Dissertation
zur Erlangung des Doktorgrades
an den Naturwissenschaftlichen Fachbereichen
(Mathematik)
der Justus-Liebig-Universität Gießen

Martingale-Transformations Of Point-Processes And Their Applications

Martingaltransformationen von Punktprozessen und ihre Anwendungen

vorgelegt von
Lukas Gröber

05.04.2024

Betreuer
Prof. Dr. Dr. h.c. Winfried Stute

Für meine Großeltern

Zusammenfassung

Die nichtnegativen Martingale stellen einen bedeutsamen Teil in der Prozesstheorie dar. Der wichtigste Beitrag dieser Arbeit besteht darin, Möglichkeiten anzubieten, um einen allgemeinen Punktprozess N der durch seine bedingte Intensität λ charakterisiert wird, in genau solch ein Objekt zu transformieren. Die dadurch erhaltenen Systeme ermöglichen die Entwicklung von Tests zur Untersuchung von statistischen Fragestellungen.

Historisch gesehen sind Martingaltransformationen zum ersten Mal für Itô-Prozesse untersucht worden. Mit einer Brownschen Bewegung als triggernden Prozess hat man in diesem Fall eine rückwärtsgerichtete Wärmeleitungsgleichung zu lösen. Diese Theorie hat sich über Jahrzehnte hinweg erfolgreich weiterentwickelt und lässt sich heute in verallgemeinerter Form für Semimartingale formulieren.

In der vorliegenden Arbeit werden diese Resultate unter anderem genutzt, um Martingale zu verallgemeinern, die bereits im Zusammenhang mit der empirischen Verteilungsfunktion bekannt sind. Im Rahmen dieses Ansatzes ergeben sich verschiedene Klassen von Lösungen, deren zugehörige Prozesse verschiedene Charakteristiken aufweisen. Somit ermöglichen wir es, einen natürlichen Zugang zur Klasse der Poisson-Charlier Funktionen als Lösung der Differentialgleichung für den Fall eines Poisson-Prozesses anzugeben. Die Poisson-Charlier Funktionen werden in einem weiteren Schritt für allgemeinere Punkprozesse generalisiert. Für die Generierung der nicht-negativen Martingale bedienen wir uns zudem der Integraldarstellung für Martingale, indem wir charakteristische Eigenschaften an die erzeugenden vorhersehbaren Prozesse stellen.

Abstract

Non-negative martingales represent an important aspect of process theory. The main contribution of this thesis is to present methods to transform a general point process N, characterized by its conditional intensity λ, into a process that satisfies the properties of a non-negative martingale. The resulting systems allow the development of techniques for the investigation of statistical problems.

Historically, martingale transformations have been studied for the first time with respect to Itô processes. In this case, where the triggering process is a Brownian motion, a backward heat equation must be solved. This theory has evolved successfully over several decades and can now be formulated in a generalized form for semimartingales.

In the present work, these results are used, among other things, to generalize martingales that are already known in connection with the empirical distribution function. Within this approach, various solutions arise, whose associated processes exhibit different characteristics. This allows us to have a natural access to the class of Poisson-Charlier functions as a solution to the differential equation for the case of a Poisson process. The Poisson-Charlier functions are further generalized in a subsequent step for more general point processes. To generate non-negative martingales, we also make use of the integral representation for martingales by imposing characteristic properties on the generating predictable processes.

Contents

List of Figures

Chapter 1

Introduction

Point processes on the real line form an important class of stochastic processes that allow far-reaching applications in fields as diverse as market research, service systems, and actuarial mathematics. Within this introductory chapter the basic techniques and concepts will be explained. Since it is a summary of already known results, we will refrain from providing any proof in this part. Rather, it is a list of valuable tools that will be used in later investigations and thus form the foundation of this work.

The first question that arises when writing a treatise on point processes is the choice of the mathematical characterization. If one takes a look at the common literature, there are a few possibilities. In a nutshell, it involves constructing a discrete random set of points taking values on the real line.

One particular case of point process theory can be found in the renewal theory. Here, one has given a monotone sequence of random variables $T_1, T_2, ...$, whose intermediate times $X_i = T_{i+1} - T_i$ represent independent and identically distributed random variables. However, the iid assumption regarding the intermediate times turns out to be a limitation when describing real world phenomena.

Another method that has influenced a large part of the literature is the measure theoretical approach. A wealth of valuable resources can be found in references like [21], [27] and [29].

In a first step we want to define the class of time-continuous point processes we refer to in all investigations and implement the definition of a point process used in this work. For the upcoming definition, we assume throughout that $s_0 = 0, s_1 \in \mathbb{R}$ with $s_0 < s_1$.

Definition 1.0.1 (Simple, locally finite counting processes on $I = [s_0, s_1)$)
Let $\mathcal{T}_1(I)$ be the space of all right-continuous step functions with jump size 1 starting in 0. This means that an arbitrary $f \in \mathcal{T}_1(I)$ satisfies the following properties:

1. $f(s_0) = 0$.
2. For all $t \in I$, $\lim_{s\downarrow t} f(s)$ and $f(t) \in \mathbb{N}_0$.
3. For all $t \in I$, $f(t) - \lim_{s\uparrow t} f(s) \in \{0, 1\}$.

A stochastic process $N = (N_t)_{t\in I}$ with values in $\mathcal{T}_1(I)$ is called a simple locally finite counting process on the set $I = [s_0, s_1)$.

In this thesis a simple locally finite counting process N is just called a counting process. The collection of points given by

$$\left\{ x \in I : f(x) - \lim_{s\uparrow x} f(s) = 1 \right\}$$

constitutes the corresponding point process. Since the point process, defined as a collection of points, and the counting process contain the same information, and given that we will work with martingale transformations applied to the counting process, we will refer to the counting process N simply as a point process. Referring to the counting process as a point process follows the notation used in the classical work of Brémaud (see Chapter 2, page 19 in [5]). Note that we consider point processes characterized by unit jump sizes, with their sample paths represented as càdlàg step functions supported on the real line, starting at $s_0 = 0$. Consequently, we can define an order on each realization of the point process, where the random variable T_i denotes the i-th smallest point in the set. For notational convenience we set $T_0 = 0$.

Before considering an explicit example of a point process, we proceed with the popular Doob-Meyer-decomposition, which serves as another cornerstone of our analysis. Throughout this work we assume a given filtered, complete probability space $(\Omega, \mathcal{H}, (\mathcal{H}_t)_{t\in I}, \mathbb{P})$, where the filtration $(\mathcal{H}_t)_{t\in I}$ is assumed to be right continuous.

Definition 1.0.2 ($\mathcal{H}$-Martingale) (see Protter [31] p. 7)
Let $(M_s)_{s\in I}$ be a real-valued stochastic process adapted to the filtration $(\mathcal{H}_s)_{s\in I}$. M is called a $\mathcal{H}$-martingale if the conditions

$$\mathbb{E}\left[|X_t|\right] < \infty, \text{ for all } t \in I$$

and

$$\mathbb{E}[M_t \mid \mathcal{H}_s] = M_s,$$

are satisfied for all $s \leq t$.

Roughly said, a martingale is a collection of random variables for which, at a particular time, the conditional expectation of a future value is equal to the present value, regardless of all prior values.

An important result in martingale theory, established by Doob [9], states that any integrable discrete-time process Z can be uniquely decomposed into the sum of a martingale M and a predictable process A, with $A_0 = 0$. A continuous-time analogue of this decomposition, for processes in the so-called class D, was later established by Meyer [28]. A concise proof of the theorem can be found in Beiglboeck et al [3].

Theorem 1.0.3 (Doob-Meyer)
An adapted process $X = (X_t)_{t\in I}$ is of class D if the family of random variables X_τ where τ ranges through all stopping times is uniformly integrable. Let $X = (X_t)_{t\in I}$ be a càdlàg submartingale of class D. Then, X can be written in a unique way in the form

$$X_t = M_t + A_t,$$

where $M = (M_t)_{t\in I}$ is a martingale and $A = (A_t)_{t\in I}$ is a predictable increasing process starting at 0.

Now, since we have given a rigorous definition of a point process, we want to construct a trivial example and derive some additional properties.

Construction 1.0.4 (Single-Event Process)
Let X be a real random variable with distribution F. We define the stochastic process $N = (N_t)_{t\in\mathbb{R}_+}$ by

$$N_t = 1_{\{X\leq t\}},$$

where 1_A denotes the characteristic function of a set A. Clearly, the process N is a point process. Next we introduce the time-discrete filtration $\mathcal{H} = (\mathcal{H}_{t_n})_{n\in\mathbb{N}}$ by

$$\mathcal{H}_{t_n} = \sigma\left(\{N_{t_l} : 1 \leq l \leq n\}\right),$$

whereby $t_1, t_2, ...$ is an increasing sequence of real numbers. For all $l \in \mathbb{N}$ we can rewrite the random variable N_{t_l} as

$$\begin{aligned} N_{t_l} &= 1_{\{X\leq t_l\}} \\ &= 1_{\{X\leq t_l\}}1_{\{X\leq t_{l-1}\}} + 1_{\{X\leq t_l\}}1_{\{X>t_{l-1}\}} \\ &= 1_{\{X\leq t_{l-1}\}} + 1_{\{t_{l-1}<X\leq t_l\}}. \end{aligned}$$

Note that the first indicator is $\mathcal{H}_{t_{l-1}}$-measurable. Hence the conditional expectation $\mathbb{E}[N_{t_l} \mid \mathcal{H}_{t_{l-1}}]$ can be expressed by

$$\begin{aligned} \mathbb{E}[N_{t_l} \mid \mathcal{H}_{t_{l-1}}] &= 1_{\{X\leq t_{l-1}\}} + \mathbb{E}[1_{\{t_{l-1}<X\leq t_l\}} \mid \mathcal{H}_{t_{l-1}}] \\ &= 1_{\{X\leq t_{l-1}\}} + \frac{1_{\{t_{l-1}<X\}}\int_{\{t_{l-1}<X\}} 1_{\{t_{l-1}<X\leq t_l\}}d\mathbb{P}}{\mathbb{P}[t_{l-1}<X]} \\ &= 1_{\{X\leq t_{l-1}\}} + 1_{\{t_{l-1}<X\}}\frac{F(t_l)-F(t_{l-1})}{1-F(t_{l-1})}. \end{aligned}$$

The conditional expectation derived above is interesting since it yields the Doob-Meyer-decomposition

$$M_{t_l} = 1_{\{X\leq t_l\}} - \sum_{j=1}^{l} 1_{\{X>t_{j-1}\}}\frac{F(t_j)-F(t_{j-1})}{1-F(t_{j-1})}.$$

If one chooses a suitable grid sequence, the time-continuous version is obtained as

$$M_t = 1_{\{X\leq t\}} - \int_{(-\infty,t]} \frac{1_{\{v\leq X\}}}{1-F(v-)}dF(v)$$

and

$$A_t = \int_{(-\infty,t]} \frac{1_{\{v\leq X\}}}{1-F(v-)}dF(v),$$

with respect to the filtration $\mathcal{H}$ given by

$$\mathcal{H}_t = \sigma(1_{\{X \leq s\}} : s \leq t).$$

The previous investigations motivate the definition of the cumulative Hazard-function

$$\tilde{\Lambda}(t) = \int_{(-\infty,t]} \frac{1}{1-F(v-)} dF(v).$$

Furthermore, if F admits a Lebesgue-density f, the Hazard function is defined as

$$\tilde{\lambda}(v) := \frac{f(v)}{1-F(v)}.$$

The cumulative Hazard-function is interesting since the underlying distribution function may be obtained as

$$1 - F(t) = exp\left(-\tilde{\Lambda}(t) + \sum_{v \leq t} \Delta\tilde{\Lambda}(v)\right) \prod_{v \leq t} [1 - \Delta\tilde{\Lambda}(v)].$$

Thus, instead of characterizing the single event process via the distribution function F, one is able to characterize it using the hazard function. It will turn out, that this kind of characterization is closely related to the description of point processes via the so-called conditional intensity. Roughly speaking, the unconditional distribution F of any point X is no longer of central importance, but more its behavior conditioned by the past before t.

Construction 1.0.5 (The empirical distribution function)
One way to generalize the Single-Event process is to increase the sample size. Let $X_1, X_2, ..., X_n$ be a collection of iid random variables with a distribution function F. Furthermore, assume that F admits a Lebesgue density f. The renormalized empirical distribution function $N = n \cdot F_n$ is given by

$$N_t = \sum_{i=1}^{n} 1_{\{X_i \leq t\}}.$$

Clearly N is a point process, characterized by the iid structure and the distribution function F. For instance, the process N is adapted to the filtration $\mathcal{H}^S$ defined by

$$\mathcal{H}^S_s = \sigma\left(1_{\{X_i \leq v\}} : v \leq s, 1 \leq i \leq n\right).$$

Some straightforward calculations yield the $\mathcal{H}^S$-innovation martingale M of the Doob-Meyer decomposition

$$M_s = N_s - \int_{(-\infty,s]} \frac{n - N_{v-}}{1 - F(v-)} F(dv).$$

Clearly, $\mathcal{H}^S$ is not the only considerable filtration. The process N is for example also adapted to the natural generated filtration $\mathcal{H}$ defined by

$$\mathcal{H}_s := \sigma\left(N_v : v \leq s\right).$$

It turns out (see Hess [19], p. 26, Theorem 3.3) that in this particular case the Doob-Meyer decomposition is equal to that in the case of the filtration $\mathcal{H}^S$. Note that this may seem surprising at first glance, since the two respective filtrations contain different information.

Lastly, we want to give a brief motivation of conditional intensities which are of fundamental importance in the field of point process theory. To guarantee the existence of a conditional $\mathcal{H}$-intensity as it is described in the common literature we throughout make the assumption that F admits a Lebesgue-density f. In this context the conditional $\mathcal{H}$-intensity is given by

$$\lambda_v = \frac{n - N_{v-}}{1 - F(v)} f(v).$$

Consequently, the conditional $\mathcal{H}$-intensity is the integrator of the compensator with respect to the Lebsegue measure

$$\begin{aligned} M_s &= N_s - \int_{(-\infty,s]} \frac{n - N_{v-}}{1 - F(v-)} F(dv) \\ &= N_s - \int_{(-\infty,s]} \frac{n - N_{v-}}{1 - F(v)} f(v)dv \\ &= N_s - \int_{(-\infty,s]} \lambda_v dv. \end{aligned}$$

An important technique, which has its origin in stochastic analysis with a Brownian motion B as triggering process, consists in applying a deterministic function $\varphi : I \times \mathbb{R} \to \mathbb{R}$ to the underlying process $(B_t)_{t \in I}$ such that the predictable process A disappears.

Martingale theory is a well-known field in stochastics and offers a wide range of applications. In summary, the following steps are necessary to transform a stochastic process into a martingale with respect to a given filtration:

1. Determine for a general smooth function φ in two variables, the Doob-Meyer decomposition of the process $(\varphi(Z_t, t))_{t\in I}$.
2. Determine the set of functions φ_M, for which the predictable process in the Doob-Meyer decomposition disappears. For this purpose a differential equation has to be solved. In general, the differential equation has a stochastic character.

Let F be a distribution function that admits a Lebesgue density f. As demonstrated at the conclusion of this chapter, and in accordance with Itô's formula for semimartingales, it will be shown that the Doob-Meyer-decomposition of the process $(\varphi(n\cdot F_n(t), t))_{t\in\mathbb{R}}$ with respect to its natural generated filtration $\mathcal{H}$ is given by

$$\begin{aligned}\varphi(n \cdot F_n(s), s) = M_s + \int_{(-\infty,s]} &\frac{\partial\varphi}{\partial x_2}(F_n(v), v)\\ &+ \left(\varphi(F_n(v) + \frac{1}{n}, v) - \varphi(F_n(v), v)\right) \frac{n(1-F_n(v))}{1-F(v)} f(v)dv.\end{aligned}$$

Therefore, the process $(\varphi(F_n(t), t))_{t\in\mathbb{R}}$ turns out to be a martingale if φ is a solution of the differential equation

$$\frac{\partial\varphi}{\partial x_2}(x, s) + \left(\varphi(x + \frac{1}{n}, s) - \varphi(x, s)\right) \frac{n(1-x)}{1-F(s)} f(s) = 0.$$

As it is shown in the following result, the basis of solutions corresponding to this equation depends in the case of the empirical distribution function on the underlying sample size n.

Theorem 1.0.6 (Polynomial martingales with respect to the empirical distribution function) [19]
Let $m \in \{0, 1, ..., n\}$ be fixed but arbitrary. The stochastic process $X = (X_t)_{t\in\mathbb{R}}$ given by

$$X_s = \frac{(n-m)!n^m}{n!} \prod_{j=n-(m-1)}^{n} \frac{\frac{j}{n} - F_n(s)}{1-F(s)},$$

is a martingale with respect to the natural filtration of $(F_n(t))_{t\in\mathbb{R}}$.

In a further step one can show that each martingale generated via a transformation in time and space $\varphi(F_n(s), s)$, can be represented as a linear combination of these polynomial martingales.

Without entering into an in-depth discussion of the properties and possible applications of these processes, one matter should be emphasized. The stochastic process $(X_t)_{t\in\mathbb{R}}$ is not only a martingale, it is also nonnegative. This circumstance proves to be an advantage for the application of stopping techniques. For example, set $m = 1$ in Theorem 1.0.6. to obtain the process

$$X_s = \frac{1 - F_n(s)}{1 - F(s)}. \tag{1.1}$$

This special case is a well known process in the literature [36] (P. 269 Proposition 2). Since we assumed F to be continuous, one can show the identity

$$\mathbb{P}\left[\sup_{t\in\mathbb{R}}\left(\frac{1 - F_n(s)}{1 - F(s)}\right) > c\right] = \frac{1}{c}. \tag{1.2}$$

The objective of this work is to obtain a variety of martingales with the character of (1.2) in the context of point process theory.

We now present a concise overview of the key results concerning conditional intensities. Our definition of a point process as locally finite ensures that it is non-explosive, meaning for all $t > 0$ we have $N_t < \infty$ $\mathbb{P}$-almost surely.

Definition 1.0.7 (Stochastic Intensity) (see Brémaud [5], p. 27)
Let $N = (N_t)_{t\in I}$ be a point process adapted to some filtration $\mathcal{H}$, and let $\lambda = (\lambda_t)_{t\in I}$ be a nonnegative $\mathcal{H}$-progressive process such that for all $t \geq 0$

$$\int_0^t \lambda_s ds < \infty,$$

$\mathbb{P}$-almost sure. If for all nonnegative $\mathcal{H}$-predictable processes $H = (H_t)$, the equality

$$\mathbb{E}\left[\int_{s_0}^{s_1} H_s dN_s\right] = \mathbb{E}\left[\int_{s_0}^{s_1} H_s \lambda_s ds\right]$$

is satisfied, then we say: N admits the $\mathcal{H}$-intensity λ.

Theorem 1.0.8 (Martingale Characterization of Intensity) (see Brémaud [5], p. 28)
Let $N = (N_t)_{t \in I}$ be a non-explosive point process adapted to $(\mathcal{H}_t)_{t \in I}$, and suppose that for some nonnegative $\mathcal{H}_t$-progressive process $\lambda = (\lambda_t)_{t \in I}$ and for all $n \geq 0$,

$$N_{t \wedge T_n} - \int_0^{t \wedge T_n} \lambda_s ds,$$

is a martingale. Then λ is the $\mathcal{H}$-intensity of N.

As been highlighted in [5] the martingale characterization of the conditional intensity implies

$$\mathbb{E}[N_t - N_s \mid \mathcal{H}_s] = \mathbb{E}\left[\int_s^t \lambda_v dv \mid \mathcal{H}_s\right],$$

for $0 \leq s \leq t$. Furthermore, if λ is right continuous and bounded, it $\mathbb{P}$-almost sure follows that

$$\lim_{t \downarrow s} \frac{1}{t-s} \mathbb{E}[N_t - N_s \mid \mathcal{H}_s] = \lambda_s.$$

An Introduction to Semimartingales

Lastly, we would like to provide an excursus on the theory of semimartingales to enhance our understanding of the presented results and establish a foundation for the upcoming investigations. Throughout this section we assume $I = \mathbb{R}_+$. The results presented here are oriented on the textbook by Protter [31].

The class of semimartingales encompasses a broad range of stochastic processes for which a stochastic integral can be defined. In this section, we will reference Itô's formula for these integrals, which will aid in enhancing our understanding of the results presented in this work. To establish the theory of semimartingales we start with some basic definitions.

Note that a stopping time T is a random variable satisfying the condition that, for any $t \in I$, the set $\{T \leq t\}$ is measurable with respect to $\mathcal{H}_t$.

Definition 1.0.9 (Simple predictable processes) (see Protter [31], p. 51)
A process $(H_t)_{t \in I}$ is said to be simple predictable if the process has the representation

$$H_t = H_0 1_{\{0\}}(t) + \sum_{i=1}^{n} H_i 1_{(T_i, T_{i+1}]}(t),$$

where $0 = T_1 \leq T_2 \leq ... \leq T_{n+1} < \infty$ is a finite sequence of stopping times, $H_i \in \mathcal{F}_{T_i}$ with $|H_i| < \infty$ a.s., $0 \leq i \leq n$. The collection of simple predictable processes is denoted $\mathcal{S}$.

We can give $\mathcal{S}$ a topology by considering uniform convergence in (t, w), and we denote the topological space obtained by this procedure as $\mathcal{S}_u$. Additionally, we use L^0 to refer to the space of finite-valued random variables, which is topologized by convergence in probability. These topologies provide the framework necessary to analyze the convergence of a sequence of stochastic processes to a limiting process.

Let $X = (X_t)_{t \in I}$ be a stochastic process. To define an integral operator for the process X, two fundamental properties must naturally be satisfied to justify its classification as an integral. The operator should be linear and must satisfy a version of the Bounded Convergence Theorem. A particularly weak form of this theorem is that the uniform convergence of processes H_n to H implies at least the convergence in probability of the integral

$$\lim_{n \to \infty} \int H_s^n dX_s = \int H_s dX_s.$$

Building on this discussion, it is natural to define the integral for $H \in \mathcal{S}$ as a linear mapping from $\mathcal{S}$ to L^0:

$$\int H dX = H_0 X_0 + \sum_{i=1}^{n} H_i (X_{T_{i+1}} - X_{T_i}).$$

Definition 1.0.10 (Total semimartingale) (see Protter [31], p. 52)
A stochastic process $X = (X_t)_{t \in I}$ is said to be a total semimartingale if X is right-continuous with left limits, adapted and the mapping $\mathcal{S} \to L^0$

$$H \to \int H dX,$$

is continuous.

For a given stochastic process $X = (X_t)_{t\in I}$ we construct the stopped process $(X_t^T)_{t\in I}$ through $X_s^T = X_{s\wedge T}$.

Definition 1.0.11 (Semimartingale) (see Protter [31], p. 52)
A stochastic process $X = (X_s)_{s\in I}$ is said to be a semimartingale if, for each $t \in I$ the process $X^t = (X_s^t)_{s\in I}$ is a total semimartingale.

The class of semimartingales constitutes a broad and fundamental category of stochastic processes. Notable examples, such as the Brownian motion and the Poisson process, are encompassed within this framework. The definition of semimartingales is motivated by their role as suitable integrators for simple predictable processes. This class can be enlarged to the class of processes that are adapted with left continuous paths and right limits. (see Protter [31], p. 56)

The class of processes with finite variation plays a particularly important role in this study.

Proposition 1.0.12 (Finite variation processes) (see Protter [31], p. 55)
Each adapted process with right-continuous left limits paths of finite variation on compacts is a semimartingale.

Theorem 1.0.13 (Itô's formula) (see Protter [31], p. 81)
Let $X = (X_1, ..., X_n)$ be an n-tuple of semimartingales and let $\varphi : \mathbb{R}^n \to \mathbb{R}$ have continuous second order partial derivatives. Then $\varphi(X)$ is a semimartingale and

$$\begin{aligned}\varphi(X_t) =&\varphi(X_0) + \sum_{i=1}^{n} \int_0^t \frac{\partial \varphi}{\partial x_i}(X_{s-}) dX_s^i \\ &+ \frac{1}{2} \sum_{1\leq i,j\leq n} \int_0^t \frac{\partial^2 \varphi}{\partial x_i \partial x_j}(X_{s-}) d[X^i, X^j]_s^c \\ &+ \sum_{0\leq s\leq t} \left(\varphi(X_s) - \varphi(X_{s-}) - \sum_{i=1}^{n} \frac{\partial \varphi}{\partial x_i}(X_{s-}) \Delta X_s \right)\end{aligned}$$

Within the scope of this work, we consider processes of finite variation. Consequently, the corresponding term in Itô's general formula for semimartingales vanishes, requiring only that the function φ possesses continuous first-order derivatives. The following Lemma provides a formulation of Itô's Lemma that aligns with the requirements of this study as a direct consequence. The simplified formula can also be found in Protter, Theorem 31 on page 78, given the assumption that the point process has only jumps of size 1.

Lemma 1.0.14 (Itô's formula)
Let $N = (N_t)_{t \in I}$ be a point process and let $\varphi : \mathbb{R}^2 \to \mathbb{R}$ have continuous first order partial derivatives. Then $(\varphi(N_t, t))_{t \in I}$ is a semimartingale and

$$\varphi(N_t, t) = \int_0^t \frac{\partial \varphi}{\partial x_2}(N_{s-}, s) ds + \int_0^t \left(\varphi(N_{s-} + 1, s) - \varphi(N_{s-}, s)\right) dN_s.$$

Proof.

Using the formula from Theorem 1.0.11 under the consideration that N is a process of finite variation with $\Delta N_s \in \{0, 1\}$ we obtain

$$\begin{aligned}
\varphi(N_t, t) &= \varphi(N_0, 0) + \int_0^t \frac{\partial \varphi}{\partial x_1}(N_{s-}, s) dN_s + \int_0^t \frac{\partial \varphi}{\partial x_2}(N_{s-}, s) ds \\
&\quad + \sum_{0 \le s \le t} \left(\varphi(N_s, s) - \varphi(N_{s-}, s) - \frac{\partial \varphi}{\partial x_1}(N_{s-}, s) \right) \\
&= \varphi(N_0, 0) + \int_0^t \frac{\partial \varphi}{\partial x_2}(N_{s-}, s) dN_s^i + \sum_{0 \le s \le t} \left(\varphi(N_s, s) - \varphi(N_{s-}, s)\right) \\
&= \varphi(N_0, 0) + \int_0^t \frac{\partial \varphi}{\partial x_2}(N_{s-}, s) ds + \int_0^t \left(\varphi(N_s, s) - \varphi(N_{s-}, s)\right) dN_s.
\end{aligned}$$

□

Example 1.0.15 We continue our discussion on the empirical distribution function F_n in order to derive the Doob-Meyer decomposition of the process $\varphi(F_n(t), t)$. Let F be a distribution function with support on the positive real axis and define the point process $N = n \cdot F_n$. Furthermore assume that F admits a density f. Define the stochastic process

$\lambda = (\lambda_t)_{t\in I}$ by

$$\lambda_t = \frac{n - N_{t-}}{1 - F(t)} f(t).$$

Let $M^0 = (M_t^0)_{t\in I}$ be the martingale

$$M_t^0 = N_t - \int_0^t \lambda_s ds.$$

Using Itô's formula it follows that

$$\begin{aligned}
\varphi(N_t, t) =& \varphi(N_0, 0) + \int_0^t \frac{\partial\varphi}{\partial x_2}(N_{s-}, s) ds + \int_0^t \left(\varphi(N_{s-} + 1, s) - \varphi(N_{s-}, s)\right) dN_s \\
&= \varphi(N_0, 0) + \int_0^t \frac{\partial\varphi}{\partial x_2}(N_{s-}, s) ds - \left(\varphi(N_{s-} + 1, s) - \varphi(N_{s-}, s)\right) \lambda_s ds \\
&+ \int_0^t \left(\varphi(N_s, s) - \varphi(N_{s-}, s)\right) dM_s^0.
\end{aligned}$$

Note that the integrator of the first integral is predictable. If we add some integrability assumptions to the function φ the integral process becomes a martingale. We will investigate this equation in a more general context in chapter 3 and provide the necessary additional assumptions for the process to become a martingale.

Chapter 2

Point processes in discrete-time

In this chapter we want to investigate point processes in discrete time. We start our discussion by recalling some basic properties of the Single-Event Process which we introduced in the past chapter. In one further step we provide a definition of a general point process in discrete time and figure out how we can uniquely determine the law in distribution of such a stochastic process. Lastly, we will give an overview of some popular martingales and also highlight out a preview on the techniques we figured out in this work. These techniques will be further discussed for the continuous-time case in the upcoming chapter and some beneficial properties will be derived.

Single-Event Process: Let $X \sim F$ be an arbitrary random variable. As mentioned before, we want to deal with discrete lattices within this chapter. Fix $d \in \mathbb{N}$ and choose an arbitrary sequence $t_1 < t_2 < .. < t_d$ of length $d \in \mathbb{N}$. The associated time-discrete Single-Event process $N = (N_{t_l})_{l=1,2,...,d}$ defined by

$$N_{t_l} = 1_{\{X \leq t_l\}},$$

was already discussed in the first chapter yielding the $\mathcal{H}$-innovation martingale M,

$$M_{t_l} = 1_{\{X \leq t_l\}} - \sum_{j=1}^{l} 1_{\{X > t_{j-1}\}} \frac{F(t_j) - F(t_{j-1})}{1 - F(t_{j-1})}.$$

We make the convention $0 \cdot \infty = 0$. Let $t \in I$ be arbitrary with $F(t) = 1$. Note that $N = \{\omega \in \Omega \mid 1_{\{X > t\}} = 1\}$ is a null set. Hence the martingale

M is well defined. Recall that $\mathcal{H}$ is the natural filtration of the underlying process N. To define the associated point process in discrete time we can introduce the process $Z = (Z_{t_l})_{l=1,2,...,d}$,

$$Z_{t_l} := 1_{\{X \in [t_l, t_{l+1})\}} = N_{t_l} - N_{t_{l-1}} =: \Delta N_l.$$

Note that each realization of Z is a standard unit vector on $\mathbb{R}^d$. The aim of this chapter is to generalize the results in the sense that we admit the process Z to take values in the set $\{0,1\}^d$. Hence Z is a collection of not necessarily independent Bernoulli variables. One obvious but perhaps not always useful way to determine the distribution of Z uniquely is via the common distribution function $G : \mathbb{R}^d \to [0,1]$,

$$G(z_1, z_2, ..., z_d) := \mathbb{P}[Z_1 \leq z_1, Z_2 \leq z_2, ..., Z_d \leq z_d].$$

To circumvent the potentially intricate structure of G, we will utilize conditional intensities. In this section, we shall dedicate significant focus to this critical aspect and explore the various ways in which it impacts our analysis.

2.1 Basics

Definition 2.1.1 (Point processes in finite time)
Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, $d \in \mathbb{N}$ and $A_1, A_2, ..., A_d \in \mathcal{F}$ be arbitrary. A random vector Z defined by

$$Z := (Z_1, Z_2, ..., Z_d) = (1_{A_1}, 1_{A_2}, ..., 1_{A_d})$$

is called a point process on $\{1, 2, ..., d\}$. The random vector N defined by

$$N := (1_{A_1}, 1_{A_1} + 1_{A_2}, ..., \sum_{i=1}^{d} 1_{A_i})$$

is called the counting process associated with the point process Z. Furthermore, we define $\mathcal{H} \equiv \mathcal{H}^Z$ as the filtration generated by the point process Z.

Basically, a point process is a random vector $Z = (Z_1, ..., Z_d)$ whose components Z_l are Bernoulli-variables. Each component N_l of the associated counting process is nothing but the cumulative sum of the process Z up to time l. If we know the realization of Z, we also know the realization of N and vice versa. This implies that the filtration $\mathcal{H}^Z$ is tantamount to the filtration generated by the associated counting process N.

From time to time we will admit the time-space to be countably infinite. This results in the following definition of a point process, which is nothing more than a slight modification of the space in which the point process takes its values.

Definition 2.1.2 (Point processes on a countable time-space)
Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $(A_l)_{l\in\mathbb{N}}$ be an arbitrary sequence in $\mathcal{F}$, that is for all $i \in \mathbb{N}$ we have $A_i \in \mathcal{F}$. A random sequence $Z = (Z_l)_{l\in\mathbb{N}}$ defined by

$$Z_l := 1_{A_l}$$

is called a point process on $\mathbb{N}$. The stochastic process $N = (N_l)_{l\in\mathbb{N}}$ defined by

$$N_l := \sum_{i=1}^{l} 1_{A_i}$$

is called the counting process associated with the point process Z. Furthermore, we define $\mathcal{H} \equiv \mathcal{H}^Z$ as the filtration generated by the point process Z.

Example 2.1.3 (Renewal process)
Renewal processes are a widely studied class of point processes in continuous time. They are also a subject of interest in the literature in the discrete-time case, as can be verified in reference [2]. In the context of this example, our goal is to construct renewal point processes in discrete time and establish the concept of conditional intensity through calculations. The formalization of conditional intensities will be explained in the upcoming definition.

Let $(Y_l)_{l\in\mathbb{N}}$ be a sequence of independent and identically distributed random variables with

$$\mathbb{P}[Y_1 = l] = p_l \text{ for } l \in \mathbb{N} \text{ and } \sum_{l=1}^{\infty} p_l = 1.$$

Define the process $(S_l)_{l\in\mathbb{N}_0}$ by

$$S_l = \sum_{i=1}^{l} Y_i.$$

The point process $Z = (Z_l)_{l\in\mathbb{N}}$ defined by

$$Z_l = 1_{\{\exists k\le l: S_k = l\}},$$

is called a renewal process. Since the sequence S is strictly increasing it follows that the associated counting process $N = (N_l)_{l\in\mathbb{N}}$ equals

$$\begin{aligned} N_l &= \sum_{i=1}^{l} 1_{\{\exists k\le i: S_k = i\}} \\ &= |\{k\in\mathbb{N} \mid S_k \in \{1,2,...,l\}\}| \\ &= \max\{k\le l \mid S_k \le l\}. \end{aligned} \tag{A}$$

Representation (A) is common in the time-continuous case. We define the filtration $\mathcal{H}^S$ by setting

$$\mathcal{H}^S_l := \sigma(\{S_k 1_{\{S_k\le l\}} \mid k\in\mathbb{N}\}).$$

Note that for all $k,l\in\mathbb{N}$ it follows that

$$S_k 1_{\{S_k\le l\}} = \min(\{i\le l \mid \sum_{j=1}^{i} Z_i = k\}\cup\{0\}) \in \mathcal{H}_l.$$

On the other hand for all $1\le i\le l$ it follows that

$$Z_i = 1_{\{\exists j\le i: S_j = i\}} = 1_{\{\exists j\le i: S_j 1_{\{S_j\le l\}} = i\}} \in \mathcal{H}^S_l.$$

Consequently, the two filtrations $\mathcal{H}^S$ and $\mathcal{H}$ consist of the same information for all timepoints $l\in\mathbb{N}$. One benefical application of this fact arises if one wants to calculate the probability for the event $\{Z_l = 1\}$ given the first $(l-1)$ random variables $Z_1, Z_2, ..., Z_{l-1}$. It follows that

$$\begin{aligned}
\mathbb{P}[Z_l = 1 \mid \mathcal{H}_{l-1}] &= \mathbb{E}[Z_l \mid \mathcal{H}_{l-1}] \\
&= \mathbb{E}\left[1_{\{S_{N_l}=l\}} \mid \mathcal{H}_{l-1}\right] \\
&= \mathbb{E}\left[1_{\left\{\sum_{i=1}^{N_{l-1}+1} Y_i = l\right\}} \mid \mathcal{H}_{l-1}\right] \\
&= \sum_{j=1}^{l-1} 1_{\left\{\sum_{i=1}^{N_{l-1}} Y_i = j\right\}} \mathbb{E}\left[1_{\left\{Y_{N_{l-1}+1}=l-j\right\}} \mid \mathcal{H}_{l-1}\right] \\
&= \sum_{j=1}^{l-1} \sum_{k=0}^{l-1} 1_{\{N_{l-1}=k\}} 1_{\left\{\sum_{i=1}^{k} Y_i = j\right\}} \mathbb{E}\left[1_{\{Y_{k+1}=l-j\}} \mid \mathcal{H}_{l-1}\right] \\
&= \sum_{j=1}^{l-1} \sum_{k=0}^{l-1} 1_{\{N_{l-1}=k\}} 1_{\{S_k=j\}} \mathbb{E}\left[1_{\{Y_{k+1}=l-j\}} \mid \mathcal{H}^S_{l-1}\right].
\end{aligned}$$

Let $(s_m)_{m\in\mathbb{N}}$ be an arbitrary sequence with $s_m = 0$ for all $m \geq k+1$, $s_k = j$ and $s_m < s_{m+1}$ for all $m < k+1$. It follows that

$$\begin{aligned}
&\mathbb{P}\Big[N_{l-1} = k, S_k = j, Y_{k+1} = l-j \mid S_1 1_{\{S_1 \leq l-1\}} = s_1, S_2 1_{\{S_2 \leq l-1\}} = s_2, ...\Big] \\
&= \frac{\mathbb{P}\Big[N_{l-1} = k, S_k = j, Y_{k+1} = l-j, S_1 = s_1, ..., S_{k-1} = s_{k-1}\Big]}{\mathbb{P}\Big[S_1 = s_1, ..., S_{k-1} = s_{k-1}, S_k = j, S_{k+1} > l-1\Big]} \\
&= \frac{\mathbb{P}\Big[S_k = j, Y_{k+1} = l-j, S_1 = s_1, ..., S_{k-1} = s_{k-1}\Big]}{\mathbb{P}\Big[S_1 = s_1, ..., S_{k-1} = s_{k-1}, S_k = j, S_{k+1} > l-1\Big]} \\
&= \frac{\mathbb{P}\Big[Y_k = j - s_{k-1}, Y_{k+1} = l-j, Y_1 = s_1, ..., Y_{k-1} = s_{k-1} - s_{k-2}\Big]}{\mathbb{P}\Big[Y_1 = s_1, ..., Y_{k-1} = s_{k-1} - s_{k-2}, Y_k = j - s_{k-1}, Y_{k+1} > l-1-j\Big]} \\
&= \frac{\mathbb{P}\Big[Y_{k+1} = l-j\Big]}{\mathbb{P}\Big[Y_{k+1} > l-1-j\Big]} \\
&= \frac{p_{l-j}}{\sum_{n=l-j}^{\infty} p_n}.
\end{aligned}$$

Combining the above computations yields

$$\begin{aligned}\mathbb{P}[Z_l = 1 \mid \mathcal{H}_{l-1}] &= \sum_{j=1}^{l-1}\sum_{k=0}^{l-1} 1_{\{N_{l-1}=k\}} 1_{\{S_k=j\}} \frac{p_{l-j}}{\sum_{n=l-j}^{\infty} p_n} \\ &= \frac{p_{l-S_{N_{l-1}}}}{\sum_{n=l-S_{N_{l-1}}}^{\infty} p_n}.\end{aligned}$$

In chapter one we introduced the compensator Λ of the Doob-Meyer decomposition which uniquely determines the distribution of an underlying point process. This circumstance will now be investigated in detail for time discrete point processes. The advantage to deal with λ instead of the distribution function $G : \mathbb{R}^d \to [0,1]$ from a point process Z on a finite lattice, might be that λ has a more suitable structure in certain applications. Remember in this context G was given by

$$G(z_1, z_2, ..., z_d) = \mathbb{P}[Z_1 \le z_1, Z_2 \le z_2, ..., Z_d \le z_d].$$

Definition 2.1.4 (Conditional $\mathcal{H}$-intensity)
Let $d \in \mathbb{N} \cup \{\infty\}$ be arbitrary and set $I := \{1, 2, ..., d\}$. For a point process $(Z_l)_{l\in I}$ define the predictable process $\lambda = (\lambda(1 \mid \mathcal{H}_{l-1}))_{l\in I}$ by

$$\begin{aligned}\lambda(1 \mid \mathcal{H}_{l-1}) &:= \mathbb{P}[Z_l = 1 \mid \mathcal{H}_{l-1}] \\ &= \mathbb{E}[Z_l \mid \mathcal{H}_{l-1}].\end{aligned}$$

The process is called the $\mathcal{H}$-conditional intensity of the point process Z.

Let $l \in I$ be arbitrary. Note that, since the filtration $\mathcal{H}$ is generated naturally by the point process N the random variable $\lambda(1 \mid \mathcal{H}_{l-1})$ is the conditinal expectation of the Bernoulli variable Z_l, given the first $(l-1)$ random variables $Z_1, Z_2, ..., Z_{l-1}$. Sometimes we will consider $\lambda(1 \mid \mathcal{H}_{l-1})$ as a regression function $\lambda_l : \{0,1\}^{l-1} \to [0,1]$, mapping an arbitrary element $x = (x_1, x_2, ..., x_{l-1}) \in \{0,1\}^{l-1}$ to the respective conditional probability of the event $\{Z_l = 1\}$, that is

$$\lambda_l(1 \mid x) = \mathbb{P}[Z_l = 1 \mid Z_1 = x_1, Z_2 = x_2, ..., Z_{l-1} = x_{l-1}].$$

The usefulness of conditional intensities in regard to point processes is expressed in the following theorem, which states that λ uniquely determines the distribution of a point process.

Theorem 2.1.5 (Uniqueness)
Let $(Z_l^1)_{l\in\{1,2,...,d\}}$ and $(Z_l^2)_{l\in\{1,2,...,d\}}$ be two point processes on a finite time-space and let the filtration $\mathcal{H}^i$ be generated naturally by Z^i respectively. If the conditional $\mathcal{H}^1$-intensity λ^{Z^1} and conditional $\mathcal{H}^2$-intensity λ^{Z^2} satisfy the condition

$$\lambda^{Z^1} =_{\mathcal{L}} \lambda^{Z^2}.$$

It follows that

$$Z^1 =_{\mathcal{L}} Z^2.$$

Proof.

Let $i_1, i_2, ..., i_d \in \{0,1\}$ be arbitrary but fixed. It follows that

$$\begin{aligned}
&\mathbb{P}[Z_1^1 = i_1, Z_2^1 = i_2, ..., Z_d^1 = i_d] \\
&= \mathbb{P}[Z_1^1 = i_1] \cdot \mathbb{P}[Z_2^1 = i_2 \mid Z_1^1 = i_1] \cdots \mathbb{P}[Z_d^1 = i_d \mid Z_1^1 = i_1, ..., Z_{d-1}^1 = i_{d-1}] \\
&= \mathbb{P}[Z_1^2 = i_1] \cdot \mathbb{P}[Z_2^2 = i_2 \mid Z_1^2 = i_1] \cdots \mathbb{P}[Z_d^2 = i_M \mid Z_1^2 = i_1, ..., Z_{d-1} = i_{d-1}] \\
&= \mathbb{P}[Z_1^2 = i_1, Z_2^2 = i_2, ..., Z_d^2 = i_d].
\end{aligned}$$

□

As a result, there is an opportunity to characterize point processes through their respective conditional intensities. A highly straightforward scenario arises when the conditional intensity assumes the form of a deterministic function that is entirely independent of the underlying process history.

Definition 2.1.6 (Bernoulli process)
Let $d \in \mathbb{N} \cup \{\infty\}$ be arbitrary and define $I := \{1, 2, ..., d\}$. A point process $Z = (Z_l)_{l\in I}$ is called a Bernoulli process with parameter ρ, if the conditional intensity $\lambda = (\lambda(1 \mid \mathcal{H}_l))_{l\in\mathbb{N}}$ of Z satisfies

$$\lambda(1 \mid \mathcal{H}_{l-1}) = \rho,$$

for some parameter $\rho \in [0, 1]$.

Due to the deterministic character of the conditional intensity, the Bernoulli process represents the discrete-time analogue of the Poisson process. If Z is a Bernoulli process it follows that

$$\mathbb{P}[Z_l = 1 \mid \mathcal{H}_{l-1}] = \mathbb{P}[N_l - N_{l-1} = 1 \mid \mathcal{H}_{l-1}] = \rho.$$

Furthermore, it follows that the process N has independent and binomial distributed increments

$$\mathbb{P}[N_l = i] = \binom{i}{l}\rho^i(1-\rho)^{l-i}.$$

It is straightforward to compute the most common quantities of a Bernoulli process as it is shown in the following calculations:

$$\begin{aligned}
\mathbb{E}[N_l] &= \sum_{i=0}^{l} i \cdot \mathbb{P}[N_l = i] \\
&= \sum_{i=0}^{l} i \cdot \binom{l}{i}\rho^i(1-\rho)^{l-i} \\
&= \sum_{i=0}^{l} i \frac{l!}{i!(l-i)!}\rho^i(1-\rho)^{l-i} \\
&= l\rho \sum_{i=1}^{l} \frac{(l-1)!}{(i-1)!(l-i)!}\rho^{(i-1)}(1-\rho)^{(l-1)-(i-1)} \\
&= l\rho \sum_{i=0}^{l-1} \frac{(l-1)!}{i!(l-i-1)!}\rho^{i}(1-\rho)^{(l-1)-i} \\
&= l\rho \sum_{i=0}^{l-1} \binom{l-1}{i}\rho^{i}(1-\rho)^{(l-1)-i} \\
&= l\rho(\rho + (1-\rho))^{l+1} \\
&= l\rho.
\end{aligned}$$

Let $m, n \in I$ be arbitrary with $m \leq n$. It follows that

$$\begin{aligned}
Cov(N_m, N_n) &= \mathbb{E}[N_m N_n] - \mathbb{E}[N_m]\mathbb{E}[N_n] \\
&= \mathbb{E}[N_m(N_n - N_m + N_m)] - m\rho n\rho \\
&= \mathbb{E}[N_m(N_n - N_m) + N_m^2] - mn\rho^2 \\
&= \mathbb{E}[N_m]\mathbb{E}[(N_n - N_m)] + \mathbb{E}[N_m^2] - mn\rho^2 \\
&= m\rho(n-m)\rho + m^2\rho^2 - m\rho^2 + m\rho - mn\rho^2 \\
&= mn\rho^2 - m^2\rho^2 + m^2\rho^2 - m\rho^2 + m\rho - mn\rho^2 \\
&= -m\rho^2 + m\rho \\
&= m\rho(1-\rho).
\end{aligned}$$

Polynomial martingales with respect to Bernoulli processes have already been treated in the literature [34] (P.59-60). In the following investigations we will derive these martingales with our own techniques by finding a generalized result taking into account an arbitrary point process in discrete time. As discussed in earlier chapters, we are looking for functions that satisfy the equation

$$\mathbb{E}[\varphi(N_n, n) \mid \mathcal{H}_m] = \varphi(N_m, m), \forall m \leq n.$$

One fundamental case arises by using the transformation

$$\varphi(x, k) = x - \rho \cdot k, \tag{2.1}$$

as it is shown by the following computations

$$\begin{aligned}
\mathbb{E}[\varphi(N_n, n) \mid \mathcal{H}_m] &= \mathbb{E}[N_n \mid \mathcal{H}_m] - \rho n \\
&= N_m + \mathbb{E}[N_n - N_m \mid \mathcal{H}_m] - \rho n \\
&= N_m + \mathbb{E}[N_n - N_m] - \rho n \\
&= N_m + \mathbb{E}[N_{n-m}] - \rho n \\
&= N_m + \rho(n - m) - \rho n \\
&= N_m - \rho m.
\end{aligned}$$

It is important to note that in the aforementioned statement we utilized the property that the Bernoulli process can be expressed as a sum of independent Bernoulli variables with independent increments.

Proposition 2.1.7 (Doob-Meyer decomposition of the Bernoulli-process) Let Z be a Bernoulli process with parameter ρ and φ a function in two variables such that the process $\varphi(N_l, l)$ is integrable. Then the Doob-Meyer decomposition $\varphi(N_l, l) = M_l + A_l$ with respect to the internal history $\mathcal{H}$ is given by

$$M_l = \varphi(N_0, 0) + \sum_{k=1}^{l} \varphi(N_k, k) - \varphi(N_{k-1}, k) - \rho \sum_{k=1}^{l} \varphi(1 + N_{k-1}, k) - \varphi(N_{k-1}, k)$$

and

$$A_l = \sum_{k=1}^{l} \varphi(N_{k-1}, k) - \varphi(N_{k-1}, k - 1) + \rho \sum_{k=1}^{l} \varphi(1 + N_{k-1}, k) - \varphi(N_{k-1}, k).$$

Proof.

Let $j, m, i, k \in I$ be arbitrary with $j \leq m$, $m \leq k$ and $j \leq i$. It follows that

$$\mathbb{P}[N_k = m \mid N_i = j] = \mathbb{P}[N_k - N_i = m - j].$$

Since $N_k - N_i \sim Binom(\rho, k - i)$ it follows that

$$\mathbb{P}[N_k - N_i = m - j] = \binom{k-i}{m-j} \rho^{m-j}(1-\rho)^{k-i-m+j}.$$

Hence,

$$\begin{aligned}
\mathbb{E}[\varphi(N_k, k) \mid \mathcal{H}_i] &= \sum_{j=0}^{i} 1_{\{N_i=j\}} \mathbb{E}[\varphi(N_k, k) \mid N_i = j] \\
&= \sum_{j=0}^{i} \sum_{m=j}^{k-(i-N_i)} 1_{\{N_i=j\}} \varphi(m, k) \binom{k-i}{m-j} \rho^{m-j}(1-\rho)^{k-i-m+j} \\
&= \sum_{m=N_i}^{k-(i-N_i)} \varphi(m, k) \binom{k-i}{m-N_i} \rho^{m-N_i}(1-\rho)^{k-i-m+N_i}.
\end{aligned}$$

The innovation martingale M of $\varphi(N_n, n)$ is given by

$$\begin{aligned}
M_n &= M_{n-1} + \varphi(N_n, n) - \mathbb{E}[\varphi(N_n, n) \mid \mathcal{F}_{n-1}] \\
&= M_{n-1} + \varphi(N_n, n) \\
&\quad - \sum_{m=N_i}^{n-(n-1-N_{n-1})} \varphi(m, n) \binom{n-(n-1)}{m-N_{n-1}} \rho^{m-N_{n-1}}(1-\rho)^{n-(n-1)-m+N_{n-1}} \\
&= M_{n-1} + \varphi(N_n, n) - \sum_{m=N_{n-1}}^{N_{n-1}+1} \varphi(m, n) \rho^{m-N_{n-1}}(1-\rho)^{1-m+N_{n-1}} \\
&= M_{n-1} + \varphi(N_n, n) - \varphi(N_{n-1}, n)(1-\rho) - \varphi(1+N_{n-1}, n)\rho \\
&= M_{n-1} + \varphi(N_n, n) - \varphi(N_{n-1}, n) - \rho\Big(\varphi(1+N_{n-1}, n) - \varphi(N_{n-1}, n)\Big).
\end{aligned}$$

Per induction it thus follows

$$M_n = \varphi(N_0, 0) + \sum_{l=1}^{n} \varphi(N_l, l) - \varphi(N_{l-1}, l) - \rho \sum_{l=1}^{n} \varphi(1+N_{l-1}, l) - \varphi(N_{l-1}, l).$$

The compensator A is given by

$$\begin{aligned}A_n &= A_{n-1} - \varphi(N_{n-1}, n-1) + \mathbb{E}[N_n \mid \mathcal{F}_{n-1}] \\ &= A_{n-1} - \varphi(N_{n-1}, n-1) + \sum_{m=N_{n-1}}^{N_{n-1}+1} \varphi(m,n)\rho^{m-N_{n-1}}(1-\rho)^{1-m+N_{n-1}} \\ &= A_{n-1} - \varphi(N_{n-1}, n-1) + \varphi(N_{n-1}, n)(1-\rho) + \varphi(1+N_{n-1}, n)\rho \\ &= A_{n-1} + \varphi(N_{n-1}, n) - \varphi(N_{n-1}, n-1) + \rho\Big(\varphi(1+N_{n-1}, n) - \varphi(N_{n-1}, n)\Big).\end{aligned}$$

By induction it thus follows that

$$A_n = \sum_{l=1}^{n} \varphi(N_{l-1}, l) - \varphi(N_{l-1}, l-1) + \rho \sum_{l=1}^{n} \varphi(1+N_{l-1}, l) - \varphi(N_{l-1}, l).$$

□

The Doob-Meyer decomposition yields the martingale difference equation

$$\varphi(x,l) - \varphi(x, l-1) + \rho\Big(\varphi(1+x, l) - \varphi(x,l)\Big) = 0.$$

To familiarize ourselves with the techniques used, we start with the simple case where the transformation is a second order polynomial. In the next step we want to generalize this result to a general time discrete point process N with conditional intensity λ, by determining the martingales based on an arbitrary n-th order polynomial.

Proposition 2.1.8 (Polynomial martingales of second order)
Let N be a Bernoulli Process with parameter ρ and $\mathcal{H}$ be the filtration generated by N. Then the process M defined by

$$M_l := c\left(\rho^2 l^2 + \rho^2 l - \rho l - 2\rho l N_l + N_l^2\right),$$

is a martingal with respect to $\mathcal{H}$.

Proof.

Let φ be a polynomial function of second order, that is

$$\varphi(x,l) = \varphi_0(l) + \varphi_1(l) \cdot x + \varphi_2(l) \cdot x^2.$$

According to Proposition 2.1.7, the function φ needs to satisfy the difference equation

$$\varphi(x,l)-\varphi(x,l-1)+\rho\Big(\varphi(1+x,l)-\varphi(x,l)\Big)=0.$$

The left hand side equals

$$\begin{aligned}&\varphi(x,l)-\varphi(x,l-1)+\rho\Big(\varphi(1+x,l)-\varphi(x,l)\Big)\\&=\varphi_0(l)-\varphi_0(l-1)+\rho\varphi_1(l)+\rho\varphi_2(l)+(\varphi_1(l)-\varphi_1(l-1)\\&\quad+2\rho\varphi_2(l))x+(\varphi_2(l)-\varphi_2(l-1))x^2.\end{aligned}$$

The equation is fulfilled for all $x\in\mathbb{N}$ if and only if all coefficients are equal to zero. We first consider the second order term. It follows that

$$\varphi_2(l)-\varphi_2(l-1)=0,$$

if and only if $\varphi_2(l)=c$ for some $c\in\mathbb{R}$. Substituting φ_2 into the first-order term results in

$$\varphi_1(l)-\varphi_1(l-1)=-2\rho c.$$

Consequently, $\varphi_1(l)=-2\rho cl+c_1$ for some $c_1\in\mathbb{R}$. To avoid duplicate calculations, we take $c_1=0$. Plugging φ_1 and φ_2 into the 0-th order term yields

$$\varphi_0(l)-\varphi_0(l-1)=2\rho^2cl-\rho c.$$

Consequently, $\varphi_0(l)=\rho^2cl(l+1)-\rho lc+c_0$. To avoid duplicate calculations, we take $c_0=0$. By substituting φ_0,φ_1 and φ_2 into the initial representation of φ , we obtain

$$\varphi(x,l)=\rho^2l(l+1)c-\rho lc-2\rho clx+cx^2.$$

□

Note that

$$M_l=\rho^2l^2+\rho^2l-\rho l-2\rho lN_l+N_l^2=(N_l-\rho l)^2-\rho(1-\rho)l.$$

Since M is a martingale starting in $M_0=0$ it follows

$$\mathbb{E}\left[(N_l-\rho l)^2\right]=\rho(1-\rho)l.$$

2.2 Polynomial Martingales in discrete time

The aim of this section is to explore polynomial martingale transformations in discrete time for point processes of a general nature. Within this section, we unveil the expansion of Krawtchouk martingales [34], initially associated with Bernoulli processes, to encompass a wider array of discrete time point processes. In particular, the second order martingales prove to be useful for determining the covariance kernels for the compensated processes.

In contrast to the Bernoulli process, we require a function φ in three variables. We define the set of all possible λ realizations by

$$\mathbb{L} := [0, 1]^{\mathbb{N}}.$$

The general martingale transformation is expected to be path dependent, requiring the history of the conditional intensity to be incorporated as the third component of the transformation.

Proposition 2.2.1 (Doob-Meyer decomposition for point processes)
Let N be a point process with conditional $\mathcal{H}$-intensity λ and $\varphi : \mathbb{N} \times \mathbb{N} \times \mathbb{L} \to \mathbb{R}$ a function in three variables such that the process X defined by

$$X_l := \varphi(N_l, l, \lambda),$$

is adapted to $\mathcal{H}$ and integrable. Then the Doob-Meyer decomposition $X_t = M_t + A_t$, is given by

$$\begin{aligned} M_n = \varphi(N_0, 0, \lambda) &+ \sum_{l=1}^{n} \varphi(N_l, l, \lambda) - \varphi(N_{l-1}, l, \lambda) \\ &- \sum_{l=1}^{n} \lambda(1 \mid Z_1^{l-1}) \left(\varphi(1 + N_{l-1}, l, \lambda) - \varphi(N_{l-1}, l, \lambda)\right) \end{aligned}$$

and

$$\begin{aligned} A_n = &\sum_{l=1}^{n} \varphi(N_{l-1}, l, \lambda) - \varphi(N_{l-1}, l-1, \lambda) \\ &+ \sum_{l=1}^{n} \lambda(1 \mid Z_1^{l-1}) \Big(\varphi(1 + N_{l-1}, l, \lambda) - \varphi(N_{l-1}, l, \lambda)\Big). \end{aligned}$$

Proof.

The innovation martingale M of $\varphi(N_n, n)$ is given by

$$\begin{aligned}
M_n &= M_{n-1} + \varphi(N_n, n, \lambda) - \mathbb{E}[\varphi(N_n, n, \lambda) \mid \mathcal{F}_{n-1}] \\
&= M_{n-1} + \varphi(N_n, n, \lambda) - \varphi(N_{n-1}, n, \lambda)\lambda(0 \mid Z_1^{n-1}) \\
&\qquad - \varphi(1 + N_{n-1}, n, \lambda)\lambda(1 \mid Z_1^{n-1}) \\
&= M_{n-1} + \varphi(N_n, n, \lambda) - \varphi(N_{n-1}, n, \lambda)(1 - \lambda(1 \mid Z_1^{n-1})) \\
&\qquad - \varphi(1 + N_{n-1}, n, \lambda)\lambda(1 \mid Z_1^{n-1}) \\
&= M_{n-1} + \varphi(N_n, n, \lambda) - \varphi(N_{n-1}, n, \lambda) \\
&\qquad - \lambda(1 \mid Z_1^{n-1})\Big(\varphi(1 + N_{n-1}, n, \lambda) - \varphi(N_{n-1}, n, \lambda)\Big).
\end{aligned}$$

By induction it follows that

$$\begin{aligned}
M_n = \varphi(N_0, 0, \lambda) &+ \sum_{l=1}^{n} \varphi(N_l, l, \lambda) - \varphi(N_{l-1}, l, \lambda) \\
&- \sum_{l=1}^{n} \lambda(1 \mid Z_1^{l-1})\varphi(1 + N_{l-1}, l, \lambda) - \varphi(N_{l-1}, l, \lambda).
\end{aligned}$$

The compensator A of $\varphi(N_n, n)$ is given by

$$\begin{aligned}
A_n &= A_{n-1} - \varphi(N_{n-1}, n-1, \lambda) + \mathbb{E}[N_n \mid \mathcal{F}_{n-1}] \\
&= A_{n-1} + \varphi(N_{n-1}, n, \lambda) - \varphi(N_{n-1}, n-1) \\
&\qquad + \lambda(1 \mid Z_1^{n-1})\Big(\varphi(1 + N_{n-1}, n, \lambda) - \varphi(N_{n-1}, n, \lambda)\Big).
\end{aligned}$$

Per induction it follows that

$$\begin{aligned}
A_n = \sum_{l=1}^{n} &\varphi(N_{l-1}, l, \lambda) - \varphi(N_{l-1}, l-1, \lambda) \\
&+ \sum_{l=1}^{n} \lambda(1 \mid X_1^{l-1})\Big(\varphi(1 + N_{l-1}, l, \lambda) - \varphi(N_{l-1}, l, \lambda)\Big).
\end{aligned}$$

□

To proceed with the objective of this section, it is necessary to first define a sequence of numbers denoted by (a_i^j).

Definition 2.2.2 (Coefficient sequence of polynomial martingales)
We define the (deterministic) sequence $(a_i^{(0)})_{i\in\mathbb{N}_0}$ by

$$a_i^{(0)} := 1_{\{i=1\}}.$$

Further, we denote by $(a_i^{(1)})_{i\in\mathbb{N}_0}$ the sequence given by

$$a_i^{(1)} := \frac{1}{i!}.$$

Furthermore we recursively define for $j > 1$ the sequence $(a_i^{(j)})_{i\in\mathbb{N}_0}$ by

$$a_i^{(j)} := \sum_{k=1}^{i} a_k^{(j-1)} \cdot \frac{1}{(i+1-k)!}.$$

Theorem 2.2.3 (n-th order martingales)
Define $\Psi : \mathbb{N} \times \mathbb{N} \times \mathbb{L} \to \mathbb{R}_+$ by

$$\Psi(l,j,\lambda) := \sum_{i_1=1}^{l} \lambda(1 \mid Z_1^{i_1-1}) \sum_{i_2=1}^{i_1} \lambda(1 \mid Z_1^{i_{j-1}-1}) \cdots \sum_{i_j=1}^{i_{j-1}} \lambda(1 \mid Z_1^{i_j-1}).$$

Let N be a point process with conditional $\mathcal{H}$-intensity λ and a the coefficient sequence of polynomial martingales (Definition 5.3.2). Then the process X defined by

$$X_l := \sum_{i=0}^{n} \frac{n!}{(n-i)!} \sum_{j=0}^{i} (-1)^j \Psi(l,j,\lambda) a_{i-j+1}^{(j)} N_l^{n-i},$$

is a martingal with respect to $\mathcal{H}$.

Proof.

Let φ be a polynomial function of n-th order, that is

$$\varphi(x,l,\lambda) = \sum_{i=0}^{n} \varphi_i(l,\lambda) \cdot x^i.$$

According to Proposition 2.2.1 the function φ needs to satisfy the difference equation

$$\varphi(x,l,\lambda) - \varphi(x,l-1,\lambda) + \lambda(1 \mid X_1^{l-1})\Big(\varphi(1+x,l,\lambda) - \varphi(x,l,\lambda)\Big) = 0.$$

The left hand side equals

$$\begin{aligned}
&\varphi(x,l,\lambda)-\varphi(x,l-1,\lambda)+\lambda(1\mid X_1^{l-1})\Big(\varphi(1+x,l,\lambda)-\varphi(x,l,\lambda)\Big)\\
&=\sum_{k=0}^{n}\Delta\varphi_k(l,\lambda)\cdot x^k+\lambda(1\mid X_1^{l-1})\sum_{i=0}^{n}\varphi_i(x,\lambda)\Big((1+x)^i-x^i\Big)\\
&=\sum_{k=0}^{n}\Delta\varphi_k(l,\lambda)\cdot x^k+\sum_{i=0}^{n}\lambda(1\mid X_1^{l-1})\varphi_i(x,\lambda)\sum_{k=0}^{i-1}\binom{i}{k}x^k\\
&=\sum_{k=0}^{n}\Delta\varphi_k(l,\lambda)\cdot x^k+\sum_{k=0}^{n-1}\sum_{i=k+1}^{n}\lambda(1\mid X_1^{l-1})\varphi_i(x,\lambda)\binom{i}{k}x^k\\
&=\sum_{k=0}^{n-1}x^k\left(\Delta\varphi_k(l,\lambda)+\sum_{i=k+1}^{n}\lambda(1\mid X_1^{l-1})\varphi_i(l,\lambda)\binom{i}{k}\right)+\Delta\varphi_n(l,\lambda)\cdot x^n.
\end{aligned}$$

Assertion: For all $1\leq i\leq n$ it follows that

$$\varphi_{n-i}(t,\lambda)=\frac{n!}{(n-i)!}\sum_{j=0}^{i}(-1)^j\Psi(l,j,\lambda)a_{i+1-j}^{(j)}. \tag{2.2}$$

1. Base case: Show that the statement holds for $i=0$.

   Since $a_1^{(1)}=1$ and $\varphi_n(l)=c=c\frac{n!}{(n-1)!}a_1^{(0)}$ it follows that

$$\begin{aligned}
\Delta\varphi_{n-1}(l,\lambda)&=-\lambda(1\mid s_1^{l-1})\sum_{i=n}^{n}\varphi_i(l,\lambda)\binom{i}{n-1}\\
&=-\lambda(1\mid s_1^{l-1})\varphi_n(l,\lambda)\frac{n!}{(n-1)!}\\
&=\frac{n!}{(n-1)!}\sum_{j=0}^{1}(-1)^j\lambda(1\mid s_1^{l-1})a_{i+1-j}^{i}.
\end{aligned}$$

   The solution to this difference equation is

$$\begin{aligned}
\varphi_{n-1}(l,\lambda)&=\frac{n!}{(n-1)!}\sum_{j=0}^{1}(-1)^j\sum_{k=1}^{l}\lambda(1\mid s_1^{k-1})a_{i+1-j}^{(i)}\\
&=\frac{n!}{(n-1)!}\sum_{j=0}^{1}(-1)^j\Psi(l,j,\lambda)a_{i+1-j}^{(i)}.
\end{aligned}$$

2. Induction step: Show that if $\varphi_n, \varphi_{n-1}, ..., \varphi_{n-i+1}$ have the representation from (2.2), then the representation also holds for φ_{n-i}.

   It holds true that

$$\begin{aligned}
&\Delta\varphi_{n-i}(l,\lambda)\\
&= -\lambda(1 \mid s_1^{l-1}) \sum_{j=n-(i-1)}^{n} \varphi_j(l,\lambda)\binom{j}{n-i}\\
&= -\lambda(1 \mid s_1^{l-1}) \sum_{j=0}^{i-1} \varphi_{n-j}(l,\lambda)\binom{n-j}{n-i}\\
&= -\lambda(1 \mid s_1^{l-1}) \frac{n!}{(n-i)!} \sum_{j=0}^{i-1}\sum_{k=0}^{j} c_n(-1)^k \Psi(l,k,\lambda) a_{j+1-k}^{(k)} \frac{1}{(i-j)!}.
\end{aligned}$$

   Since

$$\Delta\Psi(l,k+1,\lambda) = \lambda(1 \mid s_1^{l-1})\Psi(l,k,\lambda),$$

   it follows that

$$\begin{aligned}
\varphi_{n-i}(l,\lambda) &= -\frac{n!}{(n-i)!}\sum_{j=0}^{i-1}\sum_{k=0}^{j}(-1)^k\Psi(l,k+1,\lambda)a_{j+1-k}^{(k)}\frac{1}{(i-j)!}\\
&= \frac{n!}{(n-i)!}\sum_{j=0}^{i-1}(-1)^{k+1}\Psi(l,k+1,\lambda)\sum_{k=0}^{j}a_{j+1-k}^{(k)}\frac{1}{(i-j)!}\\
&= \frac{n!}{(n-i)!}\sum_{j=0}^{i-1}(-1)^{k+1}\Psi(l,k+1,\lambda)\sum_{k=0}^{j}a_{j+1-k}^{(k)}\frac{1}{(i-j)!}\\
&= \frac{n!}{(n-i)!}\sum_{k=0}^{i-1}(-1)^{k+1}\Psi(l,k+1,\lambda)\sum_{j=k}^{i-1}a_{j+1-k}^{(k)}\frac{1}{(i-j)!}\\
&= \frac{n!}{(n-i)!}\sum_{k=0}^{i-1}(-1)^{k+1}\Psi(l,k+1,\lambda)\sum_{j=1}^{i-k}a_{j}^{(k)}\frac{1}{(i-k+1-j)!}\\
&= \frac{n!}{(n-i)!}\sum_{k=0}^{i}(-1)^{k}\Psi(l,k,\lambda)a_{i-k+1}^{(k)}.
\end{aligned}$$

By substituting the representation into the equation for φ, we obtain

$$\begin{aligned}\varphi(x,l,\lambda) &:= \sum_{i=0}^{n} \varphi_i(l,\lambda)\cdot x^i \\ &= \sum_{i=0}^{n} \varphi_{n-i}(l,\lambda)\cdot x^{n-i} \\ &= \sum_{i=0}^{n} \frac{n!}{(n-i)!} \sum_{j=0}^{i} (-1)^j \Psi(l,j,\lambda) a_{i-j+1}^{(k)} \cdot x^{n-i}.\end{aligned}$$

□

Remark 2.2.4 (Krawtchouk polynomials)
Let N be a Bernoulli process with parameter ρ. Define the function $\psi : \mathbb{N}\times\mathbb{N}\to\mathbb{R}$ by

$$\begin{aligned}\psi(l,j) &= \Psi(l,j,\rho) \\ &= \sum_{i_1=1}^{l} \rho \sum_{i_2=1}^{i_1} \rho \cdots \rho \sum_{i_j=1}^{i_{j-1}} \rho \\ &= \rho^j \sum_{i_1=1}^{l} \sum_{i_2=1}^{i_1} \cdots \sum_{i_j=1}^{i_{j-1}} 1.\end{aligned}$$

Theorem 2.2.3 yields the n-th order martingale

$$X_l = \sum_{i=0}^{n} \frac{n!}{(n-i)!} \sum_{j=0}^{i} (-\rho)^j \psi(l,j) a_{i-j+1}^{(k)} N_l^{n-i}.$$

This case has already been discussed in the literature (see [34], p.59-60). The polynomials

$$\mathcal{K}(x,l) = \sum_{i=0}^{n} \frac{n!}{(n-i)!} \sum_{j=0}^{i} (-\rho)^j \psi(l,j) a_{i-j+1}^{(j)} x^{n-i},$$

are called Krawtchouk polynomials. In the context of this example, we have shown that Theorem 2.2.3 produces the Krawtchouk martingales when applied to a Bernoulli process.

Set $n = 2$ to obtain the martingale

$$\begin{aligned}
X_l^{(2)} &= \sum_{i=0}^{2} \frac{2!}{(2-i)!} \sum_{j=0}^{i} (-1)^j \Psi(l,j,\lambda) a_{i-j+1}^{(j)} N_l^{2-i} \\
&= N_l^2 - 2\Psi(l,1,\lambda)N_l - 2\Psi(l,1,\lambda)a_{2-1+1}^{(1)} + 2\Psi(l,2,\lambda)a_1^{(2)} \\
&= N_l^2 - 2\sum_{i=1}^{l} \lambda(i \mid Z_1^{i-1}) N_l - \sum_{i=1}^{l} \lambda(i \mid Z_1^{i-1}) \\
&\qquad + 2\sum_{i=1}^{l} \lambda(i \mid Z_1^{i-1}) \sum_{k=1}^{i} \lambda(k \mid Z_1^{k-1}) \\
&= \left(N_l - \sum_{i=1}^{l} \lambda(i \mid Z_1^{i-1}) \right)^2 - \sum_{i=1}^{l} \lambda(i \mid Z_1^{i-1})(1 - \lambda(i \mid Z_1^{i-1})).
\end{aligned}$$

Define the process X by

$$X_l := N_l - \sum_{i=1}^{l} \lambda(i \mid Z_1^{i-1}).$$

Since $X^{(2)}$ is a martingale starting in $X_0 = 0$ the variance of X at some arbitrary point in time equals

$$\operatorname{Var}(X_l) = \sum_{i=1}^{l} \mathbb{E}[\lambda(i \mid Z_1^{i-1})(1 - \lambda(i \mid Z_1^{i-1}))]. \tag{2.3}$$

If additionally the process C defined by

$$C_l := \lambda(l \mid Z_1^{l-1})(1 - \lambda(l \mid Z_1^{l-1})),$$

is a martingale starting in $C_0 = \rho(1-\rho)$ for some constant $\rho \in [0,1]$, equation (2.3) becomes

$$\operatorname{Var}(X_l) = \rho(1-\rho)l.$$

Since X is a martingale this implies

$$\operatorname{Cov}(X_m, X_n) = \rho(1-\rho)\min(m,n). \tag{2.4}$$

A kernel of minimum type, such as the one presented in equation (2.4), can prove to be highly beneficial in the context of Principal Component Analysis (PCA), as expounded in the forthcoming Chapter 5.

2.3 Non-negative martingales

In the first chapter of this work we introduced the martingales discovered by Hess and Stute (paper in work). These martingales have in common that the path of each realization can take only non-negative values. This additional property helps in calculating certain exceedance probabilities.

Proposition 2.3.1 (Exceedance probabilities)
Let $\mathcal{H} = (\mathcal{H}_n)_{n\in\mathbb{N}}$ be a filtration, $M = (M_n)_{n\in\mathbb{N}}$ a non-negative $\mathcal{H}$-martingale starting in $M_0 = 1$, $c > 1$ be fixed and define

$$\tau := \begin{cases} \min(\{n \in \mathbb{N} \mid M_n \geq c\}) & \text{if } \{n \in \mathbb{N} \mid M_n \geq c\} \neq \emptyset, \\ \infty & \text{else.} \end{cases}$$

Additionally assume that the following conditions are satisfied:

1. $\lim_{n\to\infty} M_n = 0$.
2. The random variable M_τ is integrable.

Then it follows that

$$\mathbb{P}\left[\max_{n\in\mathbb{N}} M_n \geq c\right] \leq \frac{1}{c}.$$

Proof.

We defined the $\mathcal{H}$-Stopping time τ by

$$\tau = \begin{cases} \min(\{n \in \mathbb{N} \mid M_n \geq c\}) & \text{if } \{n \in \mathbb{N} \mid M_n \geq c\} \neq \emptyset, \\ \infty & \text{else.} \end{cases}$$

Note that by assumption it follows that $M_\tau \geq c$ on $\{\tau < \infty\}$. Consequently,

$$\left\{\max_{n\in\mathbb{N}} M_n \geq c\right\} = \{\tau < \infty\}.$$

Stopped $\mathcal{H}$-martingales are again $\mathcal{H}$-martingales. Hence, the process $M_{n\wedge\tau}$ is again a $\mathcal{H}$-martingale satisfying

$$\mathbb{E}[M_{n\wedge\tau}] = \mathbb{E}[M_{0\wedge\tau}] = \mathbb{E}[M_0] = 1.$$

By Lebesgues dominated convergence Theorem,

$$\begin{aligned}
c\mathbb{P}[\tau < \infty] &= \int_{\{\tau<\infty\}} c d\mathbb{P} \\
&\leq \int_{\{\tau<\infty\}} M_\tau d\mathbb{P} \\
&= \int_{\{\tau<\infty\}} \lim_{n\to\infty} M_{n\wedge\tau} d\mathbb{P} \\
&= \lim_{n\to\infty} \int_{\{\tau<\infty\}} M_{n\wedge\tau} d\mathbb{P} \\
&= \lim_{n\to\infty} \left(1 - \int_{\{\tau=\infty\}} M_{n\wedge\tau} d\mathbb{P}\right) \\
&= 1 - \lim_{n\to\infty} \left(\int_{\{\tau=\infty\}} M_{n\wedge\tau} d\mathbb{P}\right) \\
&= 1 - \left(\int_{\{\tau=\infty\}} \lim_{n\to\infty} M_{n\wedge\tau} d\mathbb{P}\right) \\
&= 1 - \left(\int_{\{\tau=\infty\}} \lim_{n\to\infty} M_n d\mathbb{P}\right) \\
&= 1.
\end{aligned}$$

Hence,

$$\frac{1}{c} \geq \mathbb{P}[\tau < \infty] = \mathbb{P}\left[\max_{n\in\mathbb{N}} M_n \geq c\right].$$

□

Proposition 2.3.2 (Integral martingales in discrete time)
Let N be a point process with conditional $\mathcal{H}$-intensity λ and g a $\mathcal{H}$-predictable, integrable process. Define the process M through

$$M_n := \sum_{l=1}^{n} g_l(\Delta N_l - \lambda_l) + 1.$$

In this situation it follows that M is a $\mathcal{H}$-martingale.

Proof.

The assertion follows from

$$\begin{aligned}\mathbb{E}[M_n \mid \mathcal{H}_{n-1}] &= \sum_{l=1}^{n} \mathbb{E}[g_l(\Delta N_l - \lambda_l) \mid \mathcal{H}_{n-1}] + 1 \\ &= M_{n-1} + g_n(\mathbb{E}[\Delta N_n \mid \mathcal{H}_{n-1}] - \lambda_n) \\ &= M_{n-1}.\end{aligned}$$

□

Theorem 2.3.3 (Criteria for the martingales)
Let N be a point process with conditional $\mathcal{H}$-intensity λ, g be a $\mathcal{H}$-predictable process and $c > 1$, $0 < \epsilon < 0.5$ be two constants. Additionally assume that λ satisfies

$$\mathbb{P}[\forall k \in \mathbb{N} : \lambda_k \in (0 + \epsilon, 1 - \epsilon)] = 1.$$

Define the martingale M by

$$M_k = \sum_{l=1}^{k} g_l(\Delta N_l - \lambda_l) + 1.$$

In particular, we assume that the following conditions are satisfied:

(a) There exists an $n \in \mathbb{N}$ such that for all $l > T_{n-1} : g_l = -\frac{M_{l-1}}{1-\lambda_l}$.

(b) For all $l \in \mathbb{N} : -\frac{M_{l-1}}{1-\lambda_l} \leq g_l \leq \frac{M_{l-1}}{\lambda_l}$.

Then it follows that

$$\mathbb{P}\left[\max_{k \in \mathbb{N}} M_k \geq c\right] \leq \frac{1}{c}.$$

Proof.

We need to prove that M satisfies the conditions formulated in Proposition 2.3.1.

- M is a non-negative $\mathcal{H}$-martingale starting in $M_0 = 1$.

  Due to Proposition 2.3.2, M is a martingale. It follows that

$$M_0 = \sum_{l=1}^{0} g_l(\Delta N_l - \lambda_l) + 1 = 1.$$

It remains to prove that M is a non-negative process which will be shown by induction.

Case $k = 0$: This is satisfied by definition since $M_0 = 1 \geq 0$.

Induction step: Show, if $M_k \geq 0$ holds, then $M_{k+1} \geq 0$ also holds.

We distinguish four cases:

1. $\Delta N_{k+1} = 1$ and $g_{k+1} \leq 0$:

$$\begin{aligned} M_{k+1} &= M_k + g_{k+1} - g_{\tau_-}\lambda_{k+1} \\ &= M_k + g_{k+1}(1 - \lambda_{k+1}) \\ &\geq M_{k+1} - \frac{M_k}{1 - \lambda_{k+1}}(1 - \lambda_{k+1}) \\ &= 0. \end{aligned}$$

2. $\Delta N_{k+1} = 1$ and $g_{k+1} > 0$:

$$\begin{aligned} M_{k+1} &= M_k + g_{k+1} - g_{k+1}\lambda_{k+1} \\ &= M_k + g_{k+1}(1 - \lambda_{k+1}) \\ &\geq 0. \end{aligned}$$

3. $\Delta N_{k+1} = 0$ and $g_{k+1} \leq 0$:

$$\begin{aligned} M_{k+1} &= M_k - g_{k+1}\lambda_{k+1} \\ &\geq M_k \\ &\geq 0. \end{aligned}$$

4. $\Delta N_{k+1} = 0$ and $g_{k+1} > 0$:

$$\begin{aligned} M_{k+1} &= M_k - g_{k+1}\lambda_{\tau_-} \\ &\geq M_k - \frac{M_k}{\lambda_{k+1}}\lambda_{k+1} \\ &= 0. \end{aligned}$$

- $\lim_{n\to\infty} M_n = 0$.

By assumption (a) it follows that $g_l = -\frac{M_{l-1}}{1-\lambda_l}$ on $l > T_{n-1}$. Hence,

$$\begin{aligned} M_{T_n} &= M_{T_n-1} + g_{T_n} - g_{T_n}\lambda_{T_n} \\ &= M_{T_n-1} - \frac{M_{T_n-1}}{1-\lambda_{T_n}}(1-\lambda_{T_n}) \\ &= 0. \end{aligned}$$

Assumption (b) now recursively implies that for all $l \geq T_n$, $M_l = 0$. Thus, the claim is established.

- The random variable M_τ is integrable.

  We show that M_τ is a bounded random variable.

$$\begin{aligned} |M_\tau| &= |M_{\tau-1} + g_\tau(\Delta N_\tau - \lambda_\tau)| \\ &\leq |c| + |g_\tau| \\ &\leq c + \max\left(|\frac{M_{\tau-1}}{1-\lambda_\tau}|, |\frac{M_{\tau-1}}{\lambda_\tau}|\right) \\ &\leq c + \max\left(\frac{c}{1-\epsilon}, \frac{c}{\epsilon}\right). \end{aligned}$$

□

In the next Lemma we want to specify the subclass of test-strategies g such that the resulting inequality is sharp in the sense that $M_{\tau_c} = c$ on $\{\tau_c < T_{n-1}\}$. To do this, we only need to restrict the upper and lower bound values with respect to the test strategy g.

Theorem 2.3.4 (Test martingales)
Let N be a point process with conditional $\mathcal{H}$-intensity λ, g be a $\mathcal{H}$-predictable process and $c > 1$, $0 < \epsilon < 0.5$ be two constants. Additionally, assume that λ satisfies

$$\mathbb{P}[\forall k \in \mathbb{N} : \lambda_k \in \{0+\epsilon, 1-\epsilon\}] = 1.$$

Define the martingale M by

$$M_k := \sum_{l=1}^{k} g_l(\Delta N_l - \lambda_l) + 1.$$

In particular, we assume that the following conditions are satisfied:

(a) There exists an $n \in \mathbb{N}$ such that for all $l > T_{n-1} : g_l = -\frac{M_{l-1}}{1-\lambda_l}$.

(b) For all $l \in \mathbb{N} : -\frac{M_{l-1}}{1-\lambda_l} \leq g_l \leq \frac{M_{l-1}}{\lambda_l}$.

(c) For all $l < \tau : \frac{M_{l-1}-c}{\lambda_l} \leq g_l \leq \frac{c-M_{l-1}}{1-\lambda_l}$.

Then it follows that

$$\mathbb{P}\left[\max_{k \in \mathbb{N}} M_k \geq c\right] \leq \frac{1}{c}.$$

Proof.

This is a special case of Theorem 2.3.3.

□

Remark 2.3.5 In this remark we want to briefly discuss the conditions of Theorem 2.3.5.

(a) There exists an $n \in \mathbb{N}$ such that for all $l > T_{n-1} : g_l = -\frac{M_{l-1}}{1-\lambda_l}$. This condition is necessary to guarantee that the martingale will end in the state 0 in the limit and remain there. This property of the martingale is important to derive exceedance probabilities.

(b) For all $l \in \mathbb{N} : -\frac{M_{l-1}}{1-\lambda_l} \leq g_l \leq \frac{M_{l-1}}{\lambda_l}$. This condition guarantees that the martingale is a non-negative process. This property of the martingale is important to derive exceedance probabilities.

(c) For all $l < \tau : \frac{M_{l-1}-c}{\lambda_l} \leq g_l \leq \frac{c-M_{l-1}}{1-\lambda_l}$. If this additional condition is satisfied, the martingale has the property that if the critical value c is exceeded before T_{n-1}, it follows that $M_{\tau_c} = c$. This guarantees the martingale to be more efficient with respect to tests on the critical value c.

Example 2.3.6
Let N be a point process with conditional intensity λ and $n \in \mathbb{N}$ be fixed. Define the predictable process g by

$$g_l := -\frac{M_{l-1}}{1-\lambda_l} 1_{\{N_{l-1}=n-1\}}.$$

The strategy g satisfies the conditions from Lemma 2.3.3. Consequently, the process M defined by

$$M_k := \sum_{l=1}^{k} g_l(\Delta N_l - \lambda_l) + 1$$

is a martingale and

$$\mathbb{P}\left[\max_{n\in\mathbb{N}} M_n \geq c\right] \leq \frac{1}{c}.$$

Note that we can rewrite M to become

$$\begin{aligned}
M_k &= \sum_{l=1}^{k} g_l(\Delta N_l - \lambda_l) + 1 \\
&= \sum_{l=1}^{k} M_{l-1}\left(-1_{\{N_{l-1}=n-1\}}\frac{\Delta N_l - \lambda_l}{1-\lambda_l}\right) + 1 \\
&= \prod_{l=1}^{k}\left(1 - 1_{\{N_{l-1}=n-1\}}\frac{\Delta N_l - \lambda_l}{1-\lambda_l}\right) \\
&= 1_{\{N_k<n-1\}} + 1_{\{N_k=n-1\}}\prod_{l=T_{n-1}}^{k}\frac{\lambda_l}{1-\lambda_l}.
\end{aligned}$$

We aim to outline a decision-making scenario that resembles the one discussed in reference [11] but within the context of point processes in our work. Let N^1 and N^2 be two different point processes admitting respectively the conditional $\mathcal{H}$-intensities λ^1 and λ^2. The statistician observes, beginning at time-period $n = 0$ a process N which is either N^1 or N^2 and wishes to decide whether N is N^1 or N^2. This leads to the (simple) test problem $H_0 : N = N^1$ against the alternative $H_1 : N = N^2$.

Definition 2.3.7 (Sequential test martingales)
The set of sequential test strategies to the critical value c with respect to the alternative λ^2 is defined by

$$\begin{aligned}
&\mathcal{S}^c(\lambda^1, \lambda^2) \\
&:= \left\{ g \in \mathcal{T}^c(\lambda^1) \mid \left(\sum_{l=1}^{k} g_l(\Delta N_l - \lambda_l^2) + 1\right)_{k\in\mathbb{N}} \text{ is a submartingale} \right\}.
\end{aligned}$$

Lemma 2.3.8 (Criteria for sequential test martingales)
In the situation of this section, assume

$$\forall l > T_{n-1} : \lambda_l^1 \geq \lambda_l^2.$$

Then for all $\mathcal{T}^c(\lambda^1)$-test strategies g satisfying

(a) $\forall l \in \mathbb{N} : \lambda_l^1 > \lambda_l^2 \Rightarrow g_l \geq 0$,

(b) $\forall l \in \mathbb{N} : \lambda_l^1 < \lambda_l^2 \Rightarrow g_l \leq 0$,

follows $g \in \mathcal{S}^c(\lambda^1, \lambda^2)$.

Proof.

Let $n \in \mathbb{N}$ be arbitrary but fixed. It holds that

$$\begin{aligned}\mathbb{E}[X_n \mid \mathcal{H}_{n-1}] &= X_{n-1} + g_n(\mathbb{E}\left[\Delta N_n \mid \mathcal{H}_{n-1}\right] - \lambda_n^1) \\ &= X_{n-1} + g_n(\lambda_n^2 - \lambda_n^1) \\ &\geq X_{n-1}.\end{aligned}$$

Hence, X is a submartingale.

□

Chapter 3

Point processes in continuous time

Simple, local finite point processes offer a versatile framework for modeling stochastic phenomena over time and have achieved remarkable success in various applications. This ranges from less complicated dynamics, such as the Poisson process for the description of radioactive decay, up to the explanation of highly complex microeconomic data sets in market research. For instance, the work of Kopperschmidt and Stute [24] focuses on this class of point processes in the analysis of customer behavior and proposes an estimator for any parametric model subject to mild conditions.

In this chapter, we confine our examination to the (integrated) conditional intensity in relation to the internal history of the underlying point process. However, in theory it is plausible to extend the results to other cases provided that appropriate conditions are met.

3.1 Finding a suitable differential equation

In this chapter the aim is to find interesting martingale transformations from a general point process N characterized by its internal integrated conditional intensity Λ. As discussed in the initial chapter, the procedure entails the identification of a specific differential equation, followed by its resolution. We commence with a preliminary and informal discussion to attain a comprehensive understanding of the significant results.

When considering a conditional intensity λ in a time-discrete framework, it becomes evident that the process λ at any given time period m represents the probability of an event occurring at the subsequent time period $m+1$, taking into account all available information up to that point in time m. Similarly to the discrete-time case, the conditional intensity of a point process in continuous time is defined as the integrand with respect to the Lebesgue measure in the compensator of the Doob-Meyer decomposition

$$N_t - \int_{s_0}^{t} \lambda_v dv.$$

However, in continuous time the conditional intensity at some time point $s \in I$ is heuristically given by

$$\lambda_s = \lim_{\Delta \to 0} \frac{\mathbb{P}[N_{s+\Delta} - N_s = 1 \mid \mathcal{H}_s]}{\Delta}$$

and can be interpreted as the rate at which events take place given the information of the internal history. Note that for an arbitrary $s \in I$, the σ-algebra $\mathcal{H}_s$ has the form

$$\mathcal{H}_s = \sigma(\{N_v : v < s\}).$$

Assuming that the underlying point process N admits a conditional intensity λ, it is well known that λ uniquely determines the distribution of N. Earlier in this work, this concept was discussed in the context of a discrete time framework, and its roots can be traced back to the Doob-Meyer decomposition.

The cumulative conditional intensity, commonly referred to as the compensator, is described by the process $\Lambda = (\Lambda_t)_{t \in I}$ that is defined by

$$\Lambda_s := \int_{s_0}^{s} \lambda_v dv.$$

The process Λ shows continuous paths due to its definition as a Lebesgue integral. It is important to note that the compensator Λ may not always be representable as an integral with respect to the Lebesgue measure. Consequently, we will make an effort to work with Λ whenever feasible in order to circumvent discussions regarding the existence of λ. In the context of Poisson point processes, the compensator Λ is determined to be a deterministic function, and in the specific instance of a homogeneous process with parameter ρ, it takes the form of a linear function, $\Lambda_t =$

ρt. Stepping outside of the Poisson point process framework results in compensators with a stochastic nature. As an illustration, examine the renormalized empirical distribution function, which incorporates not only a temporal aspect but also the current, inherently random state of the system. In this context, $d\Lambda$ is represented by the equation

$$d\Lambda_s = \frac{n - nF_n(s)}{1 - F(s)} f(s)ds.$$

For a point process that admits this type of a compensator the martingale differential equation with respect to the naturally generated filtration results in

$$\frac{\partial \varphi}{\partial x_2}(x,t) + (\varphi(1+x,t) - \varphi(x,t)) \frac{n - nF_n(t)}{1 - F(t)} f(t) = 0. \tag{3.1}$$

One contribution of Stute and Hess (paper in work) was to uncover a general solution to this equation. To accomplish this goal, they discovered polynomial martingales and demonstrated that every adapted martingale derived from a transformation of time and state of the system, denoted by $\varphi(x,t)$, can be expressed as a linear combination of these polynomial martingales.

An equivalent differential equation as in (3.1) is also known for the homogeneous Poisson process with parameter ρ [32] and is given by

$$\frac{\partial \varphi}{\partial x_2}(x,t) + (\varphi(1+x,t) - \varphi(x,t))\, \rho = 0. \tag{3.2}$$

Known solutions of this equation are for instance the geometric Poisson process and the Poisson process subtracted by its compensator $\Lambda_t = \rho t$. A detailed analysis of (3.2) is covered in chapter five.

The field in question is characterized by the presence of two fundamental statements. On the one hand, we have a theorem which states that for every predictable process $(h_t)_{t\in I}$, the process $(X_t)_{t\in I}$ defined by

$$X_t := \int_0^t h_v(dN_v - d\Lambda_v), \tag{3.3}$$

is a martingale. For instance, setting $h = 1$ results in the martingale $X_t = N_t - \Lambda_t$. Conversely, there exists a theorem of existence related to the representation in equation (3.3). Namely, for each adapted martingale X with respect to the filtration $\mathcal{H}$ exists a predictable process h so that X can be represented by the integral form given in (3.3). The question on how to generate interesting martingales from this integral representation remains unanswered with the existence statement and is about to be answered in a certain sense in the upcoming investigations.

Heuristically, considering equation (3.2) and (3.3) it can already be presumed that the general equation has the form

$$\frac{\partial \varphi}{\partial x_2}(x,t) + (\varphi(1+x,t) - \varphi(x,t))\,\lambda_t = 0. \tag{3.4}$$

In the subsequent steps, the derivation of equation (3.4) will be performed. Throughout this analysis, we make the following assumptions:

1. It $\mathbb{P}$-almost sure holds that
$$\lim_{t \downarrow v} \frac{\mathbb{P}[N_t > N_v \mid \mathcal{H}_v]}{t-v} = \lambda_v$$
uniformly.
2. For all $x \in \mathbb{R}, i \in \mathbb{N}$ it holds $\mathbb{P}$-almost sure that
$$\int_{T_i}^{T_{i+1}} |\varphi(x,v)| dv < \infty.$$
3. For all $s \in I$ it $\mathbb{P}$-almost sure holds that
$$\int_0^s |\varphi(N_{v-}+1,v) - \varphi(N_{v-},v)| \lambda_v dv < \infty.$$

For ease of notation we will refer the above assumptions as the usual assumptions. Our examination is focused on finite point processes on an interval $I \subseteq \mathbb{R}$, and thus, we set $s_0 = \inf I$ for clarity.

Proposition 3.1.1 (Innovation martingale)
Let $N = (N_t)_{t \in I}$ be a point process with compensator $d\Lambda = \lambda dt$. We define the adapted process $(X_t)_{t \in I}$ by setting

$$X_s := \varphi(N_s, s), \forall s \in I.$$

If φ meets the usual conditions, the innovation martingale of X with respect to $\mathcal{H}$ is given by the process $M = (M_t)_{t \in I}$ defined by

$$M_s := \varphi(0, s_0) + \int_{s_0}^{s} \big(\varphi(1+N_{v-},v) - \varphi(N_{v-},v)\big)(dN_v - \lambda_v dv).$$

Proof.

The proof is oriented towards the limiting procedure used in Beiglboeck [3]. N is a process on the intervall $[0, s_1)$. To make use of our results from Chapter two, we first define a sequence of discrete point processes $(N^k)_{k\in\mathbb{N}}$ on the dyadic numbers $\mathcal{D}_t^k$. Let $s \in I$ be fixed but arbitrary. For $k \in \mathbb{N}$ the set $\mathcal{D}_s^k$ is given by

$$\mathcal{D}_s^k = \left\{ \frac{ms}{2^k} \mid m \in \{1, ..., 2^k\} \right\}.$$

For $l \in \{1, ..., 2^k\}$ we define the l-th element in $\mathcal{D}_s^k$ by

$$t_l^k = \frac{ls}{2^k}.$$

The discrete point process $(N_t^k)_{t\in\{t_1^k,...,t_{2^k-1}^k\}}$ is defined by

$$N_0^k = N_0$$

and for $0 < n \in \mathcal{D}_s^k$ by

$$N_{t_n^k}^k = N_{t_{n-1}^k}^k + 1_{\{N_{t_n^k} > N_{t_{n-1}^k}\}}.$$

Note that, the random variable $N_{t_n^k}^k$ is not equal to $N_{t_n^k}$ for $\omega \in \{N_{t_n}^k - N_{t_{n-1}}^k > 1\}$. But since N is a simple locally finite point process on the interval $[0, s_1)$, there exists a $\tilde{k}$ for each fixed but arbitrary $\omega \in \Omega$ such that $N_{t_n^k} = N_{t_n^k}$ for all $k \geq \tilde{k}, t_n^k \in \mathcal{D}_s^k$. N^k is a point process in discrete time with conditional intensity $(\lambda_t^k)_{t\in\{t_1^k,...t_{2^k-1}^k\}}$ given by

$$\lambda_{t_n^k}^k = \mathbb{P}\left[N_{t_n}^k > N_{t_{n-1}}^k \mid \mathcal{H}_{t_{n-1}^k} \right].$$

Define $X^k = (X^k)_{t\in\{t_1^k,...,t_{2^k-1}^k\}}$ by

$$X_{t_n^k}^k = \varphi(N_{t_n^k}, t_n^k).$$

According to Proposition 2.2.1, the martingale $(M_t^k)_{t\in\{t_1^k,...,t_{2^k-1}^k\}}$ in the discrete Doob-Meyer decomposition of the process X^k is given by

$$\begin{aligned} M_{t_n^k}^k = &\sum_{l=1}^{2^k-1} \varphi\left(N_{t_l^k}^k, t_l^k\right) - \varphi\left(N_{t_{l-1}^k}^k, t_l^k\right) \\ &- \sum_{l=1}^{n} \lambda_{t_l^k}^k \left(\varphi(N_{t_{l-1}^k} + 1, t_l^k) - \varphi(N_{t_{l-1}^k}, t_l^k)\right). \end{aligned}$$

First we show for the first integral

$$\sum_{l=1}^{2^k}\left(\varphi\left(N^k_{t^k_l},t^k_l\right)-\varphi\left(N^k_{t^k_{l-1}},t^k_l\right)\right)\xrightarrow{k\to\infty}\int_0^s\varphi\left(N_s,s\right)-\varphi\left(N_{s-},s\right)dN_s,$$

$\mathbb{P}$-almost sure. Let $M\in\mathbb{N}$ be arbitrary and fix $\omega\in\{N_s=M\}$ for the following operations. Define $\tilde{k}\in\mathbb{N}$ as the minimum such that $N_s(\omega)=N^{\tilde{k}}_s(\omega)$. For all $k\geq\tilde{k}$ we define $T^k_1,...,T^k_M$ as the ordered points of the point process N^k and $T_0=0,T_{M+1}=s$. It follows that

$$\sum_{l=1}^{2^k-1}\left(\varphi\left(N^k_{t^k_l},t^k_l\right)-\varphi\left(N^k_{t^k_{l-1}},t^k_l\right)\right)=\sum_{i=1}^{M}\varphi\left(i,T^k_i\right)-\varphi\left(i-1,T^k_i\right).$$

By construction of N^k we have $N^k\to N$ $\mathbb{P}$-almost sure. Hence $T^k_1,..,T^k_M$ converge to $T_1,..,T_M$, respectively $\mathbb{P}$-almost sure. By the continuous mapping theorem it follows that

$$\begin{aligned}\sum_{i=1}^{M}\varphi\left(i,T^k_i\right)-\varphi\left(i-1,T^k_i\right)\to&\sum_{i=1}^{M}\varphi\left(i,T_i\right)-\varphi\left(i-1,T_i\right)\\&=\int_0^s\varphi\left(N_v,v\right)-\varphi\left(N_{v-},v\right)dN_v.\end{aligned}$$

Since the union $\bigcup_{M=0}^{\infty}\{N_s=M\}$ forms a countable partition of Ω, this yields almost sure convergence of the first integral.

For the second integral we show that

$$\sum_{l=1}^{2^k}\lambda^k_{t^k_l}\left(\varphi(N_{t^k_{l-1}}+1,t^k_l)-\varphi(N_{t^k_{l-1}},t^k_l)\right)\xrightarrow{k\to\infty}\int_0^s\varphi(N_{v-}+1,v)-\varphi(N_{v-},v)\lambda_v dv.$$

We have that $N_{t^k_l-}=N_{t^k_l}$ or $N_{t^k_l-}=N_{t^k_l}-1$. Since $k\geq\tilde{k}$, it follows that $N_{t^k_l-}=N_{t^k_{l-1}}$. Consequently, per definition of the Riemann integral the limit equals

$$\begin{aligned}&\int_0^s\varphi(N_{v-}+1,v)-\varphi(N_{v-},v)\lambda_v dv\\&\quad=\lim_{k\to\infty}\sum_{l=0}^{2^k-1}\lambda_{t^k_l}(t^k_l-t^k_{l-1})\left(\varphi(N_{t^k_{l-1}}+1,t^k_l)-\varphi(N_{t^k_{l-1}},t^k_l)\right).\end{aligned}$$

Subtracting the second integral from the discrete version reveals that it is necessary to demonstrate the following:

$$\lim_{k\to\infty} \sum_{l=0}^{2^k} \lambda_{t_l^k} \left(\varphi(N_{t_{l-1}^k}+1, t_l^k) - \varphi(N_{t_{l-1}^k}, t_l^k)\right) - \lambda^k_{t_l^k} \left(\varphi(N^k_{t_{l-1}^k}+1, t_l^k) - \varphi(N^k_{t_{l-1}^k}, t_l^k)\right) = 0.$$

Fix $\omega \in \{N_s = M\}$ and let $\tilde{k} \in \mathbb{N}$ again be the minimum such that $N_s(\omega) = N_s^{\tilde{k}}(\omega)$. The equation simplifies to

$$\lim_{k\to\infty} 1_{\{k\geq\tilde{k}\}} \sum_{l=0}^{2^k-1} \left(\lambda_{t_l^k}(t_l^k - t_{l-1}^k) - \lambda^k_{t_l^k}\right) \left(\varphi(N_{t_{l-1}^k}+1, t_l^k) - \varphi(N_{t_{l-1}^k}, t_l^k)\right) = 0$$

Per assumption we have that

$$\lim_{t\to v} \frac{\mathbb{P}\left[N_t - N_v > 0 \mid \mathcal{H}_v\right]}{t-v} = \lambda_v.$$

uniformly on $[0, s)$. For $i \in \{1, ..., M\}$ it follows that

$$\begin{aligned}
&\sum_{l=0}^{2^k-1} 1_{\{T_i\leq t_l^k<T_{i+1}\}}(t_l^k - t_{l-1}^k)\left(\varphi(N_{t_{l-1}^k}+1, t_l^k)\right) \\
&\quad= \sum_{l=0}^{2^k-1} 1_{\{T_i\leq t_l^k<T_{i+1}\}}(t_l^k - t_{l-1}^k)\varphi(i+1, t_l^k) \\
&\quad\leq \sum_{l=0}^{2^k-1} 1_{\{T_i\leq t_l^k<T_{i+1}\}}(t_l^k - t_{l-1}^k)|\varphi(i+1, t_l^k)| \\
&\quad\to \int_{T_i}^{T_{i+1}} |\varphi(i+1, v)|dv.
\end{aligned}$$

Analogously, it follows that

$$\begin{aligned}
\sum_{l=0}^{2^k-1} 1_{\{T_i\leq t_l^k<T_{i+1}\}}(t_l^k - t_{l-1}^k)\varphi(N_{t_{l-1}^k}, t_l^k) &\leq \sum_{l=0}^{2^k-1} 1_{\{T_i\leq t_l^k<T_{i+1}\}}(t_l^k - t_{l-1}^k)|\varphi(i, t_l^k)| \\
&\to \int_{T_i}^{T_{i+1}} |\varphi(i, v)|dv.
\end{aligned}$$

Overall, it follows that

$$
\begin{aligned}
&\sum_{l=0}^{2^k-1}(t_l^k - t_{l-1}^k)\left(\varphi(N_{t_{l-1}^k}+1, t_l^k) - \varphi(N_{t_{l-1}^k}, t_l^k)\right)\\
&\quad = \sum_{i=0}^{M}\sum_{l=0}^{2^k-1} 1_{\{T_i \le t_l^k < T_{i+1}\}}(t_l^k - t_{l-1}^k)\left(\varphi(N_{t_{l-1}^k}+1, t_l^k) - \varphi(N_{t_{l-1}^k}, t_l^k)\right)\\
&\quad \le \sum_{i=0}^{M}\sum_{l=0}^{2^k-1} 1_{\{T_i \le t_l^k < T_{i+1}\}}(t_l^k - t_{l-1}^k)\left(|\varphi(i+1, t_l^k)| + |\varphi(i, t_l^k)|\right)\\
&\quad \to \sum_{i=0}^{M}\int_{T_i}^{T_{i+1}} |\varphi(i+1, v)| + |\varphi(i, v)| dv < \infty.
\end{aligned}
$$

Here we used the fact that the sum of M finite summands is again finite. Consequently,

$$
\begin{aligned}
&\sum_{l=0}^{2^k-1}\left(\lambda_{t_l^k}(t_l^k - t_{l-1}^k) - \lambda_{t_l^k}^k\right)\left(\varphi(N_{t_{l-1}^k}+1, t_l^k) - \varphi(N_{t_{l-1}^k}, t_l^k)\right)\\
&= \sum_{l=0}^{2^k-1}\left(\lambda_{t_l^k} - \frac{\lambda_{t_l^k}^k}{t_l^k - t_{l-1}^k}\right)(t_l^k - t_{l-1}^k)\left(\varphi(N_{t_{l-1}^k}+1, t_l^k) - \varphi(N_{t_{l-1}^k}, t_l^k)\right)\\
&\le \sup_{j\in\mathcal{D}_{s_1}^k}\left(\lambda_{t_j^k} - \frac{\lambda_{t_j^k}^k}{t_j^k - t_{j-1}^k}\right)\cdot\sum_{l=0}^{2^k-1}(t_l^k - t_{l-1}^k)\left(\varphi(N_{t_{l-1}^k}+1, t_l^k) - \varphi(N_{t_{l-1}^k}, t_l^k)\right)\\
&\to 0.
\end{aligned}
$$

Note that

$$
\int_0^{s_1}\varphi(N_{s-}+1, s) - \varphi(N_{s-}, s)\lambda_s ds = \int_0^{s_1}\varphi(N_s, s) - \varphi(N_{s-}, s)\lambda_s ds.
$$

For all $s \in I$ we have shown that

$$
\mathbb{P}\left[\lim_{k\to\infty} M_s^k - \int_0^s \varphi(N_v, v) - \varphi(N_{v-}, v)(dN_v - \lambda_v dv) = 0\right] = 1.
$$

The integrator $\varphi(N_v, v) - \varphi(N_{v-}, v)$ has left continuous paths and hence is a predictable process. Consequently, it follows that $M = (M_t)_{t\in I}$ is a (local) martingale.

□

As an example, consider $N = (N_t)_{t\in\mathbb{R}}$ to be a renormalized empirical distribution function of a sample with distribution F. Then, N can be viewed as a point process with compensator given by

$$d\Lambda_s = \frac{(n - N_s)}{1 - F(s)} dF(s).$$

As per Proposition 3.1.1, the innovation martingale M is established to be equal to

$$\begin{aligned} M_s = \lim_{x\to-\infty} \varphi(0, x) + \int_{-\infty}^{s} \big(\varphi(1 + N_{v-}, v) - \varphi(N_{v-}, v)\big) dN_v \\ - \int_{-\infty}^{s} \big(\varphi(1 + N_{v-}, v) - \varphi(N_{v-}, v)\big) \frac{(n - N_v)}{1 - F(v)} dF(v). \end{aligned} \tag{3.5}$$

Equation (3.6) can be found in Theorem 3.4 of [19].

Let $\tilde{N} = (\tilde{N}_t)_{t\in\mathbb{R}_+}$ be a point process with compensator

$$\tilde{\Lambda}_s = \rho s,$$

for some arbitrary $\rho > 0$. By invoking Proposition 3.1.1, we can express the innovation martingale $\tilde{M}$ as follows

$$\begin{aligned} \tilde{M}_s =& \varphi(0, 0) + \int_0^s \big(\varphi(1 + \tilde{N}_{v-}, v) - \varphi(\tilde{N}_{v-}, v)\big) d\tilde{N}_v \\ & - \int_0^s \big(\varphi(1 + \tilde{N}_{v-}, v) - \varphi(\tilde{N}_{v-}, v)\big) \rho dv. \end{aligned} \tag{3.6}$$

Proposition 3.1.2 (Compensator)
Let $N = (N_t)_{t\in I}$ be a point process with compensator $d\Lambda = \lambda dt$ and let $h = (h_t)_{t\in I}$ be a predictable process. We define the adapted process $(X_t)_{t\in I}$ by

$$X_s := \varphi(N_s, s), \qquad \text{for all } s \in I.$$

If φ meets the usual conditions, the compensator of X with respect to $\mathcal{H}$ is given by the process $A = (A_t)_{t\in I}$ with

$$A_s = \int_{s_0}^{s} \frac{\partial\varphi}{\partial x_2}(N_v, v) + \Big(\varphi(1 + N_{v-}, v) - \varphi(N_{v-}, v)\Big) \lambda_v dv.$$

Proof.

The proof is oriented towards the limiting procedure used in Beiglboeck [3]. N is a process on the intervall $[0, s_1)$. To make use of our results from Chapter two, we first define a sequence of discrete point processes $(N^k)_{k\in\mathbb{N}}$ on the dyadic numbers $\mathcal{D}_t^k$. Let $s \in I$ be fixed but arbitrary. For $k \in \mathbb{N}$ the set $\mathcal{D}_s^k$ is given by

$$\mathcal{D}_s^k = \left\{ \frac{ms}{2^k} \mid m \in \{1, ..., 2^k\} \right\}.$$

For $l \in \{1, ..., 2^k\}$ we define the l-th element in $\mathcal{D}_s^k$ by

$$t_l^k = \frac{ls}{2^k}.$$

The discrete point process $(N_t^k)_{t\in\{t_1^k,...,t_{2^k-1}^k\}}$ is defined by

$$N_0^k = N_0$$

and for $0 < n \in \mathcal{D}_s^k$ by

$$N_{t_n^k}^k = N_{t_{n-1}^k}^k + 1_{\{N_{t_n^k} > N_{t_{n-1}^k}\}}.$$

Note that, the random variable $N_{t_n^k}^k$ is not equal to $N_{t_n^k}$ for $\omega \in \{N_{t_n}^k - N_{t_{n-1}}^k > 1\}$. But since N is a simple locally finite point process on the interval $[0, s_1)$, there exists a $\tilde{k}$ for each fixed but arbitrary $\omega \in \Omega$ such that $N_{t_n^k} = N_{t_n^k}$ for all $k \geq \tilde{k}, t_n^k \in \mathcal{D}_s^k$. N^k is a point process in discrete time with conditional intensity $(\lambda_t^k)_{t\in\{t_1^k,...t_{2^k-1}^k\}}$ given by

$$\lambda_{t_n^k}^k = \mathbb{P}\left[N_{t_n}^k > N_{t_{n-1}}^k \mid \mathcal{H}_{t_{n-1}^k}\right].$$

Define $X^k = (X^k)_{t\in\{t_1^k,...,t_{2^k-1}^k\}}$ by

$$X_{t_n^k}^k = \varphi(N_{t_n^k}, t_n^k).$$

According to Proposition 2.2.1, the compensator $(A_t^k)_{t\in\{t_1^k,...,t_{2^k}\}}$ in the discrete Doob-Meyer decomposition of the process X^k is given by

$$\begin{aligned} A_{t_n^k}^k = &\sum_{l=1}^{2^k-1} \varphi\left(N_{t_{l-1}^k}^k, t_l^k\right) - \varphi\left(N_{t_{l-1}^k}^k, t_{l-1}^k\right) \\ &- \sum_{l=1}^{n} \lambda_{t_l^k}^k \left(\varphi(N_{t_{l-1}^k} + 1, t_l^k) - \varphi(N_{t_{l-1}^k}, t_l^k)\right). \end{aligned}$$

We show for the first sum

$$\sum_{l=1}^{2^k-1} \varphi\left(N^k_{t^k_{l-1}}, t^k_l\right) - \varphi\left(N^k_{t^k_{l-1}}, t^k_{l-1}\right) \to \int_0^s \frac{\partial\varphi}{\partial x_2}(N_{v-}, v)dv.$$

We assumed that the function φ has continuous first-order partial derivatives. Hence for all $0 \leq a \leq b \leq s$ it follows that

$$\sum_{l=1}^{2^k-1} 1_{\{a\leq t^k_l<b\}}\varphi\left(N^k_{t^k_{l-1}}, t^k_l\right) - \varphi\left(N^k_{t^k_{l-1}}, t^k_{l-1}\right) \to \int_a^b \frac{\partial\varphi}{\partial x_2}(N_{v-}, v)dv.$$

Let $M \in \mathbb{N}$ be arbitrary and fix $\omega \in \{N_s = M\}$. For all $k \geq \tilde{k}$ we define $T^k_1(\omega), ..., T^k_M(\omega)$ as the ordered points of the point process N^k and $T_0 = 0, T_{M+1} = s$. It follows that

$$\begin{aligned}
&\sum_{l=1}^{2^k-1} \varphi\left(N^k_{t^k_{l-1}}, t^k_l\right) - \varphi\left(N^k_{t^k_{l-1}}, t^k_{l-1}\right) \\
&\quad = \sum_{i=1}^{M}\sum_{l=1}^{2^k-1} 1_{\{T_i\leq t^k_l<T_{i+1}\}}\varphi\left(N^k_{t^k_{l-1}}, t^k_l\right) - \varphi\left(N^k_{t^k_{l-1}}, t^k_{l-1}\right) \\
&\quad \to \sum_{i=0}^{M}\int_{T_i}^{T_{i+1}} \frac{\partial\varphi}{\partial x_2}\varphi(N_{v-})dv = \int_0^s \frac{\partial\varphi}{\partial x_2}\varphi(N_{v-})dv.
\end{aligned}$$

The set $\bigcup_{M=0}^{\infty}\{N_s = M\}$ is a countable partition of Ω. Hence, we have shown $\mathbb{P}$-almost sure convergence for the first integral. For the second term we have shown in Proposition 3.1.1. the following:

$$\sum_{l=1}^{n} \lambda^k_{t^k_l}\left(\varphi(N_{t^k_{l-1}}+1, t^k_l) - \varphi(N_{t^k_{l-1}}, t^k_l)\right) \to \int_0^s \varphi(N_{v-}, v) - \varphi(N_{v-}, v)\lambda_v dv.$$

The first integral is a continuous function $\mathbb{P}$-almost sure. Furthermore, the second integral is left-continuous $\mathbb{P}$-almost sure. Consequently, the process $(A_t)_t$ is a predictable process.

□

The following Lemma provides a sufficient criterion for the function φ to become a martingale transformation.

Lemma 3.1.3 (Martingale differential equation)
Let φ be a function that meets the usual conditions. Let $N = (N_t)_{t\in I}$ be a point process with compensator $d\Lambda_s = \lambda_s ds$. Furthermore, assume that λ admits the representation

$$\lambda_s = \lambda(N_s, s),$$

for a deterministic function $\lambda : \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}_+$. The adapted process $X = (X_t)_{t\in I}$ given by $X_s = \varphi(N_s, s)$ is a martingale with respect to $\mathcal{H}$ if φ satisfies the differential equation

$$\frac{\partial\varphi}{\partial x_2}(x,t) + \Big(\varphi(1+x,t) - \varphi(x,t)\Big)\lambda(x,t) = 0,$$

for all $x, t \in \mathbb{R}_+$.

Proof.

This follows directly from Proposition 3.1.2.

□

To handle a point process with conditional intensity

$$\lambda(t \mid \mathcal{H}_t) = \frac{(n - N_t)f(t)}{1 - F(t)},$$

we need to solve the equation

$$\frac{\partial\varphi}{\partial F}(N_{v-}, v) + \Big(\varphi(1 + N_{v-}, v) - \varphi(N_{v-}, v)\Big)\frac{(n - N_v)}{1 - F(v)} = 0. \qquad (3.7)$$

Clearly, equation (3.7) is equivalent to the statement that for all $v \in \mathbb{R}$ and $x \in \{1, 2, ..., n\}$ the function φ satisfies

$$\frac{\partial\varphi}{\partial F}(x, v) + \Big(\varphi(1 + x, v) - \varphi(x, v)\Big)\frac{(n - x)}{1 - F(v)} = 0,$$

which can be found in Theorem 3.7 of [19].

In this section, we derived sufficient conditions under which a transformation φ yields a martingale of the form $X_t = \varphi(N_t, t)$. Within the field of point process theory, martingales play a fundamental role and have been the subject of research for several decades. A early result concerning point processes and martingales can be found by Watanabe and Kunita [25] stating that a Poisson process can be characterized by the martingale property of the process

$$N_t - \int_0^t \lambda(s) ds,$$

where λ is a deterministic function. A important work by Davis was given in 1976 [8] showing that every martingale M with respect to the natural generated filtration can be represented as

$$M_t = \int h_s(dN_s - d\Lambda_s).$$

Therefore, we can conclude that the martingales considered in this work can be represented in the integral form given above, which we will also provide explicitly in later investigations. A variant of Itô's formula tailored to our special case of point processes with jumpsize 1 was also given in chapter 1 with the help of Protter [31]. Originally formulated for Brownian motion as the driving process, the concept has since evolved into numerous variants, including those involving functional transformations of semimartingales (see Dupire [10]). A survey of the different generalizations of this formula can be found in [26]. Consequently, the framework adopted in this section is not the most general possible. However, as demonstrated in the following sections, it sufficiently meets the requirements of the results presented in this work. We will see that it is sufficient to work with transformations in two variables, while still achieving path-dependent martingale transformations through the following technique. Starting from a point process N, new point processes $\tilde{N}$ can be constructed by removing specific points from the original process. For example, consider the point process $\tilde{N}_t = 1_{\{N_t \geq n\}}$ for some arbitrary $n \in \mathbb{N}$. Processes of the form $\tilde{X} = \varphi(\tilde{N}, t)$ will be shown to be path-dependent martingale transformations of the original point process N.

3.2 Resolving the Equation

To establish the usefulness of the methods, we will initiate the discussion by examining two widely recognized solutions. This section will commence by explicitly utilizing the conditional intensity, a common approach in the literature of point processes. Subsequent examinations will rely on the integrated conditional intensity, also known as the compensator, as cumulative quantities' existence is established within a broader context.

Example 3.2.1 (Trivial martingale)
Deducting the integrated conditional intensity from a point process leads to the most popular martingale. By defining

$$\varphi(x, v, h|_v) := x - \int_0^v h(w)dw,$$

we have

$$\begin{aligned}\frac{\partial \varphi}{\partial v}(N_{v-}, v, \lambda|_v) &= -\lambda(v \mid \mathcal{H}_v) + \lambda(0 \mid \mathcal{H}_0) \\ &= -\lambda(v \mid \mathcal{H}_v)\end{aligned}$$

and

$$\Big(\varphi(1 + N_{v-}, v, \lambda|_v) - \varphi(N_{v-}, v, \lambda|_v)\Big)\lambda(v \mid \mathcal{H}_v) = \lambda(v \mid \mathcal{H}_v).$$

Hence, it can be inferred that φ satisfies the aforementioned differential equation.

Example 3.2.2 (Survival Analysis)
Let Z be a random variable with density f and $n = 1$. Define the point process $N = (N_t)_{t \in \mathbb{R}}$ by

$$N_s := 1_{\{Z \leq s\}}.$$

Then the internal conditional intensity $\lambda = (\lambda(t \mid \mathcal{H}_t))_{t \in \mathbb{R}}$ is given by

$$\lambda(s \mid \mathcal{H}_s) = \frac{(n - N_s)f(s)}{1 - F(s)}.$$

The martingale $(X_t)_{t \in \mathbb{R}}$ given by

$$X_s := (1 - N_s)e^{\int_{-\infty}^s \frac{f(v)}{1 - F(v)}dv},$$

is known from the field of survival analysis. We define the function φ as follows:

$$\varphi(x, v) := (1 - x)e^{\int_{-\infty}^v \frac{f(w)}{1 - F(w)}dw}.$$

We can conclude that

$$\begin{aligned}\frac{\partial\varphi}{\partial v}(N_{v-},v) &= (1-N_{v-})\frac{f(v)}{1-F(v)}e^{\int_{-\infty}^{t}\frac{f(v)}{1-F(v)}dv}\\ &= \lambda(v\mid\mathcal{H}_v)e^{\int_{-\infty}^{t}\frac{f(v)}{1-F(v)}dv}\end{aligned}$$

and

$$\begin{aligned}&\Big(\varphi(1+N_{v-},v)-\varphi(N_{v-},v)\Big)\lambda(v\mid\mathcal{H}_v)\\ &=\left((1-1+N_{v-})e^{\int_{-\infty}^{t}\frac{f(v)}{1-F(v)}dv}-(1-N_{v-})e^{\int_{-\infty}^{t}\frac{f(v)}{1-F(v)}dv}\right)\lambda(v\mid\mathcal{H}_v)\\ &=-\lambda(v\mid\mathcal{H}_v)e^{\int_{-\infty}^{t}\frac{f(v)}{1-F(v)}dv}.\end{aligned}$$

Hence, it can be inferred that φ satisfies the aforementioned differential equation.

In the subsequent Proposition, we will provide a slight generalization of Example 3.2.2. Note that each martingale X derived from the differential equation will be confirmed through the Martingale Representation Theorem at the end of this section by defining a predictable process h that yields

$$X_t = \int_{s_0}^{t} h_v(dN_v - d\Lambda_v).$$

By defining a predictable process h that generates the martingales obtained in the subsequent calculations, the procedures applied are validated, and at the same time a connection is established with the well-known martingale representation theorem. In summary, we will offer two distinct proofs to verify the validity of the resulting martingales. The first proof will rely on solving differential equations, while the second will involve defining a predictable process h, such that the generated integral martingale yields the martingale in question.

Proposition 3.2.3 (Non-negative martingale)
Let $N=(N_t)_{t\in I}$ be a point process with internal conditional $\mathcal{H}$-intensity $\lambda=(\lambda(t\mid\mathcal{H}_t))_{t\in I}$. Then the stochastic process $X=(X_t)_{t\in I}$ defined by

$$X_s := 1_{\{N_s=0\}}\cdot e^{\int_{s_0}^{s}\lambda(v\mid\mathcal{H}_v)dv},$$

is a non-negative martingale with respect to $\mathcal{H}$.

Proof.

Let $\tilde{N} = (\tilde{N}_t)_{t\in I}$ be a point process defined by

$$\tilde{N}_t = 1_{\{N_t \geq 1\}}.$$

Clearly, the conditional intensity of $\tilde{N}$ with respect to $\mathcal{H}$ is given by

$$\tilde{\lambda}_t = \lambda_t 1_{\{N_t=0\}}.$$

It follows that

$$X_t = (1 - \tilde{N}_t)e^{\Lambda_t}.$$

Define $\varphi(x, y) : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}$ by

$$\varphi(x, y) = (1 - x)e^y.$$

Consequently, it follows that $X_t = \varphi(N_t, \Lambda_t)$. To complete the proof, we derive a slight generalization of the results presented in Proposition 3.1.1 and 3.1.2.

By Itô's formula it follows that

$$\begin{aligned}\varphi(\tilde{N}_t, \tilde{\Lambda}_t) &= \varphi(0,0) + \int_0^t \frac{\partial\varphi}{\partial x_2}(\tilde{N}_{s-}, \tilde{\Lambda}_s)d\tilde{\Lambda}_s + \int_0^t (\varphi(1 + \tilde{N}_{s-}, \tilde{\Lambda}_s) - \varphi(\tilde{N}_{s-}, \tilde{\Lambda}_s))dN_s \\ &= \varphi(0,0) + \int_0^t \left(\frac{\partial\varphi}{\partial x_2}(\tilde{N}_{s-}, \tilde{\Lambda}_s) + (\varphi(1 + \tilde{N}_{s-}, \tilde{\Lambda}_s) - \varphi(\tilde{N}_{s-}, \tilde{\Lambda}_s))\right)\tilde{\lambda}_s ds \\ &\quad + \int_0^t (\varphi(1 + \tilde{N}_{s-}, \tilde{\Lambda}_s) - \varphi(\tilde{N}_{s-}, \tilde{\Lambda}_s))(d\tilde{N}_s - d\tilde{\Lambda}_s)\end{aligned}$$

Since φ satisfies the usual conditions, it follows that the second integral process is a martingale. Consequently, it remains to proof that the first integral vanishes. For the first factor in the integrator it follows that

$$\frac{\partial\varphi}{\partial x_2}(x_1, x_2) + (\varphi(1 + x_1, x_2) - \varphi(x_1, x_2)) \ = -x_1 e^{x_2}$$

For $t \in [0, T_1)$ it follows that

$$\int_0^t -N_s e^{\tilde{\Lambda}_s}\tilde{\lambda}_s ds = \int_0^t 0 \cdot e^{\tilde{\Lambda}_s}\tilde{\lambda}_s ds = 0$$

and for $t \geq T_1$ it follows that

$$\int_0^t -N_s e^{\tilde{\Lambda}_s}\tilde{\lambda}_s ds = \int_0^t -N_s e^{\tilde{\Lambda}_s} \cdot 0 ds = 0.$$

□

The following theorem presents a minor extension of Proposition 3.2.3, specifically concerning the placement of the non-trivial portion of the transformation.

Proposition 3.2.4 (General non-negative martingale)
Let $k \in \mathbb{N}_0$ be arbitrarily but fixed. Let $N = (N_t)_{t \in I}$ be a point process with internal conditional $\mathcal{H}$-intensity $\lambda = (\lambda(t \mid \mathcal{H}_t))_{t \in I}$. The stochastic process $X^k = (X^k_t)_{t \in I}$ defined by

$$X^k_s := 1_{\{N_s \leq k\}} \cdot e^{\int_{s_0}^s 1_{\{N_v = k\}} \lambda(v|\mathcal{H}_v) dv},$$

is a non-negative martingale with respect to $\mathcal{H}$.

Proof.

Define the stochastic process $\tilde{N} = (\tilde{N}_t)_{t \in I}$ by setting

$$\tilde{N}_t := 1_{\{N_t > k\}}.$$

Clearly, $\tilde{N}$ is a point process. The point process N is adapted to the filtration $\mathcal{H}$. Hence, $\tilde{N}$ has the conditional $\mathcal{H}$-intensity $\tilde{\lambda} = (\tilde{\lambda}(t \mid \mathcal{H}_t))_{t \in I}$ with

$$\tilde{\lambda}(t \mid \mathcal{H}_t) = 1_{\{N_t = k\}} \lambda(t \mid \mathcal{H}_t).$$

The rest of the proof is the same as in the previous statement.

□

For any value of k, the martingale X^k begins at $X^k_0 = 1$. If it is $\mathbb{P}$-almost sure that there exists a value $v \in I$ such that $N_v = k + 1$, then there must be a value $t \in I$ such that for all values of $w > t$, $X_w = 0$. Thus, under mild conditions, the martingales obtained always initiate at state 1 and ultimately reach state 0.

The subsequent statement will also prove to be useful.

Lemma 3.2.5 (Combining property)
Let $B \subset \mathbb{N}$ be arbitrary but fixed with $|B| = n < \infty$. The process $X^B = (X_t^B)_{t\in I}$ defined by

$$X_s^B := \frac{1}{n}\sum_{k\in B} X_s^k,$$

is a non-negative martingale with respect to $\mathcal{H}$.

Proof.

For all $k \in \mathbb{N}$ the process X^k is a martingale with respect to the filtration $\mathcal{H}$. Therefore the proof is complete, since the sum of $\mathcal{H}$-martingales is again a $\mathcal{H}$-martingale.

□

Theorem 3.2.6 (General non-negative martingale of higher order)
Let $k_1 < k_2 \in \mathbb{N}_0$ be arbitrarily but fixed. Let $N = (N_t)_{t\in I}$ be a point process with compensator Λ. The stochastic process $X^k = (X_t^k)_{t\in I}$ defined by

$$X_s^{k_1,k_2} := e^{\int_{s_0}^s \left(1-\frac{1_{\{k_1\le N_v\}}(N_v\wedge k_2-k_1)+1}{k_2-k_1}\right)1_{\{k_1\le N_v<k_2\}}d\Lambda_v} \cdot \prod_{i=1}^{1_{\{k_1\le N_s\}}(N_s\wedge k_2-k_1)} (1-i/(k_2-k_1)),$$

is a non-negative martingale with respect to $\mathcal{H}$.

Proof.

Let $\tilde{N} = (\tilde{N}_t)_{t\in I}$ be a point process defined by

$$\tilde{N}_t = 1_{\{k_1\le N_t\}}(N_t\wedge k_2 - k_1).$$

Clearly, $\tilde{N}$ is a point process. The point process is adapted to the filtration $\mathcal{H}$. Hence, $\tilde{N}$ has the conditional $\mathcal{H}$-intensity $\tilde{\lambda} = (\tilde{\lambda}_t)_{t\in I}$ with

$$\tilde{\lambda}_t = 1_{\{k_1\le N_t\le k_2\}}\lambda_t.$$

The process X can now be represented as

$$X_t = e^{\int_0^t \left(1-\frac{\tilde{N}_s}{k_2-k_1}\right)\tilde{\lambda}_s ds} \prod_{i=1}^{\tilde{N}_t}(1 - i/(k_2 - k_1)).$$

We define the stochastic process $X^1 = (X_t^1)_{t\in I}$ by

$$X_t^1 = \int_0^t \left(1 - \frac{\tilde{N}_{s-} + 1}{k_2 - k_1}\right) \tilde{\lambda}_s ds$$

and $X^2 = (X_t^2)_{t\in I}$ by

$$X_t^2 = \prod_{i=1}^{\tilde{N}_t}(1 - i/(k_2 - k_1)).$$

Define $\varphi(x_1, x_2) : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ by

$$\varphi(x_1, x_2) = e^{x_1} x_2$$

Consequently, it follows that $X_t = \varphi(X_t^1, X_t^2)$. To complete the proof, we derive a slight generalization of the results presented in Proposition 3.1.1. and 3.1.2.

Since X^1 has continuous paths and X^2 is a pure jump process, it follows from Itô's formula that

$$\begin{aligned}
&\varphi(X_t^1, X_t^2) \\
&= \int_0^t \frac{\partial\varphi}{\partial x_1}(X_s^1, X_{s-}^2) dX_s^1 + \prod_{s\le t} \left(\varphi(X_s^1, X_s^2) - \varphi(X_s^1, X_{s-}^2)\right) \\
&= \int_0^t e^{X_s^1} X_{s-}^2 \left(1 - \frac{\tilde{N}_{s-} + 1}{k_2 - k_1}\right) \tilde{\lambda}_s ds \\
&\quad + \int_0^t \varphi\left(X_s^1, X_{s-}^2 \left(1 - \frac{\tilde{N}_{s-} + 1}{k_2 - k_1}\right)\right) - \varphi(X_s^1, X_{s-}^2) d\tilde{N}_s \\
&= \int_0^t e^{X_s^1} X_{s-}^2 \left(1 - \frac{\tilde{N}_{s-} + 1}{k_2 - k_1}\right) \tilde{\lambda}_s \\
&\qquad + \varphi\left(X_s^1, X_{s-}^2 \left(1 - \frac{\tilde{N}_{s-} + 1}{k_2 - k_1}\right)\right) - \varphi(X_s^1, X_{s-}^2)\tilde{\lambda}_s ds \\
&\quad + \int_0^t \varphi\left(X_s^1, X_{s-}^2 \left(1 - \frac{\tilde{N}_{s-} + 1}{k_2 - k_1}\right)\right) - \varphi(X_s^1, X_{s-}^2)(d\tilde{N}_s - \tilde{\lambda}_s ds)
\end{aligned}$$

For the integrator of the first integral it follows that

$$\begin{aligned}
&e^{X_s^1} X_s^2 \left(1 - \frac{\tilde{N}_{s-}+1}{k_2-k_1}\right) + \varphi\left(X_s^1, X_{s-}^2\left(1 - \frac{\tilde{N}_{s-}+1}{k_2-k_1}\right)\right) - \varphi(X_s^1, X_{s-}^2) \\
&= e^{X_s^1} X_s^2 \left(1 - \frac{\tilde{N}_{s-}+1}{k_2-k_1}\right) \\
&\qquad + e^{X_s^1}\left(\prod_{i=1}^{\tilde{N}_{s-}}(1 - i/(k_2-k_1))\left(1 - \frac{\tilde{N}_{s-}+1}{k_2-k_1}\right) - \prod_{i=1}^{\tilde{N}_{s-}}(1-i/(k_2-k_1))\right) \\
&= 0.
\end{aligned}$$

The integrator of the first integral is equal to 0. The integrator of the second integral is a predictable process. Hence φ is a martingale transformation.

□

Example 3.2.7
Let $0 \leq k < n$ be arbitrary. Putting $k_1 = k$ and $k_2 = k+1$ it follows that

$$\begin{aligned}
&X_s^{k,k+1} \\
&= e^{\frac{1}{k_2-k_1}\int_{s_0}^s (1+1_{\{k_1 \leq N_v\}}(N_v \wedge k_2 - k_1))1_{\{k_1 \leq N_v < k_2\}} d\Lambda_v} \prod_{i=1}^{1_{\{k_1 \leq N_s\}}(N_s \wedge k_2 - k_1)} (1 - i/(k_2-k_1)) \\
&= e^{\frac{1}{1}\int_{s_0}^s (1+1_{\{k \leq N_v\}}(N_v \wedge (k+1) - k)1_{\{k \leq N_v < k+1\}} d\Lambda_v} \prod_{i=1}^{1_{\{k \leq N_s\}}(N_s \wedge (k+1) - k)} (1 - i/1) \\
&= e^{\int_{s_0}^s 1_{\{N_v = k\}} d\Lambda_v} 1_{\{N_s \leq k\}} \\
&= X_s^k,
\end{aligned}$$

where X^k is from Theorem 3.2.4. This example shows that Theorem 3.2.6 is a natural generalization of Theorem 3.2.4.

The following theorem presents an additional method for constructing non-negative martingales.

Theorem 3.2.8 (General non-negative martingale of higher order II) Let $k_1, m, k_2 \in \mathbb{N}_0$ be arbitrary but fixed with $0 \leq m \leq k_2 - k_1$ and $k_1 < k_2$. Let $N = (N_t)_{t\in I}$ be a point process with compensator Λ. The stochastic process $X^k = (X_t^k)_{t\in I}$ defined by

$$X_s^{k_1,k_2}(m) := \frac{(k_2-k_1-m)!}{(k_2-k_1)!} e^{m\int_{s_0}^t \frac{1}{k_2-k_1-1_{\{k_1\leq N_t\}}(N_t\wedge k_2-k_1)} 1_{\{k_1\leq N_s<k_2\}} d\Lambda_s} \cdot \prod_{j=k_2-k_1-(m-1)}^{k_2-k_1} (j - 1_{\{k_1\leq N_t\}}(N_t - k_1))$$

is a non-negative martingale with respect to $\mathcal{H}$.

Proof.

Let $\tilde{N} = (\tilde{N}_t)_{t\in I}$ be a point process defined by

$$\tilde{N}_t = 1_{\{k_1\leq N_t\}}(N_t \wedge k_2 - k_1).$$

Clearly, $\tilde{N}$ is a point process. The point process is adapted to the filtration $\mathcal{H}$. Hence, $\tilde{N}$ has the conditional $\mathcal{H}$-intensity $\tilde{\lambda} = (\tilde{\lambda}_t)_{t\in I}$ with

$$\tilde{\lambda}_t = 1_{\{k_1\leq N_t\leq k_2\}}\lambda_t.$$

Define $\varphi(x, y) : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}$ by

$$\varphi(x,y) = \frac{(k_2-k_1-m)!}{(k_2-k_1)!} e^y \prod_{j=k_2-k_1-(m-1)}^{k_2-k_1} (j-x).$$

Define the process $Y = (Y_t)_{t\in I}$ by

$$Y_t = m\int_0^t \frac{1_{\{k_1\leq N_s\leq k\}}}{k_2-k_1-1_{\{k_1\leq N_s\}}(N_s\wedge k_2-k_1)} \lambda_s ds$$

Consequently, it follows that $X_t = \varphi(\tilde{N}_t, Y_t)$. To complete the proof, we derive a slight generalization of the results presented in Proposition 3.1.1. and 3.1.2.

By Itô's formula it follows that

$$\begin{aligned}
&\varphi(\tilde{N}_t, Y_t)\\
&= \varphi(0,0) + \int_0^t \frac{\partial\varphi}{\partial x_2}(\tilde{N}_{s-}, Y_s)dY_s + \int_0^t (\varphi(1+\tilde{N}_{s-}, Y_s) - \varphi(\tilde{N}_{s-}, Y_s))d\tilde{N}_s\\
&= \varphi(0,0) + \int_0^t \frac{\partial\varphi}{\partial x_2}(\tilde{N}_{s-}, Y_s)\frac{m1_{\{k_1\le N_s\le k_2\}}}{k_2-k_1-1_{\{k_1\le N_s\}}(N_s-k_1)}\tilde{\lambda}_s\\
&\qquad\qquad\qquad\qquad + (\varphi(1+\tilde{N}_{s-}, Y_s) - \varphi(\tilde{N}_{s-}, Y_s))\tilde{\lambda}_s ds\\
&\qquad + \int_0^t (\varphi(1+\tilde{N}_{s-}, Y_s) - \varphi(\tilde{N}_{s-}, Y_s))(d\tilde{N}_s - d\tilde{\Lambda}_s).
\end{aligned}$$

We have

$$\frac{\partial\varphi}{\partial x_2}(x_1,x_2) = \varphi(x_1,x_2).$$

Furthermore, it holds that

$$\begin{aligned}
&\varphi(1+x_1,x_2) - \varphi(x_1,x_2)\\
&= \frac{(k_2-k_1-m)!}{(k_2-k_1)!}e^y\left(\prod_{j=k_2-k_1-(m-1)}^{k_2-k_1}(j-x_1-1) - \prod_{j=k_2-k_1-(m-1)}^{k_2-k_1}(j-x_1)\right)\\
&= \frac{(k_2-k_1-m)!}{(k_2-k_1)!}e^y\left(\prod_{j=k_2-k_1-(m-1)-1}^{k_2-k_1-1}(j-x_1) - \prod_{j=k_2-k_1-(m-1)}^{k_2-k_1}(j-x_1)\right)\\
&= \frac{(k_2-k_1-m)!}{(k_2-k_1)!}e^y\frac{-m}{k_2-k_1-x_1}\prod_{j=k_2-k_1-(m-1)}^{k_2-k_1-1}(j-x_1)\\
&= -\frac{m}{k_2-k_1-x_1}\varphi(x_1,x_2).
\end{aligned}$$

Note that $\lambda_s = 0$ for $s \in [0, T_{k_1}) \cup [T_{k_2}, \infty)$. For $s \in [T_{k_1}, T_{k_2})$ we have the following

$$\frac{m1_{\{k_1\le N_s\le k_2\}}}{k_2-k_1-1_{\{k_1\le N_s\}}(N_s-k_1)} = \frac{m}{k_2-k_1-\tilde{N}_s}.$$

Overall, it follows that the integrator of the first integral is equal to 0. The integrator of the second integral is a predictable process. Hence φ is a martingale transformation.

□

Example 3.2.9 (Obtain Hess' martingales) (paper in work)
Let N be a renormalized empirical distribution function. If we choose $k_1 = 0$ and $k_2 = n$ we obtain the martingale

$$\begin{aligned} X_s^{0,n}(m) &= \frac{(n-m)!}{n!} e^{m \int_{-\infty}^{t} \frac{1}{n - 1_{\{0 \le N_t\}}(N_t - 0)} 1_{\{0 \le N_t < n\}} d\Lambda_s} \\ &\qquad \cdot \prod_{j=n-(m-1)}^{n} (j - 1_{\{0 \le N_t\}}(N_t - 0)) \\ &= \frac{(n-m)!}{n!} e^{m \int_{-\infty}^{t} \frac{1}{n-N_t} \frac{(n-N_t)}{1-F(t)} dF(s)} \prod_{j=n-(m-1)}^{n} (j - N_t) \\ &= \frac{(n-m)!}{n!} e^{m \int_{-\infty}^{t} \frac{1}{1-F(s)} dF(s)} \prod_{j=n-(m-1)}^{n} (j - N_t) \\ &= \frac{n^m}{n^m} \frac{(n-m)!}{n!} \frac{1}{(1-F(t))^m} \prod_{j=n-(m-1)}^{n} (j - N_t) \\ &= \frac{n^m (n-m)!}{n!} \prod_{j=n-(m-1)}^{n} \frac{\frac{j}{n} - F_n(t)}{1 - F(t)}. \end{aligned}$$

The following Propositions aim at verifying the martingales obtained so far through the martingale representation theorem. It is well known that for each predictable process g the process X given by

$$X_t = \int_{s_0}^{t} g_v (dN_v - d\Lambda_v),$$

defines a martingale. Thus, if we succeed in finding a predictable process g that generates the martingales obtained in this section, we will have verified our results again. We start our journey with the special case of Hess' martingales (paper in work).

Proposition 3.2.10
Let N be a renormalized empirical distribution function. Define the predictable process g by

$$g_v := -X_{v-} \left(\frac{m}{n - N_{v-}} \right).$$

Then it follows that the process

$$X_t = \frac{(n-m)!n^m}{n!} \prod_{j=n-(m-1)}^{n} \frac{\frac{j}{n} - F_n(t)}{1-F(t)}$$

solves the equation

$$X_t = \int_{s_0}^{t} g_v(dN_v - d\Lambda_v) + 1.$$

Proof.

We need to show that the process X defined by

$$X_t := \frac{(n-m)!n^m}{n!} \prod_{j=n-(m-1)}^{n} \frac{\frac{j}{n} - F_n(t)}{1-F(t)}$$

solves the equation

$$X_t = \int_{s_0}^{t} -X_{v-} \left(\frac{m}{n - N_{v-}} \right) (dN_v - d\Lambda_v) + 1.$$

The method to demonstrate this is by using induction across the intervals $[T_k, T_{k+1})$.

Base case $k = 0$.

Note that for $t \in [s_0, T_1)$, it follows that

$$\left(\frac{1}{1-F(t)} \right)^m = \frac{(n-m)!n^m}{n!} \prod_{j=n-(m-1)}^{n} \frac{\frac{j}{n} - F_n(t)}{1-F(t)}.$$

Hence, it remains to prove

$$-\int_{s_0}^{t} \left(\frac{1}{1-F(v)} \right)^m \left(\frac{m}{n - N_{v-}} \right) (dN_v - d\Lambda_v) + 1 = \left(\frac{1}{1-F(t)} \right)^m,$$

for all $t \in [s_0, T_1)$. But since $t \in [s_0, T_1)$, the equation simplifies to

$$\frac{m}{n} \int_{s_0}^{t} \left(\frac{1}{1-F(v)} \right)^m \frac{n}{1-F(v)} dF(v) + 1 = \left(\frac{1}{1-F(t)} \right)^m$$

for all $t \in [s_0, T_1)$.

But this is satisfied since for $t < T_1$ it follows that

$$\begin{aligned}
\frac{m}{n}\int_{s_0}^{t}\left(\frac{1}{1-F(v)}\right)^{m}\frac{n}{1-F(v)}dF(v)+1 &= m\int_{s_0}^{t}\left(\frac{1}{1-F(v)}\right)^{m+1}dF(v)+1\\
&= m\int_{0}^{F(t)}\left(\frac{1}{1-x}\right)^{m+1}dx+1\\
&= \frac{m}{m}\left(\left(\frac{1}{1-F(t)}\right)^{m}-\left(\frac{1}{1-0}\right)^{m}\right)+1\\
&= \left(\frac{1}{1-F(t)}\right)^{m}.
\end{aligned}$$

IS: Assume the claim is true for $k = 1, 2, ..., i$. We proof that the statement is also true for $k = i + 1$.

We start with the jumping time $t = T_i$. Since the claim is true for all points in time before t it follows that

$$\begin{aligned}
&\int_{s_0}^{T_i} -X_{v-}\left(\frac{m}{n-N_{v-}}\right)(dN_v - d\Lambda_v)+1\\
&= \int_{[s_0,T_i)} -X_{v-}\left(\frac{m}{n-N_{v-}}\right)(dN_v - d\Lambda_v) - X_{T_i-}\left(\frac{m}{n-N_{T_i-}}\right)+1\\
&= X_{T_i-} - X_{T_i-}\left(\frac{m}{n-N_{T_i-}}\right)\\
&= X_{T_i-} + \frac{(n-m)!}{n!(1-F(T_i))^m}\prod_{j=n-(m-1)}^{n-1}(j-i+1)\frac{-m}{n-i+1}\\
&= X_{T_i-} + \frac{(n-m)!}{n!(1-F(T_i))^m}\left[\prod_{j=n-(m-1)-1}^{n-1}(j-i+1) - \prod_{j=n-(m-1)}^{n}(j-i+1)\right]\\
&= X_{T_i-} + \frac{(n-m)!n^m}{n!}\left[\prod_{j=n-(m-1)}^{n}\frac{\frac{j}{n}-\frac{i}{n}}{1-F(T_i)} - \prod_{j=n-(m-1)}^{n}\frac{\frac{j}{n}-\frac{i-1}{n}}{1-F(T_i-)}\right]\\
&= X_{T_i-} + (X_{T_i} - X_{T_i-})\\
&= X_{T_i}.
\end{aligned}$$

It remains to prove the statement for $t \in (T_i, T_{i+1})$. It follows that

$$\begin{aligned}
&\int_{s_0}^{t} -X_{v-}\left(\frac{m}{n-N_{v-}}\right)(dN_v - d\Lambda_v) + 1 \\
&= X_{T_i} + \int_{T_i}^{t} -X_{v-}\left(\frac{m}{n-N_{v-}}\right) d\Lambda_v \\
&= X_{T_i} + \int_{T_i}^{t} -X_{v-}\left(\frac{m}{n-i}\right)\frac{(n-i)}{1-F(v)} dF(v) \\
&= X_{T_i} + m\int_{T_i}^{t} -X_{v-}\frac{1}{1-F(v)} dF(v).
\end{aligned}$$

Hence, the statement follows if we show the equality

$$X_t = X_{T_i} + m\int_{T_i}^{t} -X_{v-}\frac{1}{1-F(v)} dF(v).$$

This is true since

$$
\begin{aligned}
&X_{T_i} + m\int_{T_i}^{t} -X_{v-}\frac{1}{1-F(v)}dF(v) \\
&= \frac{(n-m)!n^m}{n!}\prod_{j=n-(m-1)}^{n}\frac{\frac{j}{n}-F_n(T_i)}{1-F(T_i)} \\
&\qquad + m\int_{T_i}^{t}\frac{(n-m)!n^m}{n!}\left(\prod_{j=n-(m-1)}^{n}\frac{\frac{j}{n}-F_n(T_i)}{1-F(v)}\right)\frac{1}{1-F(v)}dF(v) \\
&= \frac{(n-m)!n^m}{n!}\prod_{j=n-(m-1)}^{n}\frac{\frac{j}{n}-\frac{i}{n}}{1-F(T_i)} \\
&\qquad + m\int_{T_i}^{t}\frac{(n-m)!n^m}{n!}\left(\prod_{j=n-(m-1)}^{n}\frac{\frac{j}{n}-\frac{i}{n}}{1-F(v)}\right)\frac{1}{1-F(v)}dF(v) \\
&= \frac{(n-m)!}{(n-1)!}\prod_{j=n-(m-1)}^{n}(j-i)\left[\left(\frac{1}{1-F(T_i)}\right)^m + m\int_{T_i}^{t}\left(\frac{1}{1-F(v)}\right)^{m+1}dF(v)\right] \\
&= \frac{(n-m)!}{(n-1)!}\prod_{j=n-(m-1)}^{n}(j-i)\left[\left(\frac{1}{1-F(T_i)}\right)^m + m\int_{F(T_i)}^{F(t)}\left(\frac{1}{1-x}\right)^{m+1}dx\right] \\
&= \frac{(n-m)!}{(n-1)!}\prod_{j=n-(m-1)}^{n}(j-i)\left[\left(\frac{1}{1-F(T_i)}\right)^m + \left(\left(\frac{1}{1-F(t)}\right)^m - \left(\frac{1}{1-F(T_i)}\right)^m\right)\right] \\
&= \frac{(n-m)!}{(n-1)!}\prod_{j=n-(m-1)}^{n}(j-i)\left(\frac{1}{1-F(t)}\right)^m \\
&= \frac{(n-m)!n^m}{n!}\prod_{j=n-(m-1)}^{n}\frac{\frac{j}{n}-F_n(t)}{1-F(t)} \\
&= X_t.
\end{aligned}
$$

□

Proposition 3.2.11
Let N be a point process and λ the conditional intensity of N with respect

to the natural filtration $\mathcal{H}$. Define the predictable process g by

$$g_v := -X_{v-}\left(\frac{m}{k_2 - N_{v-}}\right) 1_{\{k_1 \le N_{v-} < k_2\}}.$$

Then it follows that the process

$$X_t = \frac{(k_2 - k_1 - m)!}{(k_2 - k_1)!} e^{m \int_{s_0}^{t} \frac{1}{k_2 - k_1 - 1_{\{k_1 \le N_s\}}(N_s \wedge k_2 - k_1)} 1_{\{k_1 \le N_s < k_2\}} d\Lambda_s} \cdot \prod_{j=k_2-k_1-(m-1)}^{k_2-k_1} (j - 1_{\{k_1 \le N_t\}}(N_t \wedge k_2 - k_1))$$

solves the equation

$$X_t = \int_{s_0}^{t} g_v (dN_v - d\Lambda_v) + 1.$$

Proof.

We need to show that the process X defined by

$$X_t := \frac{(k_2 - k_1 - m)!}{(k_2 - k_1)!} e^{m \int_{s_0}^{t} \frac{1}{k_2 - k_1 - 1_{\{k_1 \le N_s\}}(N_s \wedge k_2 - k_1)} 1_{\{k_1 \le N_s < k_2\}} d\Lambda_s} \cdot \prod_{j=k_2-k_1-(m-1)}^{k_2-k_1} (j - 1_{\{k_1 \le N_t\}}(N_t - k_1))$$

solves the equation

$$X_t = \int_{s_0}^{t} -X_{v-}\left(\frac{m}{k_2 - N_{v-}}\right) 1_{\{k_1 \le N_{v-}\}}(dN_v - d\Lambda_v) + 1. \tag{A}$$

We show this by induction through the intervals $[T_k, T_{k+1})$ for all $k \ge k_1$. Notice, it is easy to check that (A) is satisfied for all $t \in [s_0, T_{k_1})$.

Base case $k = k_1$.

We start with the jumping time $t = T_{k_1}$. It follows that

$$X_{T_{k_1}} = X_{T_{k_1}-} - X_{T_{k_1}-}\left(\frac{m}{k_2 - (k_1 - 1)}\right) 1_{\{k_1 \le k_1 - 1\}} = 1 + 0 = 1.$$

For $t \in (T_{k_1}, T_{k_1+1})$ it follows that

$$
\begin{aligned}
&X_{T_{k_1}} + \int_{T_{k_1}}^{t} -X_{v-}\left(\frac{m}{k_2 - N_{v-}}\right) 1_{\{k_1 \le N_{v-}\}}(dN_v - d\Lambda_v) \\
&= 1 + \frac{m}{k_2 - k_1}\int_{T_{k_1}}^{t} X_{v-} d\Lambda_v.
\end{aligned}
$$

Therefore, we only need to show that the process satisfies the following equation

$$
X_t = 1 + \frac{m}{k_2 - k_1}\int_{T_{k_1}}^{t} X_{v-} d\Lambda_v,
$$

for all $t \in (T_{k_1}, T_{k_1+1})$. But this is true since

$$
\begin{aligned}
&1 + \frac{m}{k_2 - k_1}\frac{(k_2 - k_1 - m)!}{(k_2 - k_1)!}\int_{T_{k_1}}^{t} e^{\frac{m}{k_2 - k_1}\int_{T_{k_1}}^{v} d\Lambda_s} d\Lambda_v \prod_{j=k_2-k_1-(m-1)}^{k_2-k_1} j \\
&= 1 + \frac{m}{k_2 - k_1}\frac{(k_2 - k_1 - m)!}{(k_2 - k_1)!}\prod_{j=k_2-k_1-(m-1)}^{k_2-k_1} j\,\frac{k_2 - k_1}{m}\left(e^{\frac{m}{k_2 - k_1}\int_{T_{k_1}}^{t} d\Lambda_s} - 1\right) \\
&= \frac{(k_2 - k_1 - m)!}{(k_2 - k_1)!} e^{\frac{m}{k_2 - k_1}\int_{T_{k_1}}^{t} d\Lambda_s} \prod_{j=k_2-k_1-(m-1)}^{k_2-k_1} j \\
&= \frac{(k_2 - k_1 - m)!}{(k_2 - k_1)!} e^{m\int_{s_0}^{t} \frac{1}{k_2 - k_1 - 1_{\{k_1 \le N_t\}}(N_t \wedge k_2 - k_1)} 1_{\{k_1 \le N_s < k_2\}} d\Lambda_s} \\
&\qquad\qquad \cdot \prod_{j=k_2-k_1-(m-1)}^{k_2-k_1} (j - 1_{\{k_1 \le N_t\}}(N_t - k_1)) \\
&= X_t.
\end{aligned}
$$

IS: Assume the claim is true for $k = 1, 2, ..., i$. We proof that the statement is also true for $k = i + 1$.

We start with the jumping time $t = T_i$. Since the claim is true for all points in time before t it follows that

$$\begin{aligned}
X_{T_i} &= X_{T_i-} - X_{T_i-}\frac{m}{k_2-(i-1)} \\
&= X_{T_i-} + \frac{(k_2-k_1-m)!}{(k_2-k_1)!}e^{m\int_{s_0}^{T_i}\frac{1}{k_2-k_1-1_{\{k_1\le N_s\}}(N_s\wedge k_2-k_1)}1_{\{k_1\le N_s<k_2\}}d\Lambda_s} \\
&\qquad\cdot\left(\prod_{j=k_2-k_1-(m-1)-1}^{k_2-k_1-1}(j-(i-1-k_1)) - \prod_{j=k_2-k_1-(m-1)}^{k_2-k_1}(j-(i-1-k_1))\right) \\
&= X_{T_i-} + \frac{(k_2-k_1-m)!}{(k_2-k_1)!}e^{m\int_{s_0}^{T_i}\frac{1}{k_2-k_1-1_{\{k_1\le N_s\}}(N_s\wedge k_2-k_1)}1_{\{k_1\le N_s<k_2\}}d\Lambda_s} \\
&\qquad\cdot\prod_{j=k_2-k_1-(m-1)}^{k_2-k_1}(j-1_{\{k_1\le i\}}(i-k_1))) \\
&\quad - \frac{(k_2-k_1-m)!}{(k_2-k_1)!}e^{m\int_{s_0}^{T_i}\frac{1}{k_2-k_1-1_{\{k_1\le N_s\}}(N_s\wedge k_2-k_1)}1_{\{k_1\le N_s<k_2\}}d\Lambda_s} \\
&\qquad\cdot\prod_{j=k_2-k_1-(m-1)}^{k_2-k_1}(j-1_{\{k_1\le(i-1)\}}((i-1)-k_1))) \\
&= X_{T_i-} + X_{T_i} - X_{T_i-} \\
&= X_{T_i-}.
\end{aligned}$$

It remains to prove the statement for $t \in (T_i, T_{i+1})$. It follows that

$$\begin{aligned}
X_t &= X_{T_i} + \int_{T_i}^t -X_{v-}\left(\frac{m}{k_2-N_{v-}}\right)1_{\{k_1\le N_{v-}\}}(dN_v - d\Lambda_v) \\
&= X_{T_i} + \frac{m}{k_2-i}\int_{T_i}^t X_{v-}d\Lambda_v.
\end{aligned}$$

Hence, the statement follows if we show the equality

$$X_t = X_{T_i} + \frac{m}{k_2-i}\int_{T_i}^t X_{v-}d\Lambda_v.$$

This is true since

$$
\begin{aligned}
&\int_{T_i}^t X_{v-} d\Lambda_v \\
&= \int_{T_i}^t \frac{(k_2-k_1-m)!}{(k_2-k_1)!} e^{m \int_{s_0}^v \frac{1}{k_2-k_1-1_{\{k_1 \le N_s\}}(N_s \wedge k_2 - k_1)} 1_{\{k_1 \le N_s < k_2\}} d\Lambda_s} \\
&\qquad \cdot \prod_{j=k_2-k_1-(m-1)}^{k_2-k_1} (j - 1_{\{k_1 \le N_{v-}\}}(N_{v-} - k_1)) d\Lambda_v \\
&= \prod_{l=k_1}^{i-1} e^{\frac{m}{k_2-l} \int_{T_l}^{T_{l+1}} d\Lambda_s} \frac{(k_2-k_1-m)!}{(k_2-k_1)!} \int_{T_i}^t e^{\frac{m}{k_2-i} \int_{T_i}^v d\Lambda_s} \\
&\qquad \cdot \prod_{j=k_2-k_1-(m-1)}^{k_2-k_1} (j - (i-k_1)) d\Lambda_v \\
&= \prod_{l=k_1}^{i-1} e^{\frac{m}{k_2-l} \int_{T_l}^{T_{l+1}} \lambda_s ds} \frac{(k_2-k_1-m)!}{(k_2-k_1)!} \\
&\qquad \cdot \prod_{j=k_2-k_1-(m-1)}^{k_2-k_1} (j - (i-k_1)) \left[e^{\frac{m}{k_2-i} \int_{T_i}^t \lambda_s ds} - 1 \right] \frac{k_2-i}{m} \\
&= (X_t - X_{T_i-}) \frac{k_2-i}{m}.
\end{aligned}
$$

□

To conclude this section, we aim to further integrate the presented transformations into the general theory of semimartingales. We have proven the martingale property in certain variants of the Doob-Meyer decomposition. Key examples emerged from the dynamics

$$
dX_s = -X_{s-} \left(\frac{m}{k_2 - (k_2 \wedge N_{s-} - k_1) 1_{\{N_{s-} \ge k_1\}}} \right) 1_{\{k_1 \le N_{s-} \le k_2\}} (dN_s - \lambda_s ds)
$$

and $X_0 = 0$. If we define $Z = (Z_t)_{t \in I}$ by

$$
Z_t = -\int_0^t \frac{m}{k_2 - (k_2 \wedge N_{s-} - k_1) 1_{\{N_{s-} \ge k_1\}}} 1_{\{k_1 \le N_{s-} \le k_2\}} (dN_s - \lambda_s ds),
$$

the equation turns into $dX_t = X_{t-} dZ_t$. This stochastic differential equation represent the stochastic exponential and has the solution

$$\begin{aligned}
X_t &= e^{Z_t} \prod_{0\le s\le t} (1+\Delta X_s) e^{-\Delta Z_s} \\
&= e^{Z_t^c} \prod_{0\le s\le t} (1+\Delta X_s) \\
&= \frac{(k_2-k_1-m)!}{(k_2-k_1)!} e^{m\int_{s_0}^t \frac{1}{k_2-k_1-1_{\{k_1\le N_t\}}(N_t\wedge k_2-k_1)} 1_{\{k_1\le N_s<k_2\}} d\Lambda_s} \\
&\qquad \cdot \prod_{j=k_2-k_1-(m-1)}^{k_2-k_1} (j-1_{\{k_1\le N_t\}}(N_t-k_1)).
\end{aligned}$$

3.3 Exceedance probabilities

In this section, the aim is to rigorously derive the probabilities of exceedance for the martingales that were obtained in the previous investigations. To accomplish this, we will utilize the well-known Doob martingale inequality. Doob's martingale inequality states that for a non-negative supermartingale $(X_t)_{t\in[0,T]}$, and any positive constant $c>0$, we have

$$\mathbb{P}\left[\sup_{s\in[0,T]} X_s \ge c\right] \le \frac{\mathbb{E}[X_T]}{c}.$$

The objective of our findings is to place the widely used inequality into the context of point process martingales. Moreover, we strive to establish additional conditions that would render the inequality as an equality. This will be outlined further in the upcoming section, where we will offer a characterization of the point process martingales through the use of the integral representation. For this we throughout assume that $\mathcal{H}$ is the internal filtration of the point process N on a set $I=[s_0,s_1)$ admitting the $\mathcal{H}$-intensity λ. Let

$$s_0 = \min(I) \in \mathbb{R}\cup\{-\infty\},$$

and

$$s_1 = \sup(I) \in \mathbb{R}\cup\{\infty\},$$

denote the initial time and endpoint of the process, respectively. The following proposition represents a special case of Doob's martingale inequality. We demonstrate its proof employing stopping techniques similar to those referenced in [19].

Proposition 3.3.1 (Exceedance probabilities)
Let $\mathcal{H} = (\mathcal{H}_t)_{t\in I}$ be a filtration and $X = (X_t)_{t\in I}$ be a non-negative $\mathcal{H}$-martingale with right-continuous paths and left limits, starting in $X_{s_0} = 1$ and satisfying the following conditions:

1. $\lim_{s\to s_1} X_s = 0$.
2. For all $t \in I : X_t - \lim_{s\uparrow t} X_s \leq 0$.

Then for all $c \geq 1$ we obtain the identity

$$\mathbb{P}\left[\sup_{s\in I} X_s > c\right] = \frac{1}{c}.$$

Proof.

Let $c \in [1, \infty)$ be arbitrary but fixed. We define the $\mathcal{H}$-Stopping time τ by

$$\tau := \begin{cases} \inf(\{s \in I \mid X_s \geq c\}) & \text{if } \{s : s \in I, X_s \geq c\} \neq \emptyset, \\ s_1 & \text{else.} \end{cases}$$

First, we show that the set $M := \{s \in I \mid X_s \geq c\}$ has a minimum for all $\omega \in \Omega$ and $X_\tau = c$ on $\{\tau < s_1\}$. To do this we distinguish between two cases:

- $\omega \in \{\tau < s_1\} \cap \{X \text{ is continuous at } \tau\}$.

  To show the existence of a minimum, one must show $\tau \in M$. Let $(t_n)_{n\in\mathbb{N}}$ be a sequence in M with $\lim_{n\to\infty} t_n = \tau(\omega)$. Since $(t_n)_{n\in\mathbb{N}}$ is a sequence in M, we have $X_{t_n}(\omega) \geq c$ for all $n \in \mathbb{N}$. Furthermore the path $X_.(\omega)$ is continuous at the point $\tau(\omega)$. Thus, it follows in total $X_\tau(\omega) \geq c$, which is equivalent to $\tau(\omega) \in M$.

  To prove $X_\tau(\omega) = c$ it remains to show $X_\tau(\omega) \leq c$. Let $(v_n)_{n\in\mathbb{N}}$ be a sequence with $v_n \leq \tau(\omega)$, for all $n \in \mathbb{N}$ and $\lim_{n\to\infty} v_n = \tau(\omega)$. Since τ is the minimum of M, it follows $X_{v_n} \leq c$ for all $n \in \mathbb{N}$. Since $X_.(\omega)$ is continuous at $\tau(\omega)$ it follows

  $$X_\tau(\omega) = \lim_{n\to\infty} X_{v_n}(\omega) \leq c.$$

- $\omega \in \{\tau < s_1\} \cap \{X \text{ is discontinuous at } \tau\}$.

We prove $\{\tau < s_1\} \cap \{X \text{ is discontinuous at } \tau\} = \emptyset$ by contradiction. If we assume that $\omega \in \{\tau < s_1\} \cap \{X \text{ is discontinuous at } \tau\}$ exists. Per assumption $X_.(\omega)$ has right-continuous paths. Since $X_.(\omega)$ is discontinuous at $\tau(\omega)$, it follows that this is a left-sided discontinuity:

$$\lim_{t \uparrow \tau(\omega)} X_t \neq X_\tau(\omega).$$

Together with the assumption that the process has only negative jumps (2.) it follows

$$\lim_{t \uparrow \tau(\omega)} X_t > X_\tau(\omega).$$

But this is a contradiction to the assumption that τ is the infimum of the set M. Hence,

$$\{\tau < s_1\} \cap \{X \text{ is discontinuous at } \tau\} = \emptyset.$$

We are now able to conclude

$$\left\{\sup_{s \in I} X_s > c\right\} = \{\tau < s_1\}.$$

Stopped $\mathcal{H}$-martingales are again $\mathcal{H}$-martingales. Hence, the process $X_{t\wedge\tau}$ is again a $\mathcal{H}$-martingale satisfying

$$\mathbb{E}[X_{t\wedge\tau}] = \mathbb{E}[X_{s_0\wedge\tau}] = \mathbb{E}[X_{s_0}] = 1.$$

By Lebesgues dominated convergence Theorem it follows that

$$\begin{aligned}
c\mathbb{P}[\tau < s_1] &= \int_{\{\tau<s_1\}} c d\mathbb{P} \\
&= \int_{\{\tau<s_1\}} X_\tau d\mathbb{P} \\
&= \int_{\{\tau<s_1\}} \lim_{t\to s_1} X_{t\wedge\tau} d\mathbb{P} \\
&= \lim_{t\to s_1} \int_{\{\tau<s_1\}} X_{t\wedge\tau} d\mathbb{P} \\
&= \lim_{t\to s_1} \left(1 - \int_{\{\tau=s_1\}} X_{t\wedge\tau} d\mathbb{P}\right) \\
&= 1 - \lim_{t\to s_1} \left(\int_{\{\tau=s_1\}} X_{t\wedge\tau} d\mathbb{P}\right) \\
&= 1 - \left(\int_{\{\tau=s_1\}} \lim_{t\to s_1} X_{t\wedge\tau} d\mathbb{P}\right) \\
&= 1 - \left(\int_{\{\tau=s_1\}} \lim_{t\to s_1} X_t d\mathbb{P}\right) \\
&= 1.
\end{aligned}$$

Hence, we conclude

$$\frac{1}{c} = \mathbb{P}[\tau < s_1] = \mathbb{P}\left[\sup_{s\in I} X_s > c\right].$$

□

Lemma 3.3.2 (Point process martingale transformations)
Let N be a point process with compensator Λ. Let $B = \{k_1 < k_2 < ... < k_n\} \subset \mathbb{N}$ be an arbitrary, finite subset. If

$$\mathbb{P}[\exists s \in \mathbb{R} : N_s = k_n + 1] = 1,$$

then for all $c \geq 1$ we obtain the identity

$$\mathbb{P}\left[\sup_{t\in\mathbb{R}} \left(\frac{1}{n}\sum_{j=1}^{n} 1_{\{N_t\leq k_j\}} e^{\int_{-\infty}^{t} 1_{\{N_v=k_j\}} d\Lambda_v}\right) > c\right] = \frac{1}{c}.$$

Proof.

This is a direct consequence from Proposition 3.3.1.

□

Example 3.3.3 (Single-Event-Process)
Let Z be a random variable with continuous distribution F. Define the point process $N = (N_t)_{t\in\mathbb{R}}$ by

$$N_t := 1_{\{Z\le t\}}.$$

Clearly, the compensator Λ with respect to the natural filtration is given by

$$d\Lambda_v = (1 - N_v)\frac{dF(v)}{1 - F(v)}.$$

According to Theorem 2.2.4 the process $X = (X_t)_{t\in\mathbb{R}}$ given by

$$\begin{aligned} X_t &= 1_{\{N_t\le 0\}} e^{\int_{-\infty}^t 1_{\{N_v=0\}} d\Lambda_v} \\ &= 1_{\{Z>t\}} \cdot e^{\int_{-\infty}^t \frac{1}{1-F(v)} dF(v)}, \end{aligned}$$

is a martingale. Lemma 2.3.2 yields

$$\begin{aligned} \frac{1}{c} &= \mathbb{P}\left[\sup_{t\in\mathbb{R}}\left(1_{\{Z>t\}} \cdot e^{\int_{-\infty}^t \frac{1}{1-F(v)} dF(v)}\right) > c\right] \\ &= \mathbb{P}\left[e^{\int_{-\infty}^Z \frac{1}{1-F(v)} dF(v)} > c\right] \\ &= \mathbb{P}\left[\frac{1}{1 - F(Z)} > c\right] \\ &= \mathbb{P}\left[1 - \frac{1}{c} < F(Z)\right]. \end{aligned}$$

So, for this particular case, the calculations yield the distribution of the random variable $F(Z)$, which is of course uniform on the unit intervall.

Example 3.3.4 (Survival Analysis)
Let Z be a random variable with Hazard-function h. Define the point process $N = (N_t)_{t\in\mathbb{R}}$ by

$$N_t = 1_{\{Z\le t\}}.$$

Clearly, the conditional intensity of N with respect to the natural generated filtration is given by

$$\lambda(v \mid \mathcal{H}_v) = h(v)1_{\{Z<v\}}.$$

According to Theorem 3.2.4, the process $X = (X_t)_{t\in\mathbb{R}}$ given by

$$X_t := 1_{\{Z>t\}} \cdot e^{\int_{-\infty}^{t} h(v)1_{\{Z<v\}}dv},$$

holds true for the following equation

$$\begin{aligned}\frac{1}{c} &= \mathbb{P}\left[\sup_{t\in\mathbb{R}}\left(1_{\{Z>t\}} \cdot e^{\int_{-\infty}^{t} h(v)1_{Z<v}dv}\right) > c\right] \\ &= \mathbb{P}\left[e^{\int_{-\infty}^{Z} h(v)dv} > c\right] \\ &= \mathbb{P}\left[\int_{-\infty}^{Z} h(v)dv > ln(c)\right] \\ &= \mathbb{P}\left[Z > \Lambda^{-1}(ln(c))\right]. \end{aligned} \tag{3.8}$$

Rewrite equation (3.11) to obtain the popular limes-formula

$$\mathbb{P}\left[Z > x\right] = e^{-\Lambda(x)}.$$

Example 3.3.5 (Geometric Brownian motion)
Let $B = (B_t)_{t\in\mathbb{R}_+}$ be a Brownian motion and $(\mathcal{H}_t^B)_{t\in\mathbb{R}_+}$ be the Filtration generated by B. Since the function $\varphi : \mathbb{R}\times\mathbb{R}_+ \to \mathbb{R}_+$,

$$\varphi(x,t) = e^{x-\frac{1}{2}t},$$

is a solution of the (backward) heat-equation, it is well known that the process $X = (X_t)_{t\in\mathbb{R}_+}$ defined by

$$X_s := e^{B_s-\frac{1}{2}s},$$

is a $\mathcal{H}^B$-martingale. According to Theorem 3.2.4, we found for $c > 1$ the equation

$$\mathbb{P}\left[\sup_{t\in\mathbb{R}_+}\left(e^{B_t-\frac{1}{2}t}\right) > c\right] = \frac{1}{c}.$$

In the following Proposition we will offer a mild generalization of Proposition 3.3.1. The generalization leads to the observation that the jump sizes of the resulting martingale may take on positive values.

Proposition 3.3.6 (Exceedance probabilities II)
Let $\mathcal{H} = (\mathcal{H}_t)_{t\in I}$ be a filtration and $X = (X_t)_{t\in I}$ be a non-negative $\mathcal{H}$-martingale with right-continuous paths and left limits, starting in $X_{s_0} = 1$ and satisfying the condition:

1. $\lim_{s\to s_1} X_s = 0$.
2. There exists a $\tilde{c} \geq 1$ such that all realizations of X satisfy

$$\forall t \in I : X_t - \lim_{s\uparrow t} X_s > 0 \Rightarrow X_t \leq \tilde{c}.$$

Then for all $c \geq \tilde{c}$ we have the identity

$$\mathbb{P}\left[\sup_{s\in I} X_s > c\right] = \frac{1}{c}.$$

Proof.

Let $c \geq \tilde{c}$ be arbitrary but fixed. We define the $\mathcal{H}$-Stopping time τ by

$$\tau := \begin{cases} \inf(\{s \in I \mid X_s \geq c\}) & \text{if } \{s : s \in I, X_s \geq c\} \neq \emptyset, \\ s_1 & \text{else.} \end{cases}$$

First, we show that the set $M := \{s \in I \mid X_s \geq c\}$ has a minimum for all $\omega \in \Omega$ and $X_\tau = c$ on $\{\tau < s_1\}$. To do this we distinguish between three cases:

- $\omega \in \{\tau < s_1\} \cap \{X \text{ is continuous at } \tau\}$.

  This case was already treated in Proposition 3.3.1.

- $\omega \in \{\tau < s_1\} \cap \{X \text{ is discontinuous at } \tau\} \cap \{X_\tau - \lim_{s\uparrow\tau} X_s \leq 0\}$.

  This case was already treated in Proposition 3.3.1.

- $\omega \in \{\tau < s_1\} \cap \{X \text{ is discontinuous at } \tau\} \cap \{X_\tau - \lim_{s\uparrow\tau} X_s > 0\}$.

  To show the existence of a minimum, one must show $\tau \in M$. Let $(t_n)_{n\in\mathbb{N}}$ be a sequence in M with $\lim_{n\to\infty} t_n = \tau(\omega)$. Since $(t_n)_{n\in\mathbb{N}}$ is a sequence in M, we get $X_{t_n}(\omega) \geq c$ for all $n \in \mathbb{N}$. Furthermore, the path $X_.(\omega)$ is right-continuous at the point $\tau(\omega)$. Thus, it follows in total $X_\tau(\omega) \geq c$, which is equivalent to $\tau(\omega) \in M$.

To prove $X_\tau(\omega) = c$ it remains to show $X_\tau(\omega) \leq c$. But this is a direct consequence of assumption 2. since we have assumed that X is discontinuous in τ.

The rest of the proof is one to one with Proposition 3.3.1.

□

3.4 Criteria for the martingales

In the previous investigations we found several possibilities to transform N into a martingale X satisfying

$$\mathbb{P}\left[\sup_{s\in I} X_s > c\right] = \frac{1}{c}.$$

In this section, we aim at providing a criteria for the predictable process g such that the resulting martingale X satisfies

$$\mathbb{P}\left[\sup_{s\in I} X_s > c\right] = \frac{1}{c}.$$

Theorem 3.4.1 (Criteria for the martingales)
Let N be a point process with compensator Λ and g be a predictable process. Define the martingale X by

$$X_s := \int_{s_0}^{s} g_v(dN_v - d\Lambda_v) + 1.$$

If there exists $n \in \mathbb{N}$ and $c > 1$ such that

$$\mathbb{P}[\exists s \in I : N_s = n] = 1$$

and

(a) For all $v \in [T_{n-1}, s_1) : g_v = -X_{v-}$,

(b) for all $v \in I : -X_{v-} \leq g_v \leq \max(c - X_{v-}, 0)1_{\{X_{v-}>0\}}$.

Then, it follows that

$$\mathbb{P}\left[\sup_{s\in I} X_s > c\right] = \frac{1}{c}.$$

A predictable process g that satisfies these properties is referred to as a test strategy. Define the set of all test-strategies to the critical value c by

$$\mathcal{T}^c(\lambda) := \{g \mid g \text{ is a test-strategy to the critical value } c\}.$$

Proof.

Clearly X is a martingale with right-continuous paths and lefthands limits. According to Proposition 3.3.6, it remains to proof

1. $\forall s \in I : X_s \geq 0$,
2. $\lim_{s\to s_1} X_s = 0$,
3. there exists $\tilde{c} \leq c$, such that all realizations of X satisfy

$$\forall t \in I : X_t - \lim_{s\uparrow t} X_s > 0 \Rightarrow X_t \leq \tilde{c}.$$

1. To proof the non-negativity of the paths one must show $M^- := \{\omega \in \Omega \mid \exists t \in I : X_t < 0\} = \emptyset$. Assume M^- is not empty and let $\tilde{\omega} \in M^-$ be arbitrary. Let $M^-_{\tilde{\omega}} = \{t \in I \mid X_t(\tilde{\omega}) < 0\}$ be the set of negative values to the realization $X(\tilde{\omega})$ and define $t_{\min} := \inf M^-_{\tilde{\omega}}$. Since X has right-continuous paths it follows $X_{t_{\min}} \leq 0$. Considering assumption b.) yields that $X(\tilde{\omega})$ must be discontinuous in $t^-_{\min}$. The discontinuity implies $N_{t_{\min}} - N_{t_{\min}-} = 1$ and $X_{t_{\min}-} > 0$. Hence

$$X_{t_{\min}} - X_{t_{\min}-} = X_{t_{\min}-} + g_{t_{\min}} \underbrace{\geq}_{b.)} X_{t_{\min}-} - X_{t_{\min}-} = 0.$$

Now, we can conclude $X_{t_{\min}} = 0$. But this implies for all $\tilde{v} > t_{\min}$ $X_{\tilde{v}} = 0$ (assumption b.) and is a contradiction to the assumption that there exists a $t \in I$ with $X_t < 0$.

2. Assume $M^0 := \{\omega \in \Omega \mid X_{s_1}(\omega) > 0\}$ is not empty and let $\tilde{\omega} \in M^0$ be arbitrary. Since $X_{s_1} > 0$ it follows for all $s \in I$ that $X_s > 0$. But this is a contradiction to

$$\begin{aligned} X_{T_n} &= \int_{s_0}^{T_n} g_v(dN_v - d\Lambda_v) \\ &= X_{T_{n-1}} - \int_{T_{n-1}}^{T_n} X_{v-}(dN_v - d\Lambda_v) \\ &= X_{T_n-} - X_{T_n-} \\ &= 0. \end{aligned}$$

3. Assume the set $M^c := \{\omega \in \Omega \mid \exists s \in I : \lim_{v\uparrow s} X_v < X_s \wedge X_s > c\}$ is not empty and let $\tilde{\omega} \in M^c$ be arbitrary. Define $T^c := \inf M^c$. It follows

$$X_{T^c} = X_{T^c-} + g_{T^c} \leq X_{T^c-} + c - X_{T^c-} = c.$$

However, this is a contradiction to $X_{T^c} > c$.

□

Note, if g is a stragegy with non-positive paths satisfying (a)-(b), it follows that g is a test strategy for all critical values $c \geq 1$. In the subsequent part of this chapter we will explore a sequential decision problem.

Example 3.4.2
Let N be a point process with conditional intensity λ and $n \in \mathbb{N}$ be fixed. Define the predictable process g by

$$g_v := -X_{v-}1_{\{N_{v-}=n-1\}}.$$

We want to prove that g is a test-strategy for all $c > 1$, that is the process X defined by

$$X_s := \int_{s_0}^{s} g_v(dN_v - d\Lambda_v) + 1$$

is a martingale satisfying

$$\mathbb{P}\left[\sup_{s\in I} X_s > c\right] = \frac{1}{c}.$$

Clearly, X is a stochastic process with non-negative paths. Taking into account Theorem 3.4.1 one needs to proof

(a) For all $v \in [T_{n-1}, s_1) : g_v = -X_{v-}$,

(b) for all $v \in I : -X_{v-} \leq g_v \leq \max(c - X_{v-}, 0)1_{\{X_{v-}>0\}}$.

Clearly, claim (a) is satisfied. Define $M := [-X_{v-}, \max(c-X_{v-}, 0)1_{\{X_{v-}>0\}}]$. For claim (b) we distinguish between three cases:

$$\begin{aligned}
&v \in [s_0, T_{n-1}] \Rightarrow g_v = 0 \in M,\\
&v \in (T_{n-1}, T_n) \Rightarrow g_v = -X_{v-} \in M,\\
&v \in [T_n, s_1] \Rightarrow g_v = 0 \in M.
\end{aligned}$$

Hence, we can conclude that g is indeed a test-strategy for all $c > 1$. The subsequent step is to provide a specific expression for the process X. It follows that

$$\begin{aligned}
&X_s = 1, s \in [s_0, T_{n-1}]\\
&X_s = 1 + \int_{T_{n-1}}^{s} X_{v-} d\Lambda_v, s \in (T_{n-1}, T_n)\\
&X_s = 1 + \int_{T_{n-1}}^{T_n} X_{v-} d\Lambda_v - X_{T_n-} = 0, s \in [T_{n-1}, s_1).
\end{aligned}$$

Given that

$$de^{\int_{s_0}^{s} 1_{N_v=n-1} d\Lambda_v} = 1_{N_s=n-1} \lambda_s e^{\int_{s_0}^{s} 1_{N_v=n-1} d\Lambda_v} ds,$$

we can express X_s as

$$X_s = 1_{\{N_s<n\}} e^{\int_{s_0}^{s} 1_{N_v=n-1} \lambda_v dv}.$$

□

In the subsequent considerations for the development of tests from the martingales we previously derived, we drew inspiration from Wald's sequential test and the terminologies of test martingales from the work [35]. In the paper, the authors explore the concept of "test martingales" within a broader context that may not necessarily be related to point processes. The martingales used therein also possess the property of non-negativity. Since we have focused on the integrand in the integral representation in this section, we have introduced the concept of a test strategy, which pertains to the integrand.

Let N^1 and N^2 be two different point processes with respective compensators Λ^1 and Λ^2. Furthermore, assume that both of the compensators admit conditional intensities. The statistician diligently monitors a process, denoted as N, starting at time s_0, which could be either N^1 or N^2, and aims to discern which of the two processes N corresponds to. This leads to the (simple) test problem $H_0 : N = N^1$ against the alternative $H_1 : N = N^2$.

Definition 3.4.3 (Test strategies)
Let g be a test strategy to the critical value c, that is

$$\mathbb{P}\left[\sup_{s\in I}\left(\int_{s_0}^{s} g_v(dN_v^1 - d\Lambda_v^1) + 1\right) > c\right] = \frac{1}{c}.$$

Define the set of test strategies to the critical value c with respect to the alternative Λ^2 by

$$\mathcal{S}^c(\lambda^1, \lambda^2) := \left\{g \in \mathcal{T}^c(\lambda^1) \mid \mathbb{P}\left[\sup_{s\in I}\left(\int_{s_0}^{s} g_v(dN_v^2 - d\Lambda_v^1) + 1\right) > c\right] \geq \frac{1}{c}\right\}.$$

Lemma 3.4.4 (Criteria for the test martingales)
In the situation of this section, assume

$$\forall v > T_{n-1} : \lambda_v^1 \geq \lambda_v^2.$$

Then for all $g \in \mathcal{T}^c(\lambda^1)$ satisfying

(a) $\forall v \in I : \lambda_v^1 > \lambda_v^2 \Rightarrow g_v \geq 0$,

(b) $\forall v \in I : \lambda_v^1 < \lambda_v^2 \Rightarrow g_v \leq 0$,

it follows that $g \in \mathcal{S}^c(\lambda^1, \lambda^2)$.

Proof.

First, we prove that the process X is a submartingale under the alternative. It holds that

$$\begin{aligned}
&\mathbb{E}\left[X_t \mid \mathcal{H}_s\right] \\
&= X_s + \mathbb{E}\left[\int_s^t g_v(dN_v^2 - \lambda_v^1 dv) \mid \mathcal{H}_s\right] \\
&= X_s + \mathbb{E}\left[\int_s^t g_v(dN_v^2 - \lambda_v^2 dv) \mid \mathcal{H}_s\right] + \mathbb{E}\left[\int_s^t g_v(\lambda_v^1 - \lambda_v^2)dv \mid \mathcal{H}_s\right] \\
&= X_s + \mathbb{E}\left[\int_s^t g_v(\lambda_v^1 - \lambda_v^2)dv \mid \mathcal{H}_s\right] \\
&\geq X_s.
\end{aligned}$$

We define the $\mathcal{H}$-Stopping time τ by

$$\tau := \begin{cases} \inf(\{s \in I \mid X_s \geq c\}) & \text{if } \{s : s \in I, X_s \geq c\} \neq \emptyset, \\ s_1 & \text{else.} \end{cases}$$

X is a submartingale starting in $X_0 = 1$. Hence

$$\mathbb{E}[X_{t\wedge\tau}] \geq \mathbb{E}[X_{s_0\wedge\tau}] = \mathbb{E}[X_{s_0}] = 1.$$

Taking into account the argumentation for Proposition 3.3.6 it follows that the set $M := \{s \in I \mid X_s \geq c\}$ has a minimum for all $\omega \in \Omega$ and $X_\tau = c$ on $\{\tau < s_1\}$.
According to the Lebesgue Dominated Convergence Theorem, it can be

inferred that

$$\begin{aligned}
c\mathbb{P}[\tau < s_1] &= \int_{\{\tau<s_1\}} c d\mathbb{P} \\
&= \int_{\{\tau<s_1\}} X_\tau d\mathbb{P} \\
&= \int_{\{\tau<s_1\}} \lim_{t\to s_1} X_{t\wedge\tau} d\mathbb{P} \\
&= \lim_{t\to s_1} \int_{\{\tau<s_1\}} X_{t\wedge\tau} d\mathbb{P} \\
&\geq \lim_{t\to s_1} \left(1 - \int_{\{\tau=s_1\}} X_{t\wedge\tau} d\mathbb{P}\right) \\
&= 1 - \lim_{t\to s_1} \left(\int_{\{\tau=s_1\}} X_{t\wedge\tau} d\mathbb{P}\right) \\
&= 1 - \left(\int_{\{\tau=s_1\}} \lim_{t\to s_1} X_{t\wedge\tau} d\mathbb{P}\right) \\
&= 1 - \left(\int_{\{\tau=s_1\}} \lim_{t\to s_1} X_t d\mathbb{P}\right) \\
&= 1.
\end{aligned}$$

Hence, it follows that

$$\frac{1}{c} \leq \mathbb{P}[\tau < s_1] = \mathbb{P}\left[\sup_{s\in I} X_s > c\right].$$

□

Chapter 4

The empirical distribution function

In the previous chapter we have found several ways to transform a general point process N, characterized by its internal conditional intensity λ, into a non-negative martingale. These processes are particularly attractive for a variety of applications due to their ability to specify exceedance probabilities. The obtained results will now be applied to the renormalized empirical distribution function. This chapter aims to deepen our understanding of the stochastic behavior of the resulting transformations by presenting the findings in a more approachable format, as compared to the previous chapter. Additionally, we will explore several other noteworthy properties through a comprehensive simulation study.

4.1 Martingales with respect to F_n

Let $X_1, X_2, ..., X_n$ be a fixed size sequence of independent and identically distributed random variables with $X_1 \sim F$. The point process $(N_t)_{t\in[0,1)}$ defined by

$$N_s^n := \sum_{i=1}^{n} 1_{\{X_i \leq s\}},$$

is called the renormalized empirical distribution function $N^n = n \cdot F_n$. In this section, we will consistently use the notation N instead of N^n to represent the sample size, as it remains constant throughout the discussion. If we assume that F admits a Lebesgue density f, the conditional intensity of N, which is represented by $\lambda = (\lambda(v \mid \mathcal{H}_v))_{v\in[0,1)}$, is equal to

$$\lambda(v \mid \mathcal{H}_v) = \frac{(n - N_v)f(v)}{1 - F(v)}.$$

Lemma 4.1.1 (k-th non-negative single martingale)
Let $1 \le k \le n$ be arbitrary and T_k be the k-th order statistic of F_n. In the situation of this chapter, the process $X^k = (X_t^k)_{t\in[0,1)}$ defined by

$$\begin{aligned} X_s^k &:= 1_{\{N_s<k\}} + 1_{\{N_s=k\}} \left(\frac{1 - F(T_k)}{1 - F(s)}\right)^{(n-k)} \\ &= 1_{\{F_n(s)<k/n\}} + 1_{\{F_n(s)=k/n\}} \left(\frac{1 - F(T_k)}{1 - F(s)}\right)^{(n-k)}, \end{aligned}$$

is a $\mathcal{H}$-martingale.

Proof.

Define

$$T_k := \inf\{s \in \mathbb{R} \mid N_s = k\}.$$

Taking into account proposition 3.2.4, the process $X = (X_t)_{t\in[0,1)}$ defined by

$$X_s^k := 1_{\{N_s\le k\}} e^{\int_{-\infty}^s 1_{\{N_v=k\}} d\Lambda_v},$$

is a martingale. The proof is concluded by the fact that

$$\begin{aligned} X_s^k &= 1_{\{N_s\le k\}} e^{\int_{-\infty}^s 1_{\{N_v=k\}} d\Lambda_v} \\ &= 1_{\{N_s\le k\}} e^{1_{\{N_s\ge k\}} \int_{T_k}^{T_{k+1}\wedge s} \frac{(n-N_{v-})f(v)}{1-F(v)} dv} \\ &= 1_{\{N_s\le k\}} e^{1_{\{N_s\ge k\}}(n-k) \int_{T_k}^{T_{k+1}\wedge s} \frac{f(v)}{1-F(v)} dv} \\ &= 1_{\{N_s\le k\}} e^{1_{\{N_s\ge k\}}(n-k)(-log(1-F(T_{k+1}\wedge s))+log(1-F(T_k)))} \\ &= 1_{\{N_s\le k\}} e^{1_{\{N_s\ge k\}}(n-k)(log(\frac{1-F(T_k)}{1-F(T_{k+1}\wedge s)}))} \\ &= 1_{\{N_s\le k\}} (\frac{1 - F(T_k)}{1 - F(T_{k+1} \wedge s)})^{1_{\{N_s\ge k\}}(n-k)} \\ &= 1_{\{N_s\le k\}} \left(\frac{1 - F(T_k)}{1 - F(s)}\right)^{1_{\{N_s\ge k\}}(n-k)} \\ &= 1_{\{N_s<k\}} + 1_{\{N_s=k\}} \left(\frac{1 - F(T_k)}{1 - F(s)}\right)^{(n-k)}. \end{aligned}$$

□

The process X^k takes on the value 1 before the k-th jump time, increases exponentially between T^k and T^{k+1}, and transitions to the state of 0 after T^{k+1}. In essence, X^k acts as an expert for the stochastic time interval from T^k to T^{k+1}. However, this approach leads to a significant loss of information, which means that the entire realization of the point process cannot be reconstructed solely by only observing the path of this martingale.

Lemma 4.1.2 (Non-negative martingale) The process $M = (M_t)_{t\in[0,1)}$ given by

$$M_s := \frac{n-1}{n} - F_n(s) + \frac{1}{n}\left(\frac{1-F(T_{N_s})}{1-F(s)}\right)^{n(1-F_n(s))},$$

is a $\mathcal{H}$-martingale.

Proof.

Based on Lemma 3.2.5, we define

$$B = \{1, 2, ..., n\},$$

which subsequently leads to the conclusion that

$$\begin{aligned}
X_s^B &= \frac{1}{n}\sum_{k=0}^{n-1} X_s(k) \\
&= \frac{1}{n}\sum_{k=0}^{n}\left(1_{\{N_s<k\}} + 1_{\{N_s=k\}}\left(\frac{1-F(T_k)}{1-F(s)}\right)^{(n-k)}\right) \\
&= \frac{1}{n}\left(n - N_s + \left(\frac{1-F(T_{N_s})}{1-F(s)}\right)^{(n-N_s)}\right) \\
&= \frac{n-1}{n} - F_n(s) + \frac{1}{n}\left(\frac{1-F(T_{N_s})}{1-F(s)}\right)^{n(1-F_n(s))} \\
&= \frac{n-1}{n} - F_n(s) + \frac{1}{n}\left(\frac{1-T_{N_s}}{1-s}\right)^{n(1-F_n(s))}.
\end{aligned}$$

□

A notable distinction between the martingales derived by Hess [19] and our approach becomes evident at this step. The key dissimilarity lies in the fact that our transformation explicitly incorporates the past, utilizing the most recent jump time T_{N_s} as the basis for the transformation. Such a transformation, which is dependent not only on time and location but also on past history, cannot be obtained as a solution to their derived martingale differential equation.

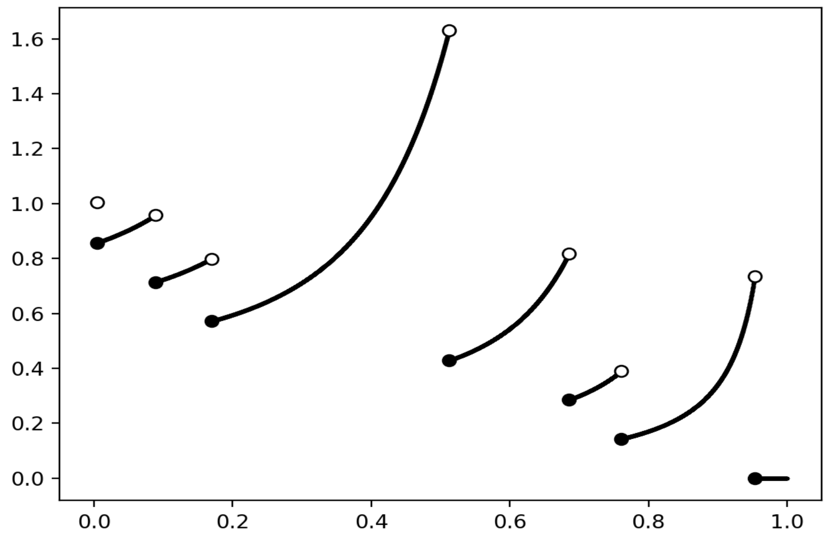

Figure 4.1: Typical path of the martingale $M_s = \frac{n-1}{n} - F_n(s) + \frac{1}{n}\left(\frac{1-T_{N_s}}{1-s}\right)^{n(1-F_n(s))}$ with sample size $n = 7$ and uniformly distributed random variables.

As previously discussed, the process commences with the value $M_0 = 1$

almost surely under probability measure $\mathbb{P}$. At the time T_k, it holds that

$$\begin{aligned} M_{T_k} &= \frac{n-1}{n} - F_n(T_k) + \frac{1}{n}\left(\frac{1-F(T_{N_{T_k}})}{1-F(T_k)}\right)^{n(1-F_n(T_k))} \\ &= \frac{n-1}{n} - \frac{k}{n} + \frac{1}{n}\left(\frac{1-F(T_k)}{1-F(T_k)}\right)^{n(1-\frac{k}{n})} \\ &= 1 - \frac{k}{n}. \end{aligned}$$

According to this, the process takes on the value 0 at the last jump time T_n and remains in this state from then on. Thus, $M_t = 0$ für all $t \geq T_n$. During the intervals between jumps, the process has the potential to be explosive. We can see this from the fact that the paths grow monotonously in each intermediate interval $[T_k, T_{k+1})$. This is explained by the special form of the quotient

$$\frac{1-F(T_{N_{T_k}})}{1-F(t)}.$$

The denominator becomes smaller and smaller between the jumps, because of the monotonous growth of the distribution F. On the other hand, the numerator is constant between the jumps. Simultaneously, the entire quotient is raised to the power of $n-k>1$. The "redemption" is brought by the next event, when the process falls to the value $1-\frac{k+1}{n}$. In applications we do not know the underlying distribution F. If we make a wrong assumption on the true distribution function F, the process M_t can become explosive. The wrongly chosen distribution function $\tilde{F}$ is implemented in the process and we get the process

$$M_t = 1 - F_n(s) + \frac{1}{n}\left(\frac{1-\tilde{F}(T_{N_s})}{1-\tilde{F}(s)}\right)^{n(1-F_n(s))}.$$

At the jumping times the process continues to behave in a non-hazardous way. The situation is different between the jumps when $\tilde{F}$ grows fast compared to the real F. If $\tilde{F}$ would be the true distribution, due to the steep growth relative to F the realization of a date, i.e. redemption, would be expected. However, the true F tends to grow more slowly and thus the probability of such an event is less than assumed by the false $\tilde{F}$.

The effect described here becomes clear even for relatively small n. In the following plot we have assumed that the data is uniformly distributed on $[0, 1]$, despite being uniformly distributed on the interval $[0.3, 1]$. The wrongly assumed $\tilde{F}$ thus grows slower on the interval $[0, 0.3]$, relative to the true F. At this point another peculiarity of the process (M_t) becomes clear. It locally gets the tendency to jump. In this case, this is the first date after the time $t = 0,3$. A clear outlier can be seen in the realization of M.

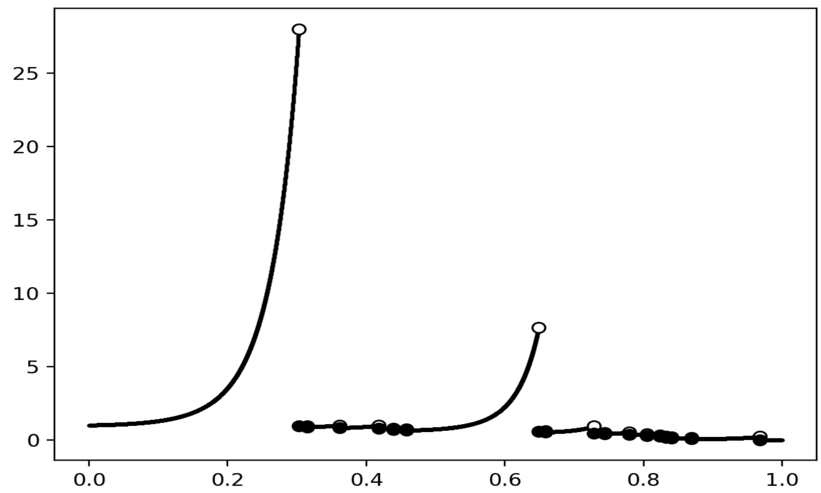

Figure 4.2: Transformation of the empirical distribution function, assuming that the data are uniform on $[0, 1]$, when they are actually uniform on $[0.3, 1]$.

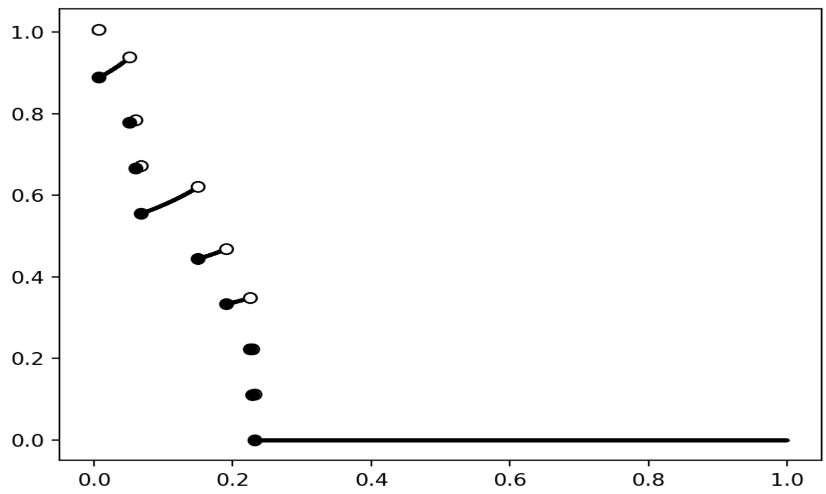

Figure 4.3: Transformation of the empirical distribution function, assuming that the data are uniform on $[0, 1]$, when they are actually uniform on $[0, 0.3]$.

The process is able to detect locally faster growth of an incorrectly assumed distribution function and indicates this through a deflection. The opposite is true if the wrongly assumed distribution function grows locally slower. In this case, the process demonstrates a milder behavior than under the actual F. This is shown in the simulation where the random variables assumed to be uniform distributed on $[0, 1]$, when in reality the data is uniformly distributed on $[0, 0.3]$. The assumed growth is therefore locally much slower at the beginning.

However, there exists a related process that exhibits precisely the opposite behavior to the previously discovered martingale, generating outliers in the opposite direction at the local level. For the derivation we need the term reverse conditional intensity, which will be discussed in the next section.

We will now proceed to investigate the higher-order transformations as presented in Theorem 3.2.6.

Lemma 4.1.3 (0-k non-negative martingale)
Set $T_0 = -\infty$. Let $1 \leq k \leq n$ be arbitrary. In the situation of this chapter, the process $X^{0,k} = (X_t^{0,k})_{t\in[0,1)}$ defined by

$$X_s^{0,k} := \prod_{j=0}^{n \cdot F_n(s) \wedge (k-1)} \left(\frac{1 - F(T_j)}{1 - F(T_{j+1} \wedge s)} \right)^{\frac{(1+j)(n-j)}{k}} \prod_{i=1}^{n \cdot F_n(s) \wedge k} (1 - i/k)$$

is a $\mathcal{H}$-martingale.

Proof.

Consider N as a renormalized empirical distribution function and $k_1 = 0$, $k := k_2 \leq n$ as arbitrary but fixed. Utilizing Theorem 3.2.6, we can construct a martingale X with

$$X_s^k := e^{\frac{1}{k} \int_{s_0}^{s} (1+N_v \wedge k) 1_{\{N_v < k\}} d\Lambda_v} \prod_{i=1}^{N_s \wedge k} (1 - i/k).$$

First, we calculate the integral in the exponent. It holds that

$$
\begin{aligned}
&\int_{s_0}^{s} (1 + N_v \wedge k) 1_{\{N_v<k\}} d\Lambda_v \\
&= \int_{s_0}^{s} (1 + N_v \wedge k) 1_{\{N_v<k\}} \frac{(n - N_v) f(v)}{1 - F(v)} dv \\
&= \sum_{j=0}^{N_s} \int_{s_0}^{s} 1_{\{N_v=j\}} (1 + N_v \wedge k) 1_{\{N_v<k\}} \frac{(n - N_v) f(v)}{1 - F(v)} dv \\
&= \sum_{j=0}^{N_s} (1 + j \wedge k) 1_{\{j<k\}} (n - j) \int_{s_0}^{s} 1_{\{N_v=j\}} \frac{f(v)}{1 - F(v)} dv \\
&= \sum_{j=0}^{N_s \wedge (k-1)} (1 + j)(n - j) \int_{T_j}^{T_{j+1} \wedge s} \frac{f(v)}{1 - F(v)} dv \\
&= \sum_{j=0}^{N_s \wedge (k-1)} (1 + j)(n - j)(-log(1 - F(T_{j+1} \wedge s)) + log(1 - F(T_j))) \\
&= log \left(\prod_{j=0}^{N_s \wedge (k-1)} \left(\frac{1 - F(T_j)}{1 - F(T_{j+1} \wedge s)} \right)^{(1+j)(n-j)} \right).
\end{aligned}
$$

Hence,

$$
\begin{aligned}
X_s^k &= e^{\frac{1}{k} \int_{s_0}^{s} (1+N_v \wedge k) 1_{\{N_v<k\}} \lambda(v|\mathcal{H}_v) dv} \prod_{i=1}^{N_s \wedge k} (1 - i/k) \\
&= e^{\log \left(\prod_{j=0}^{N_s \wedge (k-1)} \left(\frac{1-F(T_j)}{1-F(T_{j+1} \wedge s)} \right)^{\frac{(1+j)(n-j)}{k}} \right)} \prod_{i=1}^{N_s \wedge k} (1 - i/k) \\
&= \prod_{j=0}^{N_s \wedge (k-1)} \left(\frac{1 - F(T_j)}{1 - F(T_{j+1} \wedge s)} \right)^{\frac{(1+j)(n-j)}{k}} \prod_{i=1}^{N_s \wedge k} (1 - i/k).
\end{aligned}
$$

This completes the proof.

□

Note that if we choose $n = k$ the martingale $X^{0,n}$ again turns out to be a martingale without lose of information.

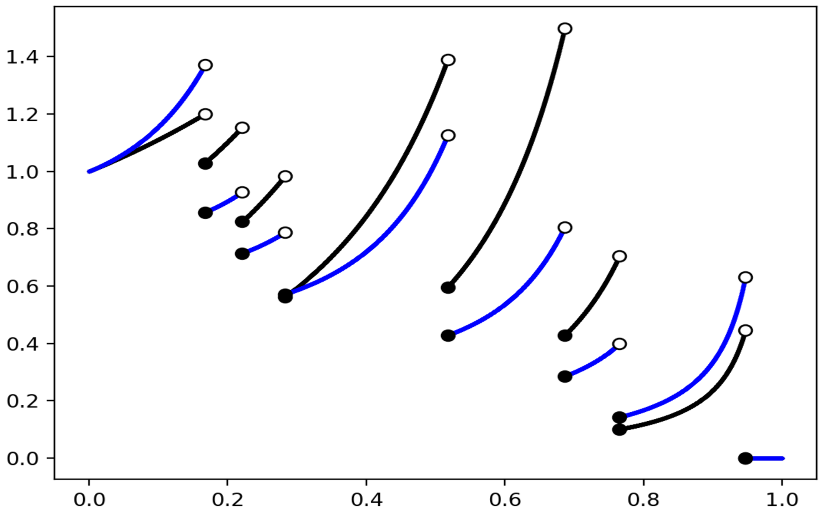

Figure 4.4: Comparison of $X^{0,n}$ (black) and M (blue) from Lemma 3.1.2.

One distinction between the martingales M and $X^{0,n}$ is rooted in their respective representations. On the one hand M has the form

$$M_t = \varphi(N_t, t, T_{N_t}),$$

so it is a process depending on the current time t, current state N_t and the most recent jumping time T_{N_t}. On the other hand the process $X^{0,n}$ has the more general form

$$X_t^n = \varphi(N_t, t, T_1 \wedge t, T_2 \wedge t, ..., T_n \wedge t).$$

Lemma 4.1.4 (k_1-k_2 non-negative martingale)
Let $1 \leq k_1 < k_2 \leq n$ be arbitrary. In the situation of this chapter, the process $X^{k_1,k_2} = (X_t^{k_1,k_2})_{t\in[0,1)}$ defined by

$$X_s^{k_1,k_2} := \prod_{j=k_1}^{n\cdot F_n(s)\wedge(k_2-1)} \left(\frac{1-F(T_j)}{1-F(T_{j+1}\wedge s)}\right)^{\frac{(1+j-k_1)(n-j)}{k_2-k_1}} \cdot \prod_{i=1}^{1_{\{k_1\leq n\cdot F_n(s)\}}(n\cdot F_n(s)\wedge k_2-k_1)} (1 - i/(k_2-k_1)),$$

is a $\mathcal{H}$-martingale.

Proof.

Let N be a renormalized empirical distribution function and $0 \leq k_1 < k_2 \leq n$ be arbitrary but fixed. We utilize Theorem 3.2.6 to derive the martingale

$$X_s^{k_1,k_2} = e^{\frac{1}{k_2-k_1}\int_{s_0}^{s}(1+1_{\{k_1\leq N_v\}}(N_v\wedge k_2-k_1))1_{\{k_1\leq N_v<k_2\}}d\Lambda_v} \cdot \prod_{i=1}^{1_{\{k_1\leq N_s\}}(N_s\wedge k_2-k_1)}(1-i/(k_2-k_1)).$$

The equation

$$\int_{s_0}^{s}(1+1_{\{k_1\leq N_v\}}(N_v\wedge k_2-k_1))1_{\{k_1\leq N_v<k_2\}}\lambda(v\mid\mathcal{H}_v)dv = log\left(\prod_{j=k_1}^{N_s\wedge(k_2-1)}\left(\frac{1-F(T_j)}{1-F(T_{j+1}\wedge s)}\right)^{(1+j-k_1)(n-j)}\right)$$

holds true. Therefore, it follows that

$$X_s^{k_1,k_2} = \prod_{j=k_1}^{N_s\wedge(k_2-1)}\left(\frac{1-F(T_j)}{1-F(T_{j+1}\wedge s)}\right)^{\frac{(1+j-k_1)(n-j)}{k_2-k_1}}\prod_{i=1}^{1_{\{k_1\leq N_s\}}(N_s\wedge k_2-k_1)}(1-i/(k_2-k_1)).$$

□

The martingales derived from Lemma 4.1.4 exhibit a property whereby the number of factors increases with time. Due to the shape $\prod_{j=k_1}^{n\cdot F_n(t)}(...)$, it is a product with a random number of factors depending on the current state of the point process. Looking at the martingales from Hess [19], the processes have the shape

$$\varphi(N_t,t) = \frac{n^m(n-m)!}{n!}\prod_{j=n-(m-1)}^{n}\frac{\frac{j}{n}-F_n(t)}{1-F(t)}.$$

Thus, we can deduce that regardless of the process state, the number of factors always equals $n-m$. The next lemma provides a generalization of these martingales.

Lemma 4.1.5 (Generalization)
Let $0 \leq k_1, m, k_2 \leq n$ be arbitrary with $0 \leq m \leq k_2 - k_1$. In the situation of this chapter the process $X^{k_1,k_2}(m) = (X_t^{k_1,k_2}(m))_{t\in[0,1)}$ defined by

$$X_s^{k_1,k_2}(m) := \frac{(k_2-k_1-m)!}{(k_2-k_1)!}\prod_{j=k_1}^{k_2-1}\left(\frac{1-F(T_j\wedge t)}{1-F(T_{j+1}\wedge t)}\right)^{\frac{m(n-j)}{k_2-j}} \cdot \prod_{j=k_2-k_1-(m-1)}^{k_2-k_1}(j-1_{\{k_1\leq N_t\}}(N_t\wedge k_2-k_1))$$

is a non-negative martingale with respect to $\mathcal{H}$.

Proof.

According to Theorem 3.2.8, the process

$$X_s^{k_1,k_2}(m) = \frac{(k_2-k_1-m)!}{(k_2-k_1)!}e^{m\int_{s0}^t \frac{1}{k_2-k_1-1_{\{k_1\leq N_t\}}(N_t\wedge k_2-k_1)}1_{\{k_1\leq N_s<k_2\}}d\Lambda_s} \cdot \prod_{j=k_2-k_1-(m-1)}^{k_2-k_1}(j-1_{\{k_1\leq N_t\}}(N_t\wedge k_2-k_1))$$

is a non-negative $\mathcal{H}$-martingale. First we calculate the integral in the exponent. It holds true that

$$\begin{aligned}
&\int_0^t \frac{1}{k_2-k_1-1_{\{k_1\leq N_s\}}(N_s-k_1)}1_{\{k_1\leq N_s<k_2\}}\frac{f(s)(n-N_s)}{1-F(s)}ds \\
&= \int_0^t \frac{(n-N_s)}{k_2-k_1-1_{\{k_1\leq N_s\}}(N_s-k_1)}1_{\{k_1\leq N_s<k_2\}}\frac{f(s)}{1-F(s)}ds \\
&= \sum_{j=k_1}^{k_2-1}\int_0^t \frac{(n-j)}{k_2-k_1-1_{\{k_1\leq j\}}(j-k_1)}1_{\{N_s=j\}}\frac{f(s)}{1-F(s)}ds \\
&= \sum_{j=k_1}^{k_2-1}\frac{(n-j)}{k_2-k_1-(j-k_1)}\int_{T_j\wedge t}^{T_{j+1}\wedge t}\frac{f(s)}{1-F(s)}ds \\
&= \sum_{j=k_1}^{k_2-1}\frac{(n-j)}{k_2-k_1-(j-k_1)}\left[-log(1-F(T_{j+1}\wedge t)+log(1-F(T_j\wedge t))\right] \\
&= log\left(\prod_{j=k_1}^{k_2-1}\left(\frac{1-F(T_j\wedge t)}{1-F(T_{j+1}\wedge t)}\right)^{\frac{(n-j)}{k_2-j)}}\right).
\end{aligned}$$

Therefore, it follows that

$$X_s^{k_1,k_2}(m) = \frac{(k_2-k_1-m)!}{(k_2-k_1)!} \prod_{j=k_1}^{k_2-1} \left(\frac{1-F(T_j \wedge t)}{1-F(T_{j+1} \wedge t)} \right)^{\frac{m(n-j)}{k_2-j}} \prod_{j=k_2-k_1-(m-1)}^{k_2-k_1} (j - 1_{\{k_1 \le N_t\}}(N_t \wedge k_2 - k_1)).$$

□

Although all these processes are the result of completely different transformations, they all have one thing in common. In all cases they are martingales with right-continuous paths and left limits, starting in $X_{s_0} = 1$ and satisfying the following conditions:

1. $\lim_{s \to s_1} X_s = 0$,
2. for all $t \in I : X_t - \lim_{s \uparrow t} X_s \le 0$.

Proposition 3.3.1 establishes that each process mentioned therein fulfills the equation

$$\mathbb{P}\left[\sup_{s \in I} X_s > c\right] = \frac{1}{c}.$$

4.2 The reverse martingale principle

This brief section will focus solely on the potential application of the results in the context of reverse martingales, where the underlying process follows a renormalized empirical distribution function. We define

$$\mathcal{F}_t := \sigma(\{N_s : s > t\}).$$

Heuristically, the reverse conditional intensity can be expressed by

$$\tilde{\lambda}(t \mid \mathcal{F}_t) = \lim_{\Delta \to 0} \frac{\mathbb{P}[N_{t-\Delta} - N_t = -1 \mid \mathcal{F}_t]}{\Delta}.$$

We define the function $\nu : [0,1] \to [0,1]$ by

$$\nu(s) := 1 - s.$$

We define the process $\tilde{N} = (\tilde{N}_t)_{t\in[0,1]}$ as

$$\tilde{N}_s := n - N_{\nu(s)-}.$$

It is a well-established fact, that the reverse conditional intensity $\tilde{\lambda} = (\tilde{\lambda}(t \mid \mathcal{F}_t))_{t\in[0,1]}$ has the form

$$\tilde{\lambda}(v \mid \mathcal{F}_v) = \frac{f(v)N_v}{F(v)},$$

with respect to the filtration $\mathcal{F} = (\mathcal{F}_t)_{t\in[0,1]}$ defined by

$$\mathcal{F}_s := \sigma(\{N_v \mid v > \rho(s)\}).$$

Lemma 4.2.1 (Reverse k-th non-negative single martingale)
Let $1 \leq k \leq n$ be arbitrary and T_k be the k-th Jump time of F_n and define $T_{n+1} = 1$. In the situation of this chapter, the process $\tilde{X}^k = (\tilde{X}^k_t)_{t\in(0,1]}$ defined by

$$\tilde{X}^k_s := 1_{\{\tilde{N}_s<k\}} + 1_{\{\tilde{N}_s=k\}} \left(\frac{F(T_k)}{F(s)}\right)^k$$

$$= 1_{\{n\cdot F_n(t-)\geq k\}} \left(\frac{F(T_{k+1})}{F(t)}\right)^{k1_{\{N_s\leq k\}}},$$

is a $\mathcal{F}$-martingale.

The point process $\tilde{N}$ is adapted to the filtration $\mathcal{F}$, and its conditional intensity is denoted by $\tilde{\lambda}$. Taking into account Theorem 3.2.4, the process $X = (X_t)_{t\in[0,1)}$ defined by

$$X^k_s =: 1_{\{\tilde{N}_s\leq k\}} e^{\int_{-\infty}^s 1_{\{\tilde{N}_v=k\}}\tilde{\lambda}(v|\mathcal{H}_v)dv},$$

is a $\mathcal{F}$-martingale. Hence,

$$1_{\{\tilde{N}_s\leq k\}} e^{\int_{-\infty}^s 1_{\{\tilde{N}_v=k\}}\tilde{\lambda}(v|\mathcal{H}_v)dv} = 1_{\{\tilde{N}_s<k\}} + 1_{\{\tilde{N}_s=k\}} \left(\frac{F(T_k)}{F(s)}\right)^k$$

$$= 1_{\{N_s\geq k\}} \left(\frac{F(T_{k+1})}{F(t)}\right)^{k1_{\{N_s\leq k\}}}$$

is a $\mathcal{F}$-martingale.

□

The subsequent figures serve to compare the reverse martingale $\tilde{X}^k$ with the martingale X^k obtained from Lemma 4.1.1, providing a visual representation of their respective behaviors over time.

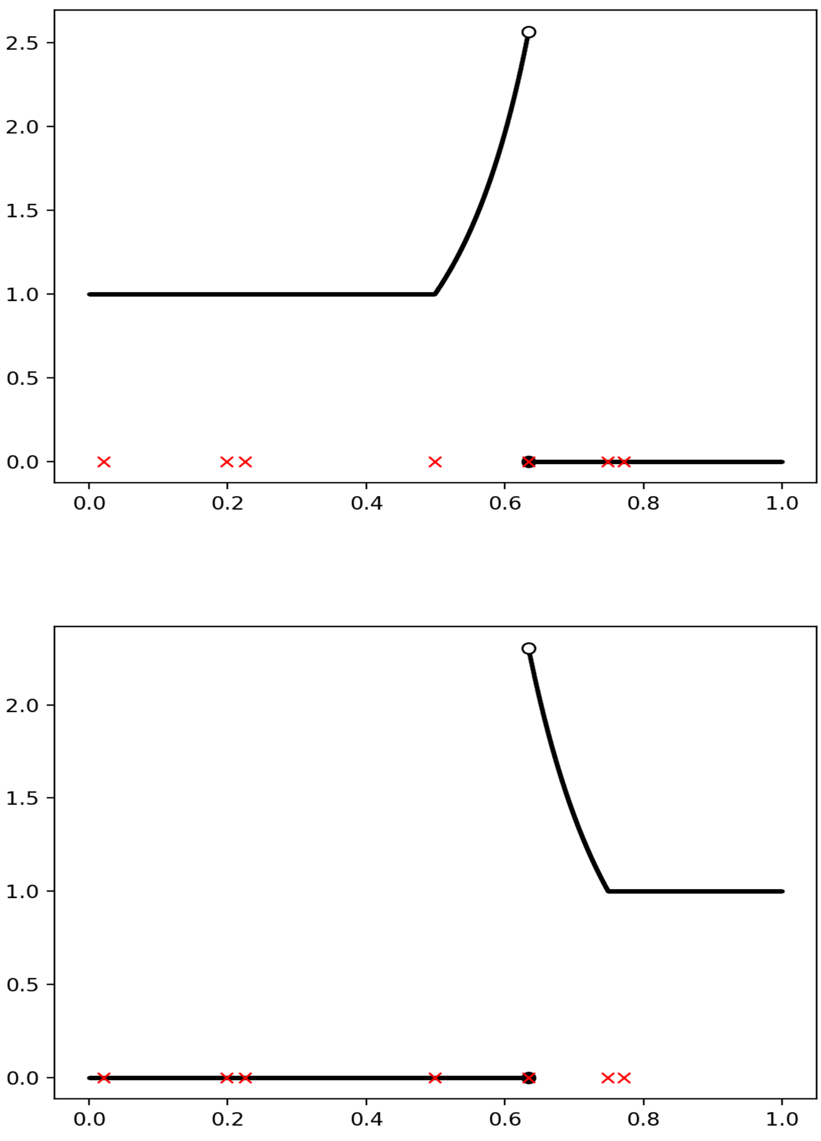

Figure 4.5: Using a uniform sample of size n=7 to transform forward X^4 (upper figure) and backward $\tilde{X}^4$ (lower figure).

Lemma 4.2.2 (Non-negative reverse martingale)
The process $\tilde{M} = (\tilde{M}_t)_{t\in[0,1)}$ defined by

$$\tilde{M}_s := F_n(s-) - \frac{1}{n} + \frac{1}{n}\left(\frac{F(T_{N_{s-}+1})}{F(s)}\right)^{nF_n(s)}$$

is a $\mathcal{F}$-martingale.

Proof.

For all $1 \le k \le n$ the process $\tilde{X}^k$ is a $\mathcal{F}$-martingale. Hence, the process $\tilde{M}$ defined by

$$\begin{aligned}\tilde{M}_s &:= \frac{1}{n}\sum_{k=0}^{n-1}\tilde{X}_s(k)\\ &= F_n(s-) + \frac{1}{n}\left(\frac{F(T_{N_{s-}})}{F(s)}\right)^{n(1-F_n(s))}\end{aligned}$$

is a $\mathcal{F}$-martingale.

□

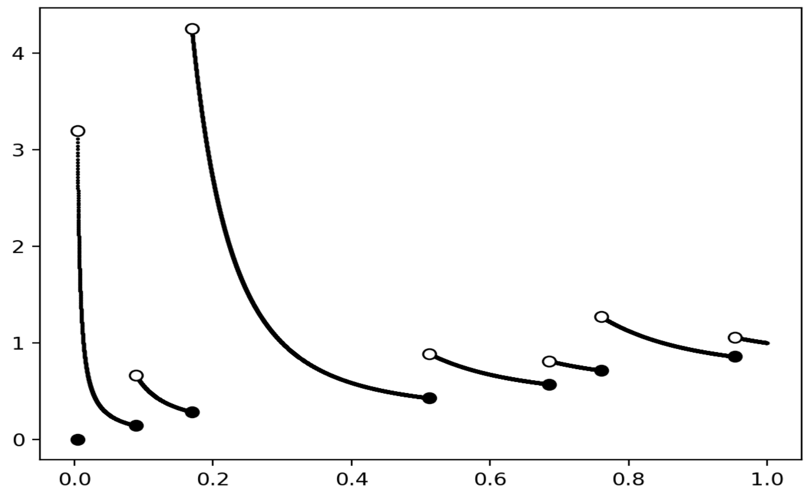

Figure 4.6: Typical path of the reverse martingale $\tilde{M}_s = F_n(s-) - \frac{1}{n} + \frac{1}{n}\left(\frac{T_{N_{s-}+1}}{s}\right)^{nF_n(s)}$ with sample size $n = 7$ and uniformly distributed random variables.

Remark 4.2.3 (The sibling-principle)
As demonstrated in Figure 4.3, the martingale M did not exhibit any outliers despite the fact that the underlying assumptions were not satisfied. However, upon applying the corresponding reverse transformation to the same data, a distinct deflection can now be observed, as depicted in the subsequent figure.

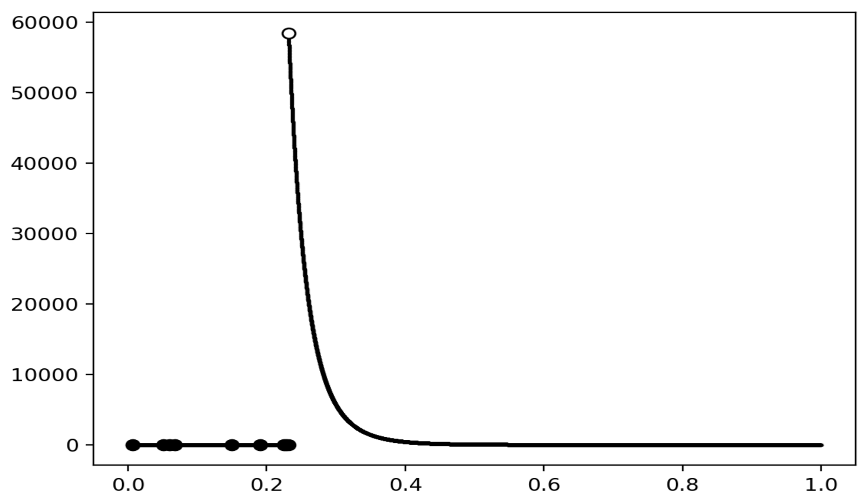

Figure 4.7: Transformation of the empirical distribution function under the assumption that the data are uniform on $[0, 1]$, but indeed are uniform on $[0, 0.3]$.

Remark 4.2.4 (Measure-preserving transformations)
In this remark we want to explain why we refrain from an in-depth analysis of reverse martingales and limit ourselves from now on exclusively to forward-facing martingales again.

Let $f : [0, 1] \to [0, 1]$ be a function such that

$$f(x) := \begin{cases} 1 - x & \text{if } x < 0.5, \\ |x - 1| & \text{if } x \geq 0.5. \end{cases}$$

So the point x is reflected around the value 0.5. Such a function f is special in the sense that it is a measure-preserving transformation with respect to the uniform distribution, if $X \sim \mathcal{U}([0,1])$ follows $f(X) \sim \mathcal{U}([0,1])$. More generaly if $X_1, X_2, ..., X_n$ are i.i.d. with $X_1 \sim \mathcal{U}([0,1])$ then the random sequence $f(X_1), f(X_2), ..., f(X_n)$ is also a uniform sample. The empirical distribution function $(F_n^f(t))_{t\in[0,1)}$ for the transformed sample is given by

$$\frac{1}{n}N_s^f := F_n^f(s) := \frac{1}{n}\sum_{i=1}^{n} 1_{\{f(X_i)\leq s\}}.$$

It is evident that the two dependent stochastic processes N^f and $N^{id} = n \cdot F_n$ have identical distributions. It follows that a deterministic martingale transformation φ in time and space for N^{id} is also a martingale transformation for N^f. It should be emphasized that while the deterministic transformation φ yields a martingale transformation for both processes, it is crucial to note that the internal filtrations of these processes are entirely distinct.

Applying Lemma 3.1.5 with $k_1 = 0$, $k_2 = n$, and $m = 1$, yields the martingale

$$\varphi(N_t, t) := \frac{1 - \frac{1}{n}N_s}{1 - F(s)} = \frac{1 - F_n(s)}{1 - F(s)}. \tag{4.1}$$

The associated reverse martingale is given by

$$\varphi^r(N_s, s) := \frac{\frac{1}{n}N_{s-}}{F(s)} = \frac{F_n(s-)}{F(s)}. \tag{4.2}$$

In the following plot we compare the $\mathcal{H}$-martingale $X = \varphi(N_\cdot^{id}, \cdot)$ with the reverse $\mathcal{F}$-martingale $\tilde{X} = \varphi^R(N_\cdot^f, \cdot)$. The reflection principle presented here, can be applied in the case of the empirical distribution function for all related transformations.

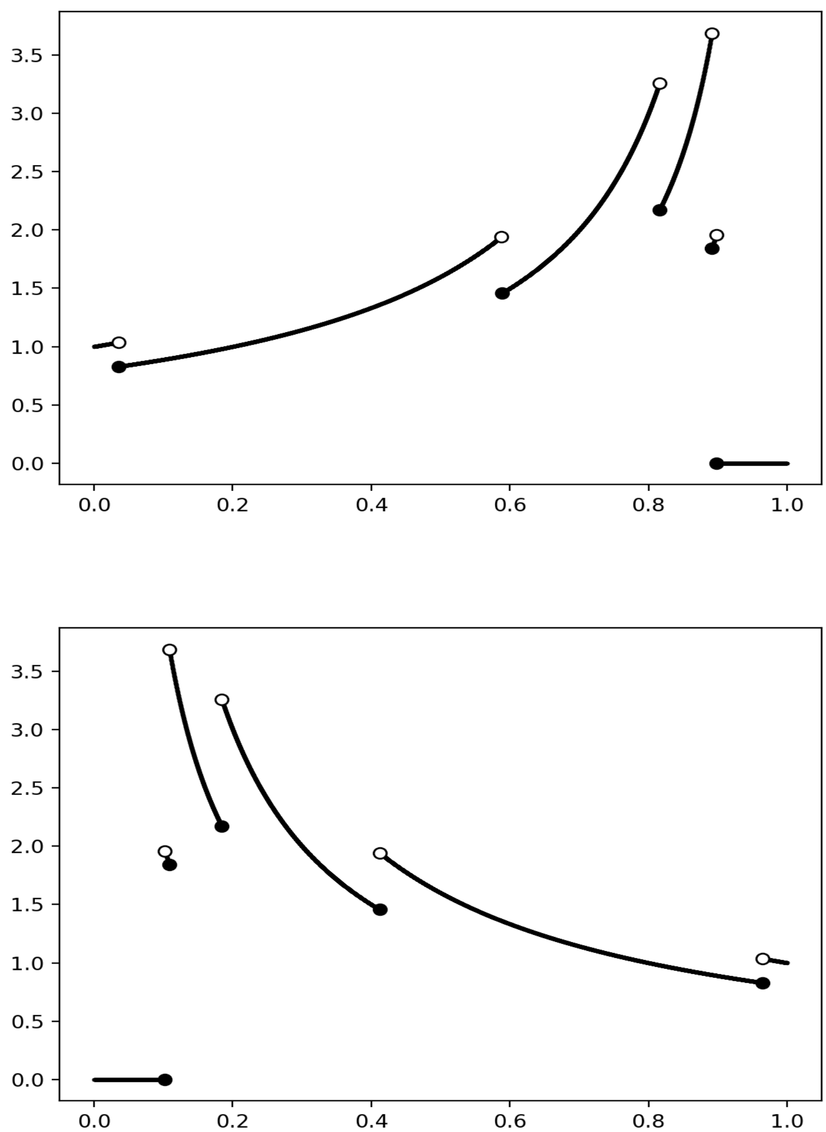

Figure 4.8: Apply a transformation to a uniform sample using $\varphi^r(N^{id}\cdot, \cdot)$ (3.1) and $\varphi(N^f\cdot, \cdot)$ (3.2).

4.3 Simulations and applications

This subsection focuses on the martingale M obtained from Lemma 4.1.2, which is defined by

$$M_s := \frac{n-1}{n} - F_n(s) + \frac{1}{n}\left(\frac{1-F(T_{N_s})}{1-F(s)}\right)^{n(1-F_n(s))}.$$

The martingale can be expressed as a transformation of the form $\varphi(t, N_t, T_{N_t})$.

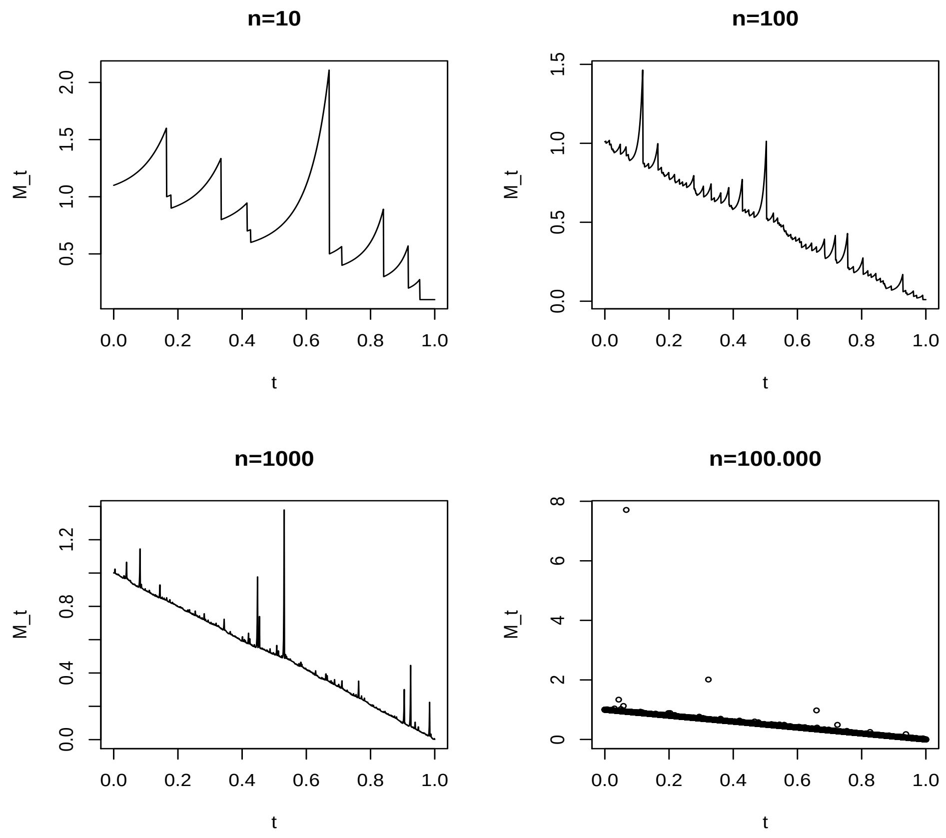

Figure 4.9: The martingale M for different sample sizes n (simplified illustration).

Remark 4.3.1 (Location of the supremum and the K–S statistic)
From previous investigations we know that the process M satisfies the equation

$$\mathbb{P}\left[\sup_{t\in[0,1)} M_t > c\right] = \frac{1}{c},$$

for all $c \geq 1$. Define the process D^n by

$$D_t^n = |F_n(t) - F(t)|.$$

The corresponding Kolmogorov–Smirnov statistic is well known in the literature (for example Noether [30]) and given by

$$\sup_{t\in[0,1)} |D_t^n| > c.$$

Inspired by the limit behavior, the mathematicians Dvoretzky, Kiefer and Wolfowitz gave a ”good” approximation in 1956 [12]. We have the inequality

$$\mathbb{P}\left[\sup_{t\in[0,1)} |D_t^n| > c\right] \leq 2e^{-2nc^2}.$$

A possible point of criticism when testing hypotheses with this test statistic is that the location where the supremum is observed is close to the median. Informally speaking, what happens at the edges is less important for the value of the supremum, since the deflections are smaller in these areas.

So we are interested in comparing the random variables

$$Z_1^n := \arg\max_{t\in[0,1)}(M_t + t)$$

and

$$Z_2^n := \arg\max_{t\in[0,1)} D_t^n.$$

Based on the previous discussion, a desirable property for a statistic is that it exhibits uniformity, i.e., we expect Z^i to be approximately uniformly distributed over the interval $[0, 1]$.

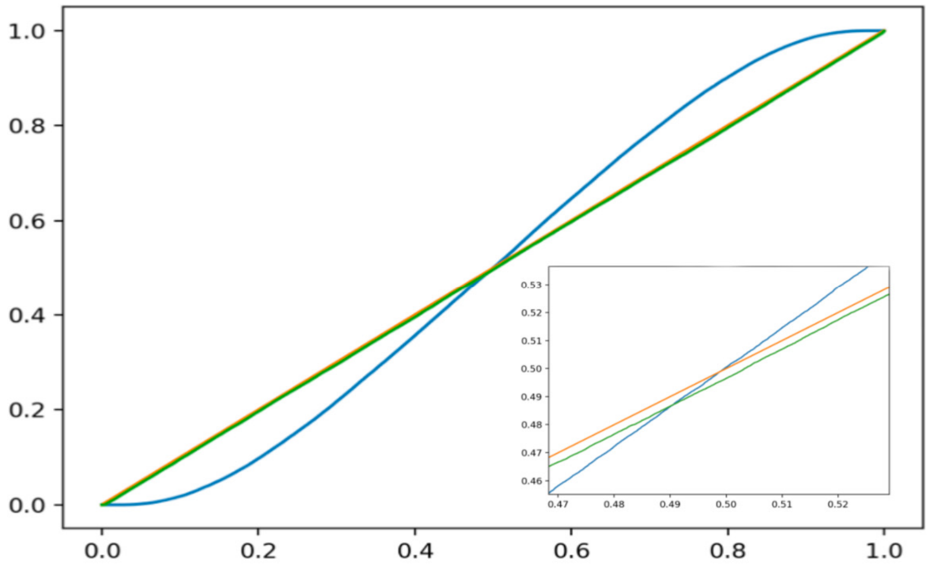

Figure 4.10: Comparison of the distributions Z_1^n (green) and Z_2^n (blue) against the uniform distribution (orange) with n=2000.

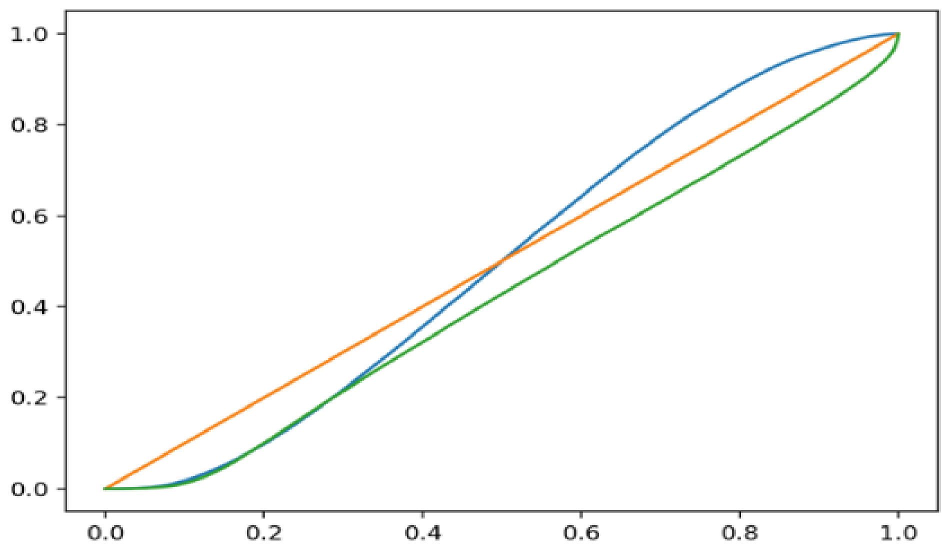

Figure 4.11: Comparison of the distributions Z_1^n (green) and Z_2^n (blue) against the uniform distribution (orange) with sample size n=20.

Overall, the simulation suggests that the distribution of Z_1^n adapts to the uniform distribution with increasing sample size n.

Remark 4.3.2 (Goodness of fit with censored data)
For an arbitrary but fixed $0 \leq m \leq n$, one can determine the complete path of the martingale X^m with

$$X_t^m := \frac{n^m(n-m)!}{n!} \prod_{j=n-(m-1)}^{n} \frac{\frac{j}{n} - F_n(t)}{1 - F(t)}, \tag{4.3}$$

by simply observing the first $n-(m-1)$ order statistics. This is one reason why Hess found an application of their work in situations involving right-censored data. The martingales presented in this thesis offer a significant extension of (4.3) and provide test statistics in a variety of scenarios involving censored data.

For instance, let $k_1 < k_2$ be arbitrary but fixed. In the context of statistical inference, suppose we are presented with a scenario wherein solely the order statistics, specifically $X_{k_1:n}, X_{k_1+1:n}, ..., X_{k_2:n}$, can be observed from a sample of uniformly and independently distributed random variables, denoted by $X_1, ..., X_n$. Lemma 4.1.5 yields the martingale $X^{k_1,k_2}(1)$ given by

$$X_s^{k_1,k_2}(1) := \frac{1}{k_2 - k_1} \prod_{j=k_1}^{k_2-1} \left(\frac{1 - F(T_j \wedge t)}{1 - F(T_{j+1} \wedge t)} \right)^{\frac{(k_2-k_1)(n-j)}{k_2-j}} \cdot (k_2 - k_1 - 1_{\{k_1 \leq n \cdot F_n(t)\}}(n \cdot F_n(t) \wedge k_2 - k_1)).$$

A sufficient condition for determining the full trajectory of the martingale $X_s^{k_1,k_2}(1)$ is to observe the random variables $X_{k_1:n}, X_{k_1+1:n}, ..., X_{k_2:n}$. As a result, the extended martingales presented in this dissertation provide a significantly broader range of potential applications in the context of censored data.

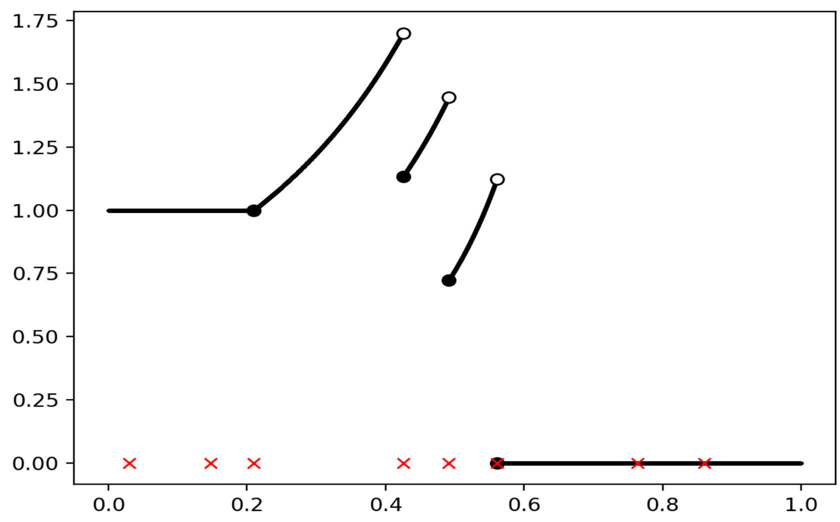

Figure 4.12: Typical path of the martingale $X^{3,6}(1)$.

Remark 4.3.3 (Power analysis)
This minor section aims to investigate the power of a test in a situation that represents an edge case of the application area. It is important to note that this test is just for illustration purposes, and tests closer to reality are usually more extensive to analyze.

Let N be a point process on the set $[0,1)$ with conditional intensity λ. For any k and $c \geq 1$ we have the equation

$$\mathbb{P}\left[\sup_{t\in[0,1)}\left(1_{\{N_s\leq k\}}e^{\int_0^t 1_{\{N_v=k\}}d\Lambda_v}.\right) > c\right] = \frac{1}{c}. \tag{4.4}$$

This test statistic represents a highly specialized test for the time between the k-th and $(k+1)$-th event. When considering the renormalized empirical distribution function with uniformly distributed random variables and $k=0$, the expression (4.4) reduces to

$$\mathbb{P}\left[\sup_{t\in[0,1)}\left(1_{\{F_n(t)=0\}}\left(\frac{1}{1-t}\right)^n\right) > c\right] = \frac{1}{c}. \tag{4.5}$$

Let $x \in (0,1)$ be fixed but arbitrary. We formulate the null hypothesis H_0 and the alternative H_1 :

1. H_0: $X_1, ..., X_n$ is i.i.d with $X_1 \sim \mathcal{U}([0,1])$.
2. H_1 : $X_1, ..., X_n$ is i.i.d with $X_1 \sim \mathcal{U}([x,1])$.

If we consider the process X given by

$$X_s := 1_{\{F_n(s)=0\}} \left(\frac{1}{1-s} \right)^n,$$

then the distribution of the supremum under the null hypothesis H_0 is given by equation (4.5). Given $0 < \alpha < 1$, we reject H_0 if the path of X exceeds the threshold value of $1/\alpha$.

Under H_1, we aim to determine the minimal sample size n at which H_0 is rejected with probability 1:

$$\mathbb{P}^{H_1} \left[\sup_{t \in [0,1)} X_t > \frac{1}{\alpha} \right] = 1. \tag{4.6}$$

Clearly, the minimum sample size will depend on the given values α and x.

Under H_1 we have $X_1, ..., X_n \geq x$. Hence,

$$\sup_{t \in [0,1)} \left(1_{\{F_n(t)=0\}} \left(\frac{1}{1-t} \right)^n \right) \geq \left(\frac{1}{1-x} \right)^n.$$

So H_1 will be rejected with probability 1 if

$$\left(\frac{1}{1-x} \right)^n > 1/\alpha,$$

which is equivalent to

$$n > \frac{log(1-x)}{log(\alpha)}. \tag{4.7}$$

All in all, we have now proven that if n satisfies condition (4.7), equation (4.6) holds true.

A variety of possibilities to transform a point process into a martingale, whose exceedance probabilities are known, have been presented in this work. This observation suggests that by selecting an appropriate martingale transformation for a given problem, tests with reasonable performance can be constructed. Our investigation shows that the search for good test statistics is equivalent to the search for martingale transformations that produce deviations under the alternative hypothesis.

Chapter 5

Principal Component Analysis

In the paper by Hackmann [17] "Karhunen–Loève expansions (KLE) of Lèvy processes", it is noted that the KLE is explicitly available for only a few processes. In the context of his work, Hackmann demonstrates the dependency of the coefficients of the Karhunen-Loève Expansion (KLE) for Levy Processes with jumps through the application of measure theoretical arguments. Within the context of this chapter, our objective is to introduce a class of point processes characterized by a minimum type kernel. More specifically, this chapter is focused on the principal component analysis of centralized point processes that have a minimum type Covariance kernel equal to that of a Brownian motion. To do so, we strive for giving a generalization of the known Poisson-Charlier martingales which are based on the Poisson process. The generalization makes the polynomials applicable to more general point processes. By calculating the higher order moments, especially the second order moments, we can derive a class of point processes with minimum type variance kernel.

It should be noted that in order to extend the applicability of the Poisson-Charlier martingales to more general point processes, we will first derive them using our own techniques. This will enable us to develop a generalization of these martingales. Hence, the focus of the initial part of this chapter is on the Poisson process. In order to obtain the Poisson-Charlier polynomials, we will employ the martingale differential equation:

$$\frac{\partial \varphi}{\partial t}(x,t) + \left(\varphi(1+x,t) - \varphi(x,t)\right)\rho = 0,$$

which is not commonly used in the literature for this purpose. Typically, the Poisson-Charlier polynomials are defined through their generating function, as demonstrated in [34]. To the best of our knowledge, the approach we have employed for the derivation of the Poisson-Charlier martingales has not been previously utilized in the literature. The calculations carried out with this methodology demonstrate an inherent link between the polynomials and the Poisson process.

5.1 Basics

The Poisson process is typically constructed by defining a sequence of independent and identically distributed intervals between events. Since the internal conditional intensity plays a crucial role in this work we will characterize the Poisson process via its internal conditional intensity.

Definition 5.1.1 (Poisson process with parameter ρ)
A point process $N = (N_t)_{t\in I}$ with natural generated filtration $\mathcal{H} = (\mathcal{H}_t)_{t\in I}$ is considered a Poisson process with parameter ρ if it has a conditional intensity $\lambda = (\lambda(t \mid \mathcal{H}_t))_{t\in I}$ such that $\lambda(s \mid \mathcal{H}_s) = \rho$ for some constant $\rho > 0$.

The Poisson process exhibits a variety of advantageous characteristics, which are summarized in the upcoming proposition. It is important to note that throughout this chapter, we consider the index set to be equivalent to the set of non-negative real numbers $I = \mathbb{R}_+$.

Proposition 5.1.2 (Properties of the Poisson process)
Let N be a Poisson Process with parameter ρ. The subsequent properties are satisfied:

1. N has independent increments, that means for all $t_1 < t_2 < ... < t_n$ the random variables $N_{t_1}, N_{t_2}-N_{t_1}, ..., N_{t_n}-N_{t_{n-1}}$ are stochastically independent.

2. For all $t \in \mathbb{R}_+$ the random variable N_t is poisson-distributed with parameter $\rho \cdot t$.

3. The process $X = (X_t)_{t\in\mathbb{R}_+}$ defined by

$$X_s := N_s - \rho \cdot s.$$

is a martingale with respect to $\mathcal{H}$. We call X the compensated process of N.

4. The compensated process X of N has the covariance kernel K with

$$K(s,t) = \min(s,t).$$

The compensated process, also known as the trivial martingale, has a widely recognized covariance structure, particularly in the context of Brownian motion. Applications arise in the theory of principal component analysis.

Remark 5.1.3 (Trivial martingale)
Let N be a point process with internal conditional $\mathcal{H}$-intensity λ. As previously established in earlier chapters, the stochastic process denoted as X given by

$$X_s := N_s - \int_0^s \lambda(v \mid \mathcal{H}_v)dv, \tag{5.1}$$

is a $\mathcal{H}$-martingale. In case of a Poisson Process with parameter ρ the process X turn out to be

$$X_s = N_s - \rho \cdot s.$$

Define the function $\varphi : \mathbb{N}_0 \times \mathbb{R}_+ \to \mathbb{R}$, $(x,s) \to x - \rho s$. Clearly, for a (PPP) we have

$$X_s = \varphi(N_s, s). \tag{5.2}$$

Therefore, it follows that X is a martingale due to a transformation that only considers the current state and time, as expressed in equation (5.2). Of course, the compensated martingale does not typically exhibit this property. For example, let $\tilde{N}$ be the renormalized empirical distribution function with a uniform sample on $[0,1]$, i.e. $\tilde{N}$ admits the internal conditional $\mathcal{H}$-intensity

$$\tilde{\lambda}(v \mid \mathcal{H}_v) = \frac{(n - \tilde{N}_v)1_{[0,1]}(v)}{1 - v}.$$

For this case, the integral in (5.1) turn out to be

$$
\begin{aligned}
&\int_0^s \frac{(n-\tilde{N}_v)1_{[0,1]}(v)}{1-v}dv \\
&= \sum_{i=0}^{T_{\tilde{N}_s}-1} \int_{T_i}^{T_{i+1}} \frac{(n-i)}{1-v}dv + \int_{T_{\tilde{N}_s}}^{s} \frac{(n-\tilde{N}_s)}{1-v}dv \\
&= \sum_{i=0}^{T_{\tilde{N}_s}-1} log\left(\left(\frac{1-T_i}{1-T_{i+1}}\right)^{(n-i)}\right) + (n-\tilde{N}_s)log\left(\left(\frac{1-T_{\tilde{N}_s}}{1-s}\right)^{n-\tilde{N}_s}\right) \\
&= \sum_{i=0}^{n-1} log\left(\left(\frac{1-(T_i\wedge s)}{1-(T_{i+1}\wedge s)}\right)^{(n-i)}\right) \\
&\equiv \varphi(\tilde{N}_s, s, T_1\wedge s, T_2\wedge s, ..., T_n\wedge s).
\end{aligned}
$$

Remark 5.1.4 (Apply the Hess' transformation [19] to a Poisson process)
Let $m, n \in \mathbb{N}$ be arbitrary but fixed with $m \leq n$ and N be a point process with compensator Λ. According to Theorem 3.2.8, the process X defined by

$$
\begin{aligned}
X_s := \frac{(n-m)!}{n!} e^{m\int_{-\infty}^{s} \frac{1}{n-1_{\{0\leq N_t\}}(N_t-0)} 1_{\{0\leq N_t<n\}} d\Lambda_t} \\
\cdot \prod_{j=n-(m-1)}^{n} (j - 1_{\{0\leq N_t\}}(N_t-0))
\end{aligned}
\tag{5.3}
$$

is a $\mathcal{H}$-martingale. In case of a renormalized empirical distribution function with uniform sample on $[0, 1]$ and samplesize n, the process in (5.3) turn out to be

$$
X_s = \frac{n^m(n-m)!}{n!} \prod_{j=n-(m-1)}^{n} \frac{\frac{j}{n} - F_n(s)}{1-F(s)} \equiv \varphi(N_s, s).
$$

On the other hand if N is a Poisson process with parameter ρ the process can be rewritten to the expression

$$\begin{aligned}
X_s &= \frac{(n-m)!}{n!} e^{m\rho \int_0^s \frac{1}{n-N_v} 1_{\{0 \le N_v < n\}} dv} \prod_{j=n-(m-1)}^{n} (j - N_s) \\
&= \frac{(n-m)!}{n!} e^{m\rho \sum_{i=0}^{N_s \wedge (n-1)} \int_0^s \frac{1_{\{N_v = i\}}}{n-i} dv} \prod_{j=n-(m-1)}^{n} (j - N_s) \\
&= \frac{(n-m)!}{n!} e^{m\rho \sum_{i=0}^{N_s \wedge (n-1)} \int_{T_i}^{T_{i+1} \wedge t} \frac{1}{n-i} ds} \prod_{j=n-(m-1)}^{n} (j - N_s) \\
&= \frac{(n-m)!}{n!} e^{m\rho \sum_{i=0}^{N_s \wedge (n-1)} \frac{1}{n-i} (T_{i+1} \wedge s - T_i \wedge s)} \prod_{j=n-(m-1)}^{n} (j - N_s) \\
&\equiv \varphi(N_s, s, T_1 \wedge s, T_2 \wedge s, ..., T_n \wedge s).
\end{aligned}$$

Remarks 5.1.3 and 5.1.4 provide valuable insights into the nature of the martingales discussed in this chapter. Specifically, the martingale transformation given in equation (5.3) is an instance of a transformation that becomes simpler when dealing with an empirical distribution function, taking the form of a function φ that depends only on time and state. On the other hand, the trivial martingale outlined in equation (5.1) follows the opposite trend.

The objective of the next section is to specify the martingales X for a Poisson process N with parameter ρ, where the transformation is a function solely in state and time, denoted as $X_s \equiv \varphi(N_s, s)$.

5.2 Martingales representable through $\varphi(N., \cdot)$

In this section, we make the assumption that N is a Poisson process with parameter ρ and we denote its internal history as $\mathcal{H}$. In this section we will characterize the set of $\mathcal{H}$-martingales X, which can be represented by a transformation of the process N and a time component $X_s \equiv \varphi(N_s, s)$.

The initial step is to derive the appropriate differential equation for this problem.

Lemma 5.2.1 (Doob-Meyer decomposition for $X \equiv \varphi(N., \cdot)$)
Let N be a Poisson process with parameter ρ and $\mathcal{H}$ be the filtration

generated by N. Let X be a $\mathcal{H}$-adapted process admitting the representation $X \equiv \varphi(N_t, t)$. The innovation martingale M and the compensator A associated with process X are given by

$$M_s := \varphi(N_0, 0) + \int_{(0,s]} \varphi(N_v, v) - \varphi(N_{v-}, v) dN_v \\ - \rho \int_{(0,s]} \varphi(N_v + 1, v) - \varphi(N_v, v) dv$$

and

$$A_s := \int_{(0,s]} \frac{\partial \varphi}{\partial t}(N_v, v) + \rho\varphi(N_v + 1, v) - \rho\varphi(N_v, v) dv.$$

Proof.

This follows directly from Propositions 3.1.1 and 3.1.2.

□

Lemma 5.2.2 (The differential equation)
Let N be a Poisson process with parameter ρ and $\mathcal{H}$ be the filtration generated by N. A process X with $X_s = \varphi(N_s, s)$ is a martingale with respect to $\mathcal{H}$, if the function $\varphi : \mathbb{N} \times \mathbb{R}_+ \to \mathbb{R}$ satisfies the equation

$$\frac{\partial \varphi}{\partial t}(x, t) + \rho\varphi(x + 1, t) - \rho\varphi(x, t) = 0,$$

for all $(x, t) \in \mathbb{N} \times \mathbb{R}_+$ Lebesgue-almost everywhere.

This statement is a direct consequence of Proposition 5.2.1.

□

As discussed in earlier chapters, the equation presented in 5.2.2 is well-known in the literature. To develop an understanding of this equation, we explore two commonly used solutions through the following examples.

Example 5.2.3 (Trivial martingale)
Define the function $\varphi(x, t) := x - \rho \cdot t$. Observe that

$$\frac{\partial \varphi}{\partial t}(x, t) = -\rho,$$

and that

$$\rho\varphi(x+1,t) - \rho\varphi(x,t) = \rho(x+1-\rho\cdot t - x + \rho\cdot t) = \rho.$$

Thus, the differential equation is satisfied, and we can conclude that the process X defined by

$$X_s = N_s - \rho \cdot s$$

is a martingale with respect to $\mathcal{H}$.

Example 5.2.4 (Geometric Poisson process)
Let $c_1 \in \mathbb{R}$ be arbitrary but fixed. Define the function $\varphi(x,t) := e^{c_1 x - (\rho e^{c_1} - \rho)t}$. Observe that

$$\frac{\partial\varphi}{\partial t}(x,t) = -\rho(e^{c_1} - 1)e^{c_1 x - (\rho e^{c_1} - \rho)t}$$

and that

$$\begin{aligned}\rho\varphi(x+1,t) - \rho\varphi(x,t) &= \rho e^{c_1(x+1) - (\rho e^{c_1} - \rho)t} - \rho e^{c_1 x - (\rho e^{c_1} - \rho)t} \\ &= \rho(e^{c_1} - 1)e^{c_1 x - (\rho e^{c_1} - \rho)t}.\end{aligned}$$

Thus, the differential equation is satisfied, and we can conclude that the process X defined by

$$X_s = e^{c_1 N_s - (\rho e^{c_1} - \rho)s}$$

is a martingale with respect to $\mathcal{H}$.

5.3 Poisson-Charlier Martingales

In this section, we aim at determining the set of martingales that can be obtained by transforming a Poisson process N_t via a function of the form $X_t = \varphi(N_t, t)$. To develop an understanding of the techniques involved, we begin by considering the case where the transformation is a second-order polynomial. Once we have gained insight into this case, we will extend our analysis to more general n-th order polynomials and determine the corresponding martingales.

Proposition 5.3.1 (Polynomial martingales of second order)
Let N be a Poisson process with paramter ρ and $\mathcal{H}$ be the filtration generated by N. Then the process X defined by

$$X_s := c\left(N_s^2 - 2\rho s N_s + \rho^2 s^2 - \rho s\right)$$

is a martingal with respect to $\mathcal{H}$.

Proof.

Let φ be a polynomial function of second order, that is

$$\varphi(x,t) = \varphi_0(t) + \varphi_1(t) \cdot x + \varphi_2(t) \cdot x^2,$$

where $\varphi_i(t)$ is a differentiable function for all $0 \leq i \leq 2$. It follows that

$$\frac{\partial \varphi}{\partial t}(x,t) = \frac{\partial \varphi_0}{\partial t}(t) + \frac{\partial \varphi_1}{\partial t}(t) \cdot x + \frac{\partial \varphi_2}{\partial t}(t) \cdot x^2.$$

According to Lemma 5.2.2, the function φ needs to satisfy the equation

$$\frac{\partial \varphi}{\partial t}(x,t) + \rho\varphi(1+x,t) - \rho\varphi(x,t) = 0. \tag{5.4}$$

By substituting the representations of φ and $\frac{\partial \varphi}{\partial t}$, we arrive at the following equation

$$\frac{\partial \varphi}{\partial t}(x,t) + \rho\varphi(1+x,t) - \rho\varphi(x,t)$$
$$= \frac{\partial \varphi_0}{\partial t}(t) + \frac{\partial \varphi_1}{\partial t}(t)x^2 + \rho\Big(\varphi_1(t) + \varphi_2(t) + \varphi_2(t)2x\Big).$$

Consequently, we can conclude that (5.4) is equivalent to

$$\frac{\partial \varphi_0}{\partial t}(t) + \rho\varphi_1(t) + \rho\varphi_2(t) + \Big(\frac{\partial \varphi_1}{\partial t}(t) + 2\rho\varphi_2(t)\Big)x + \frac{\partial \varphi_2}{\partial t}(t)x^2 = 0. \tag{5.5}$$

The equation is fulfilled for all $x \in \mathbb{N}$ if and only if all coefficients are equal to zero. We first consider the second order term in equation (5.5). The equation

$$\frac{\partial \varphi_2}{\partial t}(t) = 0$$

holds true if and only if $\varphi_2(t) = c$ for some constant $c \in \mathbb{R}$.

Substituting φ_2 into the first order term results in

$$\frac{\partial \varphi_1}{\partial t}(t) + 2\rho c = 0.$$

Consequently, it follows that

$$\varphi_1(t) = -2c\rho t + c_2,$$

for some constant $c_2 \in \mathbb{R}$. In order to eliminate any duplicate calculations, we specify that c_2 is equal to 0. Substituting φ_1 and φ_2 into the 0-th order term results in

$$\frac{\partial \varphi_0}{\partial t}(t) + \rho 2c\rho t + \rho c = 0.$$

Thus, we can conclude that

$$\varphi_0(t) = \rho^2 ct^2 - \rho ct + c_3.$$

To avoid duplicate calculations, we set $c_3 = 0$. Plugging φ_0, φ_1 and φ_2 into the representation of φ from the beginning yields

$$\begin{aligned}\varphi(x,t) &= \varphi_0(t) + \varphi_1(t)\cdot x + \varphi_2(t)\cdot x^2 \\ &= c\left(\rho^2 t^2 - \rho t - 2\rho t x + x^2\right).\end{aligned}$$

□

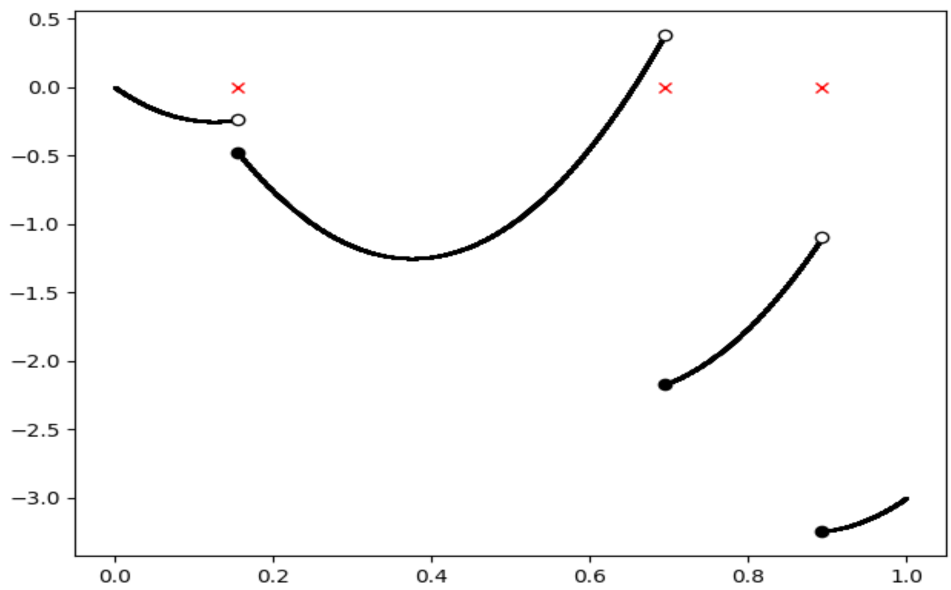

Figure 5.1: Realization of a Poisson Process with parameter $\rho = 4$ (red crosses) and the associated polynomial martingale of second order.

To proceed with the objective of this section, it is necessary to first define a sequence of numbers denoted by (a_i^j). The sequence possesses specific properties that are essential in relation to Poisson-Charlier martingales.

Definition 5.3.2 (Coefficient sequence of polynomial martingales)
We define the (deterministic) sequence $(a_i^{(0)})_{i\in\mathbb{N}_0}$ by

$$a_i^{(0)} := 1_{\{i=1\}}.$$

Further we define the sequence $(a_i^{(1)})_{i\in\mathbb{N}_0}$ by

$$a_i^{(1)} := \frac{1}{i!}.$$

Furthermore, we recursively define for $j > 1$ the sequence $(a_i^{(j)})_{i\in\mathbb{N}_0}$ by

$$a_i^{(j)} := \sum_{k=1}^{i} a_k^{(j-1)} \cdot \frac{1}{(i+1-k)!}.$$

Remark 5.3.3
By defining the sequence of coefficients as above the following results hold true:

$$\sum_{i=1}^{\infty} a_i^{(1)} = \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + ... = (e-1)$$
$$\sum_{i=1}^{\infty} a_i^{(2)} = \frac{1}{1!}\frac{1}{1!} + (\frac{1}{1!}\frac{1}{2!} + \frac{1}{2!}\frac{1}{1!}) + (\frac{1}{1!}\frac{1}{3!} + \frac{1}{2!}\frac{1}{2!} + \frac{1}{3!}\frac{1}{1!}) + ... = (e-1)^2.$$

One can prove the following generalization:

$$\sum_{i=1}^{\infty} a_i^{(k)} = (e-1)^k.$$

Theorem 5.3.4 (Poisson-Charlier polynomials)
Let N be a Poisson process with paramter ρ and $\mathcal{H}$ be the filtration generated by N. Then the process X defined by

$$X_s := \sum_{i=0}^{n} \frac{n!}{(n-i)!} \sum_{j=0}^{i} (-\rho)^j \frac{1}{j!} s^j a_{i+1-j}^{(j)} N_s^{n-i},$$

is a martingale with respect to $\mathcal{H}$.

Proof.

Let φ be a polynomial function of n-th order, that is

$$\varphi(x,t) = \sum_{k=0}^{n} \varphi_k(t) x^k.$$

Consequently, it follows that

$$\frac{\partial \varphi}{\partial t}(x,t) = \sum_{k=0}^{n} \frac{\partial \varphi_k}{\partial t}(t) x^k.$$

According to Lemma 5.2.2, the function φ needs to satisfy the equation

$$\frac{\partial \varphi}{\partial t}(x,t) + \rho\varphi(1+x,t) - \rho\varphi(x,t) = 0. \tag{5.6}$$

Through manipulation, we can transform the left-hand side of (5.6) to

$$
\begin{aligned}
&\frac{\partial\varphi}{\partial t}(x,t)+\rho\varphi(1+x,t)-\rho\varphi(x,t)\\
&=\sum_{k=0}^{n}\frac{\partial\varphi_k}{\partial t}(t)x^k+\rho\left(\sum_{k=0}^{n}\varphi_k(t)(1+x)^k-\sum_{k=0}^{n}\varphi_k(t)x^k\right)\\
&=\sum_{k=0}^{n}\frac{\partial\varphi_k}{\partial t}(t)x^k+\rho\Big(\sum_{k=0}^{n}\varphi_k(t)\sum_{j=0}^{k}\binom{k}{j}x^j-\sum_{k=0}^{n}\varphi_k(t)x^k\Big)\\
&=\sum_{k=0}^{n}\frac{\partial\varphi_k}{\partial t}(t)x^k+\rho\Big(\sum_{k=0}^{n}\varphi_k(t)\sum_{j=0}^{k-1}\binom{k}{j}x^j\Big)\\
&=\sum_{k=0}^{n}\frac{\partial\varphi_k}{\partial t}(t)x^k+\rho\left(\sum_{k=0}^{n-1}\sum_{j=k+1}^{n}\varphi_j(t)\binom{j}{k}x^k\right)\\
&=\sum_{k=0}^{n-1}x^k\left(\frac{\partial\varphi_k}{\partial t}(t)+\rho\sum_{j=k+1}^{n}\varphi_j(t)\binom{j}{k}\right)+\frac{\partial\varphi_n}{\partial t}(t).
\end{aligned}
$$

From this, we can infer that (5.6) is tantamount to

$$\sum_{k=0}^{n-1}x^k\left(\frac{\partial\varphi_k}{\partial t}(t)+\rho\sum_{j=k+1}^{n}\varphi_j(t)\binom{j}{k}\right)+\frac{\partial\varphi_n}{\partial t}(t)=0. \tag{5.7}$$

The equation is fulfilled for all $x\in\mathbb{N}$ if and only if all coefficients are equal to zero.

Assertion: The equation is fulfilled if for all $1\leq i\leq n$ the function φ_{n-i} has the representation

$$\varphi_{n-i}(t)=\frac{n!}{(n-i)!}\sum_{j=0}^{i}(-\rho)^j\frac{1}{j!}t^j a_{i+1-j}^{(j)}. \tag{5.8}$$

1. Base case: Show that the statement holds for $i=0$.

   We are setting c to be equal to 1. On the one hand (5.7) implies that

   $$\varphi_{n-0}(t)=c.$$

   On the other hand, we have

   $$\frac{n!}{(n-i)!}\sum_{j=0}^{0}(-\rho)^j\frac{1}{j!}t^j a_{0+1-j}^{(j)}=1.$$

2. Induction step: Show that for any k, if $\varphi_n, \varphi_{n-1}, ..., \varphi_{n-i+1}$ have the representation from (5.8), then the representation also holds true for φ_{n-i}.

   It follows that

$$\begin{aligned}
\frac{\partial \varphi_{n-i}}{\partial t}(t) &= -\rho \sum_{j=n-(i-1)}^{n} \varphi_j(t) \binom{j}{n-i} \\
&= -\rho \sum_{j=0}^{i-1} \varphi_{n-j}(t) \binom{n-j}{n-i} \\
&= -\rho \sum_{j=0}^{i-1} \left(c\frac{n!}{(n-j)!} \sum_{k=0}^{j} (-\rho)^k \frac{1}{k!} t^k a_{j+1-k}^{(k)} \right) \frac{(n-j)!}{(n-i)!(i-j)!} \\
&= -\rho \sum_{j=0}^{i-1} \left(c\frac{n!}{(n-i)!} \sum_{k=0}^{j} (-\rho)^k \frac{1}{k!} t^k a_{j+1-k}^{(k)} \frac{1}{(i-j)!} \right) \\
&= c\frac{n!}{(n-i)!} \sum_{j=0}^{i-1} \left(\sum_{k=0}^{j} (-\rho)^{k+1} \frac{1}{k!} t^k a_{j+1-k}^{(k)} \frac{1}{(i-j)!} \right).
\end{aligned}$$

   Consequently, it holds true that

$$\begin{aligned}
\varphi_{n-i}(t) &= \frac{n!}{(n-i)!} \sum_{j=0}^{i-1} \left(\sum_{k=1}^{j} (-\rho)^{k+1} \frac{1}{(k+1)!} t^{k+1} a_{j+1-k}^{(k)} \frac{1}{(i-j)!} \right) \\
&= \frac{n!}{(n-i)!} \sum_{j=0}^{i-1} \left(\sum_{k=1}^{j} (-\rho)^{k+1} \frac{1}{(k+1)!} t^{k+1} a_{j+1-k}^{(k)} \frac{1}{(i-j)!} \right) \\
&= \frac{n!}{(n-i)!} \sum_{k=0}^{i-1} (-\rho)^{k+1} \frac{1}{(k+1)!} t^{k+1} \sum_{j=k}^{i-1} a_{j+1-k}^{(k)} \frac{1}{(i-j)!} \\
&= \frac{n!}{(n-i)!} \sum_{k=0}^{i-1} (-\rho)^{k+1} \frac{1}{(k+1)!} t^{k+1} \sum_{j=1}^{i-1-(k-1)} a_j^{(k)} \frac{1}{(i-j+1-k)!} \\
&= \frac{n!}{(n-i)!} \sum_{k=0}^{i-1} (-\rho)^{k+1} \frac{1}{(k+1)!} t^{k+1} \sum_{j=1}^{i-k} a_j^{(k)} \frac{1}{(i-k+1-j)!} \\
&= c\frac{n!}{(n-i)!} \sum_{k=0}^{i} (-\rho)^k \frac{1}{k!} t^k a_{i-k+1}^{(k)}.
\end{aligned}$$

   Assertion (5.8) thus follows.

Plugging representation (5.8) into φ yields

$$\varphi(x,t) = \sum_{i=0}^{n} \varphi_{n-i}(t) x^{n-i} = \sum_{i=0}^{n} \frac{n!}{(n-i)!} \sum_{j=0}^{i} (-\rho)^j \frac{1}{j!} t^j a_{i+1-j}^{(j)} x^{n-i}.$$

□

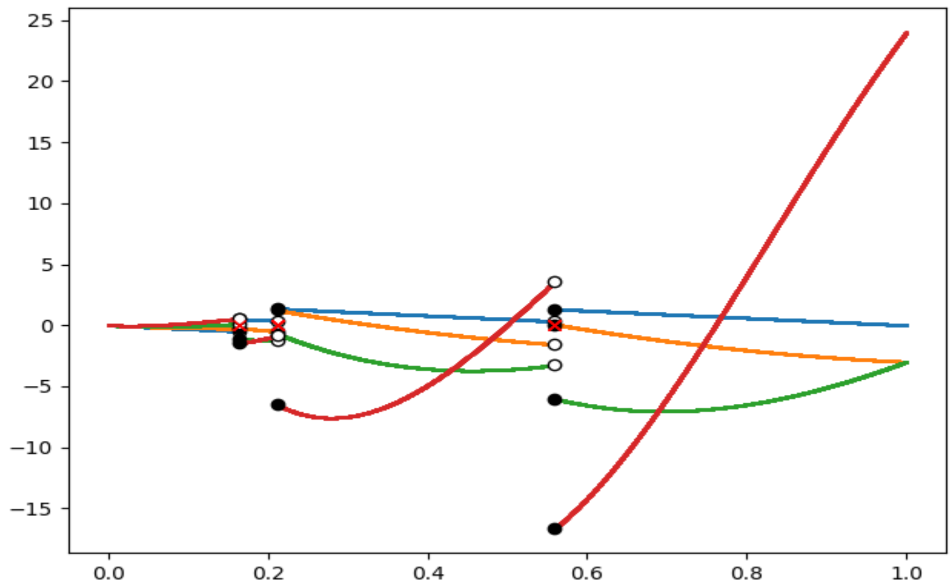

Figure 5.2: Realization of a Poisson Process with parameter $\rho = 4$ (red croesses) and the associated polynomial martingale of first (blue), second (orange), third (green) and fourth (red) order.

Remark 5.3.5 (Illustration of the functions φ_n)
In this remark, we aim to provide explicit expressions for the functions φ_{n-i} with $i = 0, 1, 2, \ldots, 5$, to facilitate a better comprehension of their properties. These are explicitly given by:

$$\begin{aligned}
\varphi_n(t) &= c \\
\varphi_{n-1}(t) &= c\frac{n!}{(n-1)!}\Big(-\rho t \underbrace{\frac{1}{1!}}_{=a_1^{(1)}}\Big) \\
\varphi_{n-2}(t) &= c\frac{n!}{(n-2)!}\Big(-\rho t \underbrace{\frac{1}{2!}}_{=a_2^{(1)}} + \rho^2\frac{1}{2}t^2 \underbrace{\frac{1}{1!}}_{=a_1^{(2)}}\Big) \\
\varphi_{n-3}(t) &= c\frac{n!}{(n-3)!}\Big(-\rho t \underbrace{\frac{1}{3!}}_{=a_3^{(1)}} + \rho^2\frac{1}{2}t^2 \underbrace{\Big(\frac{1}{2!}+\frac{1}{2!}\Big)}_{=a_2^{(2)}} - \rho^3\frac{1}{3!}t^3 \underbrace{\frac{1}{1!}}_{=a_1^{(3)}}\Big) \\
\varphi_{n-4}(t) &= c\frac{n!}{(n-4)!}\Big(-\rho t \underbrace{\frac{1}{4!}}_{=a_4^{(1)}} + \rho^2\frac{1}{2}t^2 \underbrace{\Big(\frac{1}{3!}+\frac{1}{2!2!}+\frac{1}{3!}\Big)}_{=a_3^{(2)}} - \rho^3\frac{1}{3!}t^3 \underbrace{\Big((\frac{1}{2!}+\frac{1}{2!})+\frac{1}{2!}\Big)}_{a_2^{(3)}} \\
&\qquad + \rho^4\frac{1}{4!}t^4 \underbrace{\frac{1}{1!}}_{a_1^{(4)}}\Big) \\
\varphi_{n-5}(t) &= c\frac{n!}{(n-5)!}\Big(-\rho t \underbrace{\frac{1}{5!}}_{a_5^{(1)}} + \rho^2\frac{1}{2}t^2 \underbrace{\Big(\frac{1}{4!}+\frac{1}{2!}\frac{1}{3!}+\frac{1}{3!}\frac{1}{2!}+\frac{1}{4!}\Big)}_{a_4^{(2)}} \\
&\qquad - \rho^3\frac{1}{3!}t^3 \underbrace{\Big((\frac{1}{3!}+\frac{1}{2!2!}+\frac{1}{3!})+(\frac{1}{2!}+\frac{1}{2!})\frac{1}{2!}+\frac{1}{3!}\Big)}_{a_3^{(3)}} \\
&\qquad + \rho^4\,\frac{1}{4!}t^4 \underbrace{\Big((\frac{1}{2!}+\frac{1}{2!}+\frac{1}{2!})+\frac{1}{2!}\Big)}_{a_2^{(4)}} - \rho^5\frac{1}{5!}t^5 \underbrace{\frac{1}{1!}}_{a_1^{(5)}}\Big).
\end{aligned}$$

Remark 5.3.6 (Generalization of trivial martingales)
In this remark we want to show that Theorem 5.3.4 is a generalization of the trivial martingale $(M_t)_{t\in I}$ given by

$$M_s := N_s - \rho \cdot s.$$

Consider the martingale $(X_t)_{t\in I}$ given in Theorem 5.3.4 and set n to be equal to 1. It follows that

$$\begin{aligned}
X_s &= \sum_{i=0}^{1} \frac{1!}{(1-i)!} \sum_{j=0}^{i} (-\rho)^j \frac{1}{j!} s^j a_{i+1-j}^{(j)} N_s^{1-i} \\
&= \frac{1!}{(1-0)!} \sum_{j=0}^{0} (-\rho)^j \frac{1}{j!} s^j a_{0+1-j}^{(j)} N_s^{1-0} + \frac{1!}{(1-1)!} \sum_{j=0}^{1} (-\rho)^j \frac{1}{j!} s^j a_{1+1-j}^{(j)} N_s^{1-1} \\
&= (-\rho)^0 \frac{1}{0!} s^0 a_{0+1-0}^{(0)} N_s^{1-0} + (-\rho)^0 \frac{1}{0!} s^0 a_{1+1-0}^{(0)} N_s^{1-1} + (-\rho)^1 \frac{1}{1!} s^1 a_{1+1-1}^{(1)} N_s^{1-1} \\
&= a_1^{(0)} N_s + a_2^{(0)} - \rho s a_1^{(1)} \\
&= M_s.
\end{aligned}$$

Remark 5.3.7 (Generalization of the second order martingale)
In this remark we want to state that Theorem 5.3.4 is a generalization of the martingale $(M_t)_{t\in I}$ given by

$$M_s := N_s^2 - 2\rho s N_s + \rho^2 s^2 - \rho s. \text{ (Proposition 5.3.1)}$$

Consider the martingale $(X_t)_{t\in I}$ given in Theorem 5.3.4 and set n to be equal to 2. It follows that

$$\begin{aligned}
X_s &= \sum_{i=0}^{2} \frac{2!}{(2-i)!} \sum_{j=0}^{i} (-\rho)^j \frac{1}{j!} s^j a_{i+1-j}^{(j)} N_s^{2-i} \\
&= (-\rho)^0 \frac{1}{0!} s^0 a_{0+1-0}^{(0)} N_s^{2-0} + 2 \sum_{j=0}^{1} (-\rho)^j \frac{1}{j!} s^j a_{1+1-j}^{(j)} N_s^{2-1} + 2 \sum_{j=0}^{2} (-\rho)^j \frac{1}{j!} s^j a_{2+1-j}^{(j)} N_s^{2-2} \\
&= a_1^{(0)} N_s^2 + 2a_2^{(0)} N_s - 2\rho s a_1^{(1)} N_s^1 + 2a_3^{(0)} N_s^0 - 2\rho s a_2^{(1)} + 2\rho^2 \frac{1}{2} s^2 a_1^{(2)} \\
&= N_s^2 - 2\rho s N_s - \rho s + \rho^2 s^2 \\
&= M_s.
\end{aligned}$$

A first application from Theorem 5.3.4 arises if we want to express higher moments of the Poisson distribution in a recursive form.

Lemma 5.3.8 (Moments of the Poisson-distribution)
Let N be a Poisson-distributed random variable with parameter ρ. Then the sequence of numbers $(m_n)_{n\in\mathbb{N}}$ defined recursively by

$$m_0 := 1$$

and

$$m_n := -\sum_{i=1}^{n} \frac{n!}{(n-i)!} \sum_{j=0}^{i} (-\rho)^j \frac{1}{j!} a_{i+1-j}^{(j)} m_{n-i}$$

satisfies for all $n \in \mathbb{N}_0$ the equation

$$\mathbb{E}[N^n] = m_n. \tag{5.9}$$

Proof.

1. Base case: Show that the statement holds for $n = 1$.

   This is true, since

$$\begin{aligned} m_1 &= -\sum_{i=1}^{1} \frac{1!}{(1-i)!} \sum_{j=0}^{i} (-\rho)^j \frac{1}{j!} a_{i+1-j}^{(j)} m_{1-i} \\ &= -\frac{1!}{(1-1)!} \left((-\rho)^0 \frac{1}{0!} a_{1+1-0}^{(0)} m_{1-1} + (-\rho)^1 \frac{1}{1!} a_{1+1-1}^{(1)} m_{1-1} \right) \\ &= -(0-\rho) \\ &= \mathbb{E}[N^1]. \end{aligned}$$

2. Induction step: Show that for any n, if the first $n-1$ moments satisfy equation (4.9), then the equation also holds for the n-th moment.

   Define the process $(X_t)_{t\in I}$ by

$$X_s := \sum_{i=0}^{n} \frac{n!}{(n-i)!} \sum_{j=0}^{i} (-\rho)^j \frac{1}{j!} s^j a_{i+1-j}^{(j)} N_s^{n-i}.$$

Due to Theorem 5.3.4, the process X is a martingale with $X_0 = 0$. Hence, by the martingale property, for $t = 1$ it follows

$$\mathbb{E}\left[\sum_{i=0}^{n} \frac{n!}{(n-i)!} \sum_{j=0}^{i} (-\rho)^j \frac{1}{j!} 1^j a_{i+1-j}^{(j)} N_1^{n-i}\right] = 0. \tag{5.10}$$

Define $N = N_1$. Clearly N is Poisson-distributed with parameter ρ. Rewrite equation (5.10) to

$$\mathbb{E}[N^n] = -\sum_{i=1}^{n} \frac{n!}{(n-i)!} \sum_{j=0}^{i} (-\rho)^j \frac{1}{j!} a_{i+1-j}^{(j)} \mathbb{E}\left[N^{n-i}\right].$$

Since equation (5.9) holds true for the first $n-1$-th moments, the right hand side equals

$$-\sum_{i=1}^{n} \frac{n!}{(n-i)!} \sum_{j=0}^{i} (-\rho)^j \frac{1}{j!} a_{i+1-j}^{(j)} m_{n-i}.$$

□

Next, we aim to express the centered moments of the Poisson distribution recursively in terms of its moments.

Lemma 5.3.9 (Centered moments of the Poisson-distribution)
Let N be a Poisson-distributed random variable with parameter ρ. Define the sequence of numbers $(m_n)_{n\in\mathbb{N}}$ by

$$m_n := \mathbb{E}[N^n].$$

Then the n-th centered moment of N satisfies

$$\mathbb{E}[(N-\rho)^n] = -\sum_{i=1}^{n} \frac{n!}{(n-i)!} \sum_{j=0}^{i-1} (-\rho)^j \frac{1}{j!} a_{i+1-j}^{(j)} m_{n-i}. \tag{5.11}$$

Proof.

Define the process $(X_t)_{t\in I}$ by

$$X_s := \sum_{i=0}^{n} \frac{n!}{(n-i)!} \sum_{j=0}^{i} (-\rho)^j \frac{1}{j!} s^j a_{i+1-j}^{(j)} N_s^{n-i}.$$

Due to Theorem 5.3.4, the process X is a martingale with $X_0 = 0$. Hence by the martingale property, for $t = 1$ it follows that

$$\mathbb{E}\left[\sum_{i=0}^{n} \frac{n!}{(n-i)!} \sum_{j=0}^{i} (-\rho)^j \frac{1}{j!} 1^j a_{i+1-j}^{(j)} N_1^{n-i}\right] = 0. \tag{5.12}$$

Define $N = N_1$. Clearly N is Poisson-distributed with parameter ρ.

It follows that

$$\begin{aligned}
&\sum_{i=0}^{n} \frac{n!}{(n-i)!} \sum_{j=0}^{i} (-\rho)^j \frac{1}{j!} a_{i+1-j}^{(j)} N^{n-i} \\
&= \sum_{i=0}^{n} \frac{n!}{(n-i)!} \left((-\rho)^i \frac{1}{i!} a_1^{(i)} N^{n-i} + \sum_{j=0}^{i-1} (-\rho)^j \frac{1}{j!} a_{i+1-j}^{(j)} N^{n-i} \right) \\
&= \sum_{i=0}^{n} \binom{n}{i} (-\rho)^i N^{n-i} + \sum_{i=1}^{n} \frac{n!}{(n-i)!} \sum_{j=0}^{i-1} (-\rho)^j \frac{1}{j!} a_{i+1-j}^{(j)} N^{n-i} \\
&= (N-\rho)^n + \sum_{i=1}^{n} \frac{n!}{(n-i)!} \sum_{j=0}^{i-1} (-\rho)^j \frac{1}{j!} a_{i+1-j}^{(j)} N^{n-i}.
\end{aligned}$$

Substituting (5.12) into the expression yields:

$$\mathbb{E}\left[(N-\rho)^n\right] = -\sum_{i=1}^{n} \frac{n!}{(n-i)!} \sum_{j=0}^{i-1} (-\rho)^j \frac{1}{j!} a_{i+1-j}^{(j)} m_{n-i}.$$

□

In Example 5.2.4 we introduced the geometric Poisson process. The next theorem states that this process can be represented by Poisson-Charlier martingales.

Theorem 5.3.10 (Representation of the geometric Poisson process) Let I be a closed subset of $\mathbb{R}_+$. Define the sequence $(b_i)_{i\in\mathbb{N}_0}$ by

$$b_k := \frac{c^k}{k!}.$$

Define the sequence $X(1), X(2), \ldots$ of stochastic processes by

$$X_s(n) := \sum_{i=0}^{n} \frac{n!}{(n-i)!} \sum_{j=0}^{i} (-\rho)^j \frac{1}{j!} s^j a_{i+1-j}^{(j)} N_s^{n-i}.$$

It follows that

$$\lim_{n\to\infty} \sup_{s\in I} \left| e^{cN_s-(\rho e^c-\rho)s} - \sum_{k=0}^{n} b_k X_s(k) \right| = 0.$$

Proof.

It holds true that

$$e^{cN_s-(e^c-1)\rho s} = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} \frac{(cN_t)^i}{i!}\frac{(e^c-1)^j(-\rho t)^j}{j!}$$

and that

$$\sum_{n=0}^{\infty} b_n X_s(n) = \sum_{n=0}^{\infty} \frac{c^n}{n!} \sum_{i=0}^{n}\sum_{j=0}^{i} \frac{n!}{(n-i)!}(-\rho)^j \frac{1}{j!} t^j a_{i+1-j}^{(j)} N_t^{n-i}.$$

It suffices to show that the coefficients related to $-N_t^A(\rho t)^B$ match for all A and B. Specifically,

$$-N_t^A(\rho t)^B \frac{c^A}{A!}\frac{(e^c-1)^B}{B!} = \sum_{n=0}^{\infty}\sum_{i=0}^{n}\sum_{j=0}^{i} 1_{j=B,n-i=A} \frac{c^n}{n!}\frac{n!}{(n-i)!}(-\rho)^j \frac{1}{j!} t^j a_{i+1-j}^{(j)} N_t^{n-i}.$$

Once this is shown, the proof is complete. It follows that

$$\begin{aligned}
&\sum_{n=0}^{\infty}\sum_{i=0}^{n}\sum_{j=0}^{i} 1_{\{j=B,n-i=A\}} \frac{c^n}{n!}\frac{n!}{(n-i)!}(-\rho)^j \frac{1}{j!} t^j a_{i+1-j}^{(j)} N_t^{n-i} \\
&= \sum_{n=B}^{\infty}\sum_{i=B}^{n} 1_{\{n-i=A\}} c^n \frac{1}{(n-i)!}(-\rho t)^B \frac{1}{B!} a_{i+1-B}^{(B)} N_s^{n-i} \\
&= \sum_{n=A}^{\infty}\sum_{i=0}^{n} 1_{\{i=n-A\}} c^{A+i+B} \frac{1}{A!}(-\rho t)^B \frac{1}{B!} a_{i+1}^{(B)} N_s^{A} \\
&= \sum_{n=A}^{\infty} c^{n+B} \frac{1}{A!}(-\rho t)^B \frac{1}{B!} a_{n-A+1}^{(B)} N_s^{A} \\
&= \sum_{n=0}^{\infty} c^{n+A+B} \frac{1}{A!}(-\rho t)^B \frac{1}{B!} a_{n+1}^{(B)} N_s^{A} \\
&= c^{A+B} \frac{1}{A!}(-\rho t)^B \frac{1}{B!} N_s^{A} \sum_{n=0}^{\infty} c^n a_{n+1}^{(B)} \\
&= -N_t^A(\rho t)^B \frac{c^A}{A!}\frac{(e^c-1)^B}{B!}.
\end{aligned}$$

□

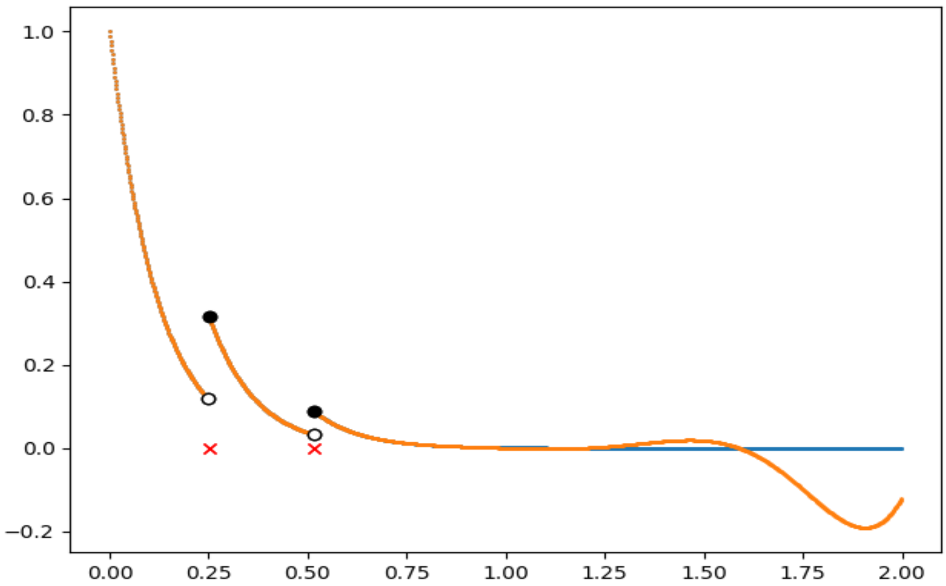

Figure 5.3: Realization of a Poisson process with parameter $\rho = 2$ (red crosses), the associated geometric Poisson process (blue) and the approximation with the martingales $X(1), X(2), ..., X(20)$.

5.4 Generalized Poisson-Charlier Martingales

The aim of this section is to extend the Poisson-Charlier martingales to a more general form. In addition to the potential application in terms of PCA, another motivation could be as follows. Let $N = (N_t)_{t\in I}$ be a Poisson process with parameter ρ and $\mathcal{H}$ be the filtration generated by N. Define for some arbitrary $i \in \mathbb{N}$ the point process $N(i) = (N_t(i))_{t\in I}$ by

$$N_s(i) := 1_{\{N_s>i\}}.$$

Clearly $N(i)$ is a transformation of the process N resulting again in a point process and the conditional intensity $\lambda^i(t \mid \mathcal{H}_t)$ of $N(i)$ is determined as

$$\lambda^i(s \mid \mathcal{H}_s) = 1_{\{N_s=i\}} \cdot \rho.$$

It is important to observe that $\mathcal{H}$ remains the filtration generated by the original Poisson process N. Clearly, the process $X = (X_t)_{t\in I}$ defined by

$$X_s := N_s(i) - \int_{s_0}^{s} \lambda^i(v \mid \mathcal{H}_v)dv$$

is a martingale with respect to $\mathcal{H}$. Also X is a martingale transformation of the underlying Poisson process N. Simultaneously, X is also a martingale transformation of $N(1)$, a point process with dependent increments and a stochastic conditional intensity with respect to $\mathcal{H}$.

Examination of the compensator shows that

$$\begin{aligned} \int_{s_0}^{s} \lambda^i(v \mid \mathcal{H}_v)dv &= \rho \int_{s_0}^{s} 1_{\{N_v=k\}}dv \\ &= \rho \int_{T_{k-1}}^{s\wedge T_k} 1_{\{T_{k-1}<s\}}dv \\ &= \rho 1_{\{T_{k-1}<s\}}(s \wedge T_k - T_{k-1}). \end{aligned}$$

Thus, X is not a transformation of N with a representation $X_s = \varphi(N_s, s)$, but is instead given by

$$X_s = \varphi(N_s, s, T_{i-1}, T_i) = 1_{\{N_s>i\}} - \rho 1_{\{T_{i-1}<s\}}(s \wedge T_i - T_{i-1}).$$

Therefore, the process $N(i)$ resulting from the plug-in mechanism described above generates a path-dependent martingale transformation of the original point process N. The introduction of a potential generalized Charlier martingale transformation for the point process $N(i)$ has the potential to lead to further path-dependent martingale transformations of the Poisson process N.

Let $\mathcal{C}$ be the space where λ takes its values. For notational purposes we define for all $j \in \mathbb{N}$ the function $I^n : \mathcal{C} \times I \to \mathbb{R}$ by

$$I^j(\lambda, s) = \int_{s_0}^{s} \lambda_{v_j} \int_{s_0}^{v_j} \cdots \int_{s_0}^{v_3} \lambda_{v_2} \int_{s_0}^{v_2} \lambda_{v_1} dv_1 dv_2 \cdots dv_j.$$

Note that for all $s \in I$ the random variable $I^n(\lambda, s)$ is measurable with respect to the σ-Algebra $\mathcal{H}_s$, $I^0(\lambda, s) = 1$ and $\lambda_s I^j(\lambda, s) ds = dI^{j+1}(\lambda, s)$.

Theorem 5.4.1 (Polynomial martingales)
Let N be a point process with conditional intensity $\lambda = (\lambda(t \mid \mathcal{H}_t))_{t \in I}$. Then the process X defined by

$$X_s := \sum_{i=0}^{n} \frac{n!}{(n-i)!} \sum_{j=0}^{i} (-1)^j I^j(\lambda, s) a_{i+1-j}^{(j)} N_s^{n-i}$$

is a martingale with respect to $\mathcal{H}$.

Proof.

Assume that φ is a polynomial function of n-th order, given by

$$\varphi(x, t, \lambda) = \sum_{k=0}^{n} \varphi_k(t, \lambda) x^k.$$

Consequently, it follows that

$$\frac{\partial \varphi}{\partial t}(x, t, \lambda) = \sum_{k=0}^{n} \frac{\partial \varphi_k}{\partial t}(t, \lambda) x^k.$$

According to Lemma 5.1.3, the function φ needs to satisfy the equation

$$\frac{\partial \varphi}{\partial t}(N_t, t, \lambda) + \Big(\varphi(1 + N_t, t, \lambda) - \varphi(N_t, t, \lambda)\Big)\lambda_t = 0. \tag{5.13}$$

Through manipulation, we can transform the left-hand side of (5.13) to

$$\begin{aligned}
&\frac{\partial \varphi}{\partial t}(N_t, t, \lambda) + \Big(\varphi(1 + N_t, t, \lambda) - \varphi(N_t, t, \lambda)\Big)\lambda_t \\
&= \sum_{k=0}^{n} \frac{\partial \varphi_k}{\partial t}(t, \lambda)x^k + \left(\sum_{k=0}^{n} \varphi_k(t, \lambda)(1 + x)^k - \sum_{k=0}^{n} \varphi_k(t)x^k\right)\lambda_t \\
&= \sum_{k=0}^{n} \frac{\partial \varphi_k}{\partial t}(t, \lambda)x^k + \Big(\sum_{k=0}^{n} \varphi_k(t, \lambda)\sum_{j=0}^{k}\binom{k}{j}x^j - \sum_{k=0}^{n} \varphi_k(t)x^k\Big)\lambda_t \\
&= \sum_{k=0}^{n} \frac{\partial \varphi_k}{\partial t}(t, \lambda)x^k + \Big(\sum_{k=0}^{n} \varphi_k(t, \lambda)\sum_{j=0}^{k-1}\binom{k}{j}x^j\Big)\lambda_t \\
&= \sum_{k=0}^{n} \frac{\partial \varphi_k}{\partial t}(t, \lambda)x^k + \left(\sum_{k=0}^{n-1}\sum_{j=k+1}^{n} \varphi_j(t, \lambda)\binom{j}{k}x^k\right)\lambda_t \\
&= \sum_{k=0}^{n-1} x^k \left(\frac{\partial \varphi_k}{\partial t}(t, \lambda) + \lambda_t \sum_{j=k+1}^{n} \varphi_j(t, \lambda)\binom{j}{k}\right) + \frac{\partial \varphi_n}{\partial t}(t, \lambda).
\end{aligned}$$

From this, we can infer that (5.13) is tantamount to

$$\sum_{k=0}^{n-1} x^k \left(\frac{\partial \varphi_k}{\partial t}(t, \lambda) + \lambda_t \sum_{j=k+1}^{n} \varphi_j(t, \lambda)\binom{j}{k}\right) + \frac{\partial \varphi_n}{\partial t}(t, \lambda) = 0. \qquad (5.14)$$

The equation is fulfilled for all $x \in \mathbb{N}$ if and only if all coefficients are equal to zero.

Assertion: The equation is fulfilled if for all $1 \leq i \leq n$ the function φ_{n-i} has the representation

$$\varphi_{n-i}(t) = \frac{n!}{(n-i)!}\sum_{j=0}^{i}(-1)^j I^j(\lambda, s)a^{(j)}_{i+1-j}.$$

1. Base case: Show that the statement holds true for $i = 0$.

   We are setting c to be equal to 1. On the one hand (5.14) implies that

   $$\varphi_{n-0}(t) = c = 1.$$

On the other hand we have

$$\frac{n!}{(n-i)!}\sum_{j=0}^{0}(-1)^j I^j(\lambda,s)a^{(j)}_{0+1-j} = 1.$$

2. Induction step: Show that for any k, if the represenation holds true for $\varphi_n, \varphi_{n-1}, ..., \varphi_{n-i+1}$ then it follows that the representation also holds true for φ_{n-i}.

   It follows that

$$\begin{aligned}
\frac{\partial \varphi_{n-i}}{\partial t}(t) &= -\lambda_t \sum_{j=n-(i-1)}^{n} \varphi_j(t)\binom{j}{n-i} \\
&= -\lambda_t \sum_{j=0}^{i-1} \varphi_{n-j}(t)\binom{n-j}{n-i} \\
&= -\lambda_t \sum_{j=0}^{i-1}\left(\frac{n!}{(n-j)!}\sum_{k=0}^{j}(-1)^k I^k(\lambda,t)a^{(k)}_{j+1-k}\right)\frac{(n-j)!}{(n-i)!(i-j)!} \\
&= -\lambda_t \sum_{j=0}^{i-1}\left(\frac{n!}{(n-i)!}\sum_{k=0}^{j}(-1)^k I^k(\lambda,t)a^{(k)}_{j+1-k}\frac{1}{(i-j)!}\right) \\
&= \frac{n!}{(n-i)!}\sum_{j=0}^{i-1}\left(\sum_{k=0}^{j}(-1)^{k+1}\lambda_t I^k(\lambda,t)a^{(k)}_{j+1-k}\frac{1}{(i-j)!}\right).
\end{aligned}$$

Consequently, it holds true that

$$\begin{aligned}
\varphi_{n-i}(t) &= \frac{n!}{(n-i)!}\sum_{j=0}^{i-1}\Big(\sum_{k=1}^{j}(-1)^{k+1}I^{k+1}(\lambda,t)a^{(k)}_{j+1-k}\frac{1}{(i-j)!}\Big)\\
&= \frac{n!}{(n-i)!}\sum_{j=0}^{i-1}\Big(\sum_{k=1}^{j}(-1)^{k+1}I^{k+1}(\lambda,t)a^{(k)}_{j+1-k}\frac{1}{(i-j)!}\Big)\\
&= \frac{n!}{(n-i)!}\sum_{k=0}^{i-1}(-1)^{k+1}I^{k+1}(\lambda,t)\sum_{j=k}^{i-1}a^{(k)}_{j+1-k}\frac{1}{(i-j)!}\\
&= \frac{n!}{(n-i)!}\sum_{k=0}^{i-1}(-1)^{k+1}I^{k+1}(\lambda,t)\sum_{j=1}^{i-1-(k-1)}a^{(k)}_{j}\frac{1}{(i-j+1-k)!}\\
&= \frac{n!}{(n-i)!}\sum_{k=0}^{i-1}(-1)^{k+1}I^{k+1}(\lambda,t)\sum_{j=1}^{i-k}a^{(k)}_{j}\frac{1}{(i-k+1-j)!}\\
&= \frac{n!}{(n-i)!}\sum_{k=0}^{i}(-1)^{k}I^{k}(\lambda,t)a^{(k)}_{i-k+1}.
\end{aligned}$$

The assertion thus follows.

Substitution of the representation into φ yields

$$\varphi(x,t) = \sum_{i=0}^{n}\varphi_{n-i}(t)x^{n-i} = \sum_{i=0}^{n}\frac{n!}{(n-i)!}\sum_{j=0}^{i}(-1)^{j}I^{j}(\lambda,s)a^{(j)}_{i+1-j}x^{n-i}.$$

□

Remark 5.4.2 (Generalization of the trivial martingale)
In this remark we want to state that Theorem 5.3.4 is a generalization of the trivial martingal $(M_t)_{t\in I}$ given by

$$M_s := N_s - \int_{s_0}^{s}\lambda_v dv.$$

Consider the martingale $(X_t)_{t\in I}$ as presented in Theorem 5.3.4, and set

$n = 1$. It follows that

$$
\begin{aligned}
X_s &= \sum_{i=0}^{1} \frac{1!}{(1-i)!} \sum_{j=0}^{i} (-1)^j I^j(\lambda, s) a_{i+1-j}^{(j)} N_s^{1-i} \\
&= \frac{1!}{(1-0)!} \sum_{j=0}^{0} (-1)^j I^j(\lambda, s) a_{0+1-j}^{(j)} N_s^{1-0} + \frac{1!}{(1-1)!} \sum_{j=0}^{1} (-1)^j I^j(\lambda, s) a_{1+1-j}^{(j)} N_s^{1-1} \\
&= (-1)^0 I^0(\lambda, s) a_{0+1-0}^{(0)} N_s^{1-0} + (-1)^0 I^0(\lambda, s) a_{1+1-0}^{(0)} N_s^{1-1} + (-1)^1 I^1(\lambda, s) a_{1+1-1}^{(1)} N_s^{1-1} \\
&= a_1^{(0)} N_s + a_2^{(0)} - I^1(\lambda, s) a_1^{(1)} \\
&= N_s - \int_{s_0}^{s} \lambda_v dv \\
&= M_s.
\end{aligned}
$$

Remark 5.4.3 (Generalization of the second order martingale)
In this remark we want to state that Theorem 5.4.1 is a generalization of the martingal $(M_t)_{t \in I}$ given by

$$
M_s := \left(N_s - \int_{s_0}^{s} \lambda_v dv \right)^2 - \int_{s_0}^{s} \lambda_v dv.
$$

Consider the martingale $(X_t)_{t \in I}$ as presented in Theorem 5.4.1, and set $n = 2$. It follows that

$$
\begin{aligned}
X_s &= \sum_{i=0}^{2} \frac{2!}{(2-i)!} \sum_{j=0}^{i} (-1)^j \frac{1}{j!} s^j a_{i+1-j}^{(j)} N_s^{2-i} \\
&= a_1^{(0)} N_s^2 + 2a_2^{(0)} N_s - 2I^1(\lambda, s) a_1^{(1)} N_s^1 + 2a_3^{(0)} N_s^0 - 2I^1(\lambda, s) a_2^{(1)} + 2I^2(\lambda, s) a_1^{(2)} \\
&= N_s^2 - 2 \int_{s_0}^{s} \lambda_v dv N_s - 2 \int_{s_0}^{s} \lambda_v dv \frac{1}{2!} + 2 \int_{s_0}^{s} \lambda_{v_1} \int_{s_0}^{v_1} \lambda_v dv dv_1.
\end{aligned}
$$

By applying Fubini's Theorem and substituting, we obtain the result that

$$
\begin{aligned}
2 \int_{s_0}^{s} \lambda_{v_1} \int_{s_0}^{v_1} \lambda_v dv dv_1 &= \int_{s_0}^{s} \int_{s_0}^{v_1} \lambda_{v_1} \lambda_v dv dv_1 + \int_{s_0}^{s} \int_{v_1}^{s} \lambda_{v_1} \lambda_v dv_1 dv \\
&= \int_{s_0}^{s} \int_{s_0}^{z} \lambda_z \lambda_v dv dz + \int_{s_0}^{s} \int_{z}^{s} \lambda_v \lambda_z dv dz \\
&= \int_{s_0}^{s} \lambda_z \left(\int_{s_0}^{z} \lambda_v dv + \int_{z}^{s} \lambda_v dv \right) dz \\
&= \left(\int_{s_0}^{s} \lambda_v dv \right)^2.
\end{aligned}
$$

Consequently, it follows that

$$\begin{aligned} M_s &= N_s^2 - 2\int_{s_0}^{s} \lambda_v dv N_s - \int_{s_0}^{s} \lambda_v dv + 2\int_{s_0}^{s} \lambda_{v_1} \int_{s_0}^{v_1} \lambda_v dv dv_1 \\ &= N_s^2 - 2\int_{s_0}^{s} \lambda_v dv N_s - \int_{s_0}^{s} \lambda_v dv + \left(\int_{s_0}^{s} \lambda_v dv\right)^2 \\ &= X_s. \end{aligned}$$

Moving forward, we will prove the generalized identity

$$\begin{aligned} &(N_s - \int_{s_0}^{s} \lambda_v dv)^n - \sum_{i=1}^{n} \frac{n!}{(n-i)!} \sum_{j=0}^{i-1} (-\rho)^j \frac{1}{j!} a_{i+1-j}^{(j)} N^{n-i} \\ &= \sum_{i=0}^{n} \frac{n!}{(n-i)!} \sum_{j=0}^{i} (-1)^j I^j(\lambda, s) a_{i+1-j}^{(j)} N_s^{n-i} = X_s. \end{aligned}$$

This was examined for the case $n = 1, 2$ in Remark 5.4.2 and 5.4.3.

Lemma 5.4.4 (Represent the n-th moment process)
Consider a point process N with conditional intensity λ. For any $n \in \mathbb{N}$ and $s \in I$, we have the identity

$$\begin{aligned} &(N_s - \int_{s_0}^{s} \lambda_v dv)^n - \sum_{i=1}^{n} \frac{n!}{(n-i)!} \sum_{j=0}^{i-1} (-\rho)^j \frac{1}{j!} a_{i+1-j}^{(j)} N^{n-i} \\ &= \sum_{i=0}^{n} \frac{n!}{(n-i)!} \sum_{j=0}^{i} (-1)^j I^j(\lambda, s) a_{i+1-j}^{(j)} N_s^{n-i} = X_s. \end{aligned}$$

Proof.

Define the process $(X_t)_{t\in I}$ by

$$X_s := \sum_{i=0}^{n} \frac{n!}{(n-i)!} \sum_{j=0}^{i} (-1)^j I^j(\lambda, s) a_{i+1-j}^{(j)} N_s^{n-i}.$$

It follows that

$$\begin{aligned} &\sum_{i=0}^{n} \frac{n!}{(n-i)!} \sum_{j=0}^{i} (-1)^j I^j(\lambda, s) a_{i+1-j}^{(j)} N^{n-i} \\ &= \sum_{i=0}^{n} \frac{n!}{(n-i)!} \left((-1)^i I^i(\lambda, s) a_1^{(i)} N^{n-i} + \sum_{j=0}^{i-1} (-1)^j I^j(\lambda, s) a_{i+1-j}^{(j)} N^{n-i} \right). \end{aligned}$$

The assertion thus follows if

$$\sum_{i=0}^{n} \frac{n!}{(n-i)!}(-1)^i I^i(\lambda, s) a_1^{(i)} N^{n-i} = \left(N_s - \int_{s_0}^{s} \lambda_v dv\right)^n .$$

Assume that

$$(i!) I^i(\lambda, s) = \left(\int_{s_0}^{s} \lambda_v dv\right)^i . \tag{5.15}$$

Then the left hand side equals

$$\begin{aligned} &\sum_{i=0}^{n} \frac{n!}{(n-i)!}(-1)^i I^i(\lambda, s) a_1^{(i)} N^{n-i} \\ &= \sum_{i=0}^{n} \binom{n}{i} (-1)^i (i!) I^i(\lambda, s) 1 N^{n-i} \\ &= \sum_{i=0}^{n} \binom{n}{i} \left(-\int_{s_0}^{s} \lambda_v dv\right)^i N^{n-i} \\ &\left(N - \int_{s_0}^{s} \lambda_v dv\right)^n . \end{aligned}$$

Thus it remains to prove assertion (5.15): For all $i \in \mathbb{N}$ it holds true that

$$(i!) I^i(\lambda, s) = \left(\int_{s_0}^{s} \lambda_v dv\right)^i .$$

1. Base case: For $i = 1$ it holds true that

$$(i!) I^i(\lambda, s) = 1! \int_{s_0}^{s} \lambda_v dv = \left(\int_{s_0}^{s} \lambda_v dv\right)^1 .$$

   Induction step: Assume that the assertion holds true for all $k =$

$1, 2, ..., i-1$. It follows that

$$
\begin{aligned}
(i!)I^i(\lambda, s) &= i! \cdot \int_{s_0}^{s} \lambda_v I^{i-1}(\lambda, v) dv \\
&= i \int_{s_0}^{s} \lambda_v (i-1)! I^{i-1}(\lambda, v) dv \\
&= i \int_{s_0}^{s} \lambda_v \left(\int_{s_0}^{v} \lambda_w dw \right)^{i-1} dv \\
&= i \left[\int_{s_0}^{v} \lambda_w dw \cdot \left(\int_{s_0}^{v} \lambda_w dw \right)^{i-1} \right]_{s_0}^{s} \\
&\qquad - i \int_{s_0}^{s} \int_{s_0}^{v} \lambda_w dw \cdot (i-1) \left(\int_{s_0}^{v} \lambda_w dw \right)^{i-2} \lambda_v dv \\
&= i \left(\int_{s_0}^{s} \lambda_w dw \right)^{i} - i(i-1) \int_{s_0}^{s} \left(\int_{s_0}^{v} \lambda_w dw \right)^{i-1} \lambda_v dv \\
&= i \left(\int_{s_0}^{s} \lambda_w dw \right)^{i} - i(i-1)(i-1)! \int_{s_0}^{s} I^{i-1}(\lambda, v) \lambda_v dv \\
&= i \left(\int_{s_0}^{s} \lambda_w dw \right)^{i} - i(i-1)(i-1)! I^i(\lambda, s) \\
&= i \left(\int_{s_0}^{s} \lambda_w dw \right)^{i} - (i-1) i! I^i(\lambda, s).
\end{aligned}
$$

□

By invoking Lemma 5.4.4, we can establish that

$$
\mathbb{E}\left[\left(N_s - \int_{s_0}^{s} \lambda_v dv \right)^n \right] = - \sum_{i=1}^{n} \frac{n!}{(n-i)!} \sum_{j=0}^{i-1} (-1)^j I^j(\lambda, s) a_{i+1-j}^{(j)} \mathbb{E}[N_s^{n-i}].
$$

Lemmas 5.3.8 and 5.3.9 were instrumental in enabling the representation of moments in the Poisson case. The equation

$$
\mathbb{E}\left[\left(N_s - \int_{s_0}^{s} \lambda_v dv \right)^2 \right] = \mathbb{E}\left[\int_{s_0}^{s} \lambda_v dv \right]
$$

will be of significant importance in analyzing the principal components of a specific class of point processes.

5.5 PCA of externally excited point processes

Let I be the interval $[0,1)$. The Hilbert-space L^2 defined by

$$L^2 := \left\{ f : I \mid \int_I f_s^2 ds < \infty, f \text{ is Lebesgue-measurable} \right\},$$

admits an inner product,

$$\langle f, g \rangle := \int_I f_s g_s ds$$

and a naturally induced norm associated with the inner product,

$$||f||_2 := \sqrt{\langle f, f \rangle}.$$

A vector f belongs to the span of a sequence of vectors $(f(i))_{i\in\mathbb{N}}$ if

$$\lim_{n\to\infty} ||f - \sum_{i=1}^{n} a_i f(i)||_2 = 0,$$

for some constants a_i. In this aforementioned situation, we simply write

$$f = \sum_{i=1}^{\infty} a_i f(i).$$

If a vector f is in the span of a sequence $(u(i))_{i\in\mathbb{N}}$ of orthonormal vectors, that is

$$\langle u(i), u(j) \rangle = 1_{\{i=j\}},$$

then it holds true that

$$f = \sum_{i=1}^{\infty} \langle f, u(i) \rangle f(i).$$

Let $X = (X_t)_{t\in I}$ be a stochastic process such that for all $\omega \in \Omega$, the path $X(\omega)$ belongs to L^2, which means that

$$\int_I X_s(\omega)^2 \mathrm{d}s < \infty.$$

Assuming that for all $\omega \in \Omega$ the process X is in the span of an orthonormal (deterministic) basis $(u(i))_{i\in I}$, yields

$$X(\omega) = \sum_{i=1}^{\infty} \langle X(\omega), u(i)\rangle u(i).$$

Thus, the stochastic process X can be represented by the sequence of random variables $(\langle X, u(i)\rangle)_{i\in\mathbb{N}}$. Let K be the covariance-kernel of the underlying process X. If the system $(u_i)_{i\in\mathbb{N}}$ are eigenvectors of the Fredholm-Integraloperator, that is

$$\int_I K(s,t)u_t(i)dt = c_i u_s(i), \tag{5.16}$$

for some constant c_i and all $s \in I$, then the random variables $(\langle X, u(i)\rangle)_{i\in\mathbb{N}}$ are called the principal components of X. The normed eigenvectors are always arranged in descending order with respect to their associated eigenvalues. It can be shown that a system defined by the equation (5.16) forms an orthonormal system. The principal components associated with a process have a number of interesting properties. They are centered and the variance of the i-th principal component is given by the eigenvalue c_i. Moreover, the components are uncorrelated. If X is a Brownian motion, the components are even stochastically independent, whereas the components of the centered Poisson process show dependencies.

Remark 5.5.1 (Principal Components of the Brownian motion)
Let X be a Brownian motion on I. Per definition the Covariance-kernel K of X is given by

$$K(s,t) = \min(s,t).$$

It is well known that the system $(u(i))_{i\in\mathbb{N}}$ defined by

$$u_t(i) := \sqrt{2}\sin\left(\frac{(2i-1)\pi t}{2}\right) \tag{5.17}$$

and the sequence

$$c_i := \frac{1}{\pi^2(i+\frac{1}{2})^2}, \tag{5.18}$$

solve the equation

$$\int_I \min(s,t)u_t(i)dt = c_i u_s(i). \tag{5.19}$$

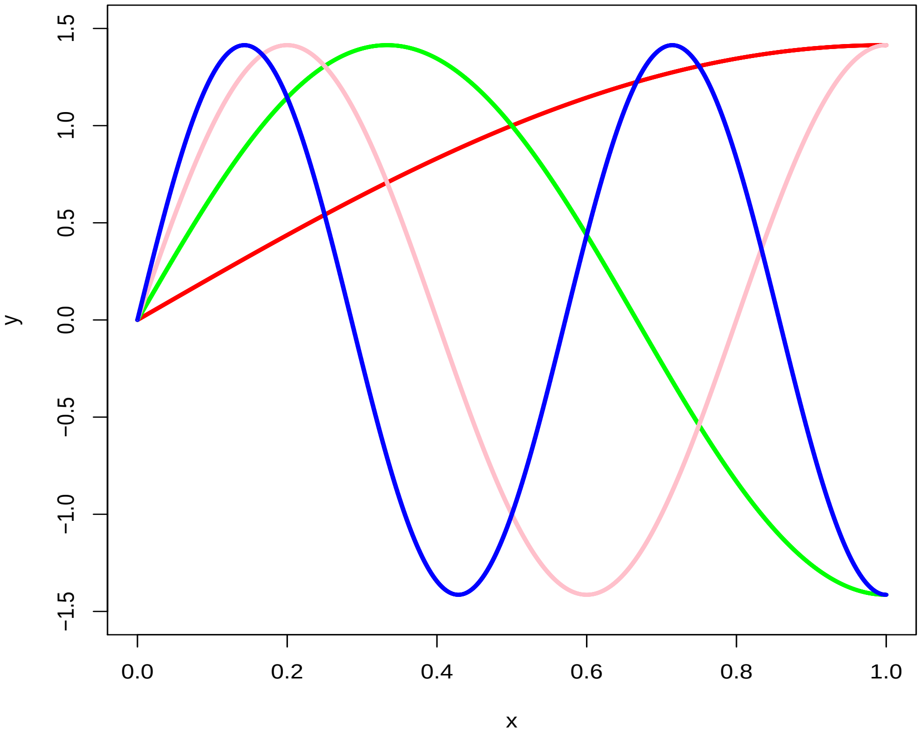

Figure 5.4: The first four eigenfunctions u_i (5.17) with respect to the operator associated with Brownian motion.

Hence with our previous discussions it follows that

$$X = \sum_{i=1}^{\infty} \langle X, u_i \rangle u_i.$$

Since the random variable

$$\langle X, u_i \rangle = \int_I X_s u_i(s) ds = \lim_{n \to \infty} \sum_{k=1}^{n} \frac{1}{n} X_{k/n} u_i(k/n)$$

is the limit of normally distributed random variables, it follows that $\langle X, u_i \rangle$ is again normal with

$$\mu = \mathbb{E}[\langle X, u_i \rangle] = 0 \text{ and } \sigma^2 = c_i.$$

Principal components theory also ensures that the random variables $(\langle X, u(i)\rangle)_{i\in\mathbb{N}}$ are uncorrelated. It should be noted that if normally distributed random variables are uncorrelated, this implies that they are independent.

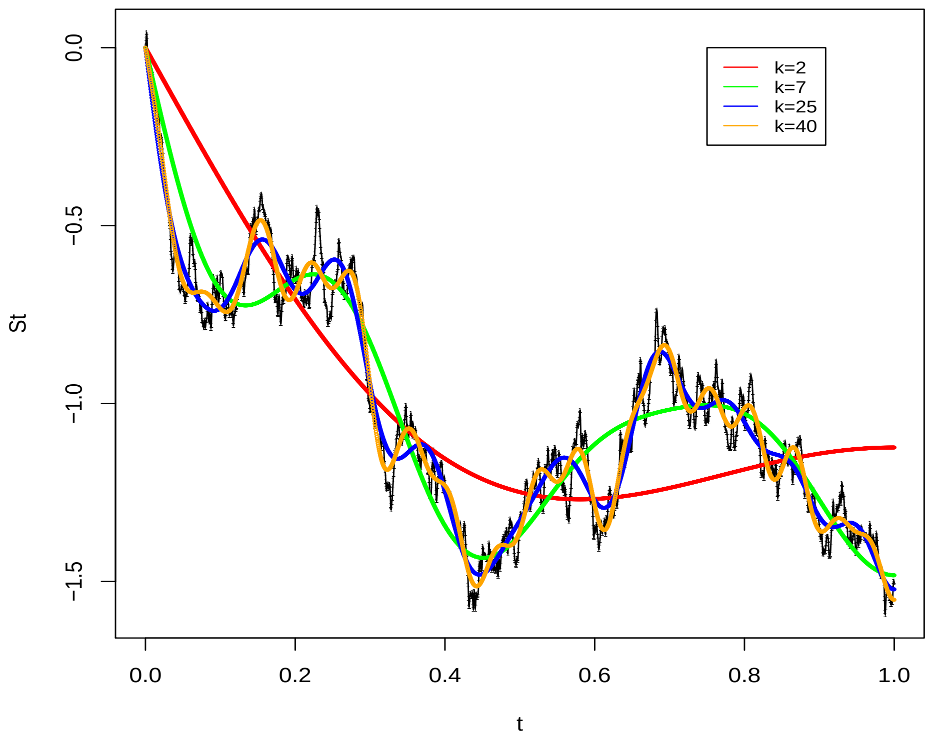

Figure 5.5: Cut the series representing a Brownian motion.

Remark 5.5.2 (Principal Components of the Poisson process)
Given a Poisson process N on the set $I = [0, 1]$ with parameter ρ, we define the process X as $X_s := \frac{N_s - \rho s}{\sqrt{\rho}}$ for all $s \in I$. For $s < t$, Remark 5.4.3 implies that the identity

$$\mathrm{Cov}(X_s, X_t) = \mathrm{Var}(X_s) = \frac{1}{\rho}\mathbb{E}[(N_s - \rho s)^2] = \frac{1}{\rho}\rho s = s$$

holds true. Consequently we need to solve the eigenvalue problem with respect to the kernel

$$K(s,t) = \min(s,t).$$

At this point, one recognizes an important commonality with Brownian motion. They have the same covariance kernel and therefore they also have the same eigenfunctions (5.17) and eigenvalues (5.18) with respect to the corresponding Fredholm operator. Some calculations yield the explicit representation

$$\langle X, u(i) \rangle = \frac{\sqrt{2}}{\sqrt{\rho}} \left[\frac{2}{(2i-1)\pi} \sum_{i=1}^{N_1} \cos\left(\frac{(2i-1)\pi T_i}{2} \right) + (-1)^i \frac{2}{(2i-1)\pi} \rho \right].$$

Like in the case of the Brownian motion, the components are centered, uncorrelated and have variance equal to the corresponding eigenvalue. A distinction becomes apparent upon analyzing the marginal distributions and dependency structures. The Fourier transformation of the i-th component equals

$$\mathbb{E}\left[e^{iz\langle X,u(i)\rangle}\right] = e^{iz(-1)^i\left[\frac{2}{(2j-1)\pi}\right]^2\sqrt{2\rho}} e^{\rho\left[J_0\left(z\frac{2\sqrt{2}}{(2i-1)\pi\sqrt{\rho}}\right)-1+i\frac{(-1)^{i-1}}{2i-1}H_0\left(z\frac{2\sqrt{2}}{(2i-1)\pi\sqrt{\rho}}\right)\right]},$$

where J_0 is the Bessel function of 0-th order

$$J_0(z) := \sum_{k=0}^{\infty} \frac{(-1)^k}{[2^k k!]^2} z^{2k}$$

and H_0 is the Struve function

$$H_0(z) := \frac{2}{\pi} \sum_{k=0}^{\infty} \frac{(-1)^k}{[(2k+1)!!]^2} z^{2k+1}.$$

Someone can rigorously prove that the Fourier transformation of the components are not integrable over $\mathbb{R}$. The lack of integrability should not come as a surprise, given that each component has an atom in its distribution. Specifically, we have that

$$\mathbb{P}\left[\langle X, u(i) \rangle = \sqrt{\rho} \frac{4\sqrt{2}}{(2i-1)^2\pi^2} (-1)^i \right] = e^{-\rho}.$$

Taking advantage of this knowledge, one can further prove that the components are dependent

$$\mathbb{P}[A^i, A^j] \neq \mathbb{P}[A^i]\mathbb{P}[A^j],$$

by defining

$$A^k := \left\{ \omega \in \Omega : \langle X, u(k) \rangle = \sqrt{\rho} \frac{4\sqrt{2}}{(2k-1)^2\pi^2}(-1)^k \right\}.$$

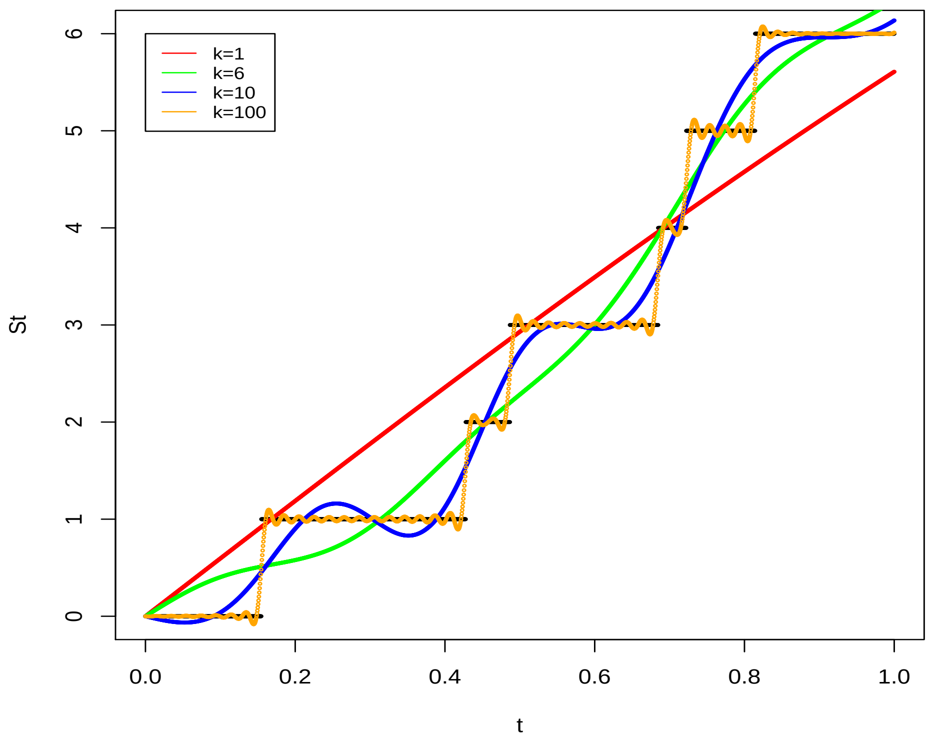

Figure 5.6: Cut the series representing a Poisson process.

Let nF_n be a renormalized empirical distribution function. That is nF_n is a point process with internal conditional intensity

$$\lambda(s \mid \mathcal{H}_s) = \frac{(n - nF_n(s-))1_{\{0 \leq s < 1\}}}{1 - s}.$$

We consider the process $\tilde{X}$ defined by

$$\tilde{X}_s := \frac{1}{\sqrt{n}}(N_s - ns).$$

It holds true that

$$K(s,t) = \min(s,t) - st.$$

Of course, $\tilde{X}$ in distribution converges to a Brownian bridge with Donsker's theorem. This case has also been treated in detail in terms of principal component analysis.

The approach to transform a process in such a way that a kernel like in the case of a Brownian bridge arises is not aimed at here. Within this section we set the goal to transform certain point processes in such a way that the transformed process has a covariance kernel K with $K(s,t) = \min(s,t)$. The following proposition utilizes the general Poisson-Charlier martingale of second order that was derived earlier.

Proposition 5.5.3 (Covarince kernel of centered point processes)
Let N be a point process with conditional intensity λ. Define the process X by

$$X_s := N_s - \int_{s_0}^{s} \lambda_v dv.$$

Then, X is a martingale starting in $X_0 = 0$ and it follows that

$$\mathrm{Cov}(X_s, X_t) = \min\left(\mathbb{E}\left[\int_0^s \lambda_v dv\right], \mathbb{E}\left[\int_0^t \lambda_v dv\right]\right). \tag{5.20}$$

Proof.

The process X is a martingale by definition. Let s, t be arbitrary with $s < t$. It follows that

$$\begin{aligned}\mathrm{Cov}(X_s, X_t) &= \mathbb{E}\left[X_s X_t\right] \\ &= \mathbb{E}\left[X_s \mathbb{E}[X_t \mid \mathcal{H}_s]\right] \\ &= \mathbb{E}\left[X_s^2\right] \\ &= \mathbb{E}\left[\left(N_s - \int_{s_0}^{s} \lambda_v dv\right)^2\right].\end{aligned}$$

Setting $n = 2$ in Theorem 5.4.1 yields the martingale M given by

$$M_s := \sum_{i=0}^{2} \frac{2!}{(2-i)!} \sum_{j=0}^{i} (-1)^j \frac{1}{j!} s^j a_{i+1-j}^{(j)} N_s^{2-i}.$$

Remark 5.4.3 states that M is equal to

$$M_s = \left(N_s - \int_0^s \lambda_v dv \right)^2 - \int_0^s \lambda_v dv.$$

Since M is a martingale with $M_0 = 0$, we have that

$$\mathbb{E}\left[\left(N_s - \int_0^s \lambda_v dv \right)^2 \right] = \mathbb{E}\left[\int_0^s \lambda_v dv \right].$$

Thus, it follows that

$$\begin{aligned} \text{Cov}(X_s, X_t) &= \min\left(\mathbb{E}\left[\left(N_s - \int_0^s \lambda_v dv \right)^2 \right], \mathbb{E}\left[\left(N_t - \int_0^t \lambda_v dv \right)^2 \right] \right) \\ &= \min\left(\mathbb{E}\left[\int_0^s \lambda_v dv \right], \mathbb{E}\left[\int_0^t \lambda_v dv \right] \right). \end{aligned}$$

□

As a direct consequence we get the following Theorem, specifying under which conditions the compensated point process $X = N - \lambda$, can be represented with the same eigenfunctions as the Brownian motion in the Karhunen-Loève decomposition.

Proposition 5.5.4 (Point processes with minimum type kernel)
Let N be a point process with conditional intensity λ. Define the process X by

$$X_s := N_s - \int_0^s \lambda_v dv.$$

If the stochastic process λ satisfies the equation

$$\mathbb{E}\left[\int_0^s \lambda_v dv \right] = s,$$

it follows that

$$K(s,t) = \mathrm{Cov}(X_s, X_t) = \min(s,t).$$

Consequently, the Karhunen Loève decomposition of X is given by

$$X = \sum_{i=1}^{\infty} \langle X, u_i \rangle u_i$$

with eigenvectors

$$u_t(i) = \sqrt{2} \sin\left(\frac{(2i-1)\pi t}{2}\right). \tag{5.21}$$

The respective eigenvalues to the associated Fredholm operator are given by

$$c_i = \frac{1}{\pi^2 (i + \frac{1}{2})^2}. \tag{5.22}$$

The i-th principal component is given by

$$\langle X, u_i \rangle = \frac{2\sqrt{2}}{(2i-1)\pi} \left(\sum_{i=1}^{N_1} \cos\left(\frac{(2i-1)\pi T_i}{2}\right) + \int_0^1 \lambda_s \cos\left(\frac{(2i-1)\pi s}{2}\right) ds \right).$$

Proof.

Considering the previous discussions, it suffices to show that the process X has the covariance kernel $K(s,t) = \min(s,t)$. This condition is satisfied by Proposition 5.5.3 since we assumed that

$$\mathbb{E}\left[\int_0^s \lambda_v dv\right] = s.$$

It remains to prove the representation of the principal components. It follows that

$$\begin{aligned}
\langle X, u_i \rangle &= \int_0^1 \left(N_s - \int_0^s \lambda_v dv \right) \sqrt{2} \sin\left(\frac{(2i-1)\pi s}{2}\right) ds \\
&= \int_0^1 N_s \sqrt{2} \sin\left(\frac{(2i-1)\pi s}{2}\right) ds - \int_0^1 \int_0^s \lambda_v dv \sqrt{2} \sin\left(\frac{(2i-1)\pi s}{2}\right) ds.
\end{aligned}$$

Standard calculations yield that for the first integral,

$$\int_0^1 N_s\sqrt{2}\sin\left(\frac{(2i-1)\pi s}{2}\right)ds = \frac{2\sqrt{2}}{(2i-1)\pi}\sum_{i=1}^{N_1}\cos\left(\frac{(2i-1)\pi T_i}{2}\right).$$

Integration by parts yields for the second integral

$$\begin{aligned}
&\int_0^1\int_0^s \lambda_v dv\sqrt{2}\sin\left(\frac{(2i-1)\pi s}{2}\right)ds \\
&= \left[\int_0^s \lambda_v dv\frac{2\sqrt{2}}{(2i-1)\pi}\cos\left(\frac{(2i-1)\pi s}{2}\right)\right]_0^1 + \frac{2\sqrt{2}}{(2i-1)\pi}\int_0^1 \lambda_s\cos\left(\frac{(2i-1)\pi s}{2}\right)ds \\
&= \frac{2\sqrt{2}}{(2i-1)\pi}\int_0^1 \lambda_s\cos\left(\frac{(2i-1)\pi s}{2}\right)ds.
\end{aligned}$$

Consequently, it follows that

$$\langle X, u(i)\rangle = \frac{2\sqrt{2}}{(2i-1)\pi}\left(\sum_{i=1}^{N_1}\cos\left(\frac{(2i-1)\pi T_i}{2}\right) + \int_0^1 \lambda_s\cos\left(\frac{(2i-1)\pi s}{2}\right)ds\right).$$

□

It is noteworthy that if N is a Poisson process with parameter ρ, then the i-th component of the compensated process X can be expressed as

$$\begin{aligned}
\langle X, u(i)\rangle &= \frac{2\sqrt{2}}{(2i-1)\pi}\left(\sum_{i=1}^{N_1}\cos\left(\frac{(2i-1)\pi T_i}{2}\right) + \int_0^1 \lambda_s\cos\left(\frac{(2i-1)\pi s}{2}\right)ds\right) \\
&= \frac{2\sqrt{2}}{(2i-1)\pi}\left(\sum_{i=1}^{N_1}\cos\left(\frac{(2i-1)\pi T_i}{2}\right) + \rho\int_0^1 \cos\left(\frac{(2i-1)\pi s}{2}\right)ds\right) \\
&= \sqrt{2}\left[\frac{2}{(2i-1)\pi}\sum_{i=1}^{N_1}\cos\left(\frac{(2i-1)\pi T_i}{2}\right) + (-1)^i\frac{2}{(2i-1)\pi}\rho\right].
\end{aligned}$$

Our next step involves applying Proposition 5.5.4 to the well-established case of an empirical distribution function. The investigation will focus on the trivial martingale (Remark 5.1.3)

$$\frac{1}{\sqrt{n}}(N_t - \Lambda_t) = \sqrt{n}F_n(s) - \frac{1}{\sqrt{n}}\sum_{i=0}^{n-1} log\left(\left(\frac{1-(T_i\wedge s)}{1-(T_{i+1}\wedge s)}\right)^{(n-i)}\right).$$

Lemma 5.5.5 (Transform F_n into a minimum type martingale)
Let nF_n be a renormalized empirical distribution function associated with a uniform distributed sample. Define the process X by

$$X_s := \sqrt{n}F_n(s) - \frac{1}{\sqrt{n}} \sum_{i=0}^{n-1} log \left(\left(\frac{1-(T_i \wedge s)}{1-(T_{i+1} \wedge s)} \right)^{(n-i)} \right). \tag{5.23}$$

Then X is a martingale with respect to the filtration generated by nF_n and it follows that

$$\mathrm{Cov}(X_s, X_t) = \min(s,t).$$

Proof.

The conditional intensity λ of the point process $N = nF_n$ is given by

$$\lambda_s := \frac{(n - N_{s-})1_{[0,1)}}{1-s}.$$

It follows that

$$\begin{aligned}
\mathbb{E}\left[\int_0^s \lambda_v dv\right] &= \mathbb{E}\left[\int_0^s \frac{(n - N_{s-})1_{[0,1)}}{1-s} dv\right] \\
&= \int_0^s \frac{(n - \mathbb{E}[N_{v-}])1_{[0,1)}}{1-v} dv \\
&= \int_0^s \frac{(n - nv)1_{[0,1)}}{1-v} dv \\
&= ns.
\end{aligned}$$

Consider the process X defined as

$$X_s := N_s - \frac{(n - N_{s-})1_{[0,1)}}{1-s}.$$

According to Proposition 5.5.3, we have

$$Cov(X_s, X_t) = n \cdot \min(s,t).$$

By rewriting the compensator, we can conclude that

$$\begin{aligned}
&\int_0^s \frac{(n-N_v)1_{[0,1]}(v)}{1-v}dv \\
&= \sum_{i=0}^{T_{N_s}-1} \int_{T_i}^{T_{i+1}} \frac{(n-i)}{1-v}dv + \int_{T_{N_s}}^{s} \frac{(n-N_s)}{1-v}dv \\
&= \sum_{i=0}^{T_{N_s}-1} log\left(\left(\frac{1-T_i}{1-T_{i+1}}\right)^{(n-i)}\right) + (n-N_s)log\left(\left(\frac{1-T_{N_s}}{1-s}\right)^{n-N_s}\right) \\
&= \sum_{i=0}^{n-1} log\left(\left(\frac{1-(T_i\wedge s)}{1-(T_{i+1}\wedge s)}\right)^{(n-i)}\right).
\end{aligned}$$

□

One can show that the minimum type martingale (5.23) converges in distribution to a Brownian motion for $n \to \infty$ as outlined in [36] on page 268.

Lemma 5.5.5 already indicates the usefulness of Proposition 5.5.4 for the empirical distribution function. We now want to construct a class of point processes to which the previous discussion is applicable.

Lemma 5.5.6 (Point processes triggered by non-negative martingales) Let N be a point process with conditional $\mathcal{H}$-intensity λ, such that λ is a non-negative $\mathcal{H}$-martingale with right-continuous paths and left limits, starting in $\lambda_0 = 1$ and satisfying the following conditions:

1. $\lim_{s\to 1} \lambda_s = 0$.
2. For all $t \in I : \lambda_t - \lim_{s\uparrow t} \lambda_s \leq 0$.

Then the martingale X defined by

$$X_s := N_s - \int_0^s \lambda_v dv,$$

satisfies

$$\mathrm{Cov}(X_s, X_t) = \min(s,t).$$

Proof.

The process λ is a martingale starting in $\lambda_0 = 1$. Thus it follows that

$$\mathbb{E}\left[\int_0^s \lambda_v dv\right] = \int_0^s \mathbb{E}[\lambda_v]dv$$
$$= \int_0^s \lambda_0 dv$$
$$= s.$$

The proof is concluded in accordance with Proposition 5.5.4.

□

Using our previous results, we are able to specify the distribution of the maximum value reached by the process λ on the observation interval:

$$\mathbb{P}\left[\sup_{s\in I} \lambda_s > c\right] = \frac{1}{c}.$$

Example 5.5.7 (Martingales as triggering processes)

Consider the renormalized empirical distribution function nF_n. Our prior investigations in chapter three revealed that the process λ can be expressed as

$$\lambda_s = \frac{n-1}{n} - F_n(s) + \frac{1}{n}\left(\frac{1-F(T_{nF_n})}{1-F(s)}\right)^{n(1-F_n(s))}. \tag{5.24}$$

λ is an adapted process with respect to the filtration generated by nF_n, and it meets the requirements stated in Lemma 5.5.6.

Let N be a point process with conditional intensity given by λ (5.24) with respect to the filtration generated by N and nF_n. We can conclude from Lemma 5.5.6 that the compensated process X, given by $X_s = N_s - \int_0^s \lambda_v dv$, is characterized by a minimum type kernel $\min(s,t)$. Utilizing Proposition 5.5.4, it follows that the Karhunen Loève decomposition of the compensated process X is represented as

$$X = \sum_{i=1}^{\infty} \langle X, u_i \rangle u_i$$

with eigenfunctions

$$u_t(i) = \sqrt{2}\sin\left(\frac{(2i-1)\pi t}{2}\right).$$

Chapter 6

Conclusion

The initial question of this work revolved around the extent to which martingale transformations of particular nature can be realized within the framework of point process theory. More specifically, we aimed to achieve two goals here.

On one hand, we were familiar with examples of non-negative martingales that converge to zero in the limit. In this regard, we already had well-known examples at our disposal, such as the geometric Poisson process (5.2.4) or the martingale based on Daniels' work (1.1) in the context of the empirical distribution function. Furthermore, we were aware of the work by Hess, which can be considered a significant extension of Daniels' martingale in the area of the empirical distribution function. We knew that these martingales share the commonality that their structure allows for the determination of exceedance probabilities. The foundation for calculating these exceedance probabilities represents a special case of Doob's martingale inequality, where this inequality becomes an equation (Proposition 3.3.1). What remained uncertain to us was the degree to which a class of martingales could be delineated to also exhibit these properties within the broader context of point processes. To effectively address this question, we pursued two approaches. In the first approach, we utilized the Doob-Meyer decomposition to establish a differential equation, culminating in the two results, Theorem 3.2.6 and 3.2.8. These two theorems allowed us to consider all previously known martingales within a more general context and recognize them henceforth as special cases. The second approach relied on the well-known representation theorem for martingales. In this context, we delve into criteria for the predictable process, ensuring that the resulting martingale precisely meets our requirements. This approach ultimately resulted in the successful establishment of Theorem 3.4.1. Possible applications of our martingales in the realm of goodness-of-fit testing are found in Lemma 3.4.4. As a final comment, we would like to note that all martingales that inspired us for this part of the work exhibited exclusively negative jump heights and were consistently monotonically increasing between the jumps. For the calculation of exceedance probabilities in this case, the simpler formula from Proposition 3.3.1 is appropriate. However, for the development of tests, we believe that we cannot do without allowing positive jump heights, leading to a second formulation in 3.3.6. A compelling question that follows from our results is how to construct efficient test martingales from the pool of possible test strategies outlined in Definition 3.4.3. In general, it is likely that approximation methods will need to be employed for this purpose.

On the other hand, we were interested in a more or less general class of point processes for which we could provide an orthogonal decomposition à la Karhunen-Loève. This proved to be extremely challenging, as an explicit representation of the eigenfunctions is only available for a handful of stochastic processes. The derivation of the Karhunen-Loève expansion (KLE) relies on the covariance structure of the underlying process, which is of crucial importance. In the context of the (KLE), we were already aware of the Poisson process, which exhibits the same covariance structure as the Brownian motion. Furthermore, we already had an explicit representation of the coefficients for the Poisson process and were also aware that these coefficients are stochastically dependent on each other in the case of the Poisson process (paper in work). Given our prior experience, we therefore set out to define a class of point processes whose covariance structure resembles that of the Poisson process, thus being of minimum type. A meaningful answer to this question ultimately resulted in Lemma 5.5.6, where the compensator of the underlying point process itself represents a martingale starting in 1. In deriving our result, we made use of the Poisson-Charlier polynomial of the second order at one point. For this, refer to the proof of Proposition 5.5.3. The Poisson-Charlier martingales represent a well-known class of transformations. We derived them through careful effort in Theorem 5.3.4 using the Doob-Meyer decomposition and extended them to general point processes in Theorem 5.4.1.

Bibliography

[1] Åkesson, F.; Lehoczky, J. (1998) Discrete eigenfunction expansion of multidimensional Brownian motion and the Ornstein-Uhlenbeck process, working paper

[2] Barbu, V.; Limnios, N. (2008) Discrete-Time Renewal Processes. In: Semi-Markov Chains and Hidden Semi-Markov Models toward Applications. Lecture Notes in Statistics, vol 191. Springer, New York, NY.

[3] Beiglboeck, M., Schachermayer, W., and Veliyev, B. (2010). A short proof of the doob-meyer theorem.

[4] Billingsley, P. (1995) Probability and measure, Wiley Series in Probability and Mathematical Statistics (Third edition of 1979 original ed.).

[5] Brémaud, P. (1981) Point Processes and Queues - Martingale Dynamics, Springer series in statistics

[6] Brémaud, P. (2021) Point Process Calculus in Time and Space An Introduction with Applications, Springer-Verlag

[7] Charlier, C. (1905) Über die Darstellung willkürlicher Funktionen, Ark. Mat. Astr. och Fysic 2, 20.

[8] Davis, M.H. (1976) The representation of martingales of jump processes, SIAM J. Control Optim., 14(4):623638.

[9] Doob, J. L. (1953) Stochastic processes. John Wiley and Sons Inc., New York

[10] Dupire, B. (2009) Functional Itô Calculus, Bloomberg Portfolio Research Paper No. 2009-04-FRONTIERS

[11] Dvoretzky, A.; Kiefer, J.; Wolfowitz J. (1953) Sequential Decision Problems for Processes with Continuous time Parameter. Testing Hypotheses, The Annals of Mathematical Statistics Vol. 24, No. 2 (Jun., 1953), pp. 254-264 (11 pages)

[12] Dvoretzky, A.; Kiefer, J.; Wolfowitz, J. (1956), "Asymptotic minimax character of the sample distribution function and of the classical multinomial estimator", Annals of Mathematical Statistics, 27 (3): 642–669

[13] Elliot, R. (1976) Stochastic integrals for martingales of a jump process with partially accessible jump times, Probability Theory and Related Fields, 36:213– 226.

[14] Foufoula-Georgiou, E.; Lettenmaier, D. P. (1986) Continuous-Time Versus Discrete-Time Point Process Models for Rainfall Occurrence Series, Water Resour Res 22(4):531–542.

[15] Glasserman, P. (2003) Monte Carlo Methods in Financial Engineering, Springer-Verlag

[16] ushchin, A. A. (2020) Single jump filtrations and local martingales. Modern Stochastics: Theory and Applications, page 135156.

[17] Hackmann, D. (2017) Karhunen-Loève expansion of Lévy processes, Communications in Statistics - Theory and Methods

[18] Hall, P.; Heyde, C. C. (1980) Martingale limit theory and its application. Probability and Mathematical Statistics, CA: Academic Press

[19] Hess, D. (2015) Martingalmethoden zur Berechnung von Überschreitungswahrscheinlichkeiten datenbasierter Sprungprozesse, Dissertation, Justus-Liebig-Universität Gießen

[20] Jacod, J. (1975). Multivariate point processes: predictable projection, radon-nikodym derivatives, representation of martingales. Zeitschrift fu r Wahrscheinlichkeitstheorie und verwandte Gebiete, 31(3):235–253.

[21] Kallenberg, O. (1986) Random Measures. 4th edition. Academic Press, New York, London; Akademie-Verlag, Berlin

[22] Kallenberg, O. (2021) Foundations of modern probability. Probability Theory and Stochastic Modelling. Vol. 99 (Third edition of 1997 original ed.). Springer,

[23] Karatzas, I.; Shreve, S. E. (1991) Brownian motion and stochastic calculus. Graduate Texts in Mathematics. Vol. 113 (Second edition of 1988 original ed.), New York: Springer-Verlag.

[24] Kopperschmidt, S. (2005) Statistische Analyse von Kaufverhaltensmodellen, Dissertation, Justus-Liebig-Universität Gießen

[25] Kunita H.; Watanabe S. (1967) On square integrable martingales, Nagoya Math. J. 30, pp. 209-245

[26] Kunita, H. (2010). Itô's stochastic calculus: Its surprising power for applications. Stochastic Processes and their Applications, 120(5):622–652. A tribute to Kiyosi It.

[27] Last, G. (2018) Lectures on the Poisson-Process, Cambridge University Press

[28] Meyer, P.-A. (1962). A decomposition theorem for supermartingales. Illinois J. Math., 6:193– 205

[29] Møller, J.; Waagepetersen R. P. (2004) Statistical Inference and Simulation for Spatial Point Processes

[30] Noether G.E. (1978), "A brief survey of nonparametric statistics" R.V. Hogg (ed.) , Studies in statistics , Math. Assoc. Amer. pp. 39–65

[31] Protter, P. (2005) Stochastic Integration and Differential Equations, Springer, Berlin, Heidelberg.

[32] Privault, N. (2014) Stochastic Finance - An Introduction with Market Examples, Chapman & Hall/CRC Financial Mathematics Series

[33] Rutherford, D.E. (1947) Some Continuant Determinants arising in Physics and Chemistry, Proceedings of the Royal Society of Edingburgh

[34] Schoutens, W. (2000) Stochastic Processes and Orthogonal Polynomials, Springer-Verlag

[35] Shafer G., Shen A., Vereshchagin N., Vovk V., (2011) Test Martingales, Bayes Factors and p-Values, Statistical Science 2011, Vol. 26, No. 1, 84–101

[36] Shorack, G. R.; Wellner, J. A. (1986) Empirical Processes with Applications to Statistics, Society for Industrial and Applied Mathematics

[37] Shiryaev, A. N. (2019) Probability - 2. Graduate Texts in Mathematics. Vol. 95. Translated by Boas, R. P.; Chibisov, D. M. (Third edition of 1980 original ed.), New York: Springer.

[38] Wald A. (1947) Sequential Analysis. John Wiley and Sons, 1. Aufl.

Declaration of Authorship

I declare that I have completed this dissertation single-handedly without the unauthorized help of a second party and only with the assistance acknowledged therein. I have appropriately acknowledged and cited all text passages that are derived verbatim from or are based on the content of published work of others, and all information relating to verbal communications. I consent to the use of an anti-plagiarism software to check my thesis. I have abided by the principles of good scientific conduct laid down in the charter of the Justus Liebig University Giessen 'Satzung der Justus-Liebig-Universitat Giessen zur Sicherung guter wissenschaftlicher Praxis' in carrying out the investigations described in the dissertation.

____

Giessen, April 4, 2024