INSA
TOULOUSE

Actuariat Project

A simple model of a self-exciting counting process in the Cramér-Lundberg type model

Minh Duy Nguyen - Qian Wang - Maxime Moshfeghi

INSA - Toulouse
France

February 19, 2025

Contents

Introduction

In France, healthcare coverage is reputed to be one of the best in the world, due firstly to the public social security system, which covers part of the cost of healthcare, and secondly to the private complementary health insurance provided by the "mutuelles", the coverage of which varies from one to another but which, overall, provides French households with a considerable financial aid. However, a report shows that "mutuelles" are not reducing social inequalities, and that remaining costs (after public insurance and private "mutuelle"), for some customer such as elderly people, represents a significant cost [1].

The aim of this paper is to propose a third type of insurance—an universal, mandatory, pooled private insurance. Consequently, the paper outlines two main objectives:

- To construct a model capable of covering the remaining healthcare costs, ensuring that the customer's bill is reduced to zero.
- To guarantee the financial stability of the insurance company, ensuring it avoids bankruptcy.

Both objectives converge on the necessity of a robust model to manage the company's risk. This model must accurately account for potential claims (sinisters) and ensure the company can cover them effectively. Drawing on prior financial research, the classical Cramer-Lundberg model emerges as a simple yet effective tool for modeling risk processes [2]. In this report, we extend the classical model by incorporating self-exciting processes to better capture financial risk dynamics and demonstrate its application in an insurance case study.

To model financial risks, we utilize two mathematical tools: Poisson processes and self-exciting processes. Poisson processes are commonly used to model events that occur independently, meaning that the occurrence of one event does not influence the likelihood of subsequent events. However, this framework does not account for the interdependence of events, where certain events can increase the likelihood of others occurring. To address this limitation, the objective is to develop a model that introduces an excitation effect, where the occurrence of an event creates a zone of increased probability, making related events more likely to occur shortly afterward.

In the extended Cramer-Lundberg model, the total wealth R_t (or asset) of the company at time t is given by:

$$R_t = u + ct - \sum_{i=1}^{N_t} Y_i$$

where the counting process N_t is defined as :

$$N_t = N_t^{(1)} + N_t^{(2)}$$

where $N_t^{(1)}$ is the Poisson process counting part and $N_t^{(2)}$ is the self-exciting process counting part. To provide an intuitive understanding of the self-exciting process, Figure 1 illustrates a typical example of the counting process. The visualization demonstrates how the excitation effect increases the likelihood of related events following an initial event.

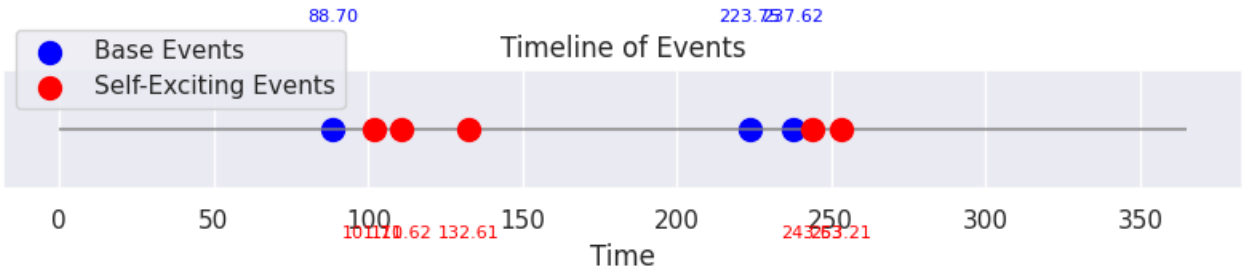


Figure 1: Simulation of event occurrences for self-exciting processes

This report is structured to provide a comprehensive understanding of the extended Cramer-Lundberg model with a self-exciting counting process. More specifically, in the first chapter we explain the principle of self-exciting counting process, by developing equations and plotting graphics. In the second chapter, we focus on the computation of expectations for the counting processes and the total assets R_t. To finish, in the last chapter, we presents an experimental simulation of the ruin probability, defined as the likelihood of the company's wealth falling below zero.

1 Fundamentals of Self-exciting process

1.1 Definitions and prerequisites

Before studying the self-exciting process, we need to define necessry basic concepts.

Definition : Point Process and Interarrival Time [3]

Let $(\Omega, \mathcal{F}, P)$ be a probability space. **A point process** on $\mathbb{R}^+$ is a sequence of non-negative random variables $\{T_k\}_{k=1,2,...}$ such that for all $k, T_k \leq T_{k+1}$ and $T_0 = 0$. T_k represents the occurence times of events.
The elapsed time from an event to the next is called the **interarrival time**, denoted by the sequence $\{W_k\}_{k=1,2,..}$ where $W_k = T_k - T_{k-1}$.

Definition : Counting process [3]

A counting process $N(t)$ is a non-decreasing function counting the number of events that have occurred at time t. The counting process associated to the point process $\{T_k\}_{k=1,2,...}$ is defined by

$$N(t) = \sum_{k=1,2,..} 1_{\{T_k \leq t\}}$$

The changes in $N(t)$ occur in discrete steps, representing individual events. This process is continuous to the right. It can also be denoted by the notation of $\{N_t\}_{t\in\mathbb{R}}$. A well-known example of counting processes are the Poisson processes.

Definition : Poisson process [4]

A counting process N is called a Poisson process with rate λ if

- $N_0 = 0$.
- The process has independent increments, i.e. the numbers of points belonging to disjoint intervals are independent.
- The number of point in any interval of length $t > 0$ follows a Poisson distribution, i.e.

$$\forall s \geq 0, N_{s+t} - N_s \sim \mathcal{P}\left(\int_s^{s+t} \lambda(t)dt\right)$$

Base on the property of the intensity λ, we can categorize Poisson processes into 2 types:

- **Homogeneous Poisson processes**: When λ is constant, N is call a homogeneous Poisson process and we thus have :

$$\forall k \in \mathbb{N}, \forall t \in [0, T], \mathbb{P}(N(t) = k) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}$$

- **Inhomogenous Poisson processes**: When λ changes in function of time, we will call N a inhomogeneous Poisson process.

One main property of Poisson processes is that the time separating two events are independent, meaning that the occurrence of an event is not triggering anything for some next events. However, in real life, there are many cases that one happened events can cause many follow-up ones to be more likely, for example an earthquake can caused many little ones in a series, or like in social network, one share of a tweet can propagate this tweet to be more viral and so more shares. To take into account those properties, we want to introduce an extended kind of processes whose occurrence gives a greater chance of a second event occurring, named **self-exciting processes**.

1.2 Self-exciting processes

Definition : Self-exciting processes

A point process N is called self-exciting if the intensity $\lambda(.)$ depends not only on time t but also the entire past of the point process. In a self-exciting process, one event can trigger another to happen in a near future.

The entire past, or history of the point process, is mathematically defined by $\mathcal{H}_t$ containing all events up to time t. We characterize the process by its **conditional intensity** defined as:

$$\lambda(t|\mathcal{H}_t) = \lim_{\Delta t \to 0} \frac{\mathbb{E}\left[N([t, t+\Delta t])|\mathcal{H}_t\right]}{\Delta t}$$

This function can count the expected rate of occurrence of events at time t given the entire history of the process up to time t ($\mathcal{H}_t$). In the general case, this intensity function can be writen in the equivalent form :

$$\lambda(t|\mathcal{H}_t) = \mu + \int_0^t \Phi(t-u)dN(u) = \mu + \sum_{i:t_i<t} \Phi(t-t_i)$$

where μ is a constant background rate of base events and Φ is the triggering function of follow-up events. In the framework of our article, we fix the triggering function as :

$$\Phi(u) := \alpha e^{-\beta u} 1_{u>0}, \text{ with } 0 < \alpha < \beta \tag{1}$$

with α, β the parameters to adjust the amplitude and the decay rate of the exciting effect. There are several forms for the triggering function. Depending on the form chosen for the triggering function Φ, the process may depend only on the recent history (if Φ decays rapidly) or may have longer term effects. Noted that the intensity must be positive all the time, i.e. $\lambda(t|\mathcal{H}_t) \geq 0, \forall t$, we require $\Phi(u) \geq 0, \forall u \geq 0$ and $\Phi(u) = 0, \forall u < 0$.

1.3 Simulation Algorithm for Self Exciting processes

To simulate the self-exciting process, we introduce a variant of the Thinning Algorithm, which is commonly used to simulate Poisson processes [3]. Different from the classical Thinning algorithm, instead of selecting points x_k at time t_k under $\lambda(t_k)$, we introduce an additional counting mechanism that considers points between λ and the exciting function Φ:

$$\forall k \text{ such that } x_k < \lambda(t_k), \text{ for } t > t_k, \text{ count the points between } \lambda(t) \text{ and } \lambda(t) + \Phi_{t_k}(t),$$
$$\text{with } \Phi_{t_k} : t \mapsto \Phi(t - t_k)$$

This algorithm comes from the fact that we can separate the self-exciting process N as the sum of two pseudo-chaos $N^{(1)}$ and $N^{(2)}$:

$$N_t = N_t^{(1)} + N_t^{(2)}$$

where the first pseudo-chaos $N_t^{(1)}$ represents a Poisson process with intensity μ and defined by

$$N_t^{(1)} := \sum_{i=1}^{|\mathcal{P}|} 1_{\theta_i \leq \mu} 1_{t_i \leq t}, t \in [0, T]$$

and $N_t^{(2)}$ the second pseudo-chaos

$$N_t^{(2)} := \sum_{i_1=1}^{|\mathcal{P}|} \sum_{i_2 : t_{i_2} > t_{i_1}}^{|\mathcal{P}|} 1_{\theta_{i_1} \leq \mu} 1_{\mu \leq \theta_{i_2} \leq \mu + \Phi(t_{i_2} - t_{i_1})} 1_{t_{i_2} \leq t}, t \in [0, T]$$

More specifically, let M, $T \in \mathbb{R}^+$, the algorithm is presented as:

Algorithm 1 Simulation of a Self-Exciting Poisson Process

Require: T: time horizon, M: maximum value of θ, μ: base intensity parameter, α: self-excitation parameter, β: decay rate parameter

Ensure: `event_times`: List of event times, `first_event_times`: Times of base Poisson process events, `second_event_times`: Times of self-exciting events:

$n \leftarrow \texttt{Poisson}(T \cdot M)$ ▷ Simulate the random Poisson measure

$\texttt{times} \leftarrow \texttt{sort}(\texttt{Uniform}(0, T, n))$ ▷ Simulate and sort uniform times

$\texttt{thetas} \leftarrow \texttt{Uniform}(0, M, n)$ ▷ Simulate θ values

$\texttt{first\_event\_times} \leftarrow \{t \in \texttt{times} : \theta \leq \mu \text{ and } t \leq T\}$ ▷ Times of base events

$\texttt{second\_event\_times} \leftarrow \emptyset$ ▷ Initialize the list of self-exciting events

for $i \leftarrow 1$ to n **do**

  for $j \leftarrow i + 1$ to n **do**

    if $t_j > t_i$ **and** $\theta_i \leq \mu$ **and** $\mu < \theta_j \leq \mu + \alpha \cdot \exp(-\beta(t_j - t_i))$ **and** $t_j \leq T$ **then**

      $\texttt{second\_event\_times} \leftarrow \texttt{second\_event\_times} \cup \{t_j\}$

    end if

  end for

end for

$\texttt{event\_times} \leftarrow \texttt{sort}(\texttt{first\_event\_times} \cup \texttt{second\_event\_times})$ **return** `event_times`, `first_event_times`, `second_event_times`

1.4 Visualization of the Thinning-type simulation algorithm

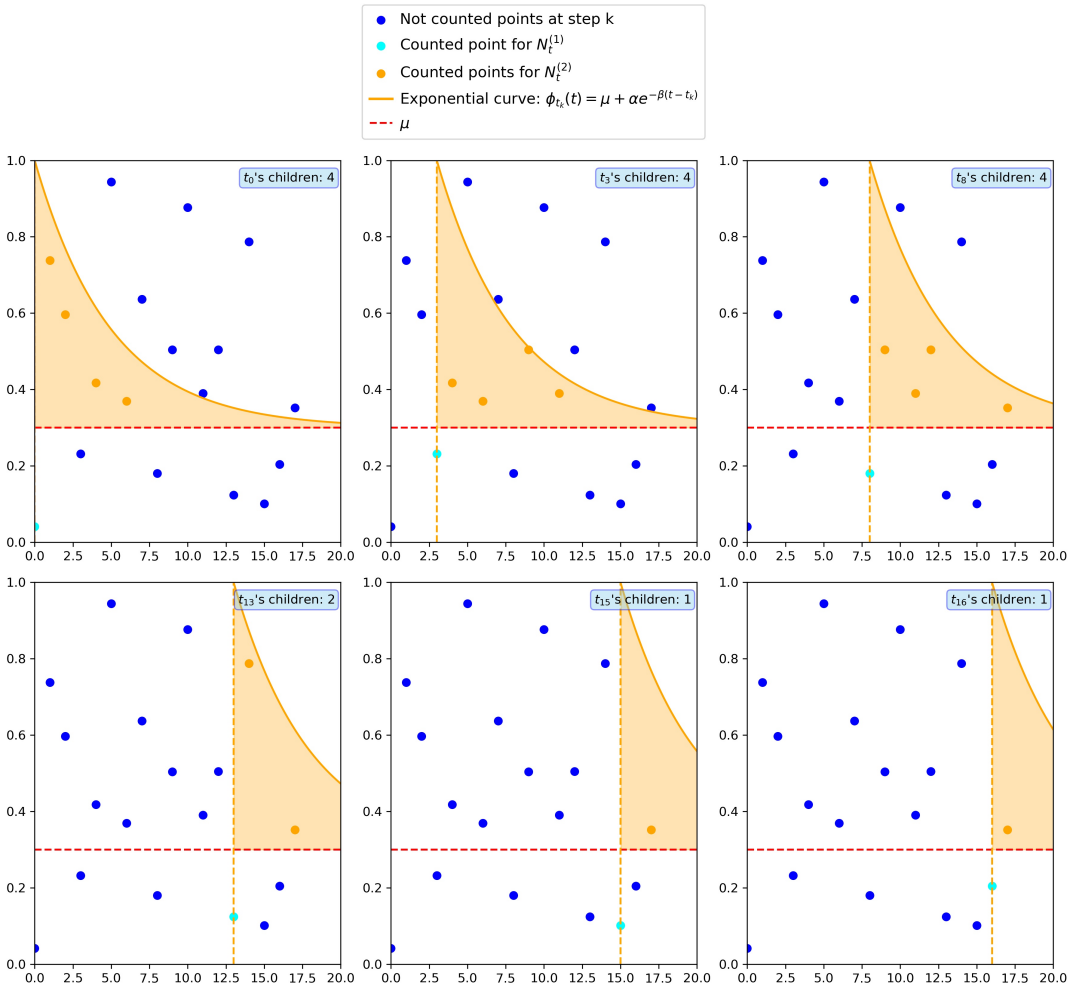


Figure 2: Illustration of the counting process $N_t^{(2)}$

Figure 2 explains visually the method of counting processes $N_t^{(2)}$. For each step, we name the cyan point the "parent" point, and the orange points the "children" points. The total number of "parents" points $N_t^{(1)}$ is equal to the number of points below the μ threshold (represented by the red dotted line), while the total number of "children" points $N_t^{(2)}$ is the sum of children points of each graph. In this case, we have:

$$N_T^{(1)} = 6 \text{ and } N_T^{(2)} = 4 + 4 + 4 + 2 + 1 + 1 = 16$$

Thus, the total counting process is:

$$N_T = N_T^{(1)} + N_T^{(2)} = 6 + 16 = 22$$

2 Mathematical analysis of Self-exciting processes

To begin with, we define the formula for the expectation of $N_t^{(2)}$ as follows:

$$\mathbb{E}\left[N_t^{(2)}\right] = \int_0^t \int_{t_1}^t \int_0^M \int_0^M \mathbb{1}_{\theta_1 \leq \mu} \mathbb{1}_{\mu < \theta_2 \leq \mu + \Phi(t_2 - t_1)} d\theta_2 d\theta_1 dt_2 dt_1$$

Next, we develop this expectation as follows:

$$\begin{aligned}
\mathbb{E}\left[N_t^{(2)}\right] &= \int_0^t \int_{t_1}^t \int_0^M \mathbb{1}_{\theta_1 \leq \mu} \underbrace{\int_0^M \mathbb{1}_{\mu < \theta_2 \leq \mu + \Phi(t_2 - t_1)} d\theta_2}_{\Phi(t_2 - t_1)} d\theta_1 dt_2 dt_1 \\
&= \int_0^t \int_{t_1}^t \int_0^M \mathbb{1}_{\theta_1 \leq \mu} \Phi(t_2 - t_1) d\theta_1 dt_2 dt_1 \\
&= \int_0^t \int_{t_1}^t \Phi(t_2 - t_1) \underbrace{\int_0^M \mathbb{1}_{\theta_1 \leq \mu} d\theta_1}_{\mu} dt_2 dt_1 \\
&= \mu \int_0^t \int_{t_1}^t \underbrace{\Phi(t_2 - t_1)}_{\alpha e^{-\beta(t_2 - t_1)}} dt_2 dt_1 \\
&= \alpha\mu \int_0^t e^{\beta t_1} \int_{t_1}^t e^{-\beta t_2} dt_2 dt_1 \\
&= -\frac{\alpha\mu}{\beta} \int_0^t e^{\beta t_1} \left(e^{-\beta t} - e^{-\beta t_1}\right) dt_1 \\
&= -\frac{\alpha\mu}{\beta} \left(\int_0^t e^{\beta(t_1 - t)} dt_1 - \underbrace{\int_0^t 1 dt_1}_{t} \right)
\end{aligned}$$

And then :

$$\mathbb{E}\left[N_t^{(2)}\right] = -\frac{\alpha\mu}{\beta} \left(\left[\frac{1}{\beta} e^{\beta(t_1 - t)} \right]_{t_1 = 0}^{t_1 = t} - t \right) = -\frac{\alpha\mu}{\beta} \left(\frac{1}{\beta} \left(1 - e^{-\beta t}\right) - t \right) = \frac{\alpha\mu}{\beta} \left(t + \frac{1}{\beta} e^{-\beta t} - \frac{1}{\beta} \right)$$

After performing the necessary integrations, we arrive at the final expression for the expectation of $N_t^{(2)}$:

Expectation of $N_t^{(2)}$

$$\mathbb{E}\left[N_t^{(2)}\right] = \frac{\alpha\mu}{\beta} \left(t + \frac{1}{\beta} e^{-\beta t} - \frac{1}{\beta} \right)$$

Now we use the expectation of $N_t^{(1)}$ and $N_t^{(2)}$ to estimate the values of u and c in the Cramer-Lundberg type equation through the expectation of R_t:

$$\begin{aligned}
\mathbb{E}[R_t] &= \mathbb{E}\left[u + ct - \sum_{i=1}^{N_t} Y_i\right] = u + ct - \mathbb{E}\left[\sum_{i=1}^{N_t} Y_i\right] = u + ct - \mathbb{E}\left[\mathbb{E}\left[\sum_{i=1}^{N_t} Y_i | N_t\right]\right] \text{ (Tower property)} \\
&= u + ct - \mathbb{E}\left[\sum_{i=1}^{N_t} \mathbb{E}[Y_i|N_t]\right] \text{ (linearity of } \mathbb{E}) \\
&= u + ct - \mathbb{E}[N_t \times \mathbb{E}[Y_i]] \text{ } (Y_i \text{ iid. and } N_t \perp\!\!\!\perp Y_i \text{ }) \\
&= u + ct - \mathbb{E}[N_t]\,\mathbb{E}[Y_i]
\end{aligned}$$

By denoting $\mathbb{E}[N_t] = m$ and $\mathbb{E}[N_t] = \mathbb{E}\left[N_t^{(1)}\right] + \mathbb{E}\left[N_t^{(2)}\right] = \mu t + \frac{\alpha\mu}{\beta}\left(t + \frac{1}{\beta}e^{-\beta t} - \frac{1}{\beta}\right)$, we deduce

Expectation of R_t

$$\mathbb{E}[R_t] = \left(u + \frac{m\alpha\mu}{\beta^2}\right) + t\left[c - m\mu\left(1 + \frac{\alpha}{\beta} + \frac{\alpha}{\beta}.\frac{e^{-\beta t}}{\beta t}\right)\right] \tag{2}$$

We always want R_t to be a positive number, i.e. we never go bankrupt. By that we can choose u and p to make profits for us in such a way that

$$u + \frac{m\alpha\mu}{\beta^2} > 0 \text{ and } c - m\mu\left(1 + \frac{\alpha}{\beta} + \frac{\alpha}{\beta}.\frac{e^{-\beta t}}{\beta t}\right) > 0$$

We can easily choose $u > 0$ and $c > m\mu\left(1 + \frac{\alpha}{\beta} + \frac{\alpha}{\beta}.\frac{e^{-\beta t}}{\beta t}\right), \forall t > 0$, which means $c > \sup_{t>0}\left[m\mu\left(1 + \frac{\alpha}{\beta} + \frac{\alpha}{\beta}.\frac{e^{-\beta t}}{\beta t}\right)\right]$. The function $f(t) = \frac{e^{-\beta t}}{\beta t}$ is monotonically decreasing on $]0, \infty[$ and its maximum value is ∞ at $t = 0$, which is an impossible value for p. Moreover, this is the suffisant condition, not the necessary condition. We can conduct another strategy to choose these parameters, that is to choose u larger and then decrease c, basing on the fact that our contract is term contract, so the constrain on c is loosen. We can choose $p > m\mu\left(1 + \frac{\alpha}{\beta} + \frac{\alpha}{\beta}.\frac{e^{-120\beta}}{120\beta}\right)$ with $t = 120$. Through our experiments, we found that these values give the most satisfying values for the simulation. More details are found in chapter 3.

3 Application of Self-exciting processes in Case Study of Insurance

3.1 Context

In France, all citizens are reimbursed 70% for services covered by the government-supported AMELI insurance system. The remaining 30% is typically covered by private complementary insurance, provided the patient has purchased such coverage. However, the report "Complementary Health Care: Players, Beneficiaries, Coverage – 2024 Edition" [1] highlights several challenges:

- Rising premiums with age for individual contracts.
- Limited access for vulnerable populations, with 12% of the most disadvantaged still uninsured despite the implementation of the Solidarity Complementary Health Insurance (CSS) reform.
- Persistent income inequality, as private insurance does not reduce disparities in access to healthcare like public insurance does.

The Figure 3 shows the distribution of excess healthcare costs in 2021 across 10 equal population groups (D1 to D10). The first six groups (D1 to D6) pay almost nothing, while costs increase significantly in the higher groups. The top 10% (D10) pay the most, with an average of €1,440 annually, highlighting a clear disparity in healthcare expenses among the population.

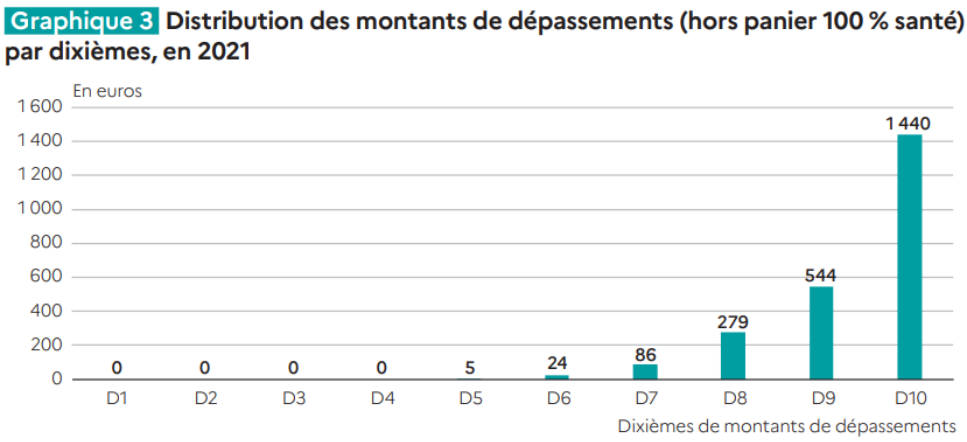


Figure 3: Average of total payment annually by 10 equal parts in 2021. Source : [1]

This disparity underscores the need for reforms to improve equity in healthcare costs. As discussed in [1], one key suggestion is the introduction of a universal, compulsory, and pooled private insurance system. Such a system aims to reduce financial inequality and ensure that excessive healthcare expenses are more evenly distributed across the population, addressing the significant burden currently borne by the top deciles.

In light of these issues, we have chosen to analyze and model out-of-pocket healthcare expenses as part of a broader analysis. Our approach ues the Cramér-Lundberg model with a self-exciting process to better understand cost dynamics and financial risk.

3.2 Asset modelling

General model

In this first part, we simulate the number of doctor visits for each customer using a counting process $(N_t)_{t\in\mathbb{R}}$, where t is measured in days. We model this counting process as a self-excting process, reflecting the fact that a patient who visits a doctor is likely to return for a follow-up appointemetn or be referred to a specialist within a short period of time. This period may range from a few days to a month or a few months. Therefore, each subsequent visit is triggered by the previous one, making the assumption of self-exciting process appropriate for capturing this interdependence.

Each time a patient consults a doctor or undergoes any medical intervention, the insurance company incurs a cost, denoted by Y_i with i standing for his/her i-th visit. The total cost of doctor visits for the insurance company depends on the number of visits, and these costs are added over time. This dynamic process can be modeled using the Cramer-Lundberg equation, a fundamental equation in risk theory, which models the financial position of an insurance company over time. To apply this Cramer-Lundberg model into our application, we need to redefine several parameters:

- t: The "wealth" of the insurance company at time t, representing the accumulated funds at any given moment.
- R_t: The "wealth" of the insurance company at time t, representing the accumulated funds at any given moment.
- u: The initial amount paid by the customer to the insurance company at the start of the insurance contract.
- c: The regular monthly premium (or subscription) paid by the customer.

The wealth of the insurance company at time t is described by the following equation:

$$R_t = u + ct - \sum_{i=1}^{N_t} Y_i \tag{3}$$

In this study, our objective is to utilize this model to simulate the "wealth" of the insurance company for a specific customer by modeling the process R_t. To achieve this, it is necessary to determine appropriate values for the parameters u, c, N_t, and Y_i. This raises a critical question: what are the optimal values for each of these parameters? In the subsequent sections, we address this question by establishing a set of hypotheses and assumptions to guide the selection of these parameters.

Counting process modelling

In this second part, we model the counting process $(N_t)_{t\in\mathbb{R}}$, representing the number of doctor visits over time for an individual. Based on data collected from Google, we observe that, on average, a French person visits a doctor for a consultation three times per year, excluding follow-up visits. To quantify this observation, we define the intensity of primary doctor visits, denoted by μ, as the average number of visits per day. Given that there are 365 days in a year, we calculate:

$$\mu = \frac{3}{365} \approx 0.00082$$

This represents the rate of base medical appointments for an individual on any given day, without accounting for follow-up visits.

To model follow-up visits or specialist referrals that are triggered by base events, we incorporate the "self-exciting" function $\Phi(t) = \alpha e^{-\beta t}$ along with two additional parameters, α and β, to capture the dynamics of these secondary events. Typically, follow-up appointments or referrals to specialists occur within one week after the initial consultation. Based on this observation, we set

$$\alpha = 0.03 \text{ and } \beta = 0.06$$

reflecting the rate at which follow-up visits or referrals occur in relation to the primary doctor visits.

To simulate this counting process, we generate a representation of doctor visits for a single individual over the course of one year. The resulting simulation, shown in Figure 4, plots the occurrences of these visits over time. The time axis is measured in days, with blue points indicating the primary doctor visits for a specific condition. Red points represent follow-up visits, which may include return visits for the same condition or referrals to specialists for further tests or consultations.

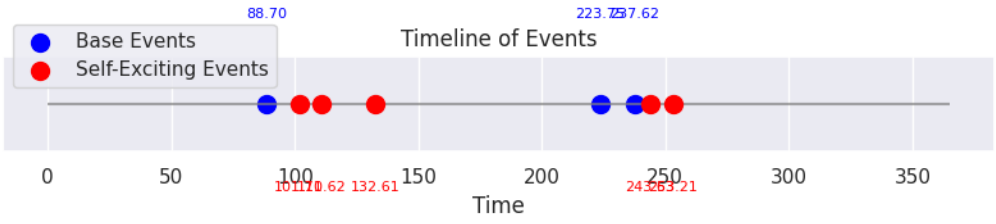


Figure 4: An exemple of self-exciting counting process

Claims modelling

Based on observations of the out-of-pocket costs ("reste-à-charge") for medical consultations in the market, after deducting the 70% coverage provided by AMELI [5], we present the table 1. This table outlines the various possible values of the out-of-pocket cost Y_i that the insurance company would cover for its customers. The cost varies depending on the category of the medical consultation, as well as the probability associated with Y_i taking on each corresponding value within these categories.

Table 1: Payment Categories for Medical Consultations

Probability	Category	Cost (Euros)
30%	General practitioner consultation	0
35%	Medication purchase	7
17%	Specialist consultation	30
8%	Radiology or dental care	120
6%	Pair of glasses	240
3%	Higher medical expenses	500
1%	Exceptional expenses	1450

Following the category table, we can calculate on average the cost of one sinister (denoted by m) is 61.05€.

Parameter selection

In this part, we need to choose parameters for the Cramer-Lundberg model, particularly u and c. These parameters need to satisfy a double objectives:

- provide an affordable price to attract customers.
- keep the company not to go bankrupt.

Our goal is to prevent the bankruptcy of the company, meaning we aim to cover all risks without focusing on generating profits just yet. Therefore, we choose u and c to meet the minimum required demand. If we want to make a profit, we can increase the initial capital u as needed. To achieve this, we need to select values for u and c that are as small as possible from the clients' perspective, and the least requirements from the company's perspective. Based on the mathematical computations in the second section, we suggest setting u to the amount the company wants to earn from the contract, saying 10€ in our simulation and $c = m\mu\left(1 + \frac{\alpha}{\beta} + \frac{\alpha}{\beta} \cdot \frac{e^{-120\beta}}{120\beta}\right)$. Theoretically, these values flatten the mean value of R_t as t increases, since c is chosen to adjust the slope of the linear term in $N_t^{(2)}$ close to 0.

3.3 Simulation of the asset using self-exciting parameters

After establishing all the needed parameters, we can simulate the asset of the company for a given customer using the stochastic process R_t, which is modeled in the figure 5.

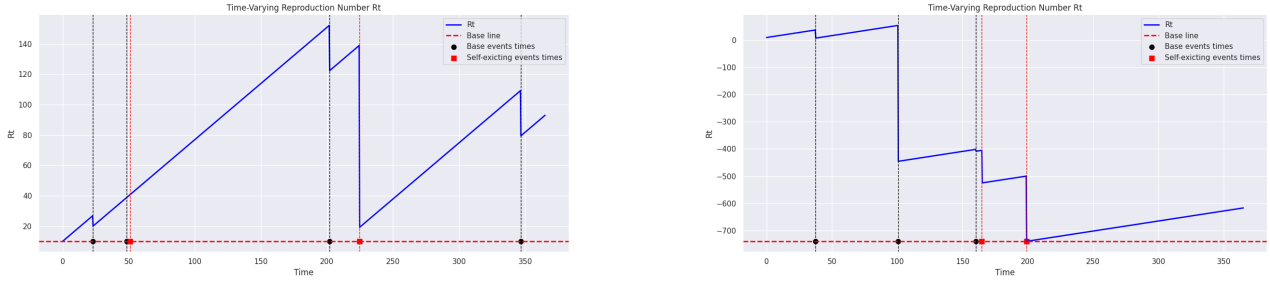


Figure 5: Simulation of risk process R_t

Estimator of expected value R_t

Before going to explain our simulation, we introduce first the empirical mean of the wealth R_t. We would like to use this estimator to approximate our theoretical mean value of R_t, which are presented in the equation 2. Here we introduce an abuse of notation that we use R_T to represent the function $(R_t)_{t\in[0,T]}$.

//

Estimator of expected value Rt

$$\begin{aligned} \bar{R}_T \quad : \quad & [0,T] \longrightarrow \mathbb{R} \\ & t \longmapsto \sum_{i=1}^{n} R_T^i(t) \end{aligned}$$

Using the Central Limit Theorem and Slutsky's Lemma, we have for all $t \in [0,T]$

$$\frac{\sqrt{n}(\bar{R}_T(t) - \mathbb{E}\left[R_T(t)\right])}{S_T(t)} \underset{n\to\infty}{\overset{\mathcal{L}}{\longrightarrow}} t_{n-1}$$

where $S_T(t)$ is the empirical variance given by

$$S_T^2(t) = \frac{1}{n-1} \sum_{i=1}^{n} \left(R_T^i(t) - \bar{R}_T(t)\right)^2$$

With n big we have:

$$\bar{R}_T(t) \underset{n\to\infty}{\overset{\mathcal{L}}{\longrightarrow}} \mathcal{N}(\mathbb{E}[R_t], \frac{S_T(t)}{n})$$

In Figure 6, we present the simulation of the mean value of R_t (blue line) using the Monte Carlo method, along with its standard deviation (represented by the shaded blue region). These results are compared to the theoretical mean value of R_t (red line). With the selected values of c and u, the simulated mean R_t aligns closely with the theoretical mean during the first 50 days. However, as time progresses, a growing divergence between the two values is observed. Despite this, the theoretical R_t consistently remains within the confidence interval of the empirical R_t, as indicated by the shaded region.

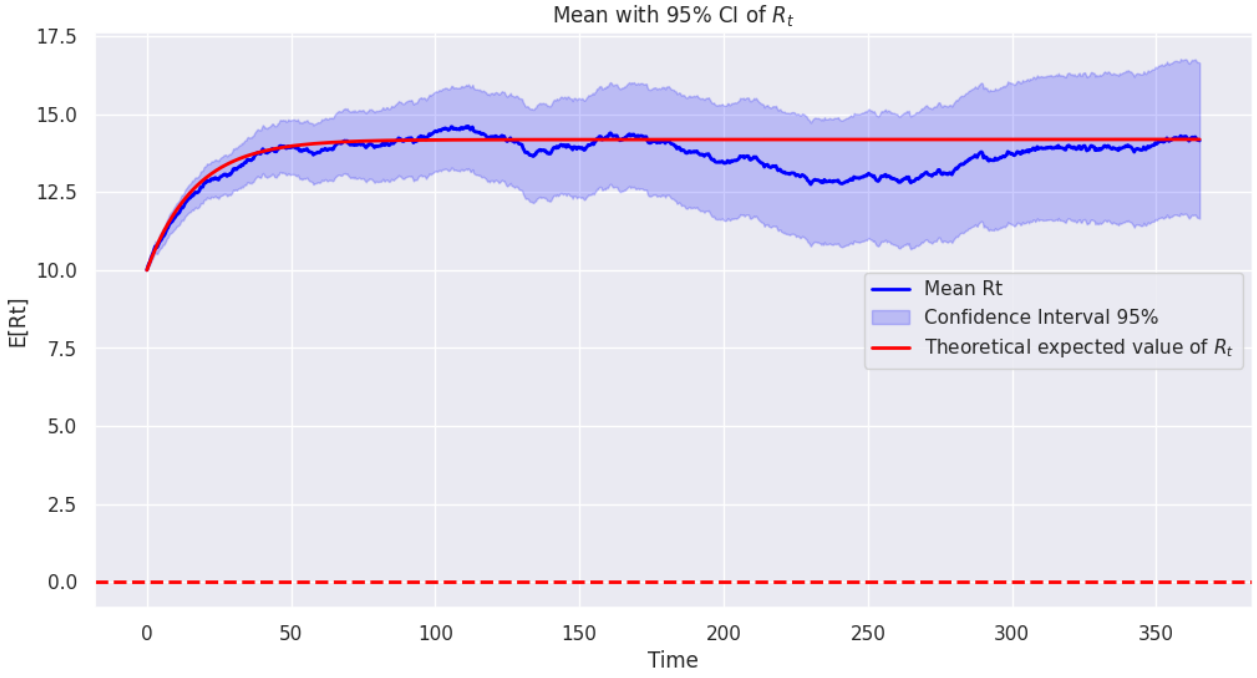


Figure 6: Mean with his confidence interval of R_t over time theorically and by Monte Carlo

The parameters c and u are determined based on the findings of Chapter 2 to ensure that R_t does not fall below the zero line, which would signify bankruptcy or ruin. However, as emphasized in Chapter 2, no explicit upper bounds are established for c and u, prompting the need to identify their optimal values. To address this, we aim to balance risk management, financial stability, and customer retention. Specifically, these parameters are chosen to minimize service prices and avoid discouraging clients. We select u and c such that R_t remains approximately flat or exhibits a very slight positive slope, ensuring that service costs remain competitive and accessible. This choice raises the concern of whether such a strategy eliminates profitability.

However, profitability is still achieved through the initial capital u, collected at the start of the contract, and through a modest increase in wealth toward the end of the contract.

3.4 Ruin probability simulation

Definition : Ruin probability

The ruin probability is defined as the probability that R_t falls below the zero line at least once within an established time interval $[0, T]$. Mathematically, if we call $\tau(u) := \inf t > 0 | R_t < 0$ the ruin time, then we set the ruin probability this quantity

$$\Psi(u) := \mathbb{P}\left[\inf_{t>0} R_t < 0\right] = \mathbb{P}\left[\tau(u) < +\infty\right] \tag{4}$$

Method : Estimating ruin probability

This probability is often very difficult to compute mathematically. We usually do a simulation of it. To do so, we can simulate n curves of total assets R_T (we can consider that it would be the simulation for n customers). Then, we can compute an empirical mean curve $\bar{R_T}$. Please note that this is a function and not a scalar. Indeed, assuming that we have the n total assets $\left(R_T^{(i)}\right)_{i\in[\![1,n]\!]} \in \left(\mathbb{R}^{[0,T]}\right)^n$, we have : We will thus say that we have ruin if there exists a time t on $[0, T]$ where $\bar{R_T}$ is negative, i.e :

$$\min_{t\in[0,T]} \bar{R_T}(t) < 0$$

We can now consider a new function :

$$\begin{array}{rccl} \zeta & : & \left(\mathbb{R}^{[0,T]}\right)^n & \longrightarrow \{0,1\} \\ & & \left(R_T^{(i)}\right)_{i\in[\![1,n]\!]} & \longmapsto \begin{cases} 1, & \text{if } \min\limits_{t\in[0,T]} \bar{R_T}(t) < 0, \\ 0, & \text{else.} \end{cases} \end{array}$$

Then, we can do this simulation N_{exp} times. Let $\left(\left(R_T^{(i)}\right)_{i\in[\![1,n]\!]}^{(k)}\right)_{k\in[\![1,N_{exp}]\!]} \in \left(\left(\mathbb{R}^{[0,T]}\right)^n\right)^{N_{exp}}$ (we perform N_{exp} times the simulation of $\bar{R_T}$) and count how many times (N_{ruin}) we have ruin. i.e :

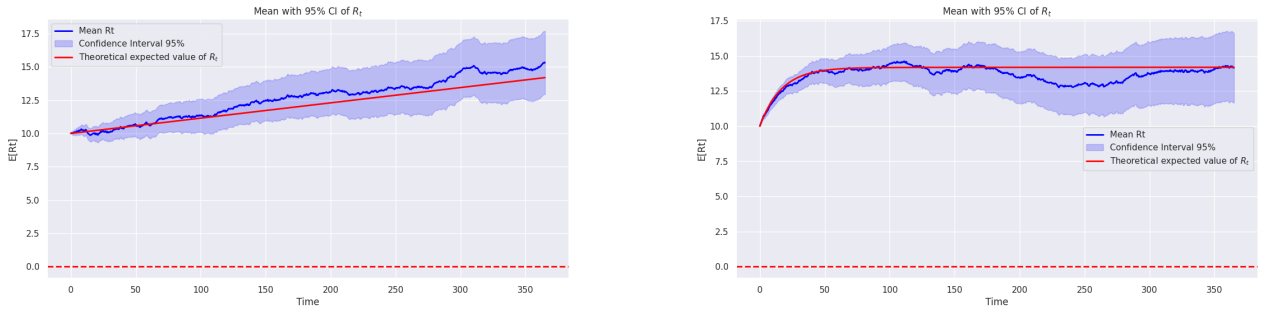
$$N_{\text{ruin}} = \sum_{k=1}^{N_{exp}} \zeta\left(\left(R_T^{(i)}\right)_{i\in[\![1,n]\!]}^{(k)}\right)$$

The probability of ruin will be thus given by :

$$p_{\text{ruin}} = \frac{N_{\text{ruin}}}{N_{\text{exp}}}$$

Using this strategy, we find that the ruin probability in this case is around 0.03, which is acceptable.. When playing on this ruin probability, we found that this probability is very sensitive to u. That means, a little increase in the initial capital can lower down the ruin much more effectively. Howerver, we did not conduct experiments to verify that. We suggest a future work using Sensitivity Analysis and important sampling for this function to find out the impact of u and c on this probability.

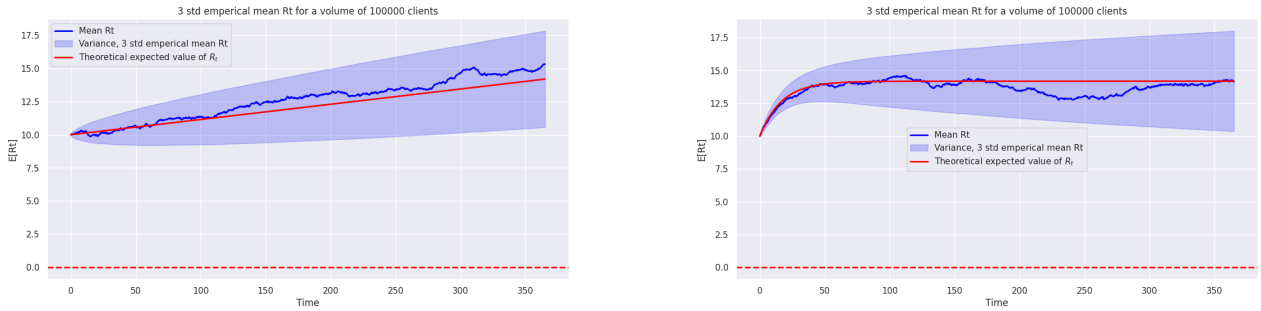
3.5 Comparison between Self-exciting process in Cramer-Lundberg type model and Poisson process in Cramer-Lundberg model



(a) Poisson process (b) Self-exciting process

Figure 7: Empirical means with their confidence interval of R_t over time

We have fixed the value $\mu_{poisson} = \frac{\mathbb{E}[N_T]}{365}$ to ensure the same yearly frequency of patient visits to the doctor between the two models. We notice that the mean obtained with the self-exciting process initially increases and then decreases sharply, while the Poisson process remains generally constant (or only slightly decreasing).



(a) Poisson process (b) Self-exciting process

Figure 8: Empirical mean with his std of R_t over time

If we were to consider a contract of longer duration, we would have to re-evaluate the value of c to straighten out the curve for self-exciting process, ideally to a horizontal asymptote when $t \to +\infty$, whereas the Poisson process model requires no adjustments. We could also lower the value of u slightly, since there is short but considerable growth at the start of the contract.

Since it is not fully constant, the curve of the empirical mean $\bar{R}_t$ obtained with Poisson processes should also be straightened by slightly increasing the value of c.

Conclusion

This study aimed to extend the Cramer-Lundberg model by incorporating self-exciting processes to better model financial risks in the context of insurance. The goal was to balance affordable healthcare costs for customers while ensuring the financial stability of the insurance company. Key methods involved modeling doctor visits using a self-exciting counting process, simulating the company's wealth over time, and estimating ruin probabilities. Results showed that self-exciting processes captured event interdependencies more effectively than Poisson processes, providing valuable insights into asset dynamics. The ruin probability was sensitive to initial capital, with small changes to u having a significant impact on financial stability. However, the model has limitations. The triggering function $\Phi(t)$, which assumes a peak immediately after an event, may not be entirely appropriate. Patients could have follow-up appointments after several days, so the current form of the triggering function may not accurately reflect such delays. Additionally, parameter choices were heuristic, and the model's assumptions about patient behavior and healthcare costs may limit its generalizability. Therefore, we suggest future studies focusing on refining the triggering function to account for delayed follow-ups, conducting sensitivity analyses on the parameters, and validating the model with real-world data. These improvements could enhance predictive accuracy and the model's applicability across different insurance scenarios.

References

[1] Drees. La complémentaire santé - acteurs, bénéficiaires, garanties. Technical report, République française, France, Édition 2024. URL https://drees.solidarites-sante.gouv.fr/sites/default/files/2024-09/CS2024MAJ040924.pdf.

[2] Anthony Réveillac. An introduction to the mathematics of actuarial sciences. Lecture notes at INSA Toulouse, 2024. Unpublished manuscript.

[3] Yuanda Chen. Thinning algorithms for simulating point processes. *Florida State University, Tallahassee, FL*, 2016.

[4] Mélisande Albert. Poisson processes and application to reliability theory and actuarial science. Lecture notes at INSA Toulouse, 2024. Unpublished manuscript.

[5] Service-Public.fr. Reimbursement of medical expenses by social security, 2024. URL https://www.service-public.fr/particuliers/vosdroits/F1069?lang=en. Accessed: 2024-12-13.