (2018), Swishchuk (2018)...
  - Contagion in Credit risk: Errais, Giesecke and Goldberg (2010), Embrechts et al. (2011), Bielecki et al. (2020)

RESEARCH INITIATIVE
Cyber Risk: Actuarial Modeling

Our framework

Goal : Computation of quantity of the form : $\mathbb{E}\left[K_T h(L_T)\right]$, where

- Claims arrivals modeled by a Hawkes process (H_t).
- L_T is the cumulative loss that activates the contract
- $K_T = \int_0^T Z_s dH_s$ is the effective covered loss

So $$\mathbb{E}\left[K_T h(L_T)\right] = \mathbb{E}\left[\int_0^T Z_t dH_t h(L_T)\right].$$

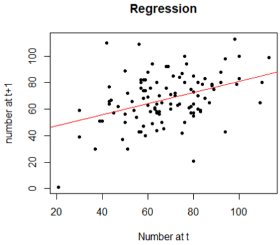
- **Provide pricing formulae for insurance contracts**
  - 2 key ingredients : Thinning algorithm + Malliavin calculus
  - expansion formula for the premium.
  - bounds on the premium

RESEARCH INITIATIVE
Cyber Risk: Actuarial Modeling

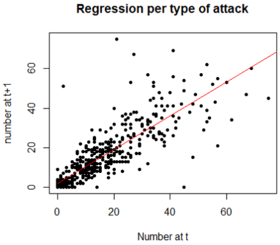
Outline

Autocorrelation of the number of cyber-events

- Privacy Rights Clearing House data-base (PRC). 8800 events over the period 2005-2019.
- Regression of the number of event during the following month $t+1$ as a function of the number of event during the current month t (should be independent for a Poisson process model to be valid)
- Autocorrelation dramatically increases when focusing on attacks of the same type



R^2=0.154 [0.030, 0.278]



R^2=0.726 [0.687, 0.766]

Hawkes model

- Taking into account autocorrelation
  - Cox model : Poisson model with stochastic intensity $\rightarrow$ difficulty to specify the stochastic intensity dynamics
  - Shot noise model: extends Cox model (Schmidt et al., Dassios and Jang for catastrophe insurance, credit risk...)
  - Hawkes model : Self-exciting model with stochastic intensity, fully specified by the point process itself
- H Hawkes process with (deterministic) excitation kernel Φ and base intensity λ_0 is the counting process ($H_0 = 0$) with intensity process

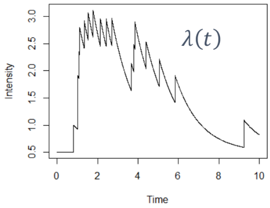
$$\lambda(t) := \lambda_0(t) + \int_{(0,t)} \Phi(t-s)dH_s, \quad t \in [0, T],$$

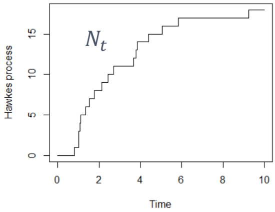
that is for $0 \leq s \leq t$ and $A \in \mathcal{F}_s$, $\mathbb{E}\left[\mathbf{1}_A(H_t - H_s)\right] = \mathbb{E}\left[\int_{(s,t]} \mathbf{1}_A \lambda(r)dr\right]$.

Toy example of Hawkes process with exponential kernel

- $(H_t)_{t\geq 0}$ counting process with jump times $(\tau_n)_{n\geq 1}$
- Intensity process of the counting process with exponential kernel

$$\lambda(t) = \mu + \sum_{\tau_n < t} \alpha \exp\left(-\beta(t - \tau_n)\right)$$





- Each jump represents an attack
- Intensity decreases exponentially between jumps

Research Initiative Cyber Risk: Actuarial Modeling

Cumulative Loss processes

$$L_t := \sum_{i=1}^{H_t} f(\eta_i)\, e^{-\kappa(t-\tau_i)}, \quad K_t := \sum_{i=1}^{H_t} g(\eta_i, \vartheta_i) e^{-\kappa(t-\tau_i)}, \quad t \in [0, T]$$

- K_T: effective loss covered by the reinsurance company,
- L_T: the loss quantity that activates the contract.
- $(\eta_i, \vartheta_i)_{i \geq 1}$ is a sequence of iid rv (independent of H),
- f and g are bounded deterministic functions,
- $\kappa \geq 0$ is a discount factor,
- $\tau_i := \inf\{t > 0,\ H_t = i\}$.

Goal : Computation of quantity of the form : $\mathbb{E}[K_T h(L_T)]$.

Some contracts in (Re-)insurance

$$L_T := \sum_{i=1}^{H_T} f(\eta_i)\, e^{-\kappa(T-\tau_i)}, \quad K_T := \sum_{i=1}^{H_T} g(\eta_i, \vartheta_i) e^{-\kappa(T-\tau_i)}$$

Generalized Stop-loss Contrats : Stop-loss Contrats provide to its buyer (another insurance company), the protection against losses which are larger than a given level K and its payoff function is given by a "call" function. Consider for example a contract where the reinsurance company pays

$$\text{Payoff}\,(L_T, K_T) = \begin{cases} 0, & \text{if } L_T \leq K \\ K_T - K, & \text{if } K \leq L_T \leq M \\ M - K, & \text{if } L_T \geq M \end{cases}$$

More precisely, when the insurance contract is triggered by the loss process L, the compensation amount can depend on some other exogenous factors $(\vartheta_i)_{i\in\mathbb{N}}$.

RESEARCH INITIATIVE
Cyber Risk: Actuarial Modeling

A related quantity

$$L_T := \sum_{i=1}^{H_T} f(\eta_i)\, e^{-\kappa(T-\tau_i)}$$

Expected Shortfall (risk measure) : The expected shortfall is a useful risk measure, that takes into account the size of the expected loss above the value at risk.

$$ES_\alpha(L_T) = \mathbb{E}[L_T | L_T > V@R_\alpha(L_T)], \quad \alpha \in (0,1).$$

$$ES_\alpha(L_T) = AV@R(L_T) := \frac{1}{1-\alpha}\int_\alpha^1 V@R_s(L_T) ds,$$

if the law of L_T is continuous, which is NOT the case here. The latter property fails already in the case where the size claims X_i are constant. So one needs an explicit computation of

$$ES_\alpha(L_T) = \frac{\mathbb{E}[L_T \mathbf{1}_{\{L_T > \beta\}}]}{\mathbb{P}(L_T > \beta)}, \quad \beta := V@R_\alpha(L_T)$$

Outline

RESEARCH INITIATIVE
Cyber Risk: Actuarial Modeling

Payoffs summary

Computation of quantity of the form

$$\mathbb{E}[K_T h(L_T)] \tag{1}$$

- K_T: effective loss covered by the reinsurance company,
- L_T: the loss quantity that activates the contract.
  - with $K_T := \sum_{i=1}^{H_T} g(\eta_i, \vartheta_i) e^{-\kappa(s-\tau_i)} = \int_{(0,T]} Z_t dH_t$ where the claim sizes $(g(\eta_i, \vartheta_i))$ are iid (and independent of H) and Z is the $\mathbb{F}$-predictable process

$$Z_s := \sum_{i=1}^{+\infty} g(\eta_i, \vartheta_i) e^{-\kappa(T-s)} \mathbf{1}_{(\tau_{i-1}, \tau_i]}(s), \quad s \in [0, T]$$

  - $F := h(L_T)$ is a functional of the Hawkes process.
- Expectation (1) can be expressed as

$$\mathbb{E}[K_T h(L_T)] = \mathbb{E}\left[\int_{(0,T]} Z_t dH_t F\right]. \tag{2}$$

Malliavin IPP formula (Mecke formula)

Aim: transformation of Equation (2) ("$dH_t \to dt$").

- If $H = N$ is an homogeneous Poisson process with intensity $\mu > 0$

$$\mathbb{E}\left[\int_{(0,T]} Z_t dN_t F\right] = \mu \int_0^T \mathbb{E}\left[Z_v F \circ \varepsilon_v^+\right] dv \qquad (3)$$

  - $F \circ \varepsilon_v^+ =: F^v$ denotes the functional on the Poisson space where a deterministic jump is added to the paths of N at time v
  - **adding a jump at some time v = adding "artificially" a claim at time v** (stress test).
- In case of a Poisson process N: the additional jump at some time v only impacts the payoff of the contract by adding a new claim in the contract
- **In case of a Hawkes process H: it also impacts the dynamic (after time v) of the counting process H.**

RESEARCH INITIATIVE
Cyber Risk: Actuarial Modeling

Main contributions

- **Generalization** of Equation (3) **for a Hawkes process** H, with self-exciting intensity process

$$\lambda_t := \mu + \int_{(0,t)} \Phi(t-s)dH_s$$

where $\mu > 0$ and $\Phi : [0,T] \to \mathbb{R}_+$ bounded self-excitating kernel with $\|\Phi\|_1 < 1$.

- Main ingredient : a representation of a Hawkes process in terms of a Poisson measure N on $[0,T] \times \mathbb{R}_+$ (known as "**Poisson embedding**" or "Thinning Algorithm")

$$\begin{cases} H_t = \int_{(0,t]} \int_{\mathbb{R}_+} \mathbf{1}_{\{\theta \leq \lambda_s\}} N(ds, d\theta), \\ \lambda_t = \mu + \int_{(0,t)} \Phi(t-u) dH_u. \end{cases} \quad (4)$$

Thinning Algorithm

$$\begin{cases} H_t = \int_{(0,t]} \int_{\mathbb{R}_+} \mathbf{1}_{\{\theta \le \lambda_s\}} N(ds, d\theta), \\ \lambda_t = \mu + \int_{(0,t)} \Phi(t-u) dH_u. \end{cases}$$

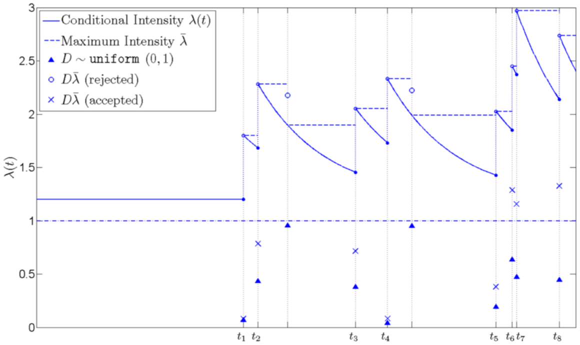


Illustration from Ogata (1981)

Shifted Hawkes process

- Our expansion formula involves "**shifted Hawkes processes**" $H^{v_n,\cdots,v_1}$ for which jumps at deterministic times $0 < v_n < \cdots < v_1$ are added to the process accordingly to the self-exciting kernel Φ.

- One shift Hawkes process at time v in $(0, T)$.

$$\begin{cases} H_t^v = \mathbf{1}_{[0,v)}(t) H_t + \mathbf{1}_{[v,T]}(t) \left(H_{v-}^v + 1 + \int_{(v,t]} \int_{\mathbb{R}_+} \mathbf{1}_{\{\theta \leq \lambda_s^v\}} N(ds, d\theta) \right) \\ \lambda_t^v = \mathbf{1}_{(0,v]}(t) \lambda_t + \mathbf{1}_{(v,T]}(t) \left(\mu^{v,1}(t) + \int_{(v,t)} \Phi(t-u) dH_u^v \right), \end{cases}$$

$$\mu^{v,1}(t) := \mu + \int_{(0,v]} \Phi(t-u) dH_u^v = \mu + \int_{(0,v)} \Phi(t-u) dH_u + \Phi(t-v).$$

First Step in the expansion

Mecke's formula for Poisson functionals gives that

$$\begin{aligned}
\mathbb{E}\left[F\int_{[0,T]} Z_t dH_t\right] &= \mathbb{E}\left[F\int_{[0,T]}\int_{\mathbb{R}_+} Z_t \mathbf{1}_{\{\theta\le\lambda_t\}} N(dt,d\theta)\right] \\
&= \mathbb{E}\left[\int_{[0,T]}\int_{\mathbb{R}_+} Z_t (F\circ\varepsilon^+_{(t,\theta)})\mathbf{1}_{\{\theta\le\lambda_t\}} dt d\theta\right] \\
&= \int_{[0,T]} \mathbb{E}\left[Z_t\int_{\mathbb{R}_+} (F\circ\varepsilon^+_{(t,\lambda_t)})\mathbf{1}_{\{\theta\le\lambda_t\}} d\theta\right] dt \\
&= \int_{[0,T]} \mathbb{E}\left[Z_t (F\circ\varepsilon^+_{(t,\lambda_t)})\lambda_t\right] dt = \mu\ m_1 + I_1,
\end{aligned}$$

with

$$m_1 := \int_{[0,T]} \mathbb{E}\left[Z_t (F\circ\varepsilon^+_{(t,\lambda_t)})\right] dt,$$

$$I_1 := \int_{[0,T]} \mathbb{E}\left[Z_t (F\circ\varepsilon^+_{(t,\lambda_t)})\int_{(0,t)} \Phi(t-u) dH_u\right] dt.$$

Research Initiative Cyber Risk: Actuarial Modeling

Expansion formula for the Hawkes process

Assume Z bounded $\mathbb{F}^H$-predictable process, F bounded $\mathcal{F}_T^N$-measurable r.v.

Theorem

$$\mathbb{E}\left[F\int_{[0,T]} Z_t dH_t\right] = \mu\int_0^T \mathbb{E}\left[Z_v F^v\right] dv$$
$$+\mu\sum_{n=2}^{+\infty}\int_0^T\int_0^{v_1}\cdots\int_0^{v_{n-1}}\prod_{i=2}^{n}\Phi(v_{i-1}-v_i)\mathbb{E}\left[Z_{v_1}^{v_n,\dots,v_2}F^{v_n,\dots,v_1}\right] dv_n\cdots dv_1.$$

- the first term corresponds to the formula for a Poisson process (setting Φ at zero)
- the sum in the second term can be interpreted as a **correcting term due to the self-exciting property** of the counting process.

Lower bounds for the premium

- Assume $L_T := \sum_{i=1}^{H_T} f(\eta_i)$ and $K_T := \sum_{i=1}^{H_T} g(\eta_i, \vartheta_i)$
- Assume that h non-decreasing and denote

$$m_\Phi(\Delta^n) := \int_0^T \cdots \int_0^{v_{n-1}} \prod_{i=2}^{n} \Phi(v_{i-1} - v_i) dv_n \cdots dv_1, \quad m_\Phi(\Delta^1) = T.$$

- **A Lower bound** (adding only the deterministic jumps)

$$\mathbb{E}[K_T h(L_T)] \geq \mu \sum_{n=1}^{+\infty} m_\Phi(\Delta^n) \mathbb{E}\left[g(\bar{\eta}_1, \bar{\vartheta}_1) \mathbb{E}\left[h\left(\sum_{k=1}^{n} f(\bar{\eta}_k) \right) \Big| \bar{\eta}_1 \right] \right]$$

  where $(\bar{\eta}_i, \bar{\vartheta}_i)_{i \geq 1}$ are iid copies of (η_1, ϑ_1)
- **More accurate lower bound**

$$\mathbb{E}[K_T h(L_T)] \geq \mu \sum_{n=1}^{+\infty} m_\Phi(\Delta^n) \sum_{p=0}^{+\infty} e^{-(T\mu)} \frac{(T\mu)^p}{p!}$$
$$\mathbb{E}\left[g(\bar{\eta}_1, \bar{\vartheta}_1) \mathbb{E}\left[h\left(\sum_{k=1}^{n} f(\bar{\eta}_k) + \sum_{i=1}^{p} f(\eta_i) \right) \Big| \bar{\eta}_1 \right] \right]$$

RESEARCH INITIATIVE Cyber Risk: Actuarial Modeling

Upper bound for the premium

- **Upper Bound**

$$\mathbb{E}[K_T h(L_T)] \leq \mu \sum_{n=1}^{+\infty} m_\Phi(\Delta^n)\beta_n \leq \mu T \sum_{n=1}^{+\infty} \beta_n \, ||\Phi||_1^{n-1} \quad where$$

$$\beta_n := e^{-T(\mu+n||\Phi||_\infty)} \mathbb{E}\left[g(\bar{\eta}_1, \bar{\vartheta}_1) \mathbb{E}\left[h\left(\sum_{k=1}^{n} f(\bar{\eta}_k) \right) \Big| \bar{\eta}_1 \right] \right] + \sum_{p=1}^{+\infty} \frac{c_n}{p^2} \mathbb{E}\left[g(\bar{\eta}_1, \bar{\vartheta}_1) \mathbb{E}\left[h\left(\sum_{k=1}^{n} f(\bar{\eta}_k) + \sum_{i=1}^{p} f(\eta_i) \right) \Big| \bar{\eta}_1 \right] \right].$$

and c_n explicit constant that depends only on the kernel Φ.

RESEARCH INITIATIVE
Cyber Risk: Actuarial Modeling

Summary

- Closed-form and efficient formula for the pricing of Stop-Loss contracts
- Cumulative loss indexed by a Hawkes process: correcting term due to the self-exciting property.
- It allows to handle general dependencies and self exciting features.
- Extension:
  - Berry Esseen bounds Central Limit Theorems for the compound Hawkes process (using Malliavin-Stein method).
  - Computations of probability of ruin and related quantities
  - Extension to intensity process depending of the claims' sizes.

Based on the joint works

- ***"An expansion formula for Hawkes processes applied to insurance derivatives"***, *Hillairet, Réveillac, Rosenbaum.*
  Submitted (2021)

- And also
  - **"Pricing formulae for derivatives in insurance using Malliavin calculus"**, *Hillairet, Jiao, Réveillac.*
    Probability, Uncertainty and Quantitative Risk, volume 3 (2018)
  - **"Multivariate Hawkes process for cyber insurance"**, *Bessy-Roland, Boumezoued, Hillairet.*
    Annals of Actuarial Science (2020)
  - ***"The Malliavin-Stein method for Hawkes functionals"***, *Hillairet, Huang, Khabou, Réveillac.*
    Submitted (2021)

Thank you for your attention !

RESEARCH INITIATIVE
Cyber Risk: Actuarial Modeling