POLITECNICO
MILANO 1863

Characterization of insurance premium principles: the extension to dynamic risk models

AY 2022-2023

Candidate: Luca Russo

Supervisor: Matteo Brachetta

Ringraziamenti

Innanzitutto desidero ringraziare il professor Brachetta per la disponibilità e l'entusiasmo con cui mi ha seguito durante la scrittura di questa tesi. Ogni momento di confronto con lui è stato illuminante, rendendo questo lavoro una grande occasione per la mia crescita professionale.

Vorrei poi ringraziare i miei genitori per avermi accompagnato e sostenuto lungo tutto il mio percorso di studi, in particolar modo nei momenti più difficili, e mio fratello Roberto, che con il suo esempio mi ha fatto appassionare alle discipline scientifiche, spronandomi sempre a migliorarmi e a dare il meglio di me.

Un grazie speciale per Alessandro, il mio migliore amico, che mi sopporta ormai da tanti anni, e che più di tutti mi accompagna sia nei miei momenti di euforia che in quelli di sconforto, la mia esperienza universitaria non sarebbe stata la stessa senza i suoi continui incoraggiamenti. Un grazie anche ad Angelo, che ha il difficile compito di sopportarmi non solo come coinquilino, ma anche come compagno di stanza.

Vorrei poi ringraziare Alice, Anna, Marcello, Camilla, Federico, Guido, che hanno reso i miei anni di magistrale molto più leggeri e divertenti. Un ringraziamento particolare va a Nicola, con il quale ho condiviso tanti momenti di gioia e disperazione durante la preparazione per gli esami, tra le cose che mi mancheranno di più dell'università ci sono sicuramente le innumerevoli ore passate insieme a lui a preparare dimostrazioni in giro per le lavagne delle aule del Politecnico.

Infine, vorrei ringraziare Tommaso, Sophie, Alice, Aurora, Elisa, Luca, Federico, Giaime, per aver reso gli ultimi mesi di scrittura della tesi infinitamente più piacevoli con la loro compagnia. Auguro a tutti loro di poter vivere il prosieguo dei loro studi con la stessa gioia e leggerezza che mi hanno sempre trasmesso.

Sono sempre stato una persona nostalgica e avversa ai cambiamenti, ma grazie a tutti voi, ogni volta che ripenserò a questi anni, lo farò sempre con un grande sorriso. Vi voglio bene.

Sommario

In letteratura attuariale è presente una diffusa trattazione della caratterizzazione matematica dei principi di premio. Tuttavia, questa è limitata al solo contesto statico, in cui i rischi sono modellizzati come singole variabili aleatorie non dipendenti dal tempo.

Lo scopo essenziale di questa tesi è di fornire una trattazione sistematica e rigorosa in un contesto dinamico. Ciò può essere potenzialmente utile per molte branche della letteratura attuariale, come la progettazione dei contratti, il controllo del rischio, strategie ottime di riassicurazione ed altre. L'analisi è svolta per due diversi modelli di rischio: il modello di Cramér-Lundberg, ampiamente usato nel settore assicurativo,e il modello di rischio basato sui processi di Hawkes, che è adatto a descrivere possibili clustering di eventi, che un portafoglio assicurativo può subire.

Parole-chiave: Assicurazioni, principi di premio, modelli di rischio.

Abstract

In actuarial literature there is a widespread treatment about the mathematical characterization of the premium principles. However, this is limited to the static context, in which risks are modeled as singles random variables, that are not time-dependent.

The essential aim of this thesis is to provide a systematic and rigorous treatment of premium principles in a dynamic context. This can be potentially useful for many branches of actuarial literature, such as contracts design, risk control, optimal reinsurance policies and others. The analysis is carried out for two different risk models: the Cramér-Lundberg model, widely used in the insurance business, and the risk model based on Hawkes processes, which is suitable for describing possible clustering of events, which an insurance portfolio can suffer.

Key words: Insurance, premium principles, risk models.

Contents

Contents **v**

Introduction **vii**

1 Premiums in static context **1**
1.1 Properties of the premium functional 1
1.2 Catalog of premium principles 6
1.2.1 Net premium principle and expected value premium principle . . . 7
1.2.2 Variance and standard deviation premium principles 8
1.2.3 Wang's premium principles 12
1.2.4 Table of properties 14
1.3 Decomposition of the premium functional into risk and deviation measures 15

2 Theoretical elements on jump processes **19**
2.1 Brief review of stochastic calculus 19
2.2 Point processes 20
2.3 Marked point processes 30

3 Premiums in dynamical context **35**
3.1 The Cramér-Lundberg model 35
3.1.1 Net premium and expected value premium 41
3.1.2 Variance premium and Standard deviation premium 44

4 Hawkes processes **47**
4.1 Review of theory 47
4.2 Risk model with Compound Hawkes 54
4.3 Estimation of the error of a Cramér-Lundberg model in a context with clustering 59
4.4 Another choice of the premium rate in the risk model with Hawkes 62

Bibliography **65**

Introduction

One of the most relevant topic in actuarial literature is the characterization of the premium principle, namely the rule for assigning the correct premium to an insurance risk. Indeed, in a perfect market, with infinitely many players, perfect information and complete diversifiability of risks, the insurance price should be the expected value of the loss, since the market should not provide rewards for diversifiable risks (see [23]). However in more realistic contexts, markets present prices larger than the expected value of the insured claims. Therefore, actuaries need more sophisticated mathematical tools for developing premium principles.

The terms premium or insurance price will refer to pure premium, so the net expected loss plus a risk load charge, ignoring loadings for profit or expenses. The present thesis aims to review and extend the existing literature about premium principles by proposing a treatment both in static and dynamic context, in which risks are modeled by stochastic processes. Indeed, in order to analyze the state of an insurance company portfolio, it is required to consider its value as a time-dependent quantity, introducing therefore a counting process for describing the arrival of the claims, and the premiums collected as a rate. The results contained in this work are thus useful for having risk models in which the premium rate is a well-defined variable, with a solid mathematical basis that proves its properties and shows its connection with the loss process, underlying possible advantages and drawbacks for each proposal. Moreover, the theory developed can be used for further works in which classical problems of the actuarial sector such as contracts design, risk control (see [2]) for an example), optimal reinsurance policies (see [18]) and others (for instance the estimation of the ruin probability according to the choice of the premium rate as in [15]). Eventually, the approach used for the characterization of the premium rate can be extended to other risk models in addition to the ones dealt with here.

The thesis is structured as follows:

- In chapter 1 the desired properties of a good premium principle are listed, then it follows the treatment of different premium principles in a static context, focused in particular on providing proves or counterexamples for each property. Eventually, it is presented the result of decomposition of the premium functional into a risk measure and a deviation measure (see [19]).

- In chapter 2 theoretical elements of stochastic calculus are introduced, with a particular care with respect to point processes and marked point processes, which are fundamental in the sequel of the work in order to build actuarial models in a dynamic context.
- In chapter 3 is analyzed the Cramér-Lundberg model, a standard risk model in the actuarial framework, with the aim of underlying its features and disadvantages. Moreover, the premium principles mentioned in chapter 1 are extended to this dynamical context, with an original treatment of their properties in continuous time.
- In chapter 4 there is a theoretical introduction to mono-dimensional Hawkes processes, then it is proposed a risk model for an insurance portfolio in which the loss is distributed as a Compound Hawkes. The net and the expected value principles are treated in this context, with an analysis of the properties as done in the previous chapter. It is also present an estimation of the error committed by the Cramér-Lundberg model with an expected value premium with respect to the new risk model, in a situation with clustering of claims. Eventually, a brief section shows that for the model provided, a premium based on the stochastic intensity of the loss allows to keep the portfolio surplus positive on average.

1 | Premiums in static context

1.1. Properties of the premium functional

Let $(\Omega, \mathcal{F}, P)$ be a probability space in which Ω is the set of states of the world or possible outcomes, $\mathcal{F} \subseteq 2^{\Omega}$ is a σ-algebra that is the collection of events in Ω and P is a probability measure. Let χ be the set of non-negative random variables measurable in the probability space defined before, which represents the set of possible losses due to insurance risks. In this mathematical environment, a premium is described as a functional from the set of insurances to the set of real non-negative numbers, therefore:

$$H : \chi \to [0, +\infty) \tag{1.1}$$

Moreover, in order to characterize the distribution of a risk $X \in \chi$, it is often used the concept of decumulative distribution function (or survival function).

Definition 1.1 (Decumulative distribution function or survival function). *The ddf of a random variable X is denoted as:*

$$S_X(t) = P(\omega : X(\omega) > t) \tag{1.2}$$

Using the definition of ddf it is possible to introduce the stochastic dominance as follows:

Definition 1.2 (First stochastic dominance). *Let X and Y in χ, X precedes Y in first stochastic dominance (FSD) if $S_X(t) \leq S_Y(t)$ for all $t \geq 0$.*

This concept is useful for comparing two risks and the potential losses they entail. Lastly, it is important to introduce the definition of comonotonicity:

Definition 1.3 (Comonotonicity). *Let X and Y in χ, they are comonotonic if and only if*

$$[X(\omega_1) - X(\omega_2)][Y(\omega_1) - Y(\omega_2)] \geq 0 \ \text{a.s. for } \omega_1, \omega_2 \in \Omega \tag{1.3}$$

Comonotonic literally means "Common monotonic", thus their outcomes move on the same direction. The following result about comonotonicity, presented in [10], gives another useful characterization of the concept of comonotonicity:

Theorem 1.1. *$X = [X_1, X_2]$ is a random vector comonotonic $\iff$ $\exists Z$ random variable and $\exists f_1, f_2$ non-decreasing functions such that $X \overset{d}{=} [f_1(Z), f_2(Z)]$;*

Proof. ($\Rightarrow$) Let U be a random variable distributed as a $Uniform[0,1]$. Now it suffices do the following observation, due to the integral transform of probability:

$$F_X(\underline{x}) = P(U \leq F_{X1}(x_1), U \leq F_{X2}(x_2)) = P(F_{X_1}^{-1}(U) \leq x_1, F_{X_2}^{-1}(U) \leq x_2)$$

The inverse functions of the cumulative distribution functions satisfy all the properties requested.

($\Leftarrow$) Trivial, since the two functions f_1, f_2 are non-decreasing.

□

Remark 1.1. The equation (1.3) may resemble the formula of covariance and leads to think that two comonotonic random variables should be positively correlated. This can be shown by developing the first expected value in the definition of covariance; Let X, Y two well-defined and comonotonic random variables in a probability space $(\Omega, \mathcal{F}, P)$ and $X, Y \in L^2(\Omega, \mathcal{F}, P)$. Then:

$$Cov(X,Y) = E[(X - E[X])(Y - E[Y])] = \sum_{i=1}^{\infty}(X(\omega_i) - E[X])(Y(\omega_i) - E[Y])P(\omega_i).$$

Since the two variables are comonotonic, the sum above is composed only by non-negative terms if the mean of each random variable belongs to its support, therefore it is non-negative as well. The result can be easily extended to the general case by defining a set $\widetilde{\Omega} = \Omega \cup \widetilde{\omega}$ and two random variables $\widetilde{X}(\omega) = X(\omega)$ for each $\omega \in \widetilde{\Omega}$, $\omega \neq \widetilde{\omega}$ and $\widetilde{Y}(\omega) = Y(\omega)$ for each $\omega \in \widetilde{\Omega}$, $\omega \neq \widetilde{\omega}$ such that $P(\widetilde{\omega}) = 0$. Noticing that the covariance of the two new random variables is equal to $Cov(X,Y)$, the result is obtained. Moreover, it can be noticed that the covariance can be null only if at least one of the two random variables is constant a.s. Eventually, if two variables are perfect correlated, it means that one is a linear and non-decreasing function of the other one, and for the characterization given in theorem 1.1, it follows that they are comonotonic.

Having set these definitions, now it is possible to start asking how a "good" premium principle can be defined. There are different methods to achieve this goal. The "ad hoc" one consists in defining the functional and then determine its properties, the characterization one instead consists in looking for a principle starting from the properties it should have (see [24]).

This section presents and discusses a set of conditions which can be desirable for H:

1. Conditional state dependence: For a given market condition, the premium for a risk X depends only on its ddf.

   This property is quite reasonable since two risks with the same distribution in the same market should have the same insurance price.

2. Monotonicity: Let X and Y be in χ, if $X(\omega) \leq Y(\omega)$ a.s., then $H(x) \leq H(y)$

   Trivially, if a risk is dominated by another almost surely, its insurance price should be smaller than the other one.

3. Comonotonic additivity: If X and Y are in χ and comonotonic,then: $H(X+Y) = H(X) + H(Y)$

4. Continuity: If $X \in \chi$ and $d \geq 0$, then: $\lim_{d \to 0+} H((X-d)_+) = H(X)$ and $\lim_{d \to \infty} H(\min(X,d)) = H(X)$

5. Risk loading: $H(X) \geq E[X]$ for all $X \in \chi$.

   An insurer generally wants to receive premiums greater than the expected value of the risk in order to gain money on average.

6. No unjustified risk loading: if $X = c$, $X \in \chi$ and c constant, then $H(X) = c$.

   If the loss is deterministic there is no reason for adding a risk loading to the premium since it is known the outcome for the insurer with probability 1

7. Maximal loss: $H(X) \leq esssup(X)$ for all $X \in \chi$

   The premium cannot be greater than the maximal possible outcome for the insurer.

8. Translation equivariance or invariance: $H(X+a) = H(X) + a$ for all $X \in \chi$ and $a \geq 0$.

   If the risk is increased by a fixed deterministic amount, then the premium should be increased by the same amount.

9. Scale equivariance or invariance: $H(bX) = bH(X)$ for all $X \in \chi$ and $b \geq 0$.

   This property guarantees no arbitrage.

10. Additivity: $H(X+Y) = H(X) + H(Y)$ for all $X, Y \in \chi$.

    Also this property is useful to prevents arbitrage opportunities.

11. Subadditivity: $H(X+Y) \leq H(X) + H(Y)$ for all $X, Y \in \chi$.
12. Superadditivity: $H(X+Y) \geq H(X) + H(Y)$ for all $X, Y \in \chi$.
13. Additivity for independent risks:$H(X+Y) = H(X) + H(Y)$ for all $X, Y \in \chi$ independents.
14. Preserves FSD ordering: If $S_X(t) \leq S_Y(t)$ for all $t \geq 0$, then $H(X) \leq H(Y)$.
15. Preserves stop-loss ordering:$E[X-d]_+ \leq E[Y-d]_+$ for all $d \geq 0$, then $H(X) \leq H(Y)$.

Approaching the problem with the characterization method does not entail that H must have all the features listed above. For instance, in [23], a premium principle is defined as the integral of the ddf of the risk X, by imposing that just the first four properties holds, however, one can show that many other properties such as sub-additivity or scale and translation equivariance can be derived by the definition. Furthermore, many properties can be justified with a "no arbitrage" argument, such as the scale equivariance or the additivity ones; indeed, if the former does not hold and, for example, the premium for $2X$ were greater than twice the premium of X, one could buy insurance for $2X$ and sell two different policies for X, making an arbitrage profit; a same reasoning holds also for the latter one. However, this is a questionable argument to justify some properties, in fact one can argue that buying an insurance for $X+Y$ and then selling the two risk separately is not allowed by the market itself, and therefore, it should be sufficient to require just the weaker property of subadditivity.

In general, it is interesting to notice how the first and the second properties are related to the 14-th.

Theorem 1.2. *If X precedes Y in FSD, then there exist a random variable Z with $S_Y = S_Z$ such that $X \leq Z$ a.s.*

Proof. [1] From hypothesis:

$$S_X(t) \leq S_Y(t)\ \forall t \Rightarrow \exists c(t) : \mathbb{R} \longrightarrow [1, +\infty)$$

[1] I am grateful to Jacopo Somaglia for his precious help in the development of the proof and in the construction of the next counterexample.

such that $S_X(t/c(t)) = S_Y(t)$, by continuity of the survival functions. $c(t)$ is a continuous function in $(0, \infty)$, and it exists its continuous extension by imposing $c(0) = 1$, since, by considering two non-negative random variable X, Y, $S_X(0) = S_Y(0) = 1$.

Let define

$$Z(\omega) = c(t)X(\omega) \ \forall \omega \in X^{-1}(t).$$

By definition, $Z \geq X$ a.s. It remains to prove that it has the same distribution of Y. Considering the survival function of Z:

$$S_Z(t) = P(Z \geq t) = P(c(t)X \geq t) = P(X \geq t/c(t)) = S_x(t/c(t)) = S_Y(t);$$

This concludes the proof. □

Remark 1.2. The function $c(t)$ used in the proof exists if the two random variables X, Y are continuous, indeed another possible definition for Z can be: $Z = S_Y^{-1}(S_X(t)) = c(t)X$. By looking at this alternative writing, it is obvious that the survival functions must be strictly monotone, and therefore the random variables continuous.

Remark 1.3. The original result provided in [23] does not contain the hypothesis of continuity of the two random variables X, Y, but it was already underlined in the previous remark the importance of it for the construction of the function $c(t)$ and, thus, of the random variable Z, since it is needed the invertibility of the two survival functions. The following counterexample shows that the theorem 1.2 does not hold in the discrete case:

Let $(\Omega, \mathcal{F}, P)$ be a probability space such that $\Omega = \{\omega_1, \omega_2\}$, $\mathcal{F} = 2^{\Omega}$, P a probability measure with $P(\omega_1) = 1/10$, $P(\omega_2) = 9/10$. Let define two random variables in it with $X(\omega_1) = 1, X(\omega_2) = 0$ and $Y(\omega_1) = 0, Y(\omega_2) = 1$. X precedes Y in FSD as shown in the figure below but it is not possible to find a random variable Z with the same distribution of Y such that $Z \geq X$ a.s.

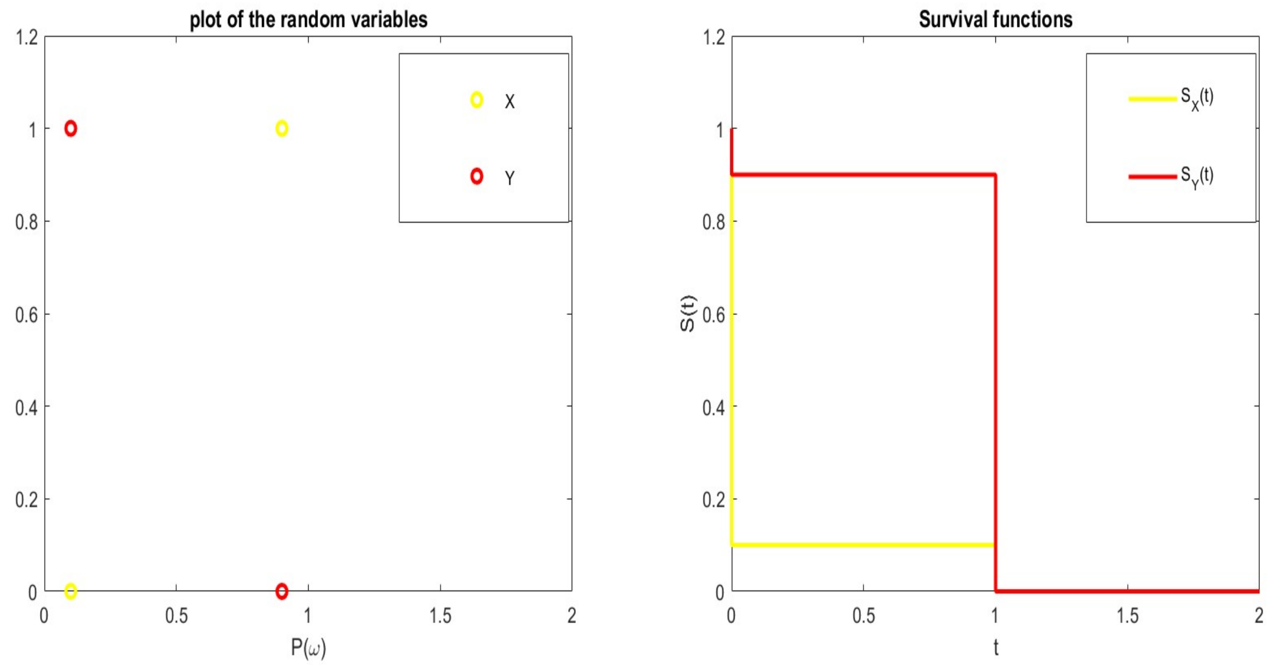

Figure 1.1: Plot of the possible realizations of X, Y and their survival functions.

As an immediate consequence of the theorem 1.2, properties of conditional state dependence and monotonicity imply that the premium functional preserves FSD, for continuous risks.(see [16]).

1.2. Catalog of premium principles

This section has the purpose of showing some premium principles and discussing their properties proving them. In particular:

- Net premium principle: $H(X) = E[X]$;
- Expected value premium principle: $H(X) = (1+\theta)E[X]$, $\theta > 0$;
- Variance premium principle: $H(X) = E[X] + \alpha Var(X)$, $\alpha > 0$;
- Standard deviation premium principle: $H(X) = E[X] + \alpha\sqrt{Var(X)}$, $\alpha > 0$;
- Wang's premium principle: $H(X) = \int_0^\infty g(S_x(t))dt$;

1.2.1. Net premium principle and expected value premium principle

These are probably the two most intuitive premium principles; the former consists just in taking as insurance price the expected value of the risk, and owns all the properties enumerated in section 1.1, the latter is built on the first one by adding a proportional risk loading, and it loses properties 6,7 and 8. First of all, it will be shown that these three ones do not hold for the expected value premium principle, then the proves of the remaining properties considered not trivial will be reported.

- Unjustified risk loading:

  $$\text{If } X = c \text{ a.s. with } c \geq 0 \text{ then } H(X) = (1+\theta)c \geq c;$$

- No maximal loss:

  It can be proven that this does not hold by many counterexamples, however, trivially, the previous reasoning is already sufficient.

- No translation equivariance:

  $$H(X+a) = (1+\theta)E[X+a] = H(X) + a + a\theta \neq H(X) + a;$$

The next proves, instead, follow for the net premium principle by considering the special case $\theta = 0$:

- Continuity:

  Let $\mathcal{I}_A$ represent the indicator function in the set A and f_x the density function of X;

  Regarding the maximum:

  $$\lim_{d\to 0+}(1+\theta)E[(X-d)_+] = \lim_{d\to 0+}(1+\theta)\int_0^{+\infty}(X-d)\mathcal{I}_{(X>d)}f_x dx = (1+\theta)E[X];$$

  Instead, for the minimum part:

  $$\begin{aligned}\lim_{d\to\infty}(1+\theta)E[\min(X,d)] &= \lim_{d\to\infty}(1+\theta)(\int_{-\infty}^{+\infty} x\mathcal{I}_{(X<d)}f_x dx + \int_{-\infty}^{+\infty} d\mathcal{I}_{(X\geq d)}dx \\ &= \lim_{d\to\infty}\int_{-\infty}^{d} x f_x dx + \int_{d}^{+\infty} d\cdot dx \\ &= (1+\theta)E[X];\end{aligned}$$

- Additivity:

  It follows by the additivity of the expected value.

- Monotonicity:

  The expected value is an operator trivially monotonic, in fact, considering the definition, if $X(\omega) \leq Y(\omega)\ \forall\omega$:

$$E[X] = \sum_{i=1}^{\infty} x_i P(x = x_i) \leq \sum_{i=1}^{\infty} y_i P(y = y_i) = E[Y].$$

  Therefore:

$$H(X) = (1+\theta)E[X] \leq (1+\theta)E[Y] = H(Y);$$

- Preserving FSD:

  It follows by theorem 1.2, since properties 1 and 2 hold.

- Preserving stop-loss ordering:

  If $X(\omega) > d$ then:

$$(1+\theta)E[X] = (1+\theta)E[(X-d)_+ + d] \leq (1+\theta)E[(Y-d)_+ + d] = H(Y);$$

  If $X(\omega) \leq d$ then:

$$(1+\theta)E[X] = (1+\theta)E[(X-d)_+ + X] \leq (1+\theta)E[(Y-d)_+ + X] \leq H(Y);$$

1.2.2. Variance and standard deviation premium principles

These two premium principles consist in adding to the net premium a risk load which is proportional to variance or standard deviation. Unfortunately, they do not have many of the properties listed in section 1.1. As done before, the proves considered too simple are omitted, in particular just the property of continuity is proved, while properties that do not hold for at least one principle are discussed more in details. It is important to notice that the following arguments proposed for variance premium principle apply also for the standard deviation one, except for property 9, which is owned only by this last principle, and property 13, owned only by variance one.

- Continuity:

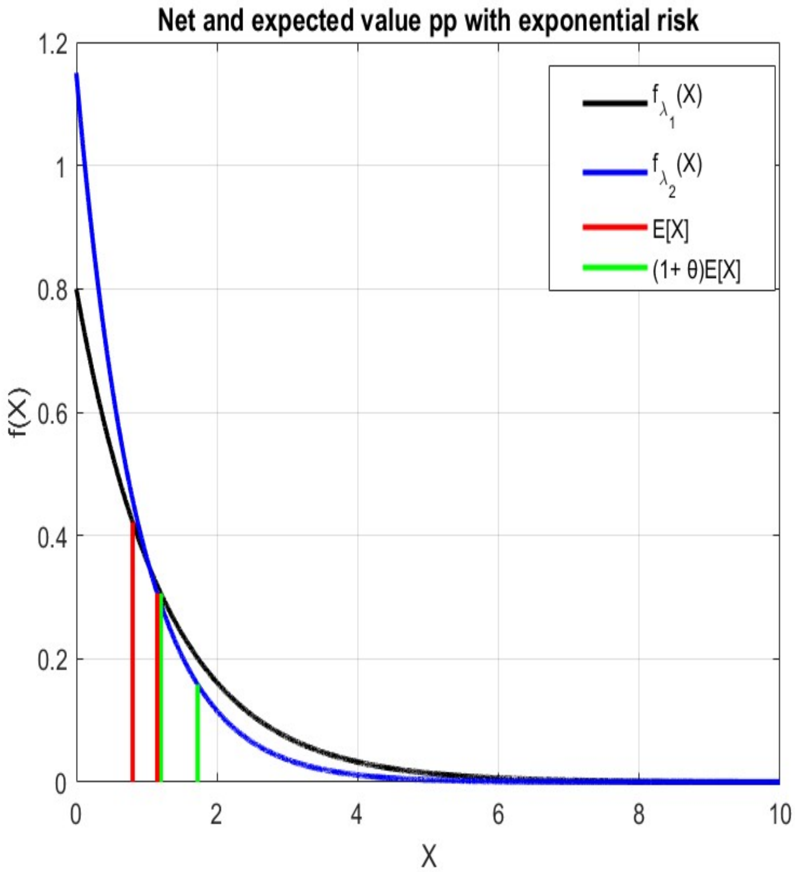

Figure 1.2: Net and expected value premium principles compared for two exponential risks with parameters $\lambda_1 = 0.8$ and $\lambda_2 = 1.15$, $\theta = 0.5$.

First of all, one can notice that, since continuity is guaranteed for the expected value, then it is sufficient to show that it is also guaranteed for the variance. However, since $Var(X) = E[X^2] - E^2[X]$, the proof consists in studying the behaviour of the term $E[X^2]$; by repeating the reasoning done for the mean, the proof can be easily completed, indeed:

$$\begin{aligned} \lim_{d \to 0+} E[(X-d)_+^2] &= \lim_{d \to 0+} \int_{-\infty}^{\infty} (x-d)^2 f_x \mathcal{I}_{(x>d)} dx \\ &= \lim_{d \to 0+} \int_{d}^{\infty} (x-d)^2 f_x dx \\ &= \int_{0}^{\infty} x^2 f_x dx = E[X^2]; \end{aligned}$$

A similar procedure can be applied to show that $\lim_{d \to \infty} E[(\min(X,d))^2] = E[X^2]$.

- No monotonicity:

  It can be proved constructing a counterexample; let consider a risk $X \sim \mathcal{B}(1/2)$ distributed as a Bernoulli and another one defined as $Y = X/2 + 1$. Then:

  $$H(X) = 1/2 + \alpha/4.$$

  $$H(Y) = 5/4 + \alpha/16.$$

  By taking, for instance, $\alpha = 5$, $H(X) > H(Y)$ even if $X(\omega) < Y(\omega)\ \forall \omega$;

- No max-loss:

  As counterexample it can be considered a risk $X \sim \mathcal{B}(1/2)$. Thus:

  $$H(X) = 1/2 + \alpha/4 \geq 1$$

  for $\alpha \geq 2$;

- No scale equivariance:

  $$H(bX) = bE[X] + b^2 Var(X) \neq bH(X);$$

  As mentioned above, this is one of the two properties that variance premium and standard deviation one have not in common.

- No additivity, sub-additivity and super-additivity:

It simply follows by the fact that variance and standard deviation are not additive, sub-additive and super-additive.

- Additivity for independent risks:

$$H(X+Y) = E[X] + E[Y] + \alpha(Var(X) + Var(Y) + 2Cov(X,Y));$$

If X, Y are independent $Cov(X,Y) = 0$ and $H(X+Y) = H(X) + H(Y)$;

Considering instead the standard deviation:

$$H(X+Y) = E[X] + E[Y] + \alpha(\sqrt{Var(X) + Var(Y)}) \neq H(X) + H(Y);$$

- No preserving FSD:

As counterexample one can take $X \sim \mathcal{B}(1/2)$ and $Y = 1$ a.s. Then:

$$H(X) = 1/2 + \alpha/4.$$

$$H(Y) = 1.$$

The property does not hold for $\alpha > 2$;

- No preserving stop-loss ordering:

The same counterexample used for denying the preserving of first stochastic dominance ordering can be applied also in this case;

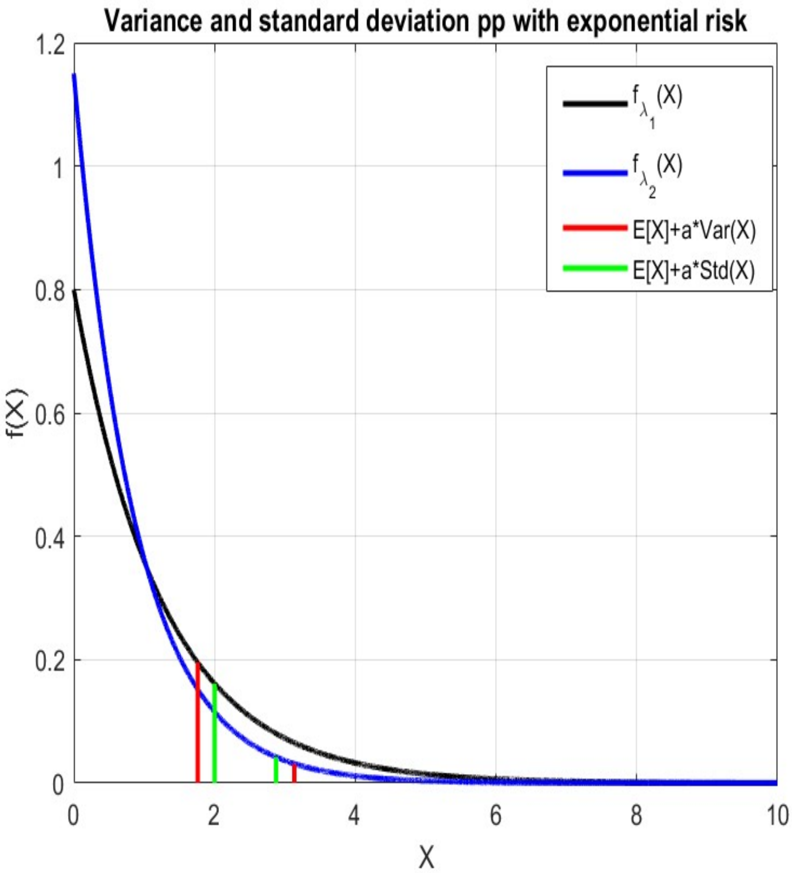

Figure 1.3: Variance and standard deviation premium principles compared for two exponential risks with parameters $\lambda_1 = 0.8$ and $\lambda_2 = 1.15$, $a = 0.8$.

1.2.3. Wang's premium principles

This part of the treatment about premiums in static context is dedicated to a brief introduction to the class of Wang's premium principles, which is well-known and widely dealt with in actuarial literature (see for instance [23], in which they are introduced). First of all it is needed a clarification about the function g in the formula of the functional H written at the beginning of the section 1.2.

Definition 1.4 (Distortion function). *A function $g : [0,1] \to [0,1]$ is called a distortion function if g is non-decreasing with $g(0) = 0$ and $g(1) = 1$.*

The main feature of Wang's class of premiums is that it is a unique representation of each premium which owns property 1-4. This is a result proved by Greco (see [9]), but here it is reported and proved in a simpler way approaching firstly risks with bounded support, and then generalizing the characterization of the class premium principles to all risks (see

[11]). Before doing this, it can be interesting see that many of the other properties listed hold for Wang's premiums; straightforwardly, by looking at the integral that define H, one can notice that properties 5,6 and 7 hold true, moreover it is easy to show also properties 8 and 9.

Theorem 1.3. *Wang's premium principle is invariant for affine transformations.*

Proof. Let $a \geq 0$ and $b \geq 0$.

Since

$$S_{aX+b}(t) = \begin{cases} 1, 0 \leq t < b \\ S_X((t-b)/a), t \geq b \end{cases}$$

then:

$$H(aX+b) = \int_0^b dt + \int_b^\infty g(S_X((t-b)/a))dt = \\ b + a\int_0^\infty g(S_X(t))dt = aH(X) + b;$$

□

Now, the main result about the characterization of Wang's class of premiums:

Theorem 1.4. *Let $H : \chi \to [0, \infty)$ be a functional such that properties from 1 to 4 hold, then $\exists!$ g distortion function concave such that $H = \int_0^\infty g(S_x(t))dt$*

Proof. Let consider a risk X with bounded support $[0, b]$ and with a decumulative distribution function piecewise constant. Thus, it exists a sequence $0 = x_0 < ... < x_n = b$ such that:

$$S_X(x) = \sum_{i=0}^{n-1} p_i \mathcal{I}_{x_i \leq x \leq x_{i+1}}$$

X can be written as:

$$X = \sum_{i=0}^{n-1} L(x_i, x_{i+1}),$$

where

$$L(x_i, x_{i+1}) = \begin{cases} 0, 0 \leq X \leq x_i \\ X - x_i, x_i < X < x_{i+1} \\ x_{i+1} - x_i, X \geq x_{i+1} \end{cases}$$

For comonotonic additivity:

$$H(X) = \sum_{i=0}^{n-1} H(L(x_i, x_{i+1})).$$

Observing that $L(x_i, x_{i+1}) \overset{d}{=} (x_{i+1} - x_i) \cdot \mathcal{B}(p_i)$, for scale invariance:

$$H(L(x_i, x_{i+1}) = (x_{i+1} - x_i) \cdot H(\mathcal{B}(p_i)) = (x_{i+1} - x_i) \cdot g(p_i).$$

Recollecting all the equations:

$$H(X) = \sum_{i=0}^{n-1} (x_{i+1} - x_i) g(p_i) = \sum_{i=0}^{n-1} \int_{x_i}^{x_{i+1}} g(S_X(t))dt = \int_0^b g(S_X(t))dt.$$

The function g, found as $g(t) = H(\mathcal{B}(t))$, has all the properties desired, and this concludes the first part of the proof. Now let X be bounded with a generic decumulative distribution function, $S_X(t)$ can be approximated as follows:

$$S_{Xn}(t) = \sum_{i=0}^{2^n-1} S_X((i+1)b/2^n)\mathcal{I}_{(ib/2^n \le t \le (i+1)b/2^n)}.$$

After some computations, one can get $|S_X^{-1}(t) - S_{X_n}^{-1}(t)| \le b/2^n$ that implies $|H(X) - H(X_n)| \le b/2^n$. Therefore, $H(X) = \lim_{n\to\infty} H(X_n)$. Applying dominated convergence theorem it follows:

$$H(X) = \int_0^b g(S_X(t))dt.$$

Arrived so far, it remains just to generalize the result for a generic unbounded risk. In order to do so, it is sufficient to notice from the last equation that the premiums for a random variable with support $[0, b]$ and the one for $\min(X, b)$ are the same. Since continuity holds for hypothesis, then the thesis follows straightforward. □

1.2.4. Table of properties

For the sake of synthesis, here there is a table resuming the principles discussed with their properties.

Properties	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Net pp	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Expected value pp	Y	Y	Y	Y	Y	N	N	N	Y	Y	Y	Y	Y	Y	Y
Variance pp	Y	N	N	Y	Y	Y	N	Y	N	N	N	N	Y	N	N
Std pp	Y	N	N	Y	Y	Y	N	Y	Y	N	N	N	N	N	N
Wang pp	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	N	N	Y	Y

1.3. Decomposition of the premium functional into risk and deviation measures

This section presents a central result, found by Nendel, Riedel and Schmeck (see [19] for further details), which clarifies the structure of premium principles providing a decomposition valid in a very general setting. First of all, it is necessary to recall the concepts of monetary risk measure and deviation measure, commonly used in financial mathematics. Let $\chi_b \subseteq \chi$ the set of non-negative and bounded measurable random variables.

Definition 1.5 (Monetary risk measure). *A functional $\rho : \chi_b \to \mathbb{R}$ is a monetary risk measure if it is:*

1. *Normed; ($\rho(0) = 0$)*
2. *Translation invariant;*
3. *Monotone.*

Definition 1.6 (Deviation measure). *A functional $\rho : \chi_b \to [0, +\infty]$ is a deviation measure if :*

1. *It is normed;*
2. *$\rho(X + m) = \rho(X)$, $\forall m \in \mathbb{R}$, $\forall X \in \chi_b$;*

With these two notions, it can be proved that a premium principle, normed and translation invariant, can be decomposed into a monetary risk measure, which catches all the risky components of the claim, and a deviation measure, which catches in some sense the possible fluctuation of the insured loss, and therefore represents the part of the premium that cannot be explained by any risk measure. Another possible interpretation is that the first term of the decomposition is a net premium, while the second one is a safety loading proportional to the variability of the risk.

Lemma 1.1. *Let $H : \chi_b \to [0, +\infty]$ be a premium principle normed and translation* *invariant. Let define:*

$$R_{max}(X) = \inf\left(H(X_0) | X_0 \in \chi_b, X_0 \geq 0\right) \tag{1.4}$$

$$D_{min}(X) = H(X) - R_{max}(X); \tag{1.5}$$

Then R_{max} is a well-defined monetary risk measure with $R_{max} \leq H(X)$ and D_{min} is a well-defined deviation measure, such that:

$$H(X) = R_{max}(X) + D_{min}(X). \tag{1.6}$$

Moreover, for every decomposition of the form $H(X) = R(X) + D(X)$ in which R and D are respectively a risk and deviation measure, $R \leq R_{max}(X)$ and $D(X) \geq D_{min}(X)$.

Proof. $R_{max} : \chi_b \to \mathbb{R}$ is well-defined, in fact $\sup(X) \in \chi_b$ and $H(X_0) \geq H(\inf(X))$ $\forall X_0 \in \chi_b$. Moreover $R_{max}(X) \leq H(X_0)$ by definition, so $D_{min}(X)$ is non-negative. Showing that R_{max} is effectively a monetary risk measure the proof is complete. Since the premium principle is normed by hypothesis, it follows that also the two measures are normed. About monotonicity, let $Y, Y_0 \in \chi_b$, $Y_0 \geq Y \geq X$. Then, by definition, $R_{max}(X) \leq H(Y_0)$ and taking the infimum over all Y_0 defined as before, it holds that $R_{max}(X) \leq R_{max}(Y)$. Lastly, it remains translation invariant property; let $X \in \chi_b$, $m \in \mathbb{R}$ and $X_0 \in \chi_b$ with $X_0 \geq X$. Then:

$$R_{max}(X+m) \leq H(X_0+m) = H(X_0) + m.$$

Taking the infimum over all X_0 implies that $R_{max}(X+m) \leq R_{max}(X)+m$. On the other hand:

$$R_{max}(X) + m = R_{max}(X + m - m) + m \leq R_{max}(X+m).$$

Eventually, the second property of the deviation measure follows straightforward from the translation invariance of the risk measure.

Moreover, if $R : \chi_b \to \mathbb{R}$ is a risk measure with $R(X) \leq H(X)$ $\forall X \in \chi_b$, then,

$$R(X) \leq R(X_0) \leq H(X_0) \ \forall X \in \chi_b, \ X_0 \in \chi_b, \ X_0 \geq X.$$

By taking the infimum over all X_0, it follows $R(X) \leq R_{max}(X)$ $\forall X \in \chi_b$. □

From this lemma it derives trivially that many different decompositions are allowed for the same premium, by considering an $R(X) \leq R_{max}(X)$ and consequently a $D(X) \geq D_{min}(X)$. Furthermore, the next theorem can be easily deduced:

Theorem 1.5. *H is a premium principle normed and translation invariant $\Longleftrightarrow H(X) = R(X) + D(X)$ such that $R(X)$ is a well-defined risk measure and $D(X)$ is a well-defined deviation measure.*

Proof. The result follows from lemma 1.1 by choosing $R(X) = R_{max}(X)$ and $D(X) = D_{min}(X)$. □

The principles listed in the previous section can be viewed under the new light given by theorem 1.5, but before it may be useful to make an observation about the property of monotonicity of H, and how it is related to the decomposition. Indeed, by adding it to the hypothesis of the theorem, it follows that H can be interpreted as a monetary risk measure, thus the decomposition trivially becomes an identity with $D_{min}(X) = 0$. This observation makes easier to identify the deviation measure and the risk one in the different principles; indeed, the expected value and Wang's principles are monetary risk measures as well, and they have null deviation part, while the variance and standard deviation principles, which are not monotonic, have the first term, namely the mean of the risk, as $R_{max}(X)$, and the second term as $D_{min}(X)$.

2 | Theoretical elements on jump processes

A jump process is a stochastic process which presents discrete movements. On the opposite, a diffusive process is characterized by continuous trajectories. Jump-based models allow to treat many problems where diffusion does not fit well, such as the realization of a set of claims in an insurance portfolio in a given time interval. Because of the different nature of the two processes, many calculus tools developed for one of them cannot be applied for the other and vice versa. Therefore, a theoretical background specific to each one of the aforementioned processes is needed for a good comprehension of both.
The whole chapter has the purpose of recalling many fundamental mathematical concepts for the future treatment of premium principles in dynamical context. All the presented topics are well-known and widely developed in literature, and, for more details, the following textbooks are suggested: [3], for a detailed review of stochastic calculus in a general diffusive framework, which is omitted in this work, and [5], which has been used as reference for the chapter.

2.1. Brief review of stochastic calculus

Let fix a probability space $(\Omega, \mathcal{F}, P)$. Let define in this space a filtration $\mathcal{F}_t$, namely an increasing family of sub-σ-algebra of $\mathcal{F}$: $\mathcal{F}_s \subset \mathcal{F}_t$ whenever $s \leq t$.

Definition 2.1 (Stopping time). *Let $(\mathcal{F}_t)_{t\in T}$ be a filtration. A random variable $\tau : \Omega \to T \cup \{+\infty\}$ is a stopping time if $\forall t \in T$, $\{\tau \leq t\} \in \mathcal{F}_t$.*

Definition 2.2 (Martingale). *A real-valued process $X = (\Omega, \mathcal{F}, \mathcal{F}_t, X_t, P)$ is a martingale if:*

$$X_t \text{ is } P\text{-integrable } \forall t \in T \ (X_t \in L^1(\Omega, \mathcal{F}, P)); \tag{2.1}$$

$$E[X_t|\mathcal{F}_s] = X_s, \ \forall 0 \leq s \leq t; \tag{2.2}$$

Definition 2.3 (Brownian motion or Wiener process)**.** *A real-valued process* $W = (\Omega, \mathcal{F}, \mathcal{F}_t, W_t, P)$ *is a Brownian motion or Wiener process if:*

$$W_0 = 0; \tag{2.3}$$

$$W_t - W_s \text{ is independent of } \mathcal{F}_s,\ \forall 0 \leq s \leq t; \tag{2.4}$$

$$W_t - W_s \sim N(0, t-s),\ \forall 0 \leq s \leq t; \tag{2.5}$$

Remark 2.1. The Brownian motion is a martingale, indeed it is P-integrable $\forall t \in T$ since it has gaussian law, and it holds:

$$E[W_t|\mathcal{F}_s] = E[W_t - W_s + W_s|\mathcal{F}_s] = E[W_t - W_s|\mathcal{F}_s] + W_s = W_s;$$

Definition 2.4 (Local martingale)**.** *A process* $X = (\Omega, \mathcal{F}, \mathcal{F}_t, X_t, P)$ *is a local martingale if there exists an increasing sequence* $(\tau_n)_{n\geq 1}$ *of stopping times such that:*

1. $\tau_n \to +\infty$ *as* $n \to \infty$ *a.s.;*
2. $(X_{t\wedge\tau_n})_t$ *is a* $\mathcal{F}_t$*-martingale for every* n*;*

Definition 2.5 (Predictability)**.** *Let* $\mathcal{F}_t$ *be a filtration on* $(\Omega, \mathcal{F}, P)$ *and define* $\mathcal{P}(\mathcal{F}_t)$ *to be the* σ*-field over* $[0, \infty] \times \Omega$ *generated by the rectangles of the form:*

$$(s, t] \times A \text{ with } 0 \leq s \leq t,\ \ A \in \mathcal{F}_s.$$

$\mathcal{P}(\mathcal{F}_t)$ *is called the* $\mathcal{F}_t$*-predictable* σ*-field over* $[0, \infty] \times \Omega$*. A real-valued process* X *such that* X_0 *is* $\mathcal{F}_0$*-measurable and the mapping* $(t, \omega) \to X_t(\omega)$ *defined from* $[0, \infty] \times \Omega$ *into* $\mathbb{R}$ *is* $\mathcal{P}(\mathcal{F}_t)$*-measurable is said to be* $\mathcal{F}_t$*-predictable.*

2.2. Point processes

The concept of point process is crucial in order to model phenomena subjected to jumps. Stochastic processes whose realizations consist of point distributed along time can be viewed in different ways, here they will be introduced via their associated counting process.

Definition 2.6 (Point process)**.** *A realization of a point process over* $[0, \infty)$ *can be described by a sequence* $\{T_n\}_{n\geq 1}$ *of random variables, which take values in* $[0, \infty)$*, and defined* *on a probability space* $(\Omega, \mathcal{F}, P)$, *such that:*

$$T_0 = 0,\ T_n < \infty,\ T_n < T_{n+1};$$

To each realization of T_n *corresponds a counting function* N_t *defined by:*

$$N_t = \begin{cases} n, & t \in [T_n, T_{n+1}], n \geq 0 \\ \infty, & t \geq T_\infty \end{cases} \tag{2.6}$$

Therefore, N_t *is a right-continuous step function such that* $N_0 = 0$ *and its jumps are upward and of magnitude 1.*

Remark 2.2. The family of counting function N_t is the counting process associated to the point process. Sometimes, N_t is also called point process. Moreover, it follows by the definition that the point processes belong to the class of "cadlag" processes (continue à droite, limitée à gauche), i.e. processes which are both right-continuous and left-limited.

Definition 2.7 (Non-explosivity). *The realization of a point process described by a sequence* T_n *in* $[0, \infty)$ *is non-explosive if and only if:*

$$T_\infty = \lim_{n \to \infty} T_n = +\infty. \tag{2.7}$$

The next definition introduces a point process that can be considered as the building block of stochastic jump framework, as well as the Brownian motion for the stochastic diffusive one.

Definition 2.8 (Poisson Process). *A* $\mathbb{N}$*-valued process* $N = (\Omega, \mathcal{F}, \mathcal{F}_t, N_t, P)$ *is a Poisson process with intensity* $\lambda > 0$ *if:*

$$N_0 = 0; \tag{2.8}$$

$$N_t - N_s \text{ is independent of } \mathcal{F}_s,\ 0 \leq s \leq t; \tag{2.9}$$

$$N_t - N_s \sim Poiss(\lambda(t-s)),\ 0 \leq s \leq t; \tag{2.10}$$

Remark 2.3. The condition (2.10) can be explicitely rewritten as:

$$P(N_t - N_s = n) = \frac{e^{\lambda(t-s)}(\lambda(t-s))}{n!}$$

Remark 2.4. It is possible to characterize the Poisson process in many different ways, the definition above was chosen among the others because it returns immediately the distribution of the process and the similarity with the definition of Brownian motion. However, for completeness, it is reported the following lemma, which provides an alternative characterization of the Poisson process, in which it is built with a sequence of random variables according to the definition 2.6.

Lemma 2.1. *Let $(\tau_i)_i$ be a sequence of random variables such that $\tau_i \sim \mathcal{E}(\lambda)$. Let $T_n = \sum_{i=1}^{n} \tau_i$ and define N_t as $N_t = \sum_{n=1}^{\infty} \mathcal{I}_{(t \geq T_n)}$. Then, N_t is a Poisson process with intensity λ.*

Proof. The property (2.8) is trivially verified. If one assumes that each increment of the form $N_t - N_s$ is distributed as a Poisson with parameter $\lambda(t-s)$, the property (2.9) holds because the Poisson distribution is stable under convolution, therefore, since $N_t - N_h = (N_t - N_s) + (N_s - N_h)\ \forall\ h \leq s \leq t$, the increments $(N_t - N_s) + (N_s - N_h)$ must be independent. Thus, proving that N_t is Poisson distributed, the statement holds. Let consider that:

$$P(N_t = n) = P(T_{n+1} > t) - P(T_n \geq t);$$

But:

$$P(T_{n+1} > t) = 1 - P(T_{n+1} < t) = 1 - \int_0^t \lambda \frac{e^{-\lambda s}(\lambda s)^n}{n!} ds.$$

The density function of T_{n+1} is known for a standard result of probability, which states that a sum of n exponential random variables independent and identically distributed defines a gamma random variable with λ as shape parameter and n as rate parameter. Integrating by parts:

$$\begin{aligned} P(T_{n+1} > t) &= 1 + \frac{e^{-\lambda t}(\lambda t)^n}{n!} - \int_0^t \frac{e^{-\lambda s}(\lambda s)^{n-1}}{n-1!} ds \\ &= 1 + \frac{e^{-\lambda t}(\lambda t)^n}{n!} - P(T_n < t) \\ &= \frac{e^{-\lambda t}(\lambda t)^n}{n!} + P(T_n \geq t). \end{aligned}$$

Therefore:

$$P(T_{n+1} > t) - P(T_n \geq t) = \frac{e^{-\lambda t}(\lambda t)^n}{n!}$$

which is the distribution of a $Poiss(\lambda t)$;

□

By the definition 2.8 and the lemma 2.1, many properties can be deduced straightforwardly with simple computations such as the moments of the Poisson process, its characteristic function, and its stationarity. In particular, by observing its mean, it is easy to notice that this process does not enjoy the martingale property. However, it can be built a martingale from any Poisson process by simply subtracting its expected value. The result of this operation is called "Compensated Poisson process".

Lemma 2.2. *The compensated version of a Poisson process $\mathcal{F}_t$-adapted with intensity λ defined as $\widetilde{N}_t = N_t - \lambda t$ is a $\mathcal{F}_t$-martingale.*

Proof. Let be $t > s$; Since:

$$E[N_t|\mathcal{F}_s] = E[N_t - N_s + N_s|\mathcal{F}_s] = E[N_t - N_s|\mathcal{F}_s] + N_s = \lambda(t-s) + N_s;$$

Then:

$$E[\widetilde{N}_t|\mathcal{F}_s] = E[N_t - \lambda t|\mathcal{F}_s] = \lambda(t-s) + N_s - \lambda t = N_s - \lambda s = \widetilde{N}_s;$$

□

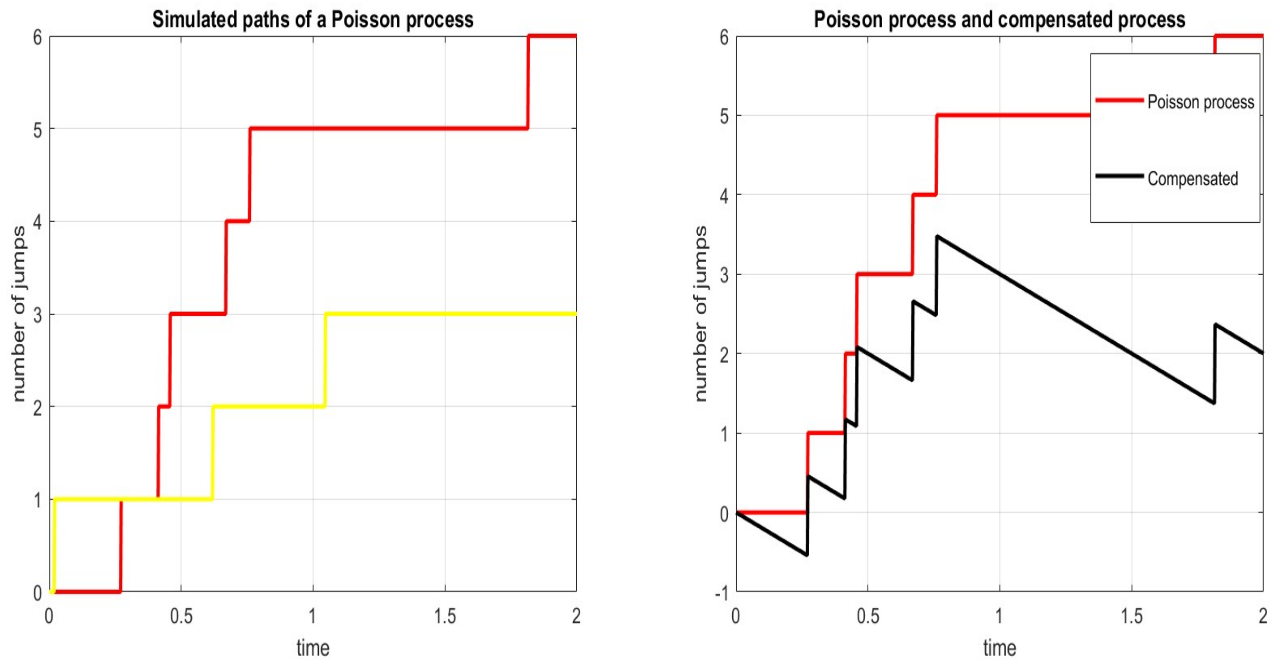

Figure 2.1: Two possible sample paths of a Poisson process with intensity equal to 2, and the paths of a compensated Poisson compared to its related process;

Up to this point, it was dealt with only the case in which the intensity is deterministic. In particular, a point process can be a Poisson one only if its intensity is constant and, obviously, the other conditions of the definition 2.8 are satisfied. However, for a general point process, it can be introduced another source of randomness by considering as intensity another stochastic process which respects the following definition:

Definition 2.9 (Stochastic intensity). *A point process N_t admits stochastic intensity λ_t if, $\forall$ C_t non-negative and $\mathcal{F}_t$-predictable:*

$$\lambda_t \geq 0 \ \forall t > 0; \tag{2.11}$$

$$\int_0^t \lambda_s ds < \infty, \ P - a.s.; \tag{2.12}$$

$$E\left[\int_0^\infty C_s dN_s\right] = E\left[\int_0^\infty C_s \lambda_s ds\right]; \tag{2.13}$$

Remark 2.5. The integral with respect to dN_s is defined as:

$$\int_0^t C_s dN_s = \sum_{n \geq 1} C_{T_n} \mathcal{I}_{(Tn \leq t)} \tag{2.14}$$

and

$$\int_0^\infty C_s dN_s = \sum_{n\geq 1} C_{T_n} \mathcal{I}_{(Tn<\infty)} \tag{2.15}$$

Remark 2.6. The stochastic intensity is always referred to a probability measure and a filtration. It is in fact correct to say that λ_t is the $(\mathcal{F}_t, P)$-intensity of the process N_t. In order to ease the notation, however, this complete writing is omitted.

It is possible to find a bound between a general point process and the concept of martingale even under the framework of stochastic intensity, as done before with the compensator of the Poisson process. In order to present this result, it is necessary to introduce the following preliminary lemma (contained in [5], appendix A1, theorem 4):

Lemma 2.3 (Verification theorem)**.** *Let Ω be a set and $\mathcal{S}$ a π-system (stable by finite intersection) on Ω. Let $\mathcal{H}$ be a vector space of real-valued functions such that:*

- $1 \in \mathcal{H}$ *and* $\mathcal{I}_A \in \mathcal{H}$ *whenever* $A \in \mathcal{S}$*;*
- *if X_n is an increasing sequence of non-negative functions of $\mathcal{H}$ such that X_n is bounded, then* $\sup X_n \in \mathcal{H}$*;*

Then $\mathcal{H}$ contains all real-valued and bounded mappings measurable with respect to $\sigma(\mathcal{S})$.

Proof. Define $\mathcal{D} = \{A|\mathcal{I}_A \in \mathcal{H}\}$. By the first hypothesis $\mathcal{S} \subset \mathcal{D}$. Also $\mathcal{D}$ owns the following three properties: $\Omega \in \mathcal{D}$, it is stable under strict difference because $\mathcal{H}$ is a vector space and it is stable by increasing sequential limit, namely, if $A_n \to A$ and $A_n \in \mathcal{D}$ then $A \in \mathcal{D}$. Therefore, by monotone convergence theorem, $\sigma(\mathcal{S}) \subset \mathcal{D}$. Thus, $\mathcal{H}$ contains all the mappings $(\mathcal{I}_A, A \in \sigma(\mathcal{S}))$. Let X be a real-valued $\sigma(\mathcal{S})$-measurable r.v. Then, X^+ and X^- are increasing limits of real-valued step r.v. with respect to $\sigma(\mathcal{S})$. They belong to $\mathcal{H}$ because $\mathcal{I}_A \in \mathcal{H}$ $\forall A$ and $\mathcal{H}$ is a vector space. Proof is concluded using the second hypothesis. □

This important lemma provides a powerful tool in order to handle the condition (2.13). Indeed, proving that it holds for a choice $C_t = \mathcal{I}_A \mathcal{I}_{(r,t]}$ $\forall A \in \mathcal{F}_r$ and $r < t$, namely for the generator class of the predictable and non-negative processes, it is proved for a generic C_s which owns these properties. Now there are all the theoretical elements for demonstrating the next two results.

Theorem 2.1. *If a point process N_t admits the $\mathcal{F}_t$-intensity λ_t, then N_t is non-explosive and the process $M_t = N_t - \int_0^t \lambda_s ds$ is an $\mathcal{F}_t$-local martingale.*

Proof. Let define a sequence of stopping times $S_n = \inf(t| \int_0^t \lambda_s ds \geq n)$. Writing (2.13) with $C_t = \mathcal{I}_{(t \leq S_n)}$, one obtains:

$$E[N_{S_n}] = E\left[\int_0^{S_n} \lambda_s ds\right] \leq n < \infty \ \forall n.$$

Therefore, $N_{S_n} < \infty$ P-a.s., which implies the non-explosivity since $S_n \to \infty$ when $n \to \infty$.

It remains to prove the local martingality. Let choose $C_t = \mathcal{I}_A \mathcal{I}_{(t \leq T_n)}$ $\forall A \in \mathcal{F}_t$ arbitrary, in which T_n is a sequence of stopping times as in definition 2.4. It holds:

$$E\left[\int_0^\infty \mathcal{I}_A \mathcal{I}_{(t\leq T_n)} dN_s\right] = E\left[\int_0^\infty \mathcal{I}_A \mathcal{I}_{(t\leq T_n)} \lambda_s ds\right]$$
$$\Longrightarrow E[\mathcal{I}_A(N_{t\wedge T_n})] = E\left[\mathcal{I}_A \int_0^{t\wedge T_n} \lambda_s ds\right]$$

Since $N_{t\wedge T_n} \leq n < \infty$, integrability is checked. By taking $r < t$, it is possible to write, applying, as done before, the condition (2.13) and an opportune choice of C_t, namely $\mathcal{I}_A(\mathcal{I}_{(t\leq T_n)} - \mathcal{I}_{(r\leq T_n)})$:

$$E\left[\mathcal{I}_A(N_{t\wedge T_n} - \int_0^{t\wedge T_n} \lambda_s ds)\right] = E\left[\mathcal{I}_A(N_{r\wedge T_n} - \int_0^{r\wedge T_n} \lambda_s ds\right]$$

By moving the term on the right to the left, for linearity of the expected value and the arbitrariness of A the second condition of martingality is verified, indeed the terms in the last equation can be recollected in the form:

$$E[\mathcal{I}_A(M_{t\wedge T_n} - M_{r\wedge T_n})] = 0 \ for \ r < t;$$

This concludes the proof.

□

Theorem 2.1 shows a relation between the stochastic intensity and the possibility of constructing a martingale starting from N_t. It is natural to ask if also the opposite implication holds, ensuring that if such a local martingale exists, then λ_t is the stochastic intensity of N_t, deducing, thus, a martingale characterization of it. This result is guaranteed by the following theorem:

Theorem 2.2. *Let N_t be a non-explosive point process $\mathcal{F}_t$-adapted, and suppose that $M_t = N_t - \int_0^t \lambda_s ds$ is a local martingale, with λ_t non-negative and $\mathcal{F}_t$-progressively measurable. Then, λ_t is the intensity of N_t.*

Proof. The condition (2.11) is guaranteed by the hypothesis, while (2.12) follows by choosing $C_s = \mathcal{I}_{(0,t]}$, therefore it is sufficient to prove (2.13). Indeed, if it holds, one can write:

$$E\left[\int_0^t dN_s\right] = E\left[\int_0^t \lambda_s ds\right]$$
$$\implies E\left[\int_0^t \lambda_s ds\right] = N_t < \infty$$

for the hypothesis of non-explosivity.

By taking an arbitrary $A \in \mathcal{F}_r$, and the sequence T_n of stopping times for which M_t is a local martingale, it holds:

$$E[\mathcal{I}_A(M_{t\wedge T_n} - M_{r\wedge T_n})] = 0 \textit{ for } r < t;$$

Making explicit the terms in function of the point process N_t and λ_t, one can get:

$$E[\mathcal{I}_A(N_{t\wedge T_n} - N_{r\wedge T_n})] = E\left[\mathcal{I}_A \int_{r\wedge T_n}^{t\wedge T_n} \lambda_s ds\right].$$

By letting T_n go to infinity, this writing is equivalent to the desired equation by taking $C_s = \mathcal{I}_A\mathcal{I}_{(r,t]}$, in fact it can be rewritten as:

$$E\left[\int_0^\infty \mathcal{I}_A\mathcal{I}_{(r,t]} dN_s\right] = E\left[\int_0^\infty \mathcal{I}_A\mathcal{I}_{(r,t]}\lambda_s ds\right]$$

By applying the verification theorem, the result can be extended to a generic process C_s predictable and non-negative, hence, λ_t is the stochastic intensity of N_t.

□

After having a characterization for the stochastic intensity, it is natural looking for some result about its uniqueness. The following theorem provides it and adds useful properties regarding the stochastic intensity stopped at a jump time, shown in the corollaries.

Theorem 2.3. *Let N_t be a point process $\mathcal{F}_t$-adapted, let λ_t and $\widetilde{\lambda}_t$ be both two intensities of N_t such that they are $\mathcal{F}_t$-predictable. Then:*

$$\lambda_t(\omega) = \widetilde{\lambda}_t(\omega),\ dPdN_t - a.e. \tag{2.16}$$

Proof. Let consider (2.13) with $C_s = \mathcal{I}_{(\lambda_s > \widetilde{\lambda}_s)} \mathcal{I}_{(s \le t)}$. The chosen C_s is $\mathcal{F}_t$-predictable because the two intensities are $\mathcal{F}_t$-predictable. Thus:

$$E\left[\int_0^t C_s dN_s\right] = E\left[\int_0^t C_s \lambda_s ds\right] = E\left[\int_0^t C_s \widetilde{\lambda}_s ds\right] \Longrightarrow$$
$$E\left[\int_0^t \mathcal{I}_{(\lambda_s > \widetilde{\lambda}_s)} \mathcal{I}_{(s \le t)} \lambda_s ds\right] = E\left[\int_0^t \mathcal{I}_{(\lambda_s > \widetilde{\lambda}_s)} \mathcal{I}_{(s \le t)} \widetilde{\lambda}_s ds\right] \Longrightarrow$$
$$E\left[\int_0^t \mathcal{I}_{(\lambda_s > \widetilde{\lambda}_s)} \lambda_s ds\right] = E\left[\int_0^t \mathcal{I}_{(\lambda_s > \widetilde{\lambda}_s)} \widetilde{\lambda}_s ds\right]$$

By looking at the last equation, it is easy to notice that it holds only if $\mathcal{I}_{(\lambda_s > \widetilde{\lambda}_s)} = 0$ $dPdN_t$-a.e. The same reasoning can be applied similarly for $\mathcal{I}_{(\lambda_s < \widetilde{\lambda}_s)}$, concluding the proof.

□

Corollary 2.1. *Let τ_n be a jump time of the point process N_t, let λ_t and $\widetilde{\lambda}_t$ be both two intensities of N_t such that they are $\mathcal{F}_t$-predictable. Then:*

$$\lambda_{\tau_n}(\omega) = \widetilde{\lambda}_{\tau_n}(\omega),\ P - a.s.,\ for\ n \ge 1.$$

Proof. Let be $C_t = \mathcal{I}_{(\lambda_{\tau n} \neq \widetilde{\lambda}_{\tau n})}$. Then:

$$\begin{aligned} 0 &= E\left[\int_0^\infty C_t dN_t\right] \\ &= E\left[\sum_{n \ge 1} C_{\tau_n} \mathcal{I}_{(\tau_n < \infty)}\right] \\ &= E\left[\sum_{n \ge 1} \mathcal{I}_{(\lambda_{\tau n} \neq \widetilde{\lambda}_{\tau n})} \mathcal{I}_{(\tau_n < \infty)}\right] \\ &= \sum_{n \ge 1} P(\lambda_{(\tau n \neq \widetilde{\lambda}_{\tau n})}; \tau_n < \infty); \end{aligned}$$

The last term shows that the equation holds if and only if $P(\lambda_{\tau_n} \neq \widetilde{\lambda}_{\tau_n}) = 0$;

□

Corollary 2.2. *Let τ_n be a jump time of the point process N_t, let λ_t an intensity of N_t such that it is $\mathcal{F}_t$-predictable. Then:*

$$\lambda_{\tau_n}(\omega) > 0 \; P - a.s., \; for \; n \geq 1.$$

Proof. By the fact that λ_t is the stochastic intensity of N_t, considering $C_t = \mathcal{I}_{(\lambda_t=0)}\mathcal{I}_{[\tau_{n-1},\tau_n]}$:

$$\begin{aligned}
&E[\int_0^\infty C_s dN_s] = E[\int_0^\infty C_s \lambda_s ds] \\
\Longrightarrow &E[\int_0^\infty \mathcal{I}_{(\lambda_t=0)}\mathcal{I}_{[\tau_{n-1},\tau_n]} dN_s] = E[\int_0^\infty \mathcal{I}_{(\lambda_t=0)}\mathcal{I}_{[\tau_{n-1},\tau_n]}\lambda_s ds] \\
\Longrightarrow &E[\mathcal{I}_{(\lambda_t=0)}] = E[\int_{\tau_{n-1}}^{\tau_n} \mathcal{I}_{(\lambda_t=0)}\lambda_s ds \\
\Longrightarrow &P(\lambda_t = 0) = 0,
\end{aligned}$$

because the term in the last integral must be null.

Since λ_t is non-negative, the proof is complete. □

Eventually, in order to finish the review of the stochastic intensity properties, the next theorem shows the form of the stochastic intensity for a given process adapted with respect to a smaller filtration of the given one.

Theorem 2.4. *Let $\mathcal{G}_t$ and $\mathcal{F}_t$ two filtrations such that $\mathcal{G}_t \subset \mathcal{F}_t \; \forall t \geq 0$. Then, if λ_t is a $\mathcal{F}_t$-intensity of the point process N_t and $E[\lambda_t|\mathcal{G}_t]$ is $\mathcal{G}_t$-progressively measurable, $E[\lambda_t|\mathcal{G}_t]$ is a $\mathcal{G}_t$-intensity for N_t.*

Proof. The hypothesis which states that λ_t is a stochastic intensity allows to write, for a generic C_s $\mathcal{F}_t$-predictable:

$$\begin{aligned}
E\Bigg[\int_0^\infty C_s dN_s\Bigg] &= E\Bigg[\int_0^\infty C_s \lambda_s ds\Bigg] \\
&= \int_0^\infty E\Bigg[C_s \lambda_s\Bigg] ds \\
&= \int_0^\infty E\Bigg[C_s E[\lambda_s|\mathcal{G}_s\Bigg] ds \\
&= E\Bigg[\int_0^\infty C_s E[\lambda_s|\mathcal{G}_s]\Bigg],
\end{aligned}$$

in which it was applied the Fubini's theorem.

The condition (2.13) for $E[\lambda_t|\mathcal{G}_t]$ was shown. It remains to prove the (2.12), since non-negativity is trivial:

$$\infty > E\Big[\int_0^t \lambda_s ds\Big]$$
$$= \int_0^t E[\lambda_s] ds$$
$$= \int_0^t E\Big[E[\lambda_s|\mathcal{G}_s]\Big] ds$$
$$= E\Big[\int_0^t E[\lambda_s|\mathcal{G}_s] ds\Big]$$

□

Remark 2.7. This proof is valid by using the extra hypothesis which guarantees that $E[\lambda_t|\mathcal{G}_t]$ is $\mathcal{G}_t$-progressively measurable. It exists a more technical result from a mathematical point of view, which avoids the use of this condition (see [5], chapter 2, theorem 14), however, for practical purposes, the issue can be ignored by taking for granted the $\mathcal{G}_t$-progressive measurability of the $\mathcal{G}_t$-intensity, that is ensured in many models.

2.3. Marked point processes

The theory developed in section 2.2 allows to model the occurrence of a jump at a certain time, it remains to introduce the part which models the random width of it.

Definition 2.10 (Marked point process)**.** *Let $(\Omega, \mathcal{F}, P)$ be a probability space and consider a measurable space $(E, \mathcal{E})$. Defining:*

1. *a point process T_n (or N_t) on $(\Omega, \mathcal{F}, P)$;*
2. *a sequence $(Z_n, n \geq 1)$ of random variables which takes values in $(E, \mathcal{E})$.*

The double sequence $(T_n, Z_n, n \geq 1)$ is called E-marked point process and $(E, \mathcal{E})$ is called mark space.

Remark 2.8. By taking $Z_n = 1$ a.s. in the definition 2.10, the result will be a counting process.

Remark 2.9. It is possible to consider, with the usual definition, a counting process for each measurable subset of E. For instance, with $A \subset E$:

$N_t(A) = \sum_{n\geq 1} \mathcal{I}_{(Z_n \in A)} \mathcal{I}_{(T_n \leq t)}$.

Since this is basically equivalent to define a new measure, another possible notation is:

$N_t(A) = \mu(\omega, [0,t] \times A)$.

In order to avoid a too heavy notation, the term ω is omitted, and it is preferred the form $\mu(dt \times dz)$.

The following definition presents a particular marked point process that, for its simplicity, is relevant in many applications.

Definition 2.11 (Compound Poisson process). *Let N_t be a Poisson process with intensity λ and Z_n a sequence of random variables i.i.d and independent with respect to N_t, such that $Z_n \sim Z$ $\forall n$. Then, the process $\sum_{n=1}^{N_t} Z_n$ is called compound Poisson process.*

Definition 2.12 (Kernel of the intensity). *Let $\mu(dt \times dz)$ be an E-marked point process with a filtration $\mathcal{F}_t$. If, $\forall A \in \mathcal{E}$, N_t admits $\lambda_t(z)$ as intensity, then $\mu(dt \times dz)$ is the $(P, \mathcal{F}_t)$-intensity kernel $\lambda_t(dz)$.*

Definition 2.13 (E-indexed process). *Any mapping $H : (0, \infty) \times \Omega \times E \to \mathbb{R}$ which is $\mathcal{F}_t$-predictable is called $\mathcal{F}_t$-predictable process indexed by E.*

Remark 2.10. The class of the E-indexed processes is generated by the following mappings $H(t, \omega, z) = C_t(\omega)\mathcal{I}_A(z)$, where C_t is a $\mathcal{F}_t$-predictable process and $A \in \mathcal{E}$. Many results are proved for the generators and then extended to the all class, as done before with the class of predictable processes. Moreover, it is usually used the smaller notation $H(t, z)$.

Remark 2.11. The integral with respect to $dN_s(A)$, $A \in \mathcal{E}$, is defined as:

$$\int_0^\infty \int_A C_s \mu(ds \times dz) = \int_0^\infty C_s dN_s(A) = \int_0^\infty C_s \lambda_s(A) ds. \tag{2.17}$$

Theorem 2.5 (Projection theorem). *Let $\mu(dt \times dz)$ be an E-marked point process with the $(P, \mathcal{F}_t)$-intensity kernel $\lambda_t(dz)$. Then, for each non-negative $\mathcal{F}_t$-predictable E-marked process H:*

$$E\left[\int_0^\infty \int_E H(s,z)\mu(ds \times dz)\right] = E\left[\int_0^\infty \int_E H(s,z)\lambda_s(dz)ds\right] \tag{2.18}$$

Proof. The proof is done on the class of generators of $H(t, z)$ and then the result is straightforwardly extended to the all class of the E-indexed processes.

Let be $H(t,z) = \mathcal{I}_A(z)C_t,\ A \in \mathcal{E}$.

$$E\bigg[\int_0^\infty \int_E H(s,z)\mu(ds \times dz)\bigg] = E\bigg[\int_0^\infty \int_A C_s\mu(ds \times dz)\bigg] = E\bigg[\int_0^\infty C_s\lambda_s(A)ds\bigg]$$

In the last equation it has been used the definition (2.17). Now the following computations leads to the statement:

$$E\bigg[\int_0^\infty C_s\lambda_s(A)ds\bigg] = E\bigg[\int_0^\infty \int_E C_s\mathcal{I}_A(z)\lambda_s(dz)ds\bigg] = E\bigg[\int_0^\infty \int_E H(s,z)\lambda_s(dz)ds\bigg].$$

□

Remark 2.12. $\lambda_s(dz)ds = \nu(ds \times dz)$ is called dual projection measure.

Corollary 2.3 (Integration theorem). *Let $\mu(dt \times dz)$ be a E-marked point process with the $(P, \mathcal{F}_t)$-intensity kernel $\lambda_t(dz)$. Let H be a $\mathcal{F}_t$-predictable E-indexed process such that, $\forall t \geq 0$, it holds:*

$$E\bigg[\int_0^t \int_E |H(s,\omega,z)|\lambda_s(dz)\bigg] < \infty;$$

Then, defining the compensated E-marked point process $\widetilde{\mu}(ds \times dz) = \mu(dt \times dz) - \lambda_s(dz)ds$,

$$M_t = \int_0^t \int_E H(s,\omega,z)\widetilde{\mu}(ds \times dz)$$

is a $(P, \mathcal{F}_t)$-martingale.

Proof. As done before, the proof is done on the class of generators of $H(t,z)$ and then extended. By taking $r < t$, let define $H(t,\omega,z) = H'(t,\omega,z)\mathcal{I}_A(\mathcal{I}_{[0,t]} - \mathcal{I}_{[0,r]})$, with $A \in \mathcal{F}_r$ and $H'(t,\omega,z)$ a $\mathcal{F}_t$-predictable E-indexed process. Now, let consider the theorem 2.5; By moving the quantity on the right to the left in (2.18), and collecting the terms in order to obtain the compensated E-marked point process, one can get:

$$E\bigg[\int_0^\infty \int_E H(s,\omega,z)\widetilde{\mu}(ds \times dz)\bigg] = 0;$$

It remains to apply the choice of the process described before:

$$E\bigg[\int_0^\infty \int_E H'(t,\omega,z)\mathcal{I}_A(\mathcal{I}_{[0,t]} - \mathcal{I}_{[0,r]})\widetilde{\mu}(ds \times dz)\bigg] = 0;$$

Exploiting the indicator functions and the Fubini's theorem, it is possible to write:

$$E[\mathcal{I}_A(M_t - M_r)] = 0;$$

This shows the martingale condition for M_t and the proof is concluded.

□

Definition 2.14 (Local characteristics). *Let $\mu(dt \times dz)$ be a E-marked point process with $(P, \mathcal{F}_t)$-intensity kernel $\lambda_t(dz)$ of the form:*

$$\lambda_t(dz) = \lambda_t \Phi_t(dz) \tag{2.19}$$

where λ_t is non-negative and $\mathcal{F}_t$-predictable process and $\Phi_t(dz)$ is a probability density. The pair $(\lambda_t, \Phi_t(dz))$ is called $(P, \mathcal{F}_t)$-local characteristics of $\mu(dt \times dz)$;

Remark 2.13. $\Phi_t(E) = 1$, therefore $\lambda_t = \lambda_t(E)$ is the $(P, \mathcal{F}_t)$-intensity of the underlying counting process N_t. Moreover, it can be useful to notice that the following writing holds:

$$\Phi_t(dz) = \frac{\lambda_t(dz)}{\lambda_t};$$

The notions developed will be useful in the sequel in order to deal with the dynamic risk models. In particular, the tools of calculus which manage marked point processes will be exploited for the computations of many quantities related to the loss process.

3 | Premiums in dynamical context

The chapter contains some proposals to model the surplus of an insurance portfolio, using the theory of jump processes developed in chapter 2. In particular, the treatment is based on the Cramér-Lundberg model, which represents a standard choice in literature (see for instance [4] for the study of the non-ruin probability of a portfolio), and it is focused on the estimation of a "fair" premium rate, extending the principles explained in the first chapter to the dynamical context. Then, for each principle proposed, it follows a discussion of its properties.

3.1. The Cramér-Lundberg model

The Cramér-Lundberg model is a risk model in which the claims have a compound Poisson distribution, and it is one of the simpler choice to deal with an insurance portfolio in a continuous time framework. Part of the results contained in this section follows the treatment of [20], chapter 5.

The surplus of an insurance portfolio is, according to the model:

$$R_t = R_0 + ct - \sum_{n=1}^{N_t} Z_n \tag{3.1}$$

in which:

- R_0 represents the starting value of the portfolio, and it is obviously non-negative.
- $c > 0$ is the premium rate, namely the model assumes that the premium income is continuous over time, and, therefore, proportional in any time interval to the interval length.
- $\sum_{n=1}^{N_t} Z_n$ is a compound Poisson as in definition 2.11. Since the sequence of $\{Z_n\}$ represent the claims, they must have positive support.

Many drawbacks of the model arises immediately by looking at its hypotheses; first of all it does not seem reasonable that the claims can arrive with an intensity constant along time. Indeed, many events can present clustering behaviours or seasonality, such as car accidents, that could happen with an higher frequency during periods with bad weather conditions. Moreover, the independence between the counting process N_t and the risks $\{Z_n\}$ is a strong hypothesis which could be unjustified in many situations of real world, again, considering the insurance contracts on car accidents, a catastrophic event like a flood will affect at the same time the intensity of the Poisson process and the size of the claims, therefore the two quantities cannot be considered independent. However, the model presents also many advantages due to its simplicity, indeed, it is possible to compute easily the mean and the variance of R_t.

Lemma 3.1. *The mean and the variance of the process R_t are:*

$$E[R_t] = R_0 + ct - \lambda t E[Z]; \tag{3.2}$$

$$Var(R_t) = \lambda t E[Z^2]; \tag{3.3}$$

Proof. Before starting with the actual proof, it can be useful to do two important remarks which will be applied during the development of the computations; first of all, let recall this standard result of calculus about the exponential series:

$$\sum_{n=0}^{\infty} \frac{x^n}{n!} = e^x.$$

Then:

$$\begin{aligned} E\left[\sum_{n=0}^{N_t} Z_n\right] &= E\left[\sum_{n=0}^{\infty}\sum_{i=0}^{n} Z_i \mathcal{I}_{(N_t=n)}\right] \\ &= \sum_{n=0}^{\infty}\sum_{i=0}^{n} E[Z_i]E[\mathcal{I}_{(N_t=n)}] \\ &= \sum_{n=0}^{\infty} n e^{-\lambda t}\frac{(\lambda t)^n}{n!} Z_n \\ &= \lambda t Z_n \sum_{n=1}^{\infty} e^{-\lambda t}\frac{(\lambda t)^{(n-1)}}{n-1!} \\ &= \lambda t Z_n; \end{aligned}$$

Moreover, in order to perform the computation of the variance, it will be useful the following:

$$
\begin{aligned}
E\left[\sum_{n=0}^{N_t} Z_n^2\right] &= E\left[\sum_{n=0}^{\infty}\sum_{i=0}^{n^2} Z_i^2 \mathcal{I}_{(N_t=n)}\right] \\
&= \sum_{n=0}^{\infty} n^2 e^{-\lambda t}\frac{(\lambda t)^n}{n!} Z_n^2 \\
&= \sum_{n=0}^{\infty}[n(n-1)+n]e^{-\lambda t}\frac{(\lambda t)^n}{n!} Z_n^2 \\
&= \lambda t Z_n^2 + (\lambda t)^2 Z_n^2 \sum_{n=2}^{\infty} e^{-\lambda t}\frac{(\lambda t)^{n-2}}{(n-2)!} \\
&= (\lambda t)^2 Z_n^2 + \lambda t Z_n^2;
\end{aligned}
$$

Furthermore, since the $\{Z_n\}$ are i.i.d., by supposing that the number N_t is known, then

$$
\sum_{n\neq m}^{N_t} Z_n Z_m = \frac{N_t(N_t-1)}{2} Z_n Z_m;
$$

Exploiting the knowledge of these quantities, the first and second moment of the Cramér-Lundberg model can be easily computed:

1. Mean:

$$
\begin{aligned}
E[R_t] &= R_0 + ct - E[\sum_{n=1}^{N_t} Z_n] \\
&= R_0 + ct - E[E[\sum_{n=1}^{N_t} Z_n | N_t = n]] \\
&= R_0 + ct - E[\lambda t Z_n] \\
&= R_0 + ct - \lambda t E[Z];
\end{aligned}
$$

2. Variance:

$$
\begin{aligned}
Var(R_t) &= Var(\sum_{n=1}^{N_t} Z_n) \\
&= E\left[(\sum_{n=1}^{N_t} Z_n)^2\right] - \lambda^2 t^2 E^2[Z];
\end{aligned}
$$

Let consider the second moment of the Compound Poisson; the following computations are performed taking into account that, by hypotheses, Z_n and Z_m with $n \neq m$

are independent and identically distributed, and, exploiting the observations above:

$$
\begin{aligned}
E\Big[(\sum_{n=1}^{N_t} Z_n)^2\Big] &= E\Big[\sum_{n=1}^{N_t} Z_n^2 + 2\sum_{n\neq m}^{N_t} Z_n Z_m\Big] \\
&= E\Big[E[\sum_{n=1}^{N_t} Z_n^2 + 2\sum_{n\neq m}^{N_t} Z_n Z_m | N_t = n]\Big] \\
&= E\Big[\lambda t E[Z_n^2] + 2\frac{(\lambda t)^2 - \lambda t + \lambda t}{2} E[Z_n Z_m]\Big] \\
&= \lambda t E[Z^2] + \lambda^2 t^2 E^2[Z];
\end{aligned}
$$

Putting this result into the original equation:

$$
\begin{aligned}
Var(R_t) &= E\Big[(\sum_{n=1}^{N_t} Z_n)^2\Big] - \lambda^2 t^2 E^2[Z_n] \\
&= \lambda t E[Z_n^2] + \lambda^2 t^2 E^2[Z_n] - \lambda^2 t^2 E^2[Z_n] \\
&= \lambda t E[Z^2];
\end{aligned}
$$

□

Remark 3.1. Since the model is characterized by a marked point process, it is possible to define it by specifying its local characteristics. Because of the simplicity of the compound Poisson, recognize its local characteristics is intuitive and straightforward, indeed, the terms in (2.19) become:

$$
\begin{aligned}
\lambda_t &= \lambda, \\
\Phi_t(dz) &= F(dz)
\end{aligned}
\tag{3.4}
$$

in which $f(dz)$ is the distribution of the random variables Z_n.

This statement is proved as follows:

Let consider a generic $\mathcal{F}_t$-predictable E-indexed process, with $\nu(ds, dz)$ the dual projection measure of the Cramér-Lundberg model. Then, indicating as usual with $\{\tau_n\}$ the jump times sequence:

$$
\int_0^t \int_E H(t,z)\nu(ds,dz) = \sum_{n\geq 1} H(\tau_n, Z_n)\mathcal{I}_{(Z_n \in E)}\mathcal{I}_{(\tau_n \leq t)}.
$$

According to the standard procedure, the proof is done on the class of generators of $H(t,z)$. In particular, let be $H(t,z) = C_t \mathcal{I}_A(Z_n)$, for an arbitrary $A \in \mathcal{E}$ and C_t an

$\mathcal{F}_t$-predictable process. The following chain of equation holds:

$$
\begin{aligned}
E\left[\int_0^t\int_0^\infty H(t,z)\nu(ds,dz)\right] &= E\left[\int_0^t\int_0^\infty C_t\mathcal{I}_A(Z_n)\nu(ds,dz)\right]\\
&= E\left[\sum_{n\geq 1} C_{\tau_n}\mathcal{I}_A(Z_n)\mathcal{I}_E(Z_n)\mathcal{I}_{\tau_n\leq\infty}\right]\\
&= \sum_{n\geq 1} E[C_{\tau_n}\mathcal{I}_{\tau_n<\infty}]E[\mathcal{I}_A(Z_n)]\\
&= F_z(A)E\left[\int_0^\infty C_{\tau_n}\mathcal{I}_{\tau_n<\infty}\right]\\
&= F_z(A)E\left[\int_0^\infty C_t dN_s\right]\\
&= E\left[\int_0^t\int_0^\infty C_t\mathcal{I}_A\lambda F(z)dzds\right];
\end{aligned}
$$

This proves that the dual projection measure can be decomposed in the local characteristics mentioned before, indeed it was obtained:

$$
E\left[\int_0^t\int_0^\infty H(t,z)\nu(ds,dz)\right] = E\left[\int_0^t\int_0^\infty H(t,z)\lambda F(z)dsdz\right] \tag{3.5}
$$

Exploiting this alternative definition, it can be computed the mean of the loss in a different way. Indeed, imposing $H(t,z)=z$, it follows:

$$
\begin{aligned}
E[L_t] &= E\left[\int_0^t\int_0^\infty z\nu(ds,dz)\right]\\
&= E\left[\int_0^t\int_0^\infty z\lambda f(dz)ds\right]\\
&= \lambda E\left[\int_0^t\int_0^\infty zf(dz)\right]\\
&= \lambda E\left[E[Z]\int_0^t ds\right] = \lambda t E[Z].
\end{aligned}
$$

The knowledge of mean and variance of the process suggests to look for an extension of the premium principles discussed in the chapter 1 suitable for the Cramér-Lundberg model.

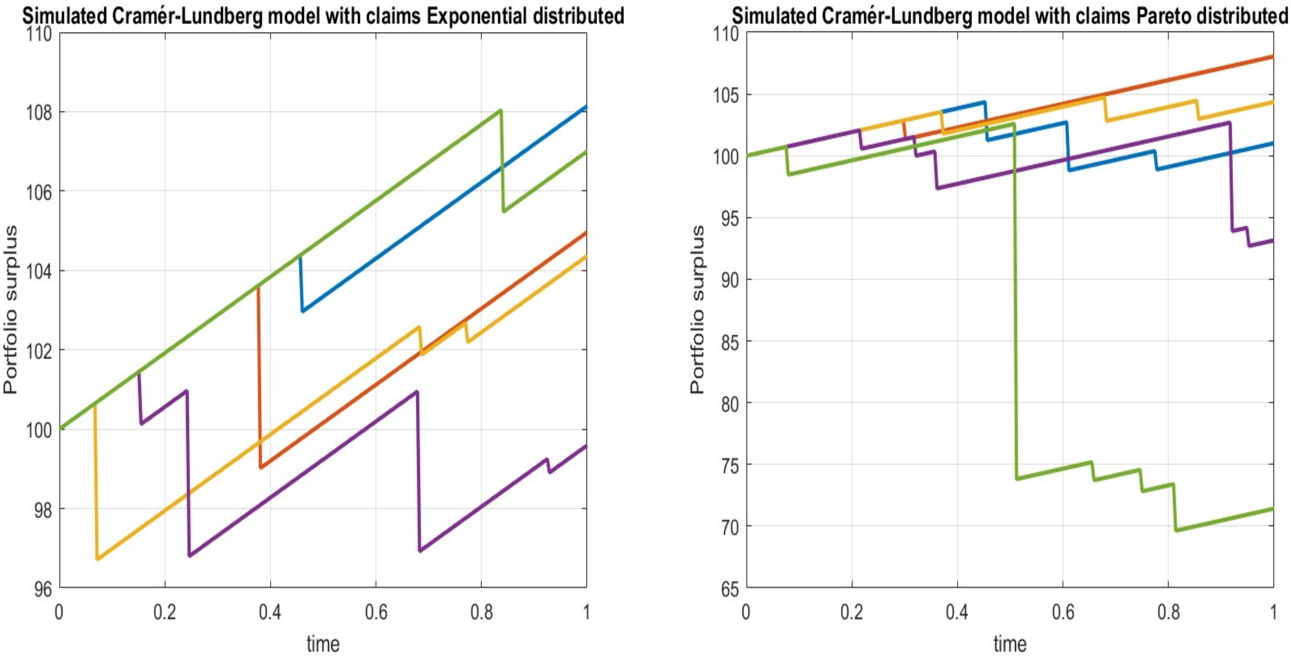

Figure 3.1: Simulated paths of a Cramér-Lundberg portfolio with net premium rate, time horizon $T = 1$, intensity $\lambda = 2$,$R_0 = 100$. In the exponential case, the parameter is $\bar{\lambda} = 0.2$, in the Pareto one the parameters are $\alpha = 1.2$,$\beta = 0.8$.

Estimation of the premium rate

The model does not provide any information about a reasonable value that c should assume in order to be a "fair" premium rate. Inspired by the premium principles adopted in a static context, it is possible to define a correspondent premium rate in the dynamic case. First of all, easing the notation, let define the loss of the portfolio as $L_t = \sum_{n=1}^{N_t} Z_n$. Moreover, it is important to notice that all the properties of the premium principles already presented must be extended for stochastic processes, therefore they must hold $\forall t \geq 0$ fixed. However, the property of continuity seems to have no meaning in the Cramér-Lundberg model, as shown in the following remark:

Remark 3.2. Let consider two different portfolios:

$$R_t = R_o + ct + \sum_{n=1}^{N_t} Z_n$$

$$\widetilde{R}_t = R_o + \widetilde{c}t + (\sum_{n=1}^{N_t} Z_n - d) = \widetilde{R}_0 + \widetilde{c}t + \sum_{n=1}^{N_t} Z_n,$$

with $\widetilde{R}_0 = R_0 + d$;

By considering $c = H(L_t)$, in which H is an suitable functional, it does not make sense the study of the behaviour of $\lim_{d\to 0^+} H((L_t - d)_+)$, since, as one can notice above, reducing the loss of a fixed quantity corresponds to a translation of the starting value, which usually has no importance in the computation of the premium rate, and therefore $c = \tilde{c} = H(L_t)$. Indeed, if R_0 contributes to c, the property of conditional state dependence is denied. However, a more reasonable extension of the property could be:

$$
\begin{aligned}
&\textit{If } d \geq 0, \ \textit{then} \ : \\
&\lim_{d\to 0^+} \ H\bigg(\sum_{n=1}^{N_t}(Z_n - d)_+\bigg) = H(L_t) \\
&\lim_{d\to \infty} H\bigg(\sum_{n=1}^{N_t}\min(Z_n, d)\bigg) = H(L_t).
\end{aligned}
\tag{3.6}
$$

3.1.1. Net premium and expected value premium

These two premium principles can be extended by imposing:

$$
ct = (1+\theta)E[L_t] \Rightarrow c = (1+\theta)\lambda E[Z]
$$

As usual, the net premium is a trivial case in which $\theta = 0$.

It is interesting to notice how the properties change with respect to the static case. Indeed, even though this particular choice for c is the simplest possible, one can notice that, for the expected value premium, the property of max loss now holds. In order to make the treatment not too heavy, just the properties considered not trivial are discussed. Moreover, for extending the definition of monotonicity for the losses, and thus for two compound Poisson, it is important to do the following observations:

Remark 3.3. A natural way to extend the concept of monotonicity to (3.1) is:

$$
c \leq \tilde{c} \ \textit{if} \ L_t(\omega) \leq \tilde{L}_t(\omega) \ \forall \omega \in \Omega, \ \ \forall t \geq 0. \tag{3.7}
$$

It is possible to distinguish five cases:

1. $\tilde{N}_t(\omega) \geq N_t(\omega)$ $\forall \omega \in \Omega$, $\forall t \geq 0$ and $Z = \tilde{Z}$ a.s.;
2. $\tilde{N}_t(\omega) = N_t(\omega)$ and $Z(\omega) \leq \tilde{Z}(\omega)$ $\forall \omega \in \Omega$, $\forall t \geq 0$;
3. $\tilde{N}_t(\omega) \geq N_t(\omega)$ and $Z(\omega) < \tilde{Z}(\omega)$ $\forall \omega \in \Omega$, $\forall t \geq 0$;
4. $\tilde{N}_t(\omega) \leq N_t(\omega)$ and $Z(\omega) > \tilde{Z}(\omega)$ $\forall \omega \in \Omega$, $\forall t \geq 0$;

5. $\widetilde{N}_t(\omega) > N_t(\omega)$ and $Z(\omega) > \widetilde{Z}(\omega \in \Omega, \forall t \geq 0$;

- **Monotonicity:**

  The main case of interest in order to define an ordering between two Poisson processes is when the greater $\widetilde{N}_t$ is built starting from the smaller N_t as $\widetilde{N}_t = N_t + \bar{N}_t$, with $N_t, \bar{N}_t$ independent.The third and fourth cases are not possible in order to have an ordering between the two losses. Let consider the following counterexample:

  Let suppose the usual probability space $(\Omega, \mathcal{F}, P)$ endowed of a filtration $\mathcal{F}_t$ in which are defined two Poisson process N_t and $\widetilde{N}_t = N_t + \bar{N}_t$ and two random variables such that $Z = 3$ a.s. and $\widetilde{Z} = 1$ a.s. It exists an $\widetilde{\omega}$ with $P(\widetilde{\omega}) > 0$ such that $N_t = 2$ and $\bar{N}_t = 1$ at a given time $t > 0$, indeed:

$$P(N_t = 2, \bar{N}_t = 1) = \frac{e^{-\lambda t}(\lambda t)^2}{2} e^{-\bar{\lambda} t} \bar{\lambda} t > 0;$$

  In this particular case $L_t = 6$ and $\widetilde{L}_t = 3$. This shows that monotonicity cannot be defined in the third case. Similar counterexamples can be constructed for the fourth one, leading to states that:

$$\begin{gathered} L_t(\omega) \leq \widetilde{L}_t(\omega) \ \forall \omega \in \Omega, \ \ \forall t \geq 0 \iff \\ N_t(\omega) \leq \widetilde{N}_t(\omega) \ and \ Z(\omega) \leq \widetilde{Z}(\omega) \ \forall \omega \in \Omega, \ \ \forall t \geq 0; \end{gathered} \tag{3.8}$$

  In the first case, indeed, it is sufficient to observe that, since $\widetilde{\lambda} \geq \lambda$, then:

$$c = (1+\theta)\lambda E[Z] \leq (1+\theta)\widetilde{\lambda} E[Z] = \widetilde{c}.$$

  In the second case, since the counting process is the same for both the two losses, the proof follows straightforwardly:

$$c = (1+\theta)\lambda E[Z] \leq (1+\theta)\lambda E[\widetilde{Z}] = \widetilde{c};$$

  Finally, since the property holds in the first two possibilities, a fortiori the property holds for the last case.

- **Unjustified risk loading:** Also for this request it is necessary to write an extension suitable for the compound Poisson which is well-formulated. Indeed it seems not reasonable to impose the loss equal to a constant. However, it can be interesting to analyze the case in which the jump size is constant, imposing that in which case the premium rate should be equal to the average number of jumps multiplied for the magnitude of them. Therefore:

$$\textit{If } Z_n = k \ a.s., \ \textit{then } c = \lambda k;$$

Trivially, this is verified only for the net premium.

- **Maximal loss:** This property holds trivially since:

$$c = (1+\theta)\lambda E[Z] \leq ess\ supL_t = \infty \ \textit{for } t > 0;$$

- **Preserving FSD:**

  By exploiting theorem 1.2, since monotonicity holds $\forall t$ and conditional state dependence is verified, this property holds true $\forall t$;

The proves of all the other properties are similar to the ones in the static case, except for the fact that now they hold for each $t \geq 0$ fixed. Eventually, it can be interesting to do the following observation;

Remark 3.4. By adopting the net premium principle for c, the process R_t enjoys the martingale property, indeed, the term $ct - L_t$ can be interpreted as a compensated version of the compound Poisson. The following equations show the result:

$$\begin{aligned} E\Big[\sum_{n=1}^{N_t} Z_n|\mathcal{F}_s\Big] &= E\Big[\sum_{n=1}^{N_t} Z_n - \sum_{n=1}^{N_s} Z_n + \sum_{n=1}^{N_s} Z_n|\mathcal{F}_s\Big] \\ &= E\Big[\sum_{n=1}^{N_t - N_s} Z_n|\mathcal{F}_s\Big] + \sum_{n=1}^{N_s} Z_n \\ &= \lambda(t-s)E[Z] + \sum_{n=1}^{N_s} Z_n. \end{aligned}$$

Thus:

$$\begin{aligned} E[R_t|\mathcal{F}_s] &= R_0 + \lambda t E[Z_n] - E\left[\sum_{n=1}^{N_t} Z_n|\mathcal{F}_s\right] \\ &= R_0 + \lambda t E[Z] - \lambda(t-s)E[Z] - \sum_{n=1}^{N_s} Z_n \\ &= R_0 + \lambda s E[Z] - \sum_{n=1}^{N_s} Z_n \\ &= R_s. \end{aligned}$$

3.1.2. Variance premium and Standard deviation premium

The variance principle can be generalized as:

$$ct = E[L_t] + \alpha Var(L_t) \implies c = \lambda E[Z] + \alpha\lambda E[Z^2];$$

Instead, the standard deviation one:

$$ct = E[L_t] + \alpha\sqrt{Var(L_t)} \implies c = \lambda E[Z] + \alpha\sqrt{\frac{\lambda}{t}E[Z^2]};$$

By taking as a choice for the premium rate one of these two functionals, and trying to check if the properties valid in the static context are preserved, one can obtain many surprising results. Indeed, not only all the properties before true keep holding, but, moreover, many others now holds. As done for the expected value premium, only the properties considered not trivial will be shown, moreover the reasoning proposed are valid for both the premium principles if not specified.

- **Monotonicity:** In the static case this does not hold since the variance does not preserve monotonicity, however, the variance of the compound Poisson is equal to a constant that multiply the second moment of the variable which represents the jump size. Since these variables are for definition non-negative and the expected value operator preserves monotonicity, then the variance and standard deviation premium principles are monotonic themselves.
- **Unjustified risk loading:** Recalling the considerations done in the discussion of the expected value premium principle, one can notice that this does not hold for both the two principles. However, by taking a time horizon much large, the standard deviation premium provides a rate which is similar to the net one, anyway in practical situations it is not reasonable to study the state of the premium rate

for $t \to \infty$, therefore, in general the property is not valid.

- **No additivity:** Let consider the following computations, with two different losses $L_t, \widetilde{L}_t$:

$$Var(L_t + \widetilde{L}_t) = Var\left(\sum_{n=1}^{N_t} Z_n + \sum_{m=1}^{\widetilde{N}_t} \widetilde{Z}_m\right)$$
$$= Var\left(\sum_{n=1}^{N_t} Z_n\right) + Var\left(\sum_{m=1}^{\widetilde{N}_t} \widetilde{Z}_m\right) + 2Cov\left(\sum_{n=1}^{N_t} Z_n, \sum_{m=1}^{\widetilde{N}_t} \widetilde{Z}_m\right).$$

Now let focus on the covariance term, exploiting the same technique used for the computation of mean and variance of the Cramér-Lundberg model:

$$Cov\left(\sum_{n=1}^{N_t} Z_n, \sum_{m=1}^{\widetilde{N}_t} \widetilde{Z}_m\right) = E\left[E[(\sum_{n=1}^{N_t} Z_n)(\sum_{m=1}^{\widetilde{N}_t} \widetilde{Z}_m)|N_t = n, \widetilde{N}_t = \widetilde{n}]\right] - \lambda\widetilde{\lambda}t^2 E[Z]E[\widetilde{Z}]$$
$$= \lambda\widetilde{\lambda}t^2 E[Z\widetilde{Z}] - \lambda\widetilde{\lambda}t^2 E[Z]E[\widetilde{Z}]$$
$$= \lambda\widetilde{\lambda}t^2 Cov(Z, \widetilde{Z})$$

Putting it into the starting equation:

$$Var(L_t + \widetilde{L}_t) = Var\left(\sum_{n=1}^{N_t} Z_n\right) + Var\left(\sum_{m=1}^{\widetilde{N}_t} \widetilde{Z}_m\right) + 2\lambda\widetilde{\lambda}t^2 Cov(Z, \widetilde{Z}).$$

It is obvious that additivity can be guaranteed only by adding a condition for which all the covariance contributes between the two losses are null. This proves the additivity for independent risks, for the variance premium principle, the standard deviation one does not enjoy the property for the same reasons of the static case. Subadditivity and superadditivity do not hold since the covariances can be both positive and negative, eventually the comonotonic additivity does not hold in fact, for the theorem 1.1, it is known that comonotonicity implies positive covariance.

- **Max loss:** The same reasoning done in the case of the expected value principle applies also in this case.

- **Preserving FSD and Stop-loss ordering**: The result obtained about the monotonicity allows to exploit the theorem 1.2, therefore the premiums preserve FSD. Moreover, it is known (see for instance [14] for further details), that the FSD ordering implies the stop-loss ordering, therefore, also this property is valid.

The other properties keep the same behaviour of the static context and their proves or counterexamples can be applied in the same way by fixing t, $\forall t \geq 0$.

Remark 3.5. Recalling the treatment presented in section 1.3, one can notice that, in the Cramér-Lundberg model, all the premium principles presented can be interpreted as monetary risk measure, since the property of monotonicity implies that the deviation measure part of the principle is null $\forall t$.

Table of properties

Here the table resuming the properties of the principles in the Cramér-Lundberg model, the numbering is the same of the chapter 1:

Properties	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Net pp	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Expected value pp	Y	Y	Y	Y	Y	N	Y	N	Y	Y	Y	Y	Y	Y	Y
Variance pp	Y	Y	N	Y	Y	N	Y	Y	N	N	N	N	Y	Y	Y
Std pp	Y	Y	N	Y	Y	N	Y	Y	Y	N	N	N	N	Y	Y

4 | Hawkes processes

As seen in chapter 3, the model (3.1) presents many drawbacks which does not properly match reality, such as the constant intensity for the claims arrival. A possible extension which improve the Cramér-Lundberg model, dealing with the aforementioned issue, is the introduction of a Hawkes as counting process instead of the Poisson one. This means using a time-dependent intensity which grows up with the occurrence of events, for this reason Hawkes processes are also called "self-exciting" processes. This chapter develops some theory about one-dimensional Hawkes processes, then it presents a new risk model and some proposal about the computation of the premium rate. The first work which introduced self-exciting processes was [12], many proposals to extend their applications in finance and insurance frameworks are contained also in [13] and [21]. Moreover, some of the theoretical results presented here are treated also in [17], [7] and [22].

4.1. Review of theory

A Hawkes process is a point process with self-exciting property, which allows to model clustering effects in the sequence of arrivals into a system. Let consider a probability space $(\Omega, \mathcal{F}, P)$ endowed with a filtration $\mathcal{F}_t$.

Definition 4.1 (Conditional intensity function). *Consider a counting process N_t with an associated filtration $\mathcal{F}_t$ for $t \geq 0$. If a non-negative function λ_t^* exists such that*

$$\lambda_t^* = \lim_{h \to 0} \frac{E[N_{t+h} - N_t | \mathcal{F}_t]}{h} \tag{4.1}$$

it is called conditional intensity function for the process N_t

According to the theory developed in chapter 2, it is possible, exploiting the fact that λ_t^* is the conditional intensity of N_t, to define the compensator of the point process with the usual quantity:

$$\Lambda_t = \int_0^t \lambda_s^* ds$$

The definition of Hawkes process, based on (4.1), is the following:

Definition 4.2 (Hawkes process). *A counting process N_t, with the associated filtration $\mathcal{F}_t$, that satisfy:*

$$P(N_{t+h} - N_t = m|\mathcal{F}_t) = \begin{cases} \lambda_t^* h + o(h), m = 1 \\ o(h), m > 1 \\ 1 - \lambda_t^* h + o(h), m = 0 \end{cases} \tag{4.2}$$

Suppose, moreover, the process conditional intensity function is of the form:

$$\lambda_t^* = \lambda + \int_{-\infty}^t \mu(t-s)dN_s \tag{4.3}$$

for some $\lambda > 0$ and $\mu : (0, \infty) \longrightarrow [0, \infty)$. Such a process N_t is a Hawkes process and $\mu(t)$ is called excitation function.

Remark 4.1. By setting the function $\mu(t) = 0 \; \forall t \geq 0$, one can obtain the trivial case of a Poisson process.

Remark 4.2. The integral in (4.3) can be rewritten according to (2.14), therefore:

$$\int_{-\infty}^t \mu(s)dN_s = \sum_{n\geq 1} \mu(T_n)\mathcal{I}_{(T_n \leq t)}$$

The choice of the function $\mu(t)$ may lead to an explosive process, therefore it is necessary a supplementary hypothesis which guarantees to avoid this event.

Lemma 4.1. *The stationarity condition:*

$$m = \int_0^\infty \mu(s)ds < 1 \tag{4.4}$$

is sufficient for non-explosivity.

Proof. Let consider the condition of non-explosivity for the Hawkes process with intensity (4.3):

$$E\left[\int_0^t \lambda_s^* ds\right] < +\infty$$

It is possible to apply the Fubini's theorem, since the conditional intensity function is non-negative and measurable by definition. Indeed, for (2.14), the integral with respect to a point process is a sum of measurable random variables. Therefore, the problem can be rewritten looking for a condition such that:

$$\int_0^t E[\lambda_s^*] ds < +\infty \tag{4.5}$$

Now let consider the mean of the conditional intensity function:

$$\begin{aligned} E[\lambda_t^*] &= \lambda + E\left[\int_0^t \mu(t-s) dN_s\right] \\ &= \lambda + E\left[\int_0^t \mu(t-s) \lambda_s^* ds\right] \\ &= \lambda + \mu(t) * E[\lambda_t^*] \end{aligned}$$

In the last equivalence it was applied again the Fubini's theorem and the definition of convolution. Hence, it holds:

$$E[\lambda_t^*] = \lambda + \mu(t) * E[\lambda_t^*]$$

This kind of equation is called in literature "Renewal equation", and it is widely treated in literature. In particular, considering the defective case, that is the case in which the condition (4.4) holds, it can be proved (see [1], proposition 7.4) that:

$$\lim_{t\to\infty} E[\lambda_t^*] = \frac{\lim_{t\to\infty} \lambda}{1-m} = \frac{\lambda}{1-m} < +\infty$$

This implies that (4.5) is verified, and the proof is concluded.

□

The definition (4.2) points out that the choice of the function $\mu(t)$ is crucial in order to characterize the Hawkes process. There are different excitation functions used in literature, a standard one is the exponentially decaying excitation function.

Definition 4.3 (Exponentially decaying Hawkes process)**.** *A point process N_t satisfying*

(4.2) and (4.3) with excitation function of the form:

$$\mu(t) = \alpha e^{-\beta t} \tag{4.6}$$

is an exponentially decaying Hawkes process, with $\alpha, \beta > 0$.

Remark 4.3. The condition (4.4), which avoid the explosion of the process, becomes:

$$\int_0^\infty \alpha e^{-\beta s} ds < 1 \Longrightarrow \alpha < \beta$$

By considering the choice (4.6), the conditional intensity function for the exponentially decaying Hawkes process follows straightforwardly:

$$\lambda_t^* = \lambda + \int_{-\infty}^t \alpha e^{-\beta s} dN_s \tag{4.7}$$

The quantity (4.7), given an initial condition $\lambda^*(0) = 0$, satisfies the stochastic differential equation

$$d\lambda_t^* = \beta(\lambda - \lambda_t^*)dt + \alpha dN_t \tag{4.8}$$

.

Lemma 4.2. *The equation (4.8) admits as unique solution*

$$\lambda_t^* = \lambda + e^{-\beta t}(\lambda_0 - \lambda) + \int_0^t \alpha e^{-\beta(t-s)} dN_s \tag{4.9}$$

Proof. Since (4.8) is linear, it seems reasonable trying to solve the equation in a similar way to the Ornstain-Uhlenbeck case in a diffusive framework (see [3], chapter 9, for further details), therefore, let derive the quantity $\lambda_t' = e^{\beta t}\lambda_t^*$:

$$\begin{aligned} d(e^{\beta t}\lambda_t^*) &= \beta e^{\beta t}\lambda_t^* dt + e^{\beta t} d\lambda_t^* \\ &= \beta e^{\beta t}\lambda_t^* dt + e^{\beta t}(\beta(\lambda - \lambda_t^*)dt + \alpha dN_t) \end{aligned}$$

Thus:

$$d\lambda_t' = e^{\beta t}\beta\lambda dt + e^{\beta t}\alpha dN_t$$

with $\lambda'(0) = \lambda_0$;

Integrating, one can obtain

$$\int_0^t d\lambda'_s = \int_0^t e^{\beta s}\beta\lambda ds + \int_0^t e^{\beta s}\alpha dN_s$$
$$\implies e^{\beta t}\lambda_t^* - \lambda_0 = e^{\beta t}\lambda - \lambda + \int_0^t e^{\beta s}\alpha dN_s$$
$$\implies \lambda_t^* = e^{-\beta t}(\lambda_0 - \lambda) + \lambda + \int_0^t e^{-\beta(t-s)}\alpha dN_s$$

This concludes the proof. □

Remark 4.4. The solution (4.9) allows to extend the treatment to the case in which an exponential decaying Hawkes process with parameters (λ, α, β) has already started in the past, but it is observed starting from the time $t = 0$, in which its conditional intensity function has value λ_0. This writing is consistent with the definition (4.7), considering as initial condition $\lambda^*_{-\infty} = \lambda$.

Given the conditional intensity function (4.9), it is possible to compute the mean of the process. First of all, let consider the following lemma:

Lemma 4.3. *Let N_t be a Hawkes process with conditional intensity as in (4.9). Then, the following equations holds:*

$$dE[N_t] = E[\lambda_t]dt \tag{4.10}$$

$$dE[\lambda_t^*] = \beta\lambda dt + (\alpha - \beta)E[\lambda_t^*]dt \tag{4.11}$$

Proof. The equation (4.10) follows straightforwardly by considering the theorem 2.1 for the Hawkes process, indeed:

$$M_t = N_t - \int_0^t \lambda_s^* ds$$
$$\implies E[M_t] = E[N_t] - \int_0^t E[\lambda_s^*]ds$$
$$\implies E[N_t] = N_0 + \int_0^t E[\lambda_s^*]ds$$

in which it was applied the Fubini's theorem. By taking the derivative in the last equation, the first part of the statement is proved.

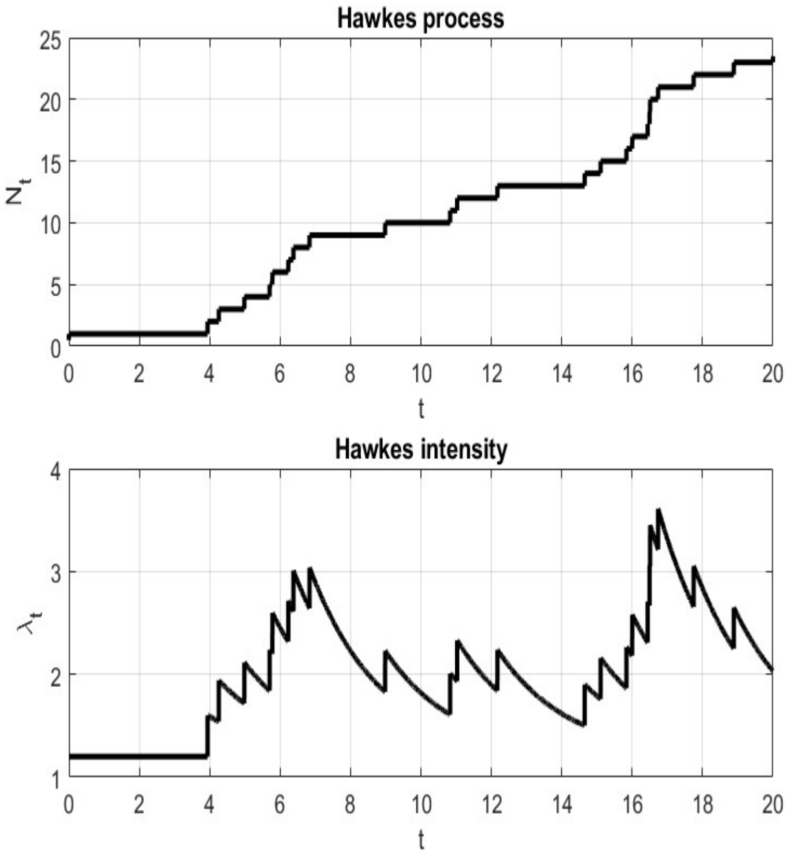

Figure 4.1: Simulated path of an exponentially decaying Hawkes process with its intensity. The parameters are $\lambda = 1$, $\alpha = 0.4$, $\beta = 0.5$ and $\lambda_0 = 1.2$; The algorithm used is the one proposed in [8].

In order to get the equation (4.11), let consider the stochastic differential equation (4.8):

$$d\lambda_t^* = \beta(\lambda - \lambda_t^*)dt + \alpha dN_t$$

By taking its integral form, one can obtain:

$$\lambda_t^* - \lambda_0 = \int_0^t \beta(\lambda - \lambda_s^*) + \int_0^t \alpha dN_s$$
$$\implies E[\lambda_t^*] = \lambda_0 + \int_0^t E[\beta(\lambda - \lambda_s^*]ds + \alpha E\left[\int_0^t dN_s\right]$$

Exploiting the fact that λ_t^* is the stochastic intensity of the process, using (2.13) with $C_s = \mathcal{I}_{(s<t)}$:

$$E[\lambda_t^*] = \lambda_0 + \beta\lambda t + (\alpha - \beta)\int_0^t E[\lambda_s]ds$$

Again, deriving, it follows (4.11), and the proof is concluded. □

Theorem 4.1. *Let N_t be a Hawkes process with conditional intensity as in (4.9). Then, its mean is:*

$$E[N_t] = N_0 + \frac{\lambda\beta(-1 + e^{(\alpha-\beta)t} - (\alpha-\beta)t)}{(\alpha-\beta)^2} + \frac{\lambda_0(-1 + e^{(\alpha-\beta)t})}{\alpha-\beta} \tag{4.12}$$

Proof. The proof is just done by solving the equations (4.11) and (4.10), considering an appropriate initial condition; These are linear ODE of the first order, so their solution follows straightforwardly by using standard tools of calculus. In particular, from (4.11) with $E[\lambda_0] = \lambda_0$, one can get:

$$E[\lambda_t] = e^{(\alpha-\beta)t}\left(\lambda_0 + \frac{\beta\lambda}{\alpha-\beta}(1 - e^{-(\alpha-\beta)t})\right) \tag{4.13}$$

Now, putting this solution into (4.10) and solving the integral, the statement is proved, considering as initial condition $E[N_0] = N_0$. □

Remark 4.5. The solution provided by the theorem above is valid in the most general case. Considering the definition (4.7) instead of (4.9), the formula can be adapted easily since it represents the situation in which $N_0 = 0$ and $\lambda_0 = \lambda$, becoming:

$$E[N_t] = \frac{\lambda\beta(-1 + e^{(\alpha-\beta)t} - (\alpha-\beta)t)}{(\alpha-\beta)^2} + \frac{\lambda(-1 + e^{(\alpha-\beta)t})}{\alpha-\beta} \tag{4.14}$$

4.2. Risk model with Compound Hawkes

Exploiting the properties of the Hawkes processes it is possible to provide a risk model which achieve the goal of avoiding a constant intensity in the arrival of the claims, that represents instead one of the main drawback of the Cramér-Lundberg model. Therefore, the surplus of an insurance portfolio can be modeled as:

$$R_t = R_0 + ct - \sum_{n=1}^{N_t} Z_n \tag{4.15}$$

in which:

- R_0 and $c > 0$ have the same meaning in (3.1)
- $\sum_{n=1}^{N_t} Z_n = L_t$ represents as usual the loss of the portfolio. $\{Z_n\}$ is a sequence of random variable i.i.d with positive support, while N_t is an exponentially decaying Hawkes process, hence its self-exciting function is as in (4.6). Again, N_t and Z_n are independent $\forall t \geq 0$ $\forall n \in \mathbb{N}$.

Even though the introduction of a compound Hawkes as loss process seems to make the model much more complicated with respect to the (3.1), the computation of its mean can be easily done exploiting the properties of the marked point processes.

Lemma 4.4. *The mean of* $L_t = \sum_{n=1}^{N_t} Z_n$ *is:*

$$E[L_t] = E[N_t]E[Z]$$

Proof. Using the equation (3.5) with $H(t,z) = z$, one can obtain:

$$
\begin{aligned}
E[L_t] &= E\left[\int_0^t \int_0^{+\infty} z\nu(ds,dz)\right] \\
&= E\left[\int_0^t \int_0^{+\infty} z\lambda_s F(z)dsdz\right] \\
&= E\left[\int_0^t \lambda_s\left(\int_0^{+\infty} zF(z)dz\right)ds\right] \\
&= E\left[E[Z]\int_0^t \lambda_s ds\right] \\
&= E[Z]E\left[\int_0^t dN_s\right] = E[Z]E[N_t]
\end{aligned}
$$

□

Remark 4.6. In the case of study presented, considering the exponentially decaying Hawkes process, the mean of the loss becomes:

$$
E[L_t] = \frac{\lambda\beta(-1+e^{(\alpha-\beta)t}-(\alpha-\beta)t)}{(\alpha-\beta)^2}E[Z] + \frac{\lambda(-1+e^{(\alpha-\beta)t})}{\alpha-\beta}E[Z]
$$

For the sake of simplicity, in fact, it will be considered the case in which the conditional intensity is given by (4.7).

As already done for the Cramér-Lundberg one, it is possible to extend the premium principles valid in the static case to the model (4.15). In particular, the treatment will be focus on the net premium principle and the expected value premium principle. Although the new risk model is more advanced, due to the presence of the Hawkes as counting process, it can be shown that the principles mentioned keep all the properties valid in (3.1).

First of all, let define the premium rate according to the principles selected:

$$
\begin{aligned}
&ct = (1+\theta)E[L_t] \\
&\Longrightarrow c = (1+\theta)\left(\frac{\lambda\beta(-1+e^{(\alpha-\beta)t}-(\alpha-\beta)t)}{(\alpha-\beta)^2 t} + \frac{\lambda(-1+e^{(\alpha-\beta)t})}{(\alpha-\beta)t}\right)E[Z]
\end{aligned}
\tag{4.16}
$$

with the net premium obtained as usual with the special case in which $\theta = 0$. Given the formula (4.16), an insurer has two possible approaches: the first consists in fixing a time horizon t for covering the risks in $[0,t]$, the second one instead is recomputing the premium rates continuously over time. Basically, this choice means that the insurance company recomputes the premium rates $\forall t \geq 0$. Obviously this does not make sense in reality, but it is interesting to study the case in which the company applies a discrete monitoring to its portfolio, changing the rate for each time interval, so updating c in t_n with $\forall n = 0, ..., N-1$ covering the risks in $[t_n, t_{n+1}]$, following the same principle with $0 = t_0 < t_1 < ... < t_n < ... < t_N = t$.

The following discussion of the properties presents only the results considered not trivial:

- **Conditional state dependence:** This is verified since the premium rate depends only on the distribution of the loss process.
- **Monotonicity:** By definition (3.7), it is possible to distinguish the same five cases of the previous treatment for the Cramér-Lundberg model. For the reader convenience, they will be reported here, moreover it is important to recall that the only ordering between two counting processes considered in this work is the one in which the greater is defined as the smaller added to another one, therefore: $\widetilde{N}_t = N_t + \bar{N}_t$.

  1. $\widetilde{N}_t(\omega) \geq N_t(\omega)\ \forall \omega \in \Omega,\ \forall t \geq 0$ and $Z = \widetilde{Z}$ a.s.;

     Since $E[\widetilde{L}_t] \geq E[L_t]$, straightforwardly one can have:

     $$c = \frac{(1+\theta)}{t} E[L_t]E[Z] \leq \frac{(1+\theta)}{t} E[\widetilde{L}_t]E[Z] = \widetilde{c}$$

  2. $\widetilde{N}_t(\omega) = N_t(\omega)$ and $Z(\omega) \leq \widetilde{Z}(\omega)\ \forall \omega \in \Omega,\ \forall t \geq 0$;

     This case is easily solved as follows:

     $$c = (1+\theta)E[L_t]E[Z_n] \leq (1+\theta)E[L_t]E[\widetilde{Z}_n] = \widetilde{c};$$

  3. $\widetilde{N}_t(\omega) \leq N_t(\omega)$ and $Z(\omega) > \widetilde{Z}(\omega)\ \forall \omega \in \Omega,\ \forall t \geq 0$;

     It can be shown, with a counterexample similar to the one presented with the compound Poisson in the previous chapter, that this does not define an ordering between the two losses.

  4. $\widetilde{N}_t(\omega) \leq N_t(\omega)$ and $Z(\omega) > \widetilde{Z}(\omega)\ \forall \omega \in \Omega,\ \forall t \geq 0$;

     As in the third case, this does not define a valid ordering.

  5. $\widetilde{N}_t(\omega) > N_t(\omega)$ and $Z(\omega) > \widetilde{Z}(\omega)\ \forall \omega \in \Omega,\ \forall t \geq 0$;

     Trivial, since the property holds in the first two cases.

- **Continuity:** Considering (3.6) it is easy to show that the property holds true if the following equations hold:

$$\lim_{d \to 0+} (1+\theta)E[(Z-d)_+] = (1+\theta)E[Z]$$
$$\lim_{d \to \infty} (1+\theta)E[\min(X,d)] = (1+\theta)E[Z]$$

  These were already proved in chapter one.

- **Unjustified risk loading:** Inspired by the extension done for the Cramér-Lundberg model, the property becomes:

$$\textit{If } Z_n = k \ \forall n \in \mathbb{N},\ k \in \mathbb{R} \Longrightarrow ct = E[N_t]k$$

  This is verified only for the net premium.

- **Maximal loss:** This one can be proved exploiting the same reasoning used for the Cramér-Lundberg.

- **Preserving FSD:** The theorem 1.2 guarantees it, since conditional state dependence and monotonicity holds. Moreover, this property implies that the principles preserve also the stop-loss ordering.

Remark 4.7. The net premium principle for the risk model proposed does not make the process R_t a martingale as in (3.1). Indeed, for $t > s \geq 0$:

$$\begin{aligned}
E[R_t|\mathcal{F}_s] &= R_0 + E[N_t]E[Z] - E[\sum_{n=1}^{N_t} Z_n|\mathcal{F}_s] \\
&= R_0 + E[N_t]E[Z] - E[\sum_{n=1}^{N_t} Z_n - \sum_{n=1}^{N_s} Z_n + \sum_{n=1}^{N_s} Z_n|\mathcal{F}_s] \\
&= R_0 + E[N_t]E[Z] - L_s + E[\sum_{n=1}^{N_t-N_s} Z_n|\mathcal{F}_s] \\
&= R_0 + E[N_t]E[Z] - L_s - E[N_t - N_s|\mathcal{F}_s]E[Z] \\
&= R_0 + (E[N_t] - E[N_t - N_s|\mathcal{F}_s])E[Z] - L_s
\end{aligned}$$

The martingale condition is respected if $(E[N_t] - E[N_t - N_s|\mathcal{F}_s]) = E[N_s]$. Unfortunately, this equation does not hold true, since, exploiting (4.12), one can observe that:

$$E[N_t - N_s|\mathcal{F}_s] = \frac{\lambda\beta(-1 + e^{(\alpha-\beta)(t-s)} - (\alpha-\beta)(t-s))}{(\alpha-\beta)^2} + \frac{\lambda_0(-1 + e^{(\alpha-\beta)(t-s)})}{\alpha-\beta}$$

Remark 4.8. Since monotonicity holds $\forall t \geq 0$, the treatment presented in section 1.3 guarantees that the premium rate with expected value principle can be interpreted as monetary risk measure as well as in the Cramér-Lundberg framework.

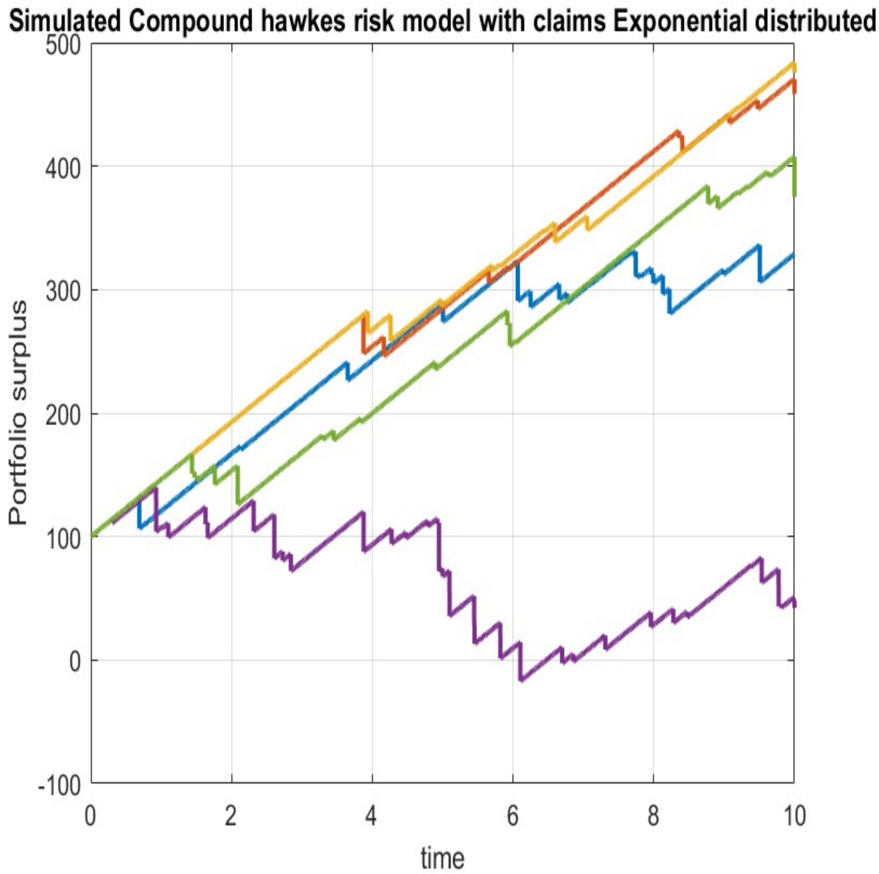

Figure 4.2: Simulated paths of the portfolio surplus in the risk model with Compound Hawkes and net premium rate, time horizon T=10, the parameters of the counting process are $\lambda = \lambda_0 = 1$, $\alpha = 0.4$, $\beta = 0.5$. The claims are exponential distributed with parameter $\bar{\lambda} = 0.08$. It is interesting to notice that there is a scenario in which the clustering of events causes a default while in the other ones, in which this effect is reduced, the portfolio surplus grows.

4.3. Estimation of the error of a Cramér-Lundberg model in a context with clustering

Let suppose that an insurance company collects premiums considering the rate provided by the model (3.1) with the expected value principle, but in a context in which the loss presents clusters of events, hence the claims should be modeled with an Hawkes as counting process. Formally, this situation is described by the following equation:

$$\widetilde{R}_t = R_0 + \widetilde{c}t - \sum_{n=1}^{N_t} Z_n \tag{4.17}$$

with N_t exponentially decaying Hawkes process with parameters λ, α, β, and $\widetilde{c} = (1 + \theta)\lambda E[Z]$. It can be interesting to study the error committed using the rate $\widetilde{c}$ instead of the "correct" one, written in (4.16). Therefore, considering the risk model (4.15), let define the error as:

$$\epsilon_t = R_t - \widetilde{R}_t = (c - \widetilde{c})t$$

The analysis will be done from the two different points of view aforementioned: for the first one it is fixed an infinite time horizon and then it is studied the error, thus, let be $T > 0$, then:

$$\epsilon_t = (1+\theta)E[Z]\left(\frac{E[N_T]}{T} - \lambda\right)t$$

As one can observe, the error grows up linearly.

The second consists in keeping "t" as a variable, thus without fixing a time horizon, therefore:

$$\begin{aligned}\epsilon_t &= (1+\theta)E[Z]\left(\frac{E[N_t]}{t} - \lambda\right)t \\ &= \lambda(1+\theta)E[Z]\left(\frac{\beta(-1+e^{(\alpha-\beta)t} - (\alpha-\beta)t)}{(\alpha-\beta)^2 t} + \frac{\lambda(-1+e^{(\alpha-\beta)t})}{(\alpha-\beta)t} - 1\right)t\end{aligned}$$

Studying the limits of the quantity above, it is possible to notice that:

$$\lim_{t\longrightarrow 0^+} \epsilon_t = \lim_{t\longrightarrow 0^+} (1+\theta)\lambda E[Z](\lambda - 1)t = 0$$
$$\lim_{t\to\infty} \epsilon_t = \lim_{t\to\infty} (1+\theta)\lambda E[Z](-\frac{\beta}{\alpha-\beta} - 1)t = +\infty$$

Remark 4.9. The limit above shows that the error grows linearly and thus explodes with $t \longrightarrow +\infty$, indeed the quantity $(-\frac{\beta}{\alpha-\beta} - 1)$ is always positive with $\beta > \alpha > 0$, because:

$$-\frac{\beta}{\alpha-\beta} > 1 \Longrightarrow \alpha > 0$$

Now, it is possible to extend the treatment to a fixed time interval $[t_1, t_2]$ with $t_2 > t_1 > 0$, considering thus the situation in which the insurance company, which uses the expected value principle on a compound Hawkes as loss, is updating the premium rate in the instant t_1. The problem becomes the study of:

$$(c - \tilde{c})(t_2 - t_1) = (1 + \theta)(E[L_{t_2}|\mathcal{F}_{t_1}] - \lambda E[Z])(t_2 - t_1)$$

First of all, let consider the conditional expected value of the loss, the following chain of equations holds:

$$\begin{aligned} E[L_{t_2}|\mathcal{F}_{t_1}] &= E[L_{t_2} - L_{t_1} + L_{t_1}|\mathcal{F}_{t_1}] \\ &= L_{t_1} + E[\sum_{n=1}^{N_{t_2}-N_{t_1}} |\mathcal{F}_{t_1}] \\ &= L_{t_1} + E[N_{t_2} - N_{t_1}|\mathcal{F}_{t_1}]E[Z] \end{aligned}$$

But, exploiting (4.12):

$$\begin{aligned} E[N_{t_2} - N_{t_1}|\mathcal{F}_{t_1}] &= E[N_{t_2}|\mathcal{F}_{t_1}] - N_{t_1} \\ &= \left(\frac{\beta(-1 + e^{(\alpha-\beta)(t_2-t_1)} - (\alpha - \beta)(t_2 - t_1))}{(\alpha - \beta)^2} + \frac{\lambda(-1 + e^{(\alpha-\beta)(t_2-t_1)})}{(\alpha - \beta)} \right) \end{aligned}$$

Therefore:

$$\begin{aligned} &(c - \tilde{c})t_2 = (1 + \theta)L_{t_1}t_2 + \\ &+ \lambda(1 + \theta)E[Z]\left(\frac{\beta(-1 + e^{(\alpha-\beta)(t_2-t_1)} - (\alpha - \beta)(t_2 - t_1))}{(\alpha - \beta)^2 t_2} + \frac{\lambda(-1 + e^{(\alpha-\beta)(t_2-t_1)})}{(\alpha - \beta)t_2} - 1 \right) t_2 \end{aligned}$$

Eventually, let compute as done before the following two limits:

$$\begin{aligned} &\lim_{t_2 \longrightarrow t_1^+} (c - \tilde{c})t_2 = (1 + \theta)(L_{t_1} - \lambda E[Z])t_1 \\ &\lim_{t_2 \to \infty} (c - \tilde{c})t_2 = \lim_{t_2 \to \infty} (1 + \theta)L_{t_1}t_2 + \lambda E[Z](-\frac{\beta}{\alpha - \beta} - 1)t_2 = +\infty \end{aligned}$$

As expected, with $t_2 \longrightarrow t_1^+$, the error assumes exactly the real loss accumulated until t_1 minus the premiums gained considering it as a compound Poisson, multiplied for the parameter $(1+\theta)$, whilst the error explodes for $t_2 \longrightarrow +\infty$ linearly, as in the first approach.

4.4. Another choice of the premium rate in the risk model with Hawkes

In literature, in risk models with stochastic intensity λ_t, it is also used as possible choice of the premium rate c a quantity proportional straightforwardly to the intensity (see [6] for an example applied to the optimal reinsurance problem in the Hawkes framework). Inspired by the net value principle, let consider as premium rate the following:

$$c = \lambda_t E[Z] \tag{4.18}$$

This is justified by the fact that the stochastic intensity of the loss is correlated to its mean and that the writing (4.18) is the one which straightforwardly adapts the net value principle of (3.1) to more complicated models. This work, however, has already shown that the actual mean of the loss process with stochastic intensity such as the Compound Hawkes is a completely different quantity with respect to (4.18). Moreover, using as estimation for c the intensity of the counting process presents some drawbacks: indeed, it is a random variable $\forall t > 0$, so, doing computations and simulations can be more expensive. Anyway, it can be shown that it is a reasonable choice in the context of the Hawkes process. Let consider, in fact, the condition:

$$E[ct - L_t] \geq 0$$

It becomes:

$$E[\lambda_t E[Z]t - L_t] = (E[\lambda_t]t - E[N_t])E[Z] \geq 0$$

Exploiting (4.13) and (4.14), considering $\lambda_0 = \lambda$, one can have:

$$\left(te^{(\alpha-\beta)t}(\lambda + \frac{\beta\lambda}{\alpha-\beta}(1 - e^{-(\alpha-\beta)t}) - \frac{\lambda\beta(-1 + e^{(\alpha-\beta)t} - (\alpha-\beta)t)}{(\alpha-\beta)^2} - \frac{\lambda(-1 + e^{(\alpha-\beta)t})}{\alpha-\beta} \right) E[Z] \geq 0$$

Studying the derivative of this quantity, after some easy computations, one can obtain:

$$\frac{d}{dt}E[ct - L_t] = \alpha t e^{(\alpha-\beta)t}\lambda E[Z] \geq 0$$

This shows that the function studied is increasingly monotone. Moreover:

$$\lim_{t \longrightarrow 0^+} E[ct - L_t] = 0$$
$$\lim_{t \to +\infty} E[ct - L_t] = -\frac{\lambda}{\alpha - \beta}E[Z] > 0$$

In conclusion, the principle (4.18) enjoys the desirable property of increasing on average the portfolio surplus without exploding for large value of t.

Bibliography

[1] Asmussen. *Applied probability and queues.* Wiley, 1989.

[2] H. Asmussen and Taksar. Optimal risk control and dividend distribution policies. example of excess-of loss reinsurance for an insurance corporation. *Finance and stochastics*, (4):299–324, 2000.

[3] Baldi. *Stochastic Calculus: An Introduction Through Theory and Exercises*. Springer, 2017.

[4] Boikov. The cramér-lundberg model with stochastic premium process. *Theory of Probability and Its Applications, vol.47*, (3):489–493, 2003.

[5] Bremaud. *Point Processes and Queues: Martingale Dynamics*. Springer, 1981.

[6] L. Cao, Landriault. Optimal reinsurance-investment strategy for a dynamic contagion claim model. *Insurance: Mathematics and Economics*, (93):206–215, 2020.

[7] da Fonseca and Zaatour. Hawkes process: fast calibration, application to trade clustering and diffusive limit. *The Journal of Futures Markets, vol. 34*, (6):548–579, 2013.

[8] Dassios and Zhao. Exact simulation of hawkes process with exponentially decaying intensity. *Electronic communication in probability, vol. 18*, (62):1–13, 2013.

[9] Denneberg. *Non-additive measure and integral.* Kluwer academic publishers, 1994.

[10] Dhaene, Denuit, Goovaerts, Kaas, and Vyncke. The concept of comonotonicity in actuarial science and finance: Theory. *Insurance: Mathematics and Economics*, (31), 2002.

[11] Goovaerts and Dhaene. On the characterization of wang's class of premium principles. *DTEW Research Report*, (9740), 1997.

[12] Hawkes. Spectra of some self-exciting and mutually exciting point processes. *Biometrika, vol. 58*, (1):83–90, 1971.

[13] Hawkes. Hawkes processes and their applications to finance: a review. *Quantitative finance, vol. 18*, (2):193–198, 2018.

[14] Hurlimann. On stop-loss order and the distortion pricing principle. *ASTIN Bulletin: The Journal of the IAA*, (28):119–134, 1998.

[15] Jasiulewicz. Probability of ruin with variable premium rate in a markovian environment. *Insurance: Mathematics and Economics*, (29):291–296, 2001.

[16] Kaas, V. Heerwaarden, and Goovaerts. *Ordering of actuarial risks*. Springer, 1994.

[17] Laub, Taimre, and Pollet. Hawkes processes. 2017.

[18] Liang and Bayraktar. Optimal reinsurance and investment with unobservable claim size and intensity. *Insurance: Mathematics and Economics*, (55):156–166, 2014.

[19] Nendel, Riedel, and Schmeck. A decomposition of general premium principles into risk and deviation. *Insurance: Mathematics and Economics*, (100):193–209, 2021.

[20] Schmidli. *Risk theory*. Springer, 2017.

[21] Swishchuk. Risk model based on general compound hawkes process. *Wilmott magazine*, (94):50–56, 2018.

[22] Swishchuk, Zagst, and Zeller. Hawkes processes in insurance: Risk model, application to empirical data and optimal investment. *Insurance: Mathematics and Economics*, (21):107–124, 2021.

[23] Wang, Young, and Panjer. Axiomatic characterization of insurance prices. *Insurance: Mathematics and Economics*, (21):173–183, 1997.

[24] Young. Premium principles. *Encyclopedia of Actuarial Science*, 2004.