FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTOCTORALES

uOttawa

L'Université canadienne
Canada's university

FACULTY OF GRADUATE AND POSDOCTORAL STUDIES

Jesse Collingwood
AUTEUR DE LA THÈSE / AUTHOR OF THESIS

M.Sc. (Mathematics)
GRADE / DEGREE

Department of Mathematics and Statistics
FACULTÉ, ÉCOLE, DÉPARTEMENT / FACULTY, SCHOOL, DEPARTMENT

Application of Martingale Methods to a Change Set Problem
TITRE DE LA THÈSE / TITLE OF THESIS

Dr. G. Ivanoff
DIRECTEUR (DIRECTRICE) DE LA THÈSE / THESIS SUPERVISOR

CO-DIRECTEUR (CO-DIRECTRICE) DE LA THÈSE / THESIS CO-SUPERVISOR

EXAMINATEURS (EXAMINATRICES) DE LA THÈSE / THESIS EXAMINERS

Dr. R. Balan

Dr. M. Csörgö

Gary W. Slater
Le Doyen de la Faculté des études supérieures et postdoctorales / Dean of the Faculty of Graduate and Postdoctoral Studies

Application of Martingale Methods to a Change Set Problem

Jesse Collingwood

Thesis Submitted to the Faculty of Graduate and Postdoctoral Studies
In partial fulfilment of the requirements for the degree of Master of Science in
Mathematics [1]

Department of Mathematics and Statistics
Faculty of Science
University of Ottawa

---

[1] The M.Sc. program is a joint program with Carleton University, administered by the Ottawa-Carleton Institute of Mathematics and Statistics

Library and
Archives Canada

Published Heritage
Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et
Archives Canada

Direction du
Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence
ISBN: 978-0-494-51640-9
Our file Notre référence
ISBN: 978-0-494-51640-9

NOTICE:
The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:
L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Abstract

A collection of random points in a bounded rectangle $[0, R]^2$ of the positive quadrant changes its intensity on an unobservable random set ξ. The goal is to detect the random change set in an optimal way, according to some reward function. Given the change set, the random points are assumed to be distributed according to a Poisson process N with intensities μ_0 on ξ^c and μ_1 on ξ. Optimal solutions are found for various reward functions and applied to some examples.

Acknowledgements

Asking a mathematician, especially one such as myself, to determine who among his acquaintances shares responsibility for his thesis can be dangerous. An entire undergrad degree spent learning the necessity of rigor and completeness can turn one into an obsessively fastidious beast that has trouble considering such a task with the common sense that he once had. Just about everyone I know gets put into a sieve, and just about everyone makes it through and earns the status "contributed in some way". This is totally unreasonable, however, and so I have devised a formula that filters out those whose contributions do not meet a tolerance level that was predetermined by a method that may serve as the subject matter for a future paper, but has no place here.

Firstly, I must thank my parents for their invaluable life lessons. My appreciation for everything, including mathematics, is a function of the values I received from them, and it is because of their hard work that I had the opportunity to enter the academe at all.

Secondly, I must thank my lifelong friend, Justin, who constantly and inadvertently taught me the need for an indominable will, and that success is a matter of hard work rather than innate ability.

I owe a serious debt to the "math crew" that made my undergrad journey a beautiful and interesting one. Chris, Matt, Andrew, Sam, Pan, Slava, Marie-Pascal, and Paul, you philosopher comedians, you kept the party going. Thank you.

My dear Lijie, your fun spirit is like an oasis in the desert, and I am very thankful that you have been here to keep me weird and laughing during this part of my life.

There have also been several Professors at the University of Ottawa that have instilled a deeper appreciation and understanding of mathematics than was required of them, or who have been superb as lecturers, or both. Most notably, Professors Raluca Balan and David McDonald, who receive my utmost gratitude.

Finally, yet foremost, I must thank my thesis supervisor, Professor Gail Ivanoff. Professor Ivanoff has a way of removing misunderstandings very calmly and with minimal effort, just as a Tai Chi master might cast away an unsound attack with great economy of movement. Her perception of my knowledge and capabilities is always much more accurate than my own, and she has been exceedingly patient with me as I have taken my first baby steps into the world of research. I could not have asked for a better guide. Thank you deeply.

Dedication

This work is dedicated to my brother Dylan, in fulfillment of a very old promise.

Contents

Chapter 1

Introduction

Imagine a cluster of random points in a region such as a forest - these points could be trees that have become infested by a certain type of beetle, for instance. If there is no discernable pattern to the points, no obvious clusters or lack thereof, it is a reasonable assumption that the points follow a Poisson process. Now imagine that, from inspection say, it is clear that some lower left sub-region is less densely populated by the points than the remainder of the region, but that these regions individually still appear to be distributed without any internal pattern. The dividing line between the region where beetle infestation is less frequent from where it is more frequent could be the habitat border for a species of birds that feeds on the beetle and diminishes its population, for instance. Among other things, finding this line and comparing it to other known information about the forest could help to identify something such as a natural predator, which could in turn help to solve the infestation problem.

Herein the model dealt with is one where an observer begins to "walk northeast" and accumulate information about these random points in an effort to determine the best halting point - the one optimal among all others with respect to some valuation or gain function. When formulated heuristically as above, the task seems quite similar to an optimal stopping problem on the line and in fact requires an extension of the theory hosting such problems. For this reason it is useful to survey some related works in that theory.

The most relevant work here is [1]. In the paper, the authors (Herberts and Jensen) consider a point process N on the line which, given a random point σ, is Poisson with intensity $\mu_0(s)$ before σ and $\mu_1(s)$ after. They assume $0 < \mu_0(s) \leq \mu_1(s)$ for all $s \in \mathbb{R}_+$, and that σ is exponentially distributed. A reward function is given to describe the gain associated with stopping at time t,

$$Z_t = c_0(t \wedge \sigma) - c_1(t-\sigma)^+ - k_0(1-Y_t) - k_1 Y_t$$

where $Y_t = I_{\{\sigma \leq t\}}$, and the authors proceed to find conditions under which $\mathbb{E}[Z_\tau]$ can be maximized for finite stopping times τ belonging to a certain class. Several information structures are considered, hence the title's reference to different observation schemes, but the only one of relevance here is the sequential observation scheme where the available information at time t is the collection of jump times of N prior to t. A stopping time referred to by the authors as the *infinitesimal-look-ahead* rule is shown to be optimal (in the sense described) under certain conditions in the paper's *monotone stopping theorem.* This is the work's main tool.

Motivated to extend the one dimensional case to two dimensions, Gail Ivanoff and Ely Merzbach in their paper ([2]) developed analogues to each of the features of the work by Herberts and Jensen just described. The fundamental issue of extending the notion of change point to change set was dealt with in their earlier work [4] in more generality and once solved provided the basis for speaking about optimal stopping in dimensions higher than one.

T.S. Hahubia and R. Mnatsakanov in their article ([11]) also deal with a generalization of the classical one-dimensional change-point problem, but their focus is on testing the null hypothesis that, given a sequence $\tilde{X}_n(A)$ of realizations of some observed random measures and independent sequences $(X_n)_{n\geq 1}$, $(\Delta_n)_{n\geq 1}$ of random measures, $X_n(A) \stackrel{d}{=} \tilde{X}_n(A)$ for some set A, versus the alternative that there exists a change set C such that $\tilde{X}_n(A) \overset{d}{=} X_n(A) + \epsilon_n \Delta_n(A \cap C)$ for some $\epsilon_n \to 0$. No gain function is considered for optimization in the article, and so there is a fundamental disparity with what is developed here.

In [10] the authors consider a "mark" Y (a random variable whose target space is suppressed) changing its conditional distribution given a random variable X taking values in $\mathbb{R}^d$ on some $G \in \mathcal{B}(\mathbb{R}^d)$, i.e.

$$P[Y \in H|X] = P_1[Y \in H] I_{\{X \notin G\}} + P_2[Y \in H] I_{\{X \in G\}}$$

for some distributions P_1 and P_2. The goal of the paper is to estimate the change set G based on observed i.i.d. sequences $(X_i, Y_i)_{i=1}^n$ of such pairs (X, Y). However, this is formulated within the context of Vapnik-Cervonenkis classes. Specifically, G is supposed to belong to such a class, and although a "score function" is considered when finding the Maximum likelihood estimator for G, the approach of using δ-nets of VC classes is quite removed from the smooth semi-martingale techniques used here.

The goal here is to describe the paper by Ivanoff and Merzbach from a graduate student's perspective, and to extend the model developed therein. To this end, a brief introduction to the theory of point processes in 1-dimension is given in the Preliminaries section of Chapter 2. Here the notion of a random measure is discussed and some key facts about this notion are proven. Some Cox processes are defined here. The discussion becomes more general in Section 2.2, Processes in Higher Dimensions, and remains so for most of the remainder of the paper. It is in this section that the n^{th}-rectangular partition, an important tool later on, is introduced. Following this, in Section 2.2.1, we give a detailed account of the space used to describe the change set of the spatial process being observed (the space of "upper layers"). Besides the space itself, the central definition in this section is that of a stopping set, the d-dimensional analogue to the 1-dimensional stopping time. Martingales indexed by $\mathbb{R}_+^d$ are defined in Section 2.2.2, and it is shown that a stopped multi-dimensional martingale still has expectation 0. The spatial point process being observed is then described in full detail in Section 2.2.3. Following this, Section 2.3 introduces the likelihood function and conditional likelihood function of a point process and contains a series of calculations ascertaining their form in the case that the point process is conditionally Poisson. Chapter 2 ends with a look at compensators and weak hazard functions, and a very useful lemma specifying the form of a compensator for a process (under certain conditions) is presented.

In Chapter 3 the groundwork laid in Chapter 2 is used to fully develop the model that will finally be used to describe the problem. The first step is to make precise what it means to detect *optimally* the change set, and this is done by introducing a gain function. The next step involves showing that the gain function can be expressed as an instance of a special type of structure, called a smooth semi-martingale, in order to apply to it lemma 3.1.5, which ensures that an optimal solution to the sequential detection problem exists (under certain conditions). The remainder of the chapter is dedicated to finding conditions sufficient for application of the lemma, and the results of this investigation are summarized in Theorem 3.2.1, the main result.

Chapter 4 details applications of the results from Chapter 3. Specifically, an alternate valuation function to the one used in the paper is investigated and shown to satisfy the hypotheses of the optimality theorems, and with a bit of calculation the example of a change set generated by the first line of a non-homogeneous Poisson process is brought within the scope of these theorems as well.

Chapter 2

Preliminaries

The spatial processes of our model are instances of random measures. The theory of random measures constitutes section 2.1 and is well expounded in the main source of the section, [6]. In section 2.2 analyses of point processes are carried into higher dimensions and many of the tools - elementary definitions and properties - pertaining to our model are explained for later use. A subsection is devoted to the space of upper layers, to which the change set of the model belongs, and it is shown how a random upper layer arises from a single line process. Stopping sets are also defined here, and in Proposition 2.2.14 it is shown that a stopped $\mathbb{R}_+^d$ indexed martingale, expressible as the difference of two increasing processes, has expectation 0. Finally, the conditional Poisson process in higher dimensions, the first line of a Poisson process, and the likelihood function of a point process are defined and discussed.

2.1 Processes in 1-dimension

Fix a probability space $(\Omega, \mathcal{F}, P)$ and let E be a locally compact, second countable, Hausdorff space. Form the σ-algebra $\mathcal{E}$ generated by the open sets in E, and let $\mathcal{B}$ be the collection of relatively compact sets (also referred to as bounded sets) from $\mathcal{E}$.

Call a measure μ on $(E, \mathcal{E})$ *radon* if $\mu(B) < \infty \; \forall B \in \mathcal{B}$, and let $\mathfrak{M}$ be the collection of radon measures on $(E, \mathcal{E})$ and $\mathfrak{N}$ the collection of measures N from $\mathfrak{M}$ such that $N(B) \in \mathbb{N} \; \forall B \in \mathcal{B}$.

Define a σ-algebra $\mathcal{M}$ on $\mathfrak{M}$ by letting $\mathcal{M}$ be generated by all maps $\mathfrak{M} \ni \mu \mapsto \mu(B)$ where $B \in \mathcal{B}$; that is

$$\mathcal{M} = \{\{\mu \in \mathfrak{M}; \mu(B) \in H\}; B \in \mathcal{B}, H \in \mathcal{B}(\mathbb{R})\}.$$

Let $\mathcal{N}$ be the σ-algebra on $\mathfrak{N}$ obtained by taking the trace of $\mathcal{M}$ over $\mathfrak{N}$ (or equivalently the σ-algebra generated by the maps $\mathfrak{N} \ni \mu \mapsto \mu(B)$ where $B \in \mathcal{B}$).

Definition 2.1.1 *A random measure (on E) is a map $M : \Omega \to \mathfrak{M}$ measurable $\mathcal{F}/\mathcal{M}$. A point process (on E) is a map $N : \Omega \to \mathfrak{N}$ measurable $\mathcal{F}/\mathcal{N}$.*

Given a point process N and $B \in \mathcal{B}$, $N(B)$ will be used to denote the map given by $\omega \mapsto N(\omega)(B)$.

Lemma 2.1.2 *Let N be a point process on $E = \mathbb{R}_+$. Then*

i) $N(A)$ is measurable for each $A \in \mathcal{E}$,

ii) If X and Y are real, non-negative random variables then so is $N[X, Y)$, and δ_X is a point process,

iii) N can be expressed as a sum

$$N = \sum_{j=1}^{\infty} \delta_{X_j}$$

for some random variables $X_j, j \geq 1$.

Proof:

(Parts of the proof can be found in lemma 2.1 of [6]). First let $A \in \mathcal{E}$, $H \in \mathcal{B}(\mathbb{R})$. The set $S := \{\mu \in \mathfrak{N}; \mu(A) \in H\}$ is in $\mathcal{N}$ by definition, so

$$N(A)^{-1}(H) \;\; = \;\; \{\omega; N(\omega)(A) \in H\}$$

$$\begin{aligned} &= N^{-1}(S) \\ &\in \mathcal{F} \end{aligned}$$

since N is a point process. This verifies part i).

Now let X, Y be non-negative random variables and let $(X_n)_{n\geq 1}, (Y_n)_{n\geq 1}$ be sequences of simple random variables converging to X and Y respectively from below, say $X_n = \sum_{i=1}^{k_n} a_i I_{A_i}, Y_n = \sum_{i=1}^{h_n} b_i I_{B_i}$. Then

$$N[X_n, Y_n) = \sum_{i=1}^{k_n} \sum_{j=1}^{h_n} N[a_i, b_j) I_{A_i \cap B_j}$$

is measurable by part i) for each n. Clearly $[X_n(\omega), Y_n(\omega)) \to [X(\omega), Y(\omega))$ as $n \to \infty$ for every $\omega \in \Omega$, and hence, as $N(\omega)$ is a measure, $N[X_n, Y_n)$ converges everywhere to $N[X, Y)$. Therefore $N[X, Y)$ is indeed a random variable.

To see that δ_X is a point process, simply note that δ_x is an integer valued measure for each $x \in \mathbb{R}_+$ and that for $A \in \mathcal{N}$, say $A = \{\mu; \mu(B) \in H\}$ for some compact $B \in \mathcal{E}$ and $H \in \mathcal{B}(\mathbb{R}_+)$,

$$\begin{aligned} \delta_X^{-1}(A) &= \{\omega; \delta_{X(\omega)}(B) \in H\} \\ &= \begin{cases} \emptyset & \text{if } 0, 1 \notin H \\ X^{-1}(B^c) & \text{if } 0 \in H, 1 \notin H \\ X^{-1}(B) & \text{if } 0 \notin H, 1 \in H \\ \Omega & \text{if } 0, 1 \in H \end{cases} \\ &\in \mathcal{F}. \end{aligned}$$

Therefore δ_X is a point process and ii) is verified.

Now let $X_k(\omega) := \inf\{s; N(\omega)([0, s]) \geq k\}$ for $k \geq 1$, $\omega \in \Omega$. Then, $\{X_k \leq t\} = \{N([0, t]) \geq k\} \in \mathcal{F}$, and so each X_k is a random variable. Furthermore, for fixed $\omega \in \Omega$, it is clear that $N(\omega)[0, X_1(\omega)) = 0$ and that, for $x \in (X_k(\omega), X_{k+1}(\omega))$ say, $k \leq N(\omega)([0, x]) < k + 1$ by definition of the X_i. As $N(\omega)$ takes only integer values, this means that $N(\omega)(X_k(\omega), X_{k+1}(\omega)) = 0$.

Therefore, if we define $\delta_\infty := 0$, we may write $N = \sum_{j=1}^{\infty} \delta_{X_j}$, verifying iii).

■

The X_j of the above lemma are the *jump points* of the process N. Note that arriving at an expression for N in this form required an ordering for the X_j, and so while the preceding arguments are valid on the line, extending them to higher dimensions is a bit tricky and some concessions have to be made (particularly regarding the measurability of the X_j).

Of particular interest is the Poisson process.

Definition 2.1.3 *A point process N on E is a Poisson process with mean measure $\mu \in \mathfrak{M}$ if*

i) whenever $A_1, A_2, ..., A_k \in \mathcal{E}$ are disjoint the random variables $N(A_1), ..., N(A_k)$ are independent

ii) for $A \in \mathcal{E}$ and $k \in \mathbb{N}$, $P(N(A) = k) = \frac{\mu(A)^k}{k!} e^{-\mu(A)}$.

When $E = \mathbb{R}_+$, μ is often taken to be a multiple of Lebesgue measure or the integral of some continuous, nonnegative function with respect to Lebesgue measure. Such multiples are called *intensities* and such functions are called *rate* or *intensity functions* for their respective processes.

There is also the notion of a process with a *change point*: The goal is to construct a (Poisson) process which, conditional upon σ, has mean measure μ_0 on $[0, \sigma]$ and μ_1 on (σ, ∞). To do this one defines a new "measure" which is some combination of μ_0 and μ_1 and serves as the intensity of the process.

Lemma 2.1.4 *Let $E = \mathbb{R}_+$, $\mu_0, \mu_1 \in \mathfrak{M}$, and let $\sigma : \Omega \to E$ be a random variable. For $A \in \mathcal{E}, \omega \in \Omega$, define*

$$\mu(\omega, A) := \mu_0(A \cap [0, \sigma(\omega)]) + \mu_1(A \cap (\sigma(\omega), \infty)).$$

Then μ is a random measure on E.

Proof:

First, fix $A \in \mathcal{E}$ and $t \in E$. Define $x = \inf\{s \in E; \mu_0(A \cap [0,s]) \geq t\}$. Then

$$\begin{aligned}\{\omega; \mu_0(A \cap [0, \sigma(\omega)]) \geq t\} &= \begin{cases} \emptyset & \text{if } x = \infty \\ \sigma^{-1}[x, \infty) & \text{otherwise} \end{cases} \\ &\in \mathcal{F}.\end{aligned}$$

Since $t \in E$ was arbitrary, it is clear that the map $\omega \mapsto \mu_0(A \cap [0, \sigma(\omega)])$ is measurable. Similar arguments show that the map $\omega \mapsto \mu_1(A \cap (\sigma(\omega), \infty))$ is measurable. Thus the map $\omega \mapsto \mu(A, \omega)$ is measurable. This map is denoted $\mu(A)$.

It is obvious that $A \mapsto \mu(A, \omega)$ is a measure for fixed ω, and that this measure is radon. Also, notice that if $M \in \mathcal{M}$, say $M = \{\rho; \rho(B) \in H\}$ for some relatively compact $B \in \mathcal{E}$ and $H \in \mathcal{B}(\mathbb{R}_+)$, then

$$\{\omega; \mu(\omega)(\cdot) \in M\} = \{\omega; \mu(\omega)(B) \in H\} \in \mathcal{F}.$$

Thus μ is a random measure.

■

Note that the above lemma shows in fact that, for each $A \in \mathcal{E}$, $\mu(A)$ is measurable $\mathcal{F}^\sigma$ (the sigma algebra generated by σ).

Definition 2.1.5 *Let $E = \mathbb{R}_+$, $\mathcal{E} = \mathcal{B}(\mathbb{R}_+)$, $\mu_0, \mu_1 \in \mathfrak{M}$, and let $\sigma : \Omega \to E$ be a random variable. Define μ as above. Then a point process N on E is a Poisson Process given σ if, with probability 1, the following hold:*

i) for each $A \in \mathcal{E}$ and $k \in \mathbb{N}$,
$P(N(A) = k|\sigma) = \frac{\mu(A)^k}{k!} e^{-\mu(A)}$

ii) whenever $A_1, ..., A_n \in \mathcal{E}$ are disjoint
$P(N(A_i) = k_i, 1 \leq i \leq n|\sigma) = \prod_{i=1}^n P(N(A_i) = k_i|\sigma)$.

The point process N of the above definition is an example of a *doubly stochastic Poisson* or *Cox* process, and σ is referred to as a *random change point*; if σ is constant it is simply referred to as a change point. Later, change points will be generalized to *change sets.*

2.2 Processes in Higher Dimensions

Consider the partial order on $\mathbb{R}_+^d$ given by $(t_1, ..., t_d) \leq (s_1, ..., s_d)$ iff $t_i \leq s_i, 1 \leq i \leq d$, and for $t \in \mathbb{R}_+^d$ define $A_t = \{s \in \mathbb{R}_+^d; s \leq t\}$, $E_t = \{s \in \mathbb{R}_+^d; s \geq t\}$. For $s, t \in \mathbb{R}_+^d$ say $s << t$ iff $s_i < t_i$ if $t_i > 0$, and $s_i = 0$ if $t_i = 0$, $i \in \{1, ..., d\}$. Recall that a *stochastic process* on a set E is a collection $(X_t)_{t\in E}$ of random variables indexed by points in E, and that a *filtration* $(\mathcal{F}_t)_{t\in E}$ on a partially ordered set E is an E-indexed family of σ-algebras such that $s \leq t \Rightarrow \mathcal{F}_s \subseteq \mathcal{F}_t$. When E is some rectangle of $\mathbb{R}_+^d$ containing 0 *outer-continuity* is typically required of a filtration; that is to say $\mathcal{F}_t = \cap_{n\geq 1}\mathcal{F}_{t_n}$ must be satisfied for all points $t \in E$ and sequences $t_n \downarrow t$. Also recall that one says X is *adapted* to the filtration if, for every $t \in E$, X_t is $\mathcal{F}_t$-measurable.

If N is some point process on E we may think of it also as a stochastic process on E by defining $N_t(\omega) = N(\omega)(A_t)$, and we may associate to it the *minimal filtration* for N, also called the filtration generated by N, denoted $\mathcal{F}^N$ and given by $\mathcal{F}_t^N := \sigma(N_s; s \leq t)$. The σ-field $\mathcal{F}_t^N$ represents the information known about N for points preceding t (with respect to the partial order). For instance, if $s \leq t$ one can determine for which ω it is that $N_s(\omega)$ is positive by examining $\mathcal{F}_t^N$, to which the set $\{\omega; N_s(\omega) > 0\} = N_s^{-1}(0, \infty)$ belongs.

Definition 2.2.1 *Let* $E = [0, R] = \prod_{i=1}^d [0, r_i]$ *for some* $R = (r_1, ..., r_d) \in \mathbb{R}_+^d$ *and let* $i_1, ..., i_d \in \{0, ..., 2^{n-1}\}$. *Let*

$$H_k = \begin{cases} (\frac{i_k r_k}{2^n}, \frac{(i_k+1)r_k}{2^n}] & \text{if } i_k > 0 \\ [0, \frac{r_k}{2^n}] & \text{if } i_k = 0 \end{cases}$$

and define $H_{i_1,\dots,i_d} = \prod_{k=1}^{d} H_k$. *The* n^{th} *rectangular partition of* E *is*

$$\mathcal{C}_n := \{H_{i_1,\dots,i_d}; 0 \le i_1, \dots, i_d \le 2^n - 1\}.$$

For $C \in \mathcal{C}_n$ *let* $t_{C\_} = \inf C, t_{C^+} = \sup C$ *be the southwest and northeast corners of* C *respectively.*

A point process (or stochastic process) N is *jointly measurable* if the map from $\Omega \times E$ to $\mathbb{R}_+$ given by $(\omega, t) \mapsto N_t(\omega)$ is measurable $\mathcal{F} \otimes \mathcal{E}$. The second part of the following lemma provides a sufficient condition for joint measurability.

Lemma 2.2.2 *Let* $E = [0, R]$ *for some* $R \in \mathbb{R}^d_+$.

i) If M *is a point process on* E *then the minimal filtration* $\mathcal{F}^M$ *for the stochastic process* $(M_t)_{t \in E}$ *is outer-continuous.*

ii) If N *is a stochastic process on* E *with outer-continuous sample paths then* N *is jointly-measurable.*

Proof:

The proof of part i) is a simplification of Proposition 5.21 of [3]. For a set $A \subseteq E$, $\text{Int}(A)$ is used to denote the interior of A. Let $t \in E$ be arbtitrary, let $(t_n)_{n \ge 1}$ be a sequence converging to t from above, and for $\omega \in \Omega$ let $\delta(\omega) := \inf\{d(s,t); s \in E \setminus A_t, M(\omega)(\{s\}) > 0\}$ (here d denotes the sup-norm). If $B_x(A_t)$ is the open ball of radius x around A_t with respect to d, then

$$\{\delta \ge x\} = \{M(B_x(A_t) \setminus A_t) = 0\} \in \mathcal{F}$$

and so δ is a random variable. Since M is finite on bounded regions, $\delta > 0$ w.p.1. Select a subsequence $(t_{n_k})_{k \ge 1}$ s.t. $d(t, t_{n_k}) < \frac{1}{k}$ for each $k \in \mathbb{N}$, and let

$$G_k = \{\omega; \delta(\omega) > \frac{1}{k}\}.$$

Note that $G_k \uparrow \{\omega; \delta(\omega) > 0\} = \Omega$.

Now if $s \leq t_{n_k}$ then $M_s(\omega) = M_{s\wedge t}(\omega)$ for each $\omega \in G_k$. That is to say that, for any Borel subset B of $\mathbb{R}_+$, $G_k \cap M_s^{-1}(B) = G_k \cap M_{s\wedge t}^{-1}(B) \in G_k \cap \mathcal{F}_t^M$. Since $\mathcal{F}_{t_{n_k}}^M$ is generated by sets of the form $M_s^{-1}(B)$ where $s \leq t_{n_k}$ and $B \in \mathcal{B}(\mathbb{R}_+)$, it follows that $G_k \cap \mathcal{F}_{t_{n_k}}^M \subseteq G_k \cap \mathcal{F}_t^M$. The reverse inclusion is immediate, and so

$$G_k \cap \mathcal{F}_{t_{n_k}}^M = G_k \cap \mathcal{F}_t^M \ \forall k \in \mathbb{N}.$$

So let $A \in \cap_{n\geq 1}\mathcal{F}_{t_n}^M = \cap_{k\geq 1}\mathcal{F}_{t_{n_k}}^M$. Then for each $k \in \mathbb{N}$ $\exists B_k \in \mathcal{F}_t^M$ such that $A \cap G_k = B_k \cap G_k$. Since the G_k are increasing one can assume that $B_k \subseteq B_{k+1}$ for each k. Then

$$\begin{aligned} A &= A \cap (\cup_{k=1}^{\infty} G_k) \\ &= \cup_{k=1}^{\infty}(A \cap G_k) \\ &= \cup_{k=1}^{\infty}(B_k \cap G_k) \\ &= \cup_{k=1}^{\infty} B_k \\ &\in \mathcal{F}_t^M. \end{aligned}$$

This means that $\cap_{n\geq 1}\mathcal{F}_{t_n}^M \subseteq \mathcal{F}_t^M$, and since the reverse inclusion is immediate it follows that $\mathcal{F}^M$ is indeed outer-continuous, verifying i).

Turning toward ii), let $\mathcal{C}_n$ be the n^{th} rectangular partition of $E = [0, R]$ and for $(\omega, t) \in \Omega \times E$ let

$$N_n(\omega, t) := \sum_{C\in\mathcal{C}_n} I_C(t) N_{t_{C^+}}(\omega).$$

Since $\bigcup_{C\in\mathcal{C}_n} C = [0, R]$, if $t \in E$ $\exists$ a sequence of rectangles $C_n \in \mathcal{C}_n$ s.t. $t_{C_n^+} \downarrow t$ as $n \to \infty$. In fact, $t_{C_n^+} = \inf\{t_{D^+} \geq t; D \in \mathcal{C}_n\}$ and $C_n \downarrow \{t\}$. Then, by the outer-continuity of $N.(\omega)$,

$$\begin{aligned} N_n(\omega, t) &= N_{t_{C_n^+}}(\omega) \\ &\xrightarrow[n\to\infty]{} N_t(\omega). \end{aligned}$$

Now, as each $N_{t_{C+}}$ is a random variable for $C \in \mathcal{C}_n$, for any $H \in \mathcal{B}(\mathbb{R}_+)$

$$\begin{aligned}\{(\omega,t); N_{t_{C+}}(\omega) \in H\} &= N^{-1}_{t_{C+}}(H) \times E \\ &\in \mathcal{F} \otimes \mathcal{E},\end{aligned}$$

and as each such C is a Borel set,

$$\begin{aligned}\{(\omega,t); I_C(t) \in H\} &= \begin{cases}\emptyset & \text{if } 0,1 \notin H \\ \Omega \times C & \text{if } 0 \notin H, 1 \in H \\ \Omega \times C^c & \text{if } 0 \in H, 1 \notin H \\ \Omega \times E & \text{if } 0,1 \in H\end{cases} \\ &\in \mathcal{F} \otimes \mathcal{E}.\end{aligned}$$

Thus N_n, as the product of jointly measurable maps, is itself jointly measurable for each $n \in \mathbb{N}$, and so too must $N = \lim_{n\to\infty} N_n$ be. This verifies ii).

■

2.2.1 The Space of Upper Layers

To begin the task of extending the notion of a change point to dimensions higher than 1 some topological constructs are needed. From this point forward assume that $E = [0, R]$ for some $R \in \mathbb{R}^d_+$.

Definition 2.2.3 *A point process L on E is called a single line process if each of its jump points are incomparable with respect to the partial order on E.*

Definition 2.2.4 *A set $D \subseteq E$ is a lower layer if $s \leq t$ and $t \in D$ together imply $s \in D$. A set U is an upper layer if it is the complement (in E) of a lower layer. $\mathcal{U}$ denotes the collection of closed upper layers on E.*

Definition 2.2.5 *Let* (X, d) *be a metric space and let* $\mathcal{C}$ *denote the set of compact subsets of* (X, d)*. For* $\epsilon > 0, U \in \mathcal{C}$ *let* $B_\epsilon(U) := \{x \in X; \exists u \in U \text{ s.t. } d(u, x) < \epsilon\}$*. The Hausdorff metric induced by* d *is the map* $d_H : \mathcal{C} \to \mathbb{R}_+$ *given by*

$$d_H(U, V) := \max\{\inf\{\epsilon > 0; V \subseteq B_\epsilon(U)\}, \inf\{\epsilon > 0; U \subseteq B_\epsilon(V)\}\}.$$

It is immediate to check that the Hausdorff metric is indeed a metric, and in [9] it is shown that $(\mathcal{C}, d_H)$ is complete if (X, d) is. Let $\mathcal{B}(\mathcal{U})$ be the σ-algebra of subsets of $\mathcal{U}$ that are Borel with respect to the Hausdorff metric induced by d, the metric on $\mathbb{R}^d_+$ induced by the sup norm. Note that this metric is equivalent to the Hausdorff metric induced by the Euclidean metric on E. $(\mathcal{U}, \mathcal{B}(\mathcal{U}))$ is a measure space that will be needed later on.

Lemma 2.2.6 *Let* L *be a single line point process on* E*. Then the map* $\xi : \Omega \to \mathcal{U}$ *given by* $\xi(\omega) = \{s \in E; L_s(\omega) > 0\}$ *is* $\mathcal{F}^L/\mathcal{B}(\mathcal{U})$*-measurable.*

Proof:

For simplicity, take d to be the sup norm. Let $O \in \mathcal{B}(\mathcal{U})$ be a basic open set, say $O = \{V; d_H(V, U) < \epsilon\}$ for some $\epsilon > 0$, $U \in \mathcal{U}$, and let $r > 0$, $\omega \in \Omega$. For $s \in [0, R]$ adopt the notation $s + r = (s_1 + r, ..., s_d + r)$.

First suppose that $\forall s \in \mathbb{Q}^d_+ \cap U$ $L_{s+r}(\omega) > 0$ and let $t \in U$. Then, as U is a closed upper layer, $\exists$ a sequence $s_n \in \mathbb{Q}^d_+ \cap U$ s.t. $s_n \downarrow t$. For each n, $L_{s_n+r}(\omega) > 0$. Since $L(\omega)$ is a measure

$$\begin{aligned} L_{t+r}(\omega) &= L(\omega)(A_{t+r}) \\ &= \lim_{n\to\infty} L(\omega)(A_{s_n+r}) \\ &\geq 1. \end{aligned}$$

As $\xi(\omega)$ is an upper layer, $L_{t+r}(\omega) > 0$ iff $t \in B_r(\xi(\omega))$. Since $t \in U$ was arbitrary, $U \subseteq B_r(\xi(\omega))$. This was deduced from the supposition that $\forall s \in \mathbb{Q}^d_+ \cap U$, $L_{s+r}(\omega) > 0$;

clearly the reverse implication holds and so

$$\{\omega; U \subseteq B_r(\xi(\omega))\} = \bigcap_{s \in \mathbb{Q}^d_+ \cap U} \{\omega; L_{s+r}(\omega) > 0\} \in \mathcal{F}^L.$$

Now suppose that, $\forall s \in \mathbb{Q}^d_+ \cap \xi(\omega)$, $s + r \in U$ and let $t \in \xi(\omega)$. As $\xi(\omega)$ is a closed upper layer, $\exists$ a sequence $(s_n)_{n \geq 1}$ in $\mathbb{Q}^d_+ \cap \xi(\omega)$ s.t. $s_n \downarrow t$. Then, since $s_n + r \in U$ for all n and U is closed, $t + r \in U$. Note that $t + r \in U$ iff $t \in B_r(U)$ since U is an upper layer. As $t \in \xi(\omega)$ was arbitrary it follows that $\xi(\omega) \subseteq B_r(U)$. This was deduced from the supposition that $\forall s \in \mathbb{Q}^d_+ \cap \xi(\omega)$, $s + r \in U$; clearly the reverse implication holds and so

$$\{\omega; \xi(\omega) \subseteq B_r(U)\} = \bigcap_{s \in \mathbb{Q}^d_+ \cap B_r(U)} \{\omega; L_s(\omega) > 0\} \in \mathcal{F}^L.$$

Now

$$\begin{aligned}
\xi^{-1}(O) &= \{d_H(\xi, U) < \epsilon\} \\
&= \{\inf\{r > 0; \xi \subseteq B_r(U)\} < \epsilon\} \cap \{\inf\{r > 0; U \subseteq B_r(\xi)\} < \epsilon\} \\
&= \{\exists r \in \mathbb{Q}_+, r < \epsilon, \xi \subseteq B_r(U)\} \cap \{\exists r \in \mathbb{Q}_+, r < \epsilon, U \subseteq B_r(\xi)\} \\
&= \bigcup_{r \in \mathbb{Q}_+, r < \epsilon} (\{\xi \subseteq B_r(U)\} \cap \{U \subseteq B_r(\xi)\}) \\
&\in \mathcal{F}^L.
\end{aligned}$$

Since O was a basic open set, the above is sufficient to ensure that ξ is measurable from $\mathcal{F}^L$ to $\mathcal{B}(\mathcal{U})$.

■

Lemma 2.2.7 *Let L be a single line point process on E, and let $\xi = \{t \in E; L_t > 0\}$. Then $\mathcal{F}^L \subseteq \mathcal{F}^\xi := \sigma(\xi)$.*

Proof:

For $n \in \mathbb{N}$ let $\mathcal{C}_n$ be the n^{th} rectangular partition of E and define $\mathcal{U}_n$ to be the collection of unions (of elements) from $\mathcal{C}_n$ which are upper layers. If $U \subseteq E \setminus A_t$ define $\delta_U := d(t, U)$. Let $F_t^n = \bigcup_{\{U \in \mathcal{U}_n; U \subseteq E \setminus A_t\}} B_{\delta_U}(U)$, where $B_{\delta_U}(U)$ represents the open ball in $\mathcal{U}$ of radius δ_U around U with respect to the Hausdorff metric d_H. It should be clear that, for large enough n, F_t^n is exactly the set $\{D \in \mathcal{U}; D \subseteq E \setminus A_t\}$. Since it is the union of open subsets of $\mathcal{U}$, it follows that this set, and hence its complement

$$\{D \in \mathcal{U}; E_t \subseteq D\},$$

is in $\mathcal{B}(\mathcal{U})$. Now, since the collection of all subsets of the form $\cap_{i=1}^{j}\{L_{s_i} = k_i\}$ for some $j, k_1, ..., k_j \in \mathbb{N}$, $s_1, ..., s_j \leq R$ is a π-system generating $\mathcal{F}^L$, it is sufficient to show that $\{L_s = k\} \in \mathcal{F}^\xi$ for any $k \in \mathbb{N}$, $s \leq R$. Now $L_s = k$ if and only if A_s contains k "corners" of ξ. This is equivalent to the condition that, for n large enough (so that each element of $\mathcal{C}_n$ contains at most one jump point), there exist distinct elements $C_1, ..., C_k \in \mathcal{C}_n$ such that $C_i \cap A_s \neq \emptyset$ and $t_{C_i}^+ \in \xi$ but $c \notin \xi$ for any of the $2^d - 1$ other corners c of C_i, for each i. Thus

$$\{L_s = k\} = \bigcup_{m \geq 1} \bigcap_{n \geq m} \bigcap_{i=1}^{k} \bigcup_{\{C_i \in \mathcal{C}_n; C_i \cap A_s \neq \emptyset\}} \{t_{C_i}^+ \in \xi\} \setminus \Big(\bigcup_{\{c \text{ a corner of } C_i\}} \{c \in \xi\}\Big),$$

and so it remains to show that $\{t \in \xi\} \in \mathcal{F}^\xi$ for any $t \in E$. However, since $\{t \in \xi\} = \xi^{-1}(\{D \in \mathcal{U}; E_t \subseteq D\})$, this follows from the argument above. Therefore $\mathcal{F}^L \subseteq \mathcal{F}^\xi$.

■

The two preceding lemmas show that $\mathcal{F}^L = \mathcal{F}^\xi$ whenever ξ is defined from the single line point process L as above. Similar arguments show that $\mathcal{F}_s^L = \mathcal{F}^{\xi_s}$, where $\xi_s : \Omega \rightarrow \mathcal{U} \cap A_s := \{U \cap A_s; U \in \mathcal{U}\}$ is the map $\xi_s(\omega) = \xi(\omega) \cap A_s$.

Definition 2.2.8 *Let $\rho : \Omega \rightarrow \mathcal{L}$ be a map to the collection of closed lower layers on E. ρ is a stopping set with respect to the filtration $(\mathcal{F}_t)_{t \in E}$ if*

$$\{t \in \rho\} \in \mathcal{F}_t \; \forall t \in E.$$

ρ is discrete if the set $\{\rho(\omega); \omega \in \Omega\}$ has finite cardinality.

The above definition generalizes the more well-known notion of a *stopping time* on the line, where it is required that $\{\tau \leq t\} \in \mathcal{F}_t$ in order for τ to be a stopping time with respect to a filtration $(\mathcal{F}_t)_{t \in \mathbb{R}_+}$. To see the connection, assume $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ is right-continuous, let $\rho = [0, \tau]$ and note that $\{\tau = t\} = \{\tau \leq t\} \backslash \cup_{s<t, s\in\mathbb{Q}} \{\tau \leq s\} \in \mathcal{F}_t$ and $\{t \leq \tau\} = \{t = \tau\} \cup \{\tau \leq t\}^c$, as well as $\{t \leq \tau\} = \{t \in \rho\}$. Hence, for all t, $\{t \in \rho\} \in \mathcal{F}_t$. Conversely, if this holds then τ is a stopping time: if $(t_n) \downarrow t$ with $t \ll t_n$ for each n, then

$$\{\tau \leq t\} = \cap_{n\geq 1}\{\tau < t_n\} = \cap_{n\geq 1}\{t_n \in \rho\}^c \in \cap_{n\geq 1}\mathcal{F}_{t_n} = \mathcal{F}_t.$$

In 1-dimension, therefore, the notions of stopping set and stopping time are equivalent (so long as the filtration is outer-continuous).

Lemma 2.2.9 *Let $\rho : \Omega \to \mathcal{L}$ be a stopping set. Then ρ is $\mathcal{B}(\mathcal{L})$-measurable.*

Proof:

For simplicity, take d to be the sup norm. Let O be a basic open set in $\mathcal{L}$, say $O = \{V; d_H(U, V) < \epsilon\}$ for some $\epsilon > 0, U \in \mathcal{L}$. Then

$$\begin{aligned}
&\{\omega; \inf\{r > 0; \rho(\omega) \subseteq B_r(U)\} < \epsilon\} \\
&= \bigcup_{\delta\in(0,\epsilon)\cap\mathbb{Q}_+} \{\omega; \rho(\omega) \subseteq \overline{B_\delta(U)}\} \\
&= \bigcup_{\delta\in(0,\epsilon)\cap\mathbb{Q}_+} \bigcap_{t\in\mathbb{Q}_+^d} \{\omega; t \in \rho(\omega) \Rightarrow t \in \overline{B_\delta(U)}\} \\
&= \bigcup_{\delta\in(0,\epsilon)\cap\mathbb{Q}_+} \bigcap_{t\in\mathbb{Q}_+^d \cap \overline{B_\delta(U)}^c} \{\omega; t \in \rho(\omega)\}^c \\
&\in \mathcal{F},
\end{aligned}$$

as well as

$$\{\omega; \inf\{r > 0; U \subseteq B_r(\rho(\omega))\} < \epsilon\}$$

$$\begin{aligned}
&= \bigcup_{\delta\in(0,\epsilon)\cap\mathbb{Q}_+} \{\omega; U \subseteq \overline{B_\delta(\rho(\omega))}\} \\
&= \bigcup_{\delta\in(0,\epsilon)\cap\mathbb{Q}_+} \bigcap_{u\in U\cap\mathbb{Q}^d_+} \{\omega; u \in \overline{B_\delta(\rho(\omega))}\} \\
&= \bigcup_{\delta\in(0,\epsilon)\cap\mathbb{Q}_+} \bigcap_{u\in U\cap\mathbb{Q}^d_+} \{\omega; u-\delta \in \rho(\omega)\} \\
&\in \mathcal{F},
\end{aligned}$$

where $u-\delta := (\max\{u_1-\delta,0\}, ..., \max\{u_d-\delta,0\})$. Since

$$\begin{aligned}
\rho^{-1}(O) &= \{\omega; \inf\{r>0; \rho(\omega)\subseteq B_r(U)\} < \epsilon\} \\
&\quad \bigcap \{\omega; \inf\{r>0; U\subseteq B_r(\rho(\omega))\} < \epsilon\},
\end{aligned}$$

and since O was an arbitrary basic open set, it follows that ρ is measurable.

■

2.2.2 Stopping and Martingales

Recall that the goal is to detect an unobservable border separating two regions with differing intensities of a conditional Poisson process. Thus, given a gain function, as we travel northeast we should be able to "stop" at some point that is optimal with respect to this gain function. Some different gain functions will be investigated later on, but a first step is to make clear what is meant by "stopping".

As has been seen, stopping sets generalize the 1-dimensional analogue of stopping times. However, it remains to be shown that, given a stopping set ρ and a stochastic process X, X_ρ is measurable. This will be done via approximations.

For a function v on E with $R \in \mathbb{R}^2$, one defines the *increment* of v over a rectangle $(s,t] = (s_1,t_1]\times(s_2,t_2]$ via $v(s,t] = v(t_1,t_2)-v(s_1,t_2)-v(t_1,s_2)+v(s_1,s_2)$. The idea is to view v as a set function defined first on the rectangles A_t emanating from the origin via $v(A_t) = v_t$, and then on $(s,t] = A_t \setminus (A_{(s_1,t_2)} \cup A_{(t_1,s_2)})$ by treating v as if it were a signed measure and using the inclusion-exclusion formula. (Increments are defined similarly in higher dimensions via the inclusion exclusion formula). Additionally, one says that v is *outer continuous with inner limits (ocil)* if at each $t \in E$, $t_n \to t \Rightarrow \lim_{n \to \infty} v(t_n)$ exists and $t_n \downarrow t \Rightarrow v(t_n) \to v(t)$.

Definition 2.2.10 *A function v on E is increasing (decreasing) if*

- *v is 0 on the axes,*
- *v is outer continuous with inner limits*
- *for every $s \leq t \in E$, $v(s,t] \geq \; (\leq) 0$.*

A stochastic process $V = (V_t)_{t \in E}$ is increasing (decreasing) if, for each $\omega \in \Omega$, the function $V.(\omega)$ is increasing (decreasing).

Let X be an increasing stochastic process on E. Theorem 12.5 of [5] implies that, for every $\omega \in \Omega$, $\exists$ a unique measure $\mu(\omega)$ on E s.t. $\mu(\omega)(s,t] = X(\omega)(s,t]$. X will also be used to denote this measure; in this way increasing processes are extended to Borel sets. Clearly a process M, if it is the difference of increasing stochastic processes, can be similarly extended with $M(\omega)$ corresponding to a signed measure for each ω. It follows that for sets A_n, A satisfying $A_n \to A$ (i.e. $A = \limsup A_n = \liminf A_n$), $X(A_n) \to X(A)$ also holds.

Lemma 2.2.11 *Let ρ be a stopping set on E. There exists a sequence $(\rho_n^+)_{n \geq 1}$ of discrete stopping sets satisfying $\rho_n^+ \downarrow \rho$.*

Proof:

Let $\mathcal{C}_n$ be the n^{th} rectangular partition of E and define the n^{th}-approximate of ρ to be $\rho_n^+ = \cup_{C \in \mathcal{C}_n, t_{C\_} \in \rho} C$. Fix $\omega \in \Omega$, let $u \in \rho_n^+(\omega)$ and let $s \leq u$. Let C_1 and C_2 be the unique elements of $\mathcal{C}_n$ containing s and u respectively. Then, as $s \leq u$, $t_{C_{1\_}} \leq t_{C_{2\_}} \in \rho(\omega)$ which implies that $t_{C_{1\_}} \in \rho(\omega)$ since $\rho(\omega)$ is a lower layer. So $s \in C_1 \subseteq \rho_n^+(\omega)$ and it follows that $\rho_n^+(\omega)$ is a lower layer.

Now let $u \in E$ be fixed, and let $C \in \mathcal{C}_n$ be unique such that $u \in C$. Then $\{u \in \rho_n^+\} = \{t_{C_-} \in \rho\} \in \mathcal{F}_{t_{C_-}} \subseteq \mathcal{F}_u$. So ρ_n^+ is a stopping set. Since each $\mathcal{C}_n$ is finite it is clear that the ρ_n^+ are discrete.

It remains to show that, for each $\omega \in \Omega$, $\rho_n^+(\omega) \longrightarrow \rho(\omega)$ with respect to d_H (where d_H is constructed from the sup-norm). Let $\epsilon > 0$ and choose n large enough so that $\frac{r_1}{2^n}, ..., \frac{r_d}{2^n} < \epsilon$. Let $m \geq n$ and $u \in \rho_m^+(\omega)$. Then $\exists C \in \mathcal{C}_m$ such that $u \in C$ and $t_{C_-} \in \rho(\omega)$. Since $t_{C_-} + \epsilon \gg t_{C_-} + (\frac{r_1}{2^m}, ..., \frac{r_d}{2^m}) \geq u$, it follows that $u \in B_\epsilon(\rho(\omega))$ and hence that $\rho_m^+(\omega) \subseteq B_\epsilon(\rho(\omega))$ as u was arbitrary. Since $\rho(\omega) \subseteq \rho_m^+(\omega)$ for all m, the conclusion is that $d_H(\rho_m^+(\omega), \rho(\omega)) \leq \epsilon$ for all $m \geq n$. Thus $(\rho_n^+)_{n \geq 1}$ is indeed a sequence of discrete stopping sets converging to ρ from above.

■

Corollary 2.2.12 *If M is a stochastic process which can be expressed as the difference of two increasing processes X and Y on $E = [0, R]$ and ρ is a stopping set, then the map $M_\rho : \Omega \longrightarrow \mathbb{R}_+$ given by $M_\rho(\omega) = M(\omega)(\rho(\omega))$ is a random variable.*

Proof:

For each $n \in \mathbb{N}$

$$X_{\rho_n^+} = \sum_{C \in \mathcal{C}_n} X(C) I_{\{t_{C_-} \in \rho\}}$$

is a random variable. Since X is increasing and $\rho_n^+(\omega) \downarrow \rho(\omega)$ for all ω, it follows that $X(\omega)(\rho_n^+(\omega)) \longrightarrow X(\omega)(\rho(\omega))$ for all ω. Thus $X_\rho = \lim_{n \to \infty} X_{\rho_n^+}$ is a random variable. Similarly $Y_\rho = \lim_{n \to \infty} Y_{\rho_n^+}$ is a random variable. Therefore M is as well (and $M_\rho = \lim_{n \to \infty} M_{\rho_n^+}$).

■

Definition 2.2.13 *Let $M = (M_t)_{t \in E}$ be an integrable stochastic process on E, adapted to a filtration $\mathcal{F} = (\mathcal{F}_t)_{t \in E}$. M is a weak $\mathcal{F}$-(sub)martingale if M is equal to 0 on the* *axes and, for every* $s \leq t \in E$,

$$\mathbb{E}[M(s,t] | \mathcal{F}_s] = (\geq) 0.$$

Proposition 2.2.14 *Let* M *be a weak (sub)martingale expressible as the difference of two increasing integrable processes. Then for any stopping set* ρ *on* E, $\mathbb{E}(M_\rho) = 0 \ (\geq 0)$.

Proof:

Assume that M is a weak martingale as only superficial changes need to be made to adapt the proof to the sub-martingale case. Let ρ_n^+ be the n^{th}-approximate of ρ. As shown in the previous corollary, $M_{\rho_n^+} \to M_\rho$ almost surely.

Let X and Y be increasing, integrable processes such that $M = X - Y$. As $\rho \subseteq [0, R]$, $X_\rho(\omega) \leq X_R(\omega)$ for each ω so that $\mathbb{E}[X_\rho] \leq \mathbb{E}[X_R]$. Similarly, $\mathbb{E}[Y_\rho] \leq \mathbb{E}[Y_R]$ and so M_ρ is integrable. Now,

$$M(\rho_n^+) = \sum_{C \in \mathcal{C}_n} M(C) I_{\{t_C\text{-} \in \rho\}}.$$

Taking the expectation yields

$$\begin{aligned} \mathbb{E}(M_{\rho_n^+}) &= \sum_{C \in \mathcal{C}_n} \mathbb{E}(M(C) \cdot I_{\{t_C\text{-} \in \rho\}}) \\ &= \sum_{C \in \mathcal{C}_n} \mathbb{E}(I_{\{t_C\text{-} \in \rho\}} \cdot \mathbb{E}[M_C | \mathcal{F}_{t_C\text{-}}]) \\ &= 0 \end{aligned}$$

since M is a weak martingale. Now, by dominated convergence ($X_R + Y_R$ dominates M_ρ and $M_{\rho_n^+}$) it follows that $\mathbb{E}(M_\rho) = 0$.

■

2.2.3 The Conditional Poisson Process

In the current model the "border" dividing the regions with differing intensities of random points corresponds to the boundary of a random upper layer ξ, and the random points correspond to the jump points of a spatial point process N that has, conditional upon ξ, a Poisson distribution with intensity μ_0 on ξ^c and μ_1 on ξ. The constructions involved are remarkably analogous to their 1-dimensional counterparts.

Lemma 2.2.15 *Let L be a single line point process on E, $\xi = \{t; L_t > 0\}$, $\mu_0, \mu_1 \in \mathbb{R}_+$ and define $\mu : \Omega \longrightarrow \mathfrak{M}$ via*

$$\mu(\omega)(A) := \mu_0|A \cap \xi^c(\omega)| + \mu_1|A \cap \xi(\omega)|$$

for $A \in \mathcal{E}$. Then μ is a random measure on E.

Proof:

First note that since L is a point process it has outer-continuous sample paths and thus is jointly-measurable by lemma 2.2.2. Fix $A \in \mathcal{E} = \mathcal{B}([0, R])$ and let $f : \Omega \times E \longrightarrow \mathbb{R}$ be the map $f(\omega, t) = I_{\{(\omega,t);L_t(\omega)=0\}}$. If $B \in \mathcal{B}(\mathbb{R}_+)$ then,

$$\{(\omega, t); f(\omega, t) \in B\} = \begin{cases} \emptyset & 0, 1 \notin B \\ \Omega \times E & 0, 1 \in B \\ L^{-1}(\{0\}^c) & 0 \in B, 1 \notin B \\ L^{-1}(\{0\}) & 0 \notin B, 1 \in B \end{cases} \tag{2.2.1}$$

and hence, by the joint-measurability of L, f is jointly-measurable. Then the section f_ω of f is $\mathcal{E}$-measurable (and so too is $f_\omega I_A$) and the map $\omega \mapsto \int f_\omega(t) I_A(t) dt$ is $\mathcal{F}$-measurable by Fubini's theorem. But this is the map $\omega \mapsto |\xi^c(\omega) \cap A|$, and similar arguments show that the map $\omega \mapsto |\xi(\omega) \cap A|$ is measurable. Hence the map $\omega \mapsto \mu(\omega)(A)$ is measurable.

To conclude, let $M \in \mathcal{M}$, say $M = \{\rho \in \mathfrak{M}; \rho(G) \in H\}$ for some $G \in \mathcal{E}$, $H \in \mathcal{B}(\mathbb{R}_+)$. Then

$$\begin{aligned} \mu^{-1}(M) &= \{\omega; \mu(\omega)(G) \in H\} \\ &= \mu(G)^{-1}(H) \\ &\in \mathcal{F}. \end{aligned}$$

Therefore μ is a random measure.

■

Note that the previous lemma shows in fact that, for every $A \in \mathcal{E}$, $\mu(A)$ is measurable $\mathcal{F}^L = \mathcal{F}^\xi$.

Definition 2.2.16 *Let N be a point process on E, L a single line point process, $\xi = \{t \in E; L_t > 0\}$, and let $\mu_0, \mu_1 \in \mathbb{R}_+$. Let μ be the random measure defined on $A \in \mathcal{E}$ by $\mu(\omega)(A) = \mu_0|A \cap \xi^c(\omega)| + \mu_1|A \cap \xi(\omega)|$. Then N has a conditional Poisson distribution given ξ with rate μ_0 on ξ^c and μ_1 on ξ if*

i) for $k \in \mathbb{N}$ and $A \in \mathcal{B}(E)$, $P[N(A) = k|\xi] = \frac{\mu(A)^k}{k!}e^{-\mu(A)}$

ii) for $A_1, ..., A_n \in \mathcal{B}(E)$ disjoint,
$P(N(A_i) = k_i, 1 \leq i \leq n|\xi) = \prod_{i=1}^n P(N(A_i) = k_i|\xi)$.

ξ is referred to as the change set.

Consider a single line process L and ξ as in the above definition, and recall that $(\mathcal{U}, \mathcal{B}(\mathcal{U}))$ is the space of upper layers on E. If $s \in E$ one may consider the space $\mathcal{U} \cap A_s = \{U \cap A_s; U \in \mathcal{U}\}$ together with the σ-field $\mathcal{B}(\mathcal{U} \cap A_s)$ traced by A_s. The map $\xi_s : \Omega \to \mathcal{U} \cap A_s$ given by $\xi_s(\omega) = \xi(\omega) \cap A_s$ projects the image of ξ onto A_s, and by arguments similar to those in lemmas 2.2.6 and 2.2.7, it is easily seen that $\sigma(\xi_s) = \mathcal{F}_s^L$.

Suppose that, given ξ, N is conditionally Poisson with mean measure μ given by definition 2.2.16 and let $G \in \sigma(\xi_s)$ be arbitrary. Since $\sigma(\xi_s) = \mathcal{F}^L_s \subseteq \mathcal{F}^L = \mathcal{F}^\xi$ by lemmas 2.2.6 and 2.2.7, it follows that for any $A \in \mathcal{E} \cap A_s$,

$$\begin{aligned}
&\int_G P[N(A) = k|\xi_s]dP \\
&= \int_G P[N(A) = k|\xi]dP \\
&= \int_G \frac{(\mu_0|\xi^c \cap A_s \cap A| + \mu_1|\xi \cap A_s \cap A|)^k}{k!} e^{-(\mu_0|\xi^c\cap A_s\cap A|+\mu_1|\xi\cap A_s\cap A|)} dP \\
&= \int_G \frac{(\mu_0|(A_s \setminus \xi_s) \cap A| + \mu_1|\xi_s \cap A|)^k}{k!} e^{-(\mu_0|(A_s\setminus\xi_s)\cap A|+\mu_1|\xi_s\cap A|)} dP.
\end{aligned}$$

One may replace, in the first term above, N with the process $N|_{A_s}$ on $(A_s, \mathcal{E} \cap A_s)$ given by $N|_{A_s}(\omega)(A_u) = N_u(\omega)$ for all $u \leq s$ (so that $\mathcal{F}^{N|_{A_s}}_u = \mathcal{F}^N_u$ for all $u \leq s$). It follows that, given ξ_s, $N|_{A_s}$ is conditionally Poisson on $(A_s, \mathcal{E} \cap A_s)$ with mean measure μ_s given by

$$\mu_s(A) := \mu_0|(A_s \setminus \xi_s) \cap A| + \mu_1|\xi_s \cap A|.$$

Following arguments similar to those in lemma 2.2.15 one can show that $\mu_s(A)$ is measurable $\mathcal{F}^{\xi_s} = \mathcal{F}^L_s$ for every $A \in \mathcal{E} \cap A_s$. Since $A \in \mathcal{E} \cap A_s$ implies

$$\int_G P[N(A) = k|\xi]dP = \int_G \frac{(\mu_0|(A_s \setminus \xi_s) \cap A| + \mu_1|\xi_s \cap A|)^k}{k!} e^{-(\mu_0|(A_s\setminus\xi_s)\cap A|+\mu_1|\xi_s\cap A|)} dP$$

for all $G \in \mathcal{F}^\xi$, it follows that $P[N(A) = k|\xi]$ is measurable $\mathcal{F}^{\xi_s}$ and hence that $P[N(A) = k|\xi] = P[N|_{A_s}(A) = k|\xi_s]$ whenever $A \in \mathcal{E} \cap A_s$. Now, if B is of the form $\bigcap_{i=1}^j \{N_{u_i} = k_i\}$ for some $u_i \leq s$ and $k_i \in \mathbb{N}$ then, since $N(\omega)$ is a measure for fixed ω, B can be written as $\bigcap_{i=1}^\ell \{N(S_i) = h_i\}$ for some disjoint sets $S_1, ..., S_\ell \in \mathcal{E} \cap A_s$ and some $h_1, ..., h_\ell \in \mathbb{N}$. Then

$$P[B|\xi] = \prod_{i=1}^\ell P[N(S_i) = h_i|\xi] = \prod_{i=1}^\ell P[N|_{A_s}(S_i) = h_i|\xi_s] = P[B|\xi_s].$$

Sets of the form B form a π-system generating $\mathcal{F}_s^N$, and so one may conclude that $P[A|\xi] = P[A|\xi_s]$ for any $A \in \mathcal{F}_s^N$. Then, for any such A,

$$\begin{aligned} P[A|\mathcal{F}_s^L, \mathcal{F}_{(s,t]}^L] &= \mathbb{E}[\mathbb{E}[I_A|\xi]|\mathcal{F}_s^L, \mathcal{F}_{(s,t]}^L] \\ &= \mathbb{E}[P[A|\xi]|\mathcal{F}_s^L, \mathcal{F}_{(s,t]}^L] \\ &= \mathbb{E}[P[A|\xi_s]|\mathcal{F}_s^L, \mathcal{F}_{(s,t]}^L] \\ &= P[A|\xi_s] \\ &= P[A|\mathcal{F}_s^L] \end{aligned}$$

since $\mathcal{F}_s^L, \mathcal{F}_{(s,t]}^L \subseteq \mathcal{F}^\xi := \sigma(\xi)$ (here $\mathcal{F}_{(s,t]}^L := \sigma(L_x; x \in (s,t])$). Therefore, if we define $\mathcal{F} \perp\!\!\!\perp_{\mathcal{G}} \mathcal{H}$ to be the condition

$$P[H|\mathcal{G}, \mathcal{F}] = P[H|\mathcal{G}] \; \forall H \in \mathcal{H}$$

for σ-fields $\mathcal{F}, \mathcal{G}, \mathcal{H}$, see [7], then one has the following:

Lemma 2.2.17 *Let L be a single line point process, let $\xi = \{t \in E; L_t > 0\}$, and let N be conditionally Poisson given ξ. Then $\mathcal{F}_s^N \perp\!\!\!\perp_{\mathcal{F}_s^L} \mathcal{F}_{(s,t]}^L$ for any $s \in E$.*

Lemma 2.2.18 *Let J be a non-explosive point process, and for each ω let $\Delta_J(\omega) = \{t \in E; J(\omega)(\{t\}) > 0\}$, $\Delta_{J_1}(\omega) = \{t \in \Delta_J(\omega); t' \not\leq t \; \forall t' \in \Delta_J$ such that $t' \neq t\}$. The map $J_1 : \Omega \times \mathcal{E} \to \mathbb{R}_+$ given by $J_1(\omega)(A) = \sum_{t \in \Delta_{J_1}(\omega)} \delta_t(A)$ is a single line point process and is referred to as the first line of J.*

Proof:

For each $n \in \mathbb{N}, A \in \mathcal{E}$ let

$$J_{1n}(A) = \sum_{C \in \mathcal{C}_n} J(C \cap A) \cdot I_{\{J_{t_C^+} \leq 1\}}.$$

Clearly $J_{1n}(A)$ is a random variable for each n and A. Furthermore, it is clear that for n large enough it is precisely the points of $\Delta_{J_1} \cap A$ that receive a mass from $J_{1m}(A)$

for $m \geq n$. Hence

$$J_1(A) = \lim_{n\to\infty} J_{1n}(A)$$

is a random variable for each $A \in \mathcal{E}$. It is also clear that $J_1(\omega)$ is a measure for each ω (it is just the sum of diracs). Finally, if $M \in \mathcal{N}$, say $M = \{\mu \in \mathfrak{N}; \mu(B) \in H\}$ for some $B \in \mathcal{E}, H \in \mathcal{B}(\mathbb{R}_+)$, then

$$\begin{aligned} \{\omega; J_1(\omega) \in M\} &= \{\omega; J_1(\omega)(B) \in H\} \\ &= J_1(B)^{-1}(H) \\ &\in \mathcal{F}. \end{aligned}$$

It follows that J_1 is a point process, and it is a single line process by construction.

■

The set Δ_{J_1} of lemma 2.2.18 is sometimes denoted $\min(\Delta_J)$.

Lemma 2.2.19 *Let L be a single line point process on E, let ξ be the random upper layer $\xi = \{t \in E; L_t > 0\}$, and let $(\mathcal{F}_t)_{t\in E}$ be a filtration satisfying $\mathcal{F}_t^L \subseteq \mathcal{F}_t$ for each $t \in E$. Then $\overline{\xi^c}$ is a stopping set with respect to $(\mathcal{F}_t)_{t\in E}$.*

Proof:

It is clear that the complement of an upper layer is a lower layer, hence $\overline{\xi(\omega)^c} \in \mathcal{L}$ for each ω and the question is whether $\{t \in \overline{\xi^c}\} \in \mathcal{F}_t$ for every $t \in E$. Note that

$$\begin{aligned} \{t \in \partial\xi\} &= \Big(\bigcap_{s\in\mathbb{Q}^d\cap\mathrm{Int}(A_t)} \{L_s = 0\}\Big) \bigcap \{L_t > 0\} \\ &\in \mathcal{F}_t^L \end{aligned}$$

(here $\partial\xi$ denotes the boundary of ξ). Then

$$\begin{aligned} \{t \in \overline{\xi^c}\} &= \{t \in \xi\}^c \bigcup \{t \in \partial\xi\} \\ &= \{L_t > 0\}^c \bigcup \{t \in \partial\xi\} \end{aligned}$$

$$\begin{aligned} &\in \mathcal{F}_t^L \\ &\subseteq \mathcal{F}_t \end{aligned}$$

and so $\overline{\xi^c}$ is an $(\mathcal{F}_t)_{t\in E}$-stopping set.

■

The change set of the conditional Poisson process N is generated by the single line process L. There are many ways that L can arise: it may consist of a single jump point, or it may be the first line of some other point process. Of particular interest is the case that L is the first line of a spatial Poisson process J, because this case is the higher dimensional analogue to the case considered in [1] where the change point σ is the first jump point of a Poisson process on the line.

2.3 The Likelihood Function

We desire that an observer in our model moving "northeast" have the capacity to decide, based on observations of the process N up to his position $t \in E$ (i.e. based on $\mathcal{F}_t^N$), whether it is likely that he has entered the change set ξ. That is, we should like for him to know $P[t \in \xi | \mathcal{F}_t^N]$. In order to manipulate this expression for the purposes of solving the optimal detection problem, formula (18) given in [2] is needed, and it will be rigorously justified in this section. For clarity of exposition and to avoid cumbersome notation, the arguments will take place in $\mathbb{R}_+^2$ although each holds in the general setting of $\mathbb{R}_+^d$.

For $t \in E, C \in \mathcal{B}(A_t)$, let $\mathbb{X}_{t,C}$ be the collection of sequences $(x_1, ..., x_k, \infty, \infty, ...)$ such that, for some $0 \leq i \leq k$, $x_1, ..., x_i \in C, x_{i+1}, ..., x_k \in A_t \setminus C$ and $x_{1,1} < ... < x_{i,1}, 0 \leq x_{i+1,1} < ... < x_{k,1} \leq R_1$. (That is, the points contained in C are ordered with respect to their first coordinate and precede the points lying outside C, which are similarly ordered amongst themselves). Let $n : \mathbb{X}_{t,C} \to \mathfrak{N}$ be the map $n(\tilde{x}) = \sum_{i=1}^{\infty} \delta_{x_i}$

(where $\delta_\infty := 0$), and let $\mathcal{X}_{t,C} = \sigma(n)$. Note that n has the same form for each $t \in E, C \in \mathcal{B}(A_t)$, and that n is injective. Furthermore, if $\mathfrak{N}(t) = \{\mu|_{A_t}; \mu \in \mathfrak{N}\}$ and $\mathfrak{N}_v(t)$ is the subset of $\mathfrak{N}(t)$ consisting of measures μ never giving mass to more than one element of any vertical line, then n is surjective onto $\mathfrak{N}_v(t)$ (see [6]). Given a point process N on E such that $N(\omega) \in \mathfrak{N}_v(R)$ for every ω, define $N^{\mathbb{X}}_{t,C} : \Omega \to \mathbb{X}_{t,C}$ via

$$N^{\mathbb{X}}_{t,C}(\omega) = (x_1, ..., x_k, \infty, \infty, ...)$$

where $x_1, ..., x_k \in A_t$ satisfy $N(\omega)(\{x_1\} \cup ... \cup \{x_k\}) = N(\omega)(A_t)$ and $(x_1, ..., x_k, \infty, \infty, ...) \in \mathbb{X}_{t,C}$. Note that $n \circ N^{\mathbb{X}}_{t,C} = N|_{A_t}$ for any $t \in E, C \in \mathcal{B}(A_t)$. Therefore, for a set $S \in \mathcal{X}_{t,C}$, say $S = \{\tilde{x} \in \mathbb{X}_{t,C}; n(\tilde{x}) \in W\}$ for some $W \in \mathcal{N}_v(t) := \mathcal{N} \cap \mathfrak{N}_v(t)$, say $W = \{\mu \in \mathfrak{N}_v(t); \mu(B) = k\}$ for some $B \in \mathcal{B}(A_t)$, $m \in \mathbb{N}$, one has

$$\begin{aligned} (N^{\mathbb{X}}_{t,C})^{-1}(S) &= \{\omega; n(N^{\mathbb{X}}_{t,C}(\omega)) \in W\} \\ &= \{\omega; N|_{A_t}(\omega)(B) = k\} \\ &= \{\omega; N(\omega)(B) = k\} \\ &= N(B)^{-1}(\{k\}) \\ &\in \mathcal{F}. \end{aligned}$$

Thus $N^{\mathbb{X}}_{t,C}$ is measurable $\mathcal{F}/\mathcal{X}_{t,C}$. Let $\nu_{t,C}$ be the measure $P(N^{\mathbb{X}}_{t,C})^{-1}$ on $(\mathbb{X}_{t,C}, \mathcal{X}_{t,C})$.

Definition 2.3.1 *(Daly, Vere-Jones [8]). Let N be a point process on E, and for $x \in E$ define*

$$\Delta^n_x := [(x_1, x_2), (x_1 + \frac{1}{n}, x_2 + \frac{1}{n})].$$

The likelihood function $\ell_N(t) : \mathfrak{N}_v(t) \to \mathbb{R}_+$ is defined as follows:

$$\ell_N(t)(\mu) := \lim_{n \to \infty} \frac{P[N(A_t \cap \Delta^n_{x_1}) = ... = N(A_t \cap \Delta^n_{x_k}) = 1, N(A_t \setminus \cup_{i=1}^k \Delta^n_{x_i}) = 0]}{|\Delta^n_{x_1} \times ... \times \Delta^n_{x_k}|}$$

where $(x_1, ..., x_k, \infty, \infty, ...) = n^{-1}(\mu) \in \mathbb{X}_{t,C}$, and $|\cdot|$ represents Lebesgue measure.

Of course the denominator in the above expression is simply $(\frac{1}{n^2})^k$, but the representation given in the definition is more suggestive of what is going on. At times, the notation $\ell_N(t)(x_1, ..., x_k)$ or $\ell_N(t)(\tilde{x})$ will be used in place of $\ell_N(t)(\mu)$ when it is clear from context that $\mu = n(x_1, ..., x_k, \infty, \infty, ...) = n(\tilde{x})$.

Definition 2.3.2 *(Daly, Vere-Jones [8]). Let N be a point process on E. Let L be a single line point process on E, and let ξ be the upper layer $\{t; L_t > 0\}$. The conditional likelihood function $\ell_{N|\xi}(t) : \mathfrak{N}_v(t) \times \Omega \to \mathbb{R}_+$ is defined as follows:*

$$\ell_{N|\xi}(t)(\mu, \omega) := \lim_{n\to\infty} \frac{P[N(A_t \cap \Delta^n_{x_1}) = ... = N(A_t \cap \Delta^n_{x_k}) = 1, N(A_t \setminus \cup^k_{i=1}\Delta^n_{x_i}) = 0|\xi](\omega)}{|\Delta^n_{x_1} \times ... \times \Delta^n_{x_k}|}$$

where $(x_1, ..., x_k, \infty, \infty, ...) = n^{-1}(\mu) \in \mathbb{X}_{t,C}$.

In order to determine the likelihood function of a Poisson process N with mean measure μ, note that for distinct x_i's and large enough n, the $\Delta^n_{x_i}$ are disjoint and hence the $N(\Delta^n_{x_i})$ and $N(A_t \setminus \cup^k_{i=1}\Delta^n_{x_i})$ are independent. Thus the likelihood function for N is

$$\begin{aligned} \ell_N(t)(x_1, ..., x_k) &= \lim_{n\to\infty} \frac{P[N(A_t \setminus \cup^k_{i=1}\Delta^n_{x_i}) = 0] \prod^k_{i=1} P[N(\Delta^n_{x_i}) = 1]}{|\Delta^n_{x_1} \times ... \times \Delta^n_{x_k}|} \\ &= \lim_{n\to\infty} (n^2)^k e^{-\mu(A_t \setminus \cup^k_{i=1}\Delta^n_{x_i})} \prod_{i=1}^k \mu(\Delta^n_{x_i}) e^{-\mu(\Delta^n_{x_i})}. \end{aligned}$$

In case μ is of the form $\mu_0|\cdot|$ for some constant $\mu_0 \geq 0$ notice that $\mu(\Delta^n_{x_i})e^{-\mu(\Delta^n_{x_i})} = \frac{\mu_0}{n^2} \sum_{j=0}^{\infty} \frac{-\mu_0^j}{n^{2j} j!} = \frac{\mu_0}{n^2} + o(\frac{1}{n^3})$ so that $n^2\mu(\Delta^n_{x_i})e^{-\mu(\Delta^n_{x_i})} = \mu_0 + o(\frac{1}{n})$. Therefore, in this case, the likelihood function takes the form $\ell_N(t)(x_1, ..., x_k) = \mu_0^k e^{-\mu_0|A_t|}$. Notice that $\ell_N(t)$ does not discern the locations of the points $x_1, ..., x_k$ - only their number. Thus $\ell_N(t)(\tilde{x})$ has the form $\mu_0^{n(\tilde{x})(A_t)} e^{-\mu_0|A_t|}$ for all $\tilde{x} \in \mathbb{X}_{t,C}$, and so one may represent the likelihood function with the more concise

$$\ell_N(t) = \mu_0^{n(A_t)} e^{-\mu_0|A_t|}.$$

Suppose now that N and L are as in definition 2.2.16: that is, L is a single line process and, given $\xi = \{t; L_t > 0\}$, N is conditionally Poisson with rate μ_0 on ξ^c and μ_1 on ξ. Fix $\omega \in \Omega$ and $x_i \in E$. According to the definition,

$$P[N(\Delta^n_{x_i}) = 1|\xi] \;=\; (\mu_0|\xi^c \cap \Delta^n_{x_i}| + \mu_1|\xi \cap \Delta^n_{x_i}|)e^{-(\mu_0|\xi^c\cap\Delta^n_{x_i}|+\mu_1|\xi\cap\Delta^n_{x_i}|)}.$$

If $x_i \in \xi^c(\omega)$ then, since $\xi^c(\omega)$ is open on its northeast border, for n large enough $\Delta^n_{x_i} \cap \xi^c(\omega) = \Delta^n_{x_i}$ and $P[N(\Delta^n_{x_i}) = 1|\xi](\omega) = \mu_0|\Delta^n_{x_i}|e^{-\mu_0|\Delta^n_{x_i}|} = \frac{\mu_0}{n^2} + o(\frac{1}{n^3})$. Also, since ξ is an upper layer, $x_i \in \xi(\omega)$ implies $\Delta^n_{x_i} \subseteq \xi(\omega)$ and hence $P[N(\Delta^n_{x_i}) = 1|\xi](\omega) = \frac{\mu_1}{n^2} + o(\frac{1}{n^3})$. Since the number of $x_i \in \xi^c(\omega)$ is precisely $n(\tilde{x})(A_t \cap \xi^c(\omega))$ and the number of $x_i \in \xi(\omega)$ is precisely $n(\tilde{x})(A_t \cap \xi(\omega))$, it follows that

$$\begin{aligned}(n^2)^k \prod_{i=1}^{k} P[N(\Delta^n_{x_i}) = 1|\xi](\omega) \;&=\; (\mu_0 + o(\frac{1}{n}))^{n(\tilde{x})(A_t\cap\xi^c(\omega))} \\ &\quad \times(\mu_1 + o(\frac{1}{n}))^{n(\tilde{x})(A_t\cap\xi(\omega))} \\ &\underset{n\to\infty}{\longrightarrow}\; \mu_0^{n(\tilde{x})(A_t\cap\xi^c(\omega))}\mu_1^{n(\tilde{x})(A_t\cap\xi(\omega))}.\end{aligned}$$

Furthermore,

$$\begin{aligned}P[N(A_t \setminus \cup_{i=1}^k \Delta^n_{x_i}) = 0|\xi] \;&=\; e^{-\mu_0|(A_t\setminus\cup_{i=1}^k\Delta^n_{x_i})\cap\xi^c|} \\ &\quad \times e^{-\mu_1|(A_t\setminus\cup_{i=1}^k\Delta^n_{x_i})\cap\xi|} \\ &\underset{n\to\infty}{\longrightarrow}\; e^{-\mu_0|A_t\cap\xi^c|}e^{-\mu_1|A_t\cap\xi|}.\end{aligned}$$

Since the $\Delta^n_{x_i}$ are disjoint for n large enough, the conditional likelihood of N given ξ is

$$\begin{aligned}\ell_{N|\xi}(t)(\tilde{x}) \;&=\; \lim_{n\to\infty}(n^2)^k P[N(A_t \setminus \cup_{i=1}^k \Delta^n_{x_i}) = 0|\xi] \\ &\quad \times \prod_{i=1}^{k} P[N(\Delta^n_{x_i}) = 1|\xi] \\ &=\; e^{-\mu_0|A_t\cap\xi^c|}e^{-\mu_1|A_t\cap\xi|}\mu_0^{n(\tilde{x})(A_t\cap\xi^c)}\mu_1^{n(\tilde{x})(A_t\cap\xi)}\end{aligned}$$

$$= e^{-\mu_0|A_t|}\mu_0^{n(\tilde{x})(A_t)}e^{-(\mu_1-\mu_0)|A_t\cap\xi|}\left(\frac{\mu_1}{\mu_0}\right)^{n(\tilde{x})(A_t\cap\xi)}.$$

Since $\ell_{N|\xi}(t)(x_1,...,x_k)$ is the same for all $\tilde{x}\in\mathbb{X}_{t,C}$, it is abbreviated to $\ell_{N|\xi}(t)$. When it is desirable to forget about Ω we will often write $\ell_{N|\xi=D}(t)$ and work over the space of upper layers. Note that

$$n(\tilde{x})(A_t\cap\xi) = \log\left(\frac{\ell_{N|\xi}(t)e^{\mu_0|A_t|}e^{(\mu_1-\mu_0)|A_t\cap\xi|}}{\mu_0^{n(\tilde{x})(A_t)}}\right)/\log\left(\frac{\mu_0}{\mu_1}\right)$$

is a measurable function of ω since $\omega\mapsto|A_t\cap\xi(\omega)|$ is (see lemma 2.2.15).

Note also that the preceding argument shows that there exists a measurable function dominating $P[N(\Delta^n_{x_1})=1,...,N(\Delta^n_{x_k})=1,N(A_t\setminus\cup_{i=1}^k\Delta^n_{x_i})=0|\xi]$ for each n, and hence that the dominated convergence theorem may be applied to conclude

$$\ell_N(t)=\int_\Omega\ell_{N|\xi}(t)dP=\int_{\mathcal{U}}\ell_{N|\xi=D}(t)d(P\xi^{-1})(D).$$

Also, if $\ell_{N|L_t=0}(t):=\frac{1}{P[L_t=0]}\int_{\{L_t=0\}}\ell_{N|\xi}(t)dP$ then

$$\ell_{N|L_t=0}(t)=\frac{1}{P[L_t=0]}\int_{\{L_t=0\}}e^{-\mu_0|A_t|}\mu_0^{n(\tilde{x})(A_t)}dP=e^{-\mu_0|A_t|}\mu_0^{n(\tilde{x})(A_t)}.$$

Now, the conditional mean measure for N is the random measure μ given by $\mu(A)=\mu_0|A\cap\xi^c|+\mu_1|A\cap\xi|=\int_A f(x)dx$ where $f(x)=\mu_0 I_{\{x\in\xi^c\}}+\mu_1 I_{\{x\in\xi\}}$. So, with probability 1,

$$\begin{aligned}&\sum_{m=0}^{\infty}\frac{1}{m!k!}e^{-\mu(A_t)}\int_{(A_t\setminus B)^m}\int_{(B)^k}\prod_{i=1}^{m}f(x_i)dx_1...dx_{k+m}\\ &= \sum_{m=0}^{\infty}\frac{1}{m!k!}e^{-\mu(A_t)}\mu(A_t\setminus B)^m\mu(B)^k\\ &= \frac{\mu(B)^k}{k!}e^{-\mu(A_t)+\mu(A_t\setminus B)}\\ &= \frac{\mu(B)^k}{k!}e^{-\mu(B)}\end{aligned}$$

$$= P[N(B) = k|\xi]$$

for all Borel sets $B \subseteq A_t$. Now notice that, for a collection $x_1, ..., x_{k+m}$ of points from A_t, $e^{-\mu(A_t)} \prod_{i=1}^{k+m} f(x_i) = e^{-\mu(A_t)} \mu_0^{n(\tilde{x})(A_t \cap \xi^c)} \mu_1^{n(\tilde{x})(A_t \cap \xi)} = \ell_{N|\xi}(n(\tilde{x}))(t)$ where $\tilde{x} = (x_1, ..., x_{k+m}, \infty, \infty, ...) \in \mathbb{X}_{t,B}$. So for a typical set $S = \{\tilde{x} \in \mathbb{X}_{t,B}; n(\tilde{x})(B) = k\} \in \mathcal{X}_{t,B}$ where k is some positive integer and $B \in \mathcal{B}(A_t)$, Fubini's theorem yields

$$\begin{aligned}
\nu_{t,B}(S) &= P[N(B) = k] \\
&= \int_\Omega P[N(B) = k|\xi] dP \\
&= \int_\Omega \sum_{m=0}^{\infty} \frac{1}{m!k!} \int_{(A_t \setminus B)^m} \int_{(B)^k} \ell_{N|\xi}(t)(\tilde{x}) dx_1 ... dx_{k+m} dP \\
&= \sum_{m=0}^{\infty} \frac{1}{m!k!} \int_{(A_t \setminus B)^m} \int_{(B)^k} \ell_N(t)(\tilde{x}) dx_1 ... dx_{k+m} \\
&= \sum_{m=0}^{\infty} \int_{S_m} \ell_N(t)(\tilde{x}) dx_1 ... dx_{k+m}
\end{aligned}$$

where $S_m = \{(x_1, ..., x_{k+m}) \in (A_t)^{k+m}; \tilde{x} = (x_1, ..., x_{k+m}, \infty, \infty, ...) \in \mathbb{X}_{t,B}, n(\tilde{x})(B) = k\}$. Note that this holds for any $B \in \mathcal{B}(A_t), S \in \mathcal{X}_{t,B}$. We shall now consider the case $B = A_u, u \leq t$.

Let $A = \{D \in \mathcal{U}; t \notin D\} \in \mathcal{B}(\mathcal{U})$, and note that $\xi^{-1}(A) = \{L_t = 0\}$. Let $h_t : \mathfrak{N} \to \mathbb{R}$ be the map

$$h_t(\mu) = \frac{\int_{\xi^{-1}(A)} \ell_{N|\xi}(t)(\mu)(\omega) dP(\omega)}{\int_\Omega \ell_{N|\xi}(t)(\mu)(\omega) dP(\omega)} = \frac{\int_A \ell_{N|\xi=D}(t)(\mu) d(P\xi^{-1})(D)}{\ell_N(t)(\mu)}.$$

Note that h_t is measurable since $\ell_N(t)$ and $\ell_{N|\xi}(t)$ are, and since $L^1(\Omega) \ni f \mapsto \int_{\xi^{-1}(A)} f dP \in \mathbb{R}$ is measurable for each $A \in \mathcal{U}$.

Let $G = [N_u = k]$ for some $u \leq t$. Using the fact that $N|_{A_t} = n \circ N_{t,A_u}^{\mathbb{X}}$ and a change of variable, one has

$$
\begin{aligned}
&\int_G h(N)dP \\
&= \int_{(N^{\mathbb{X}}_{t,A_u})^{-1}(\{\tilde{x}\in\mathbb{X}_{t,A_u};n(\tilde{x})(A_u)=k\})} h_t(n\circ N^{\mathbb{X}}_{t,A_u}(\omega))dP(\omega) \\
&= \int_{\{\tilde{x}\in\mathbb{X}_{t,A_u};n(\tilde{x})(A_u)=k\}} h_t(n(\tilde{x}))d(P(N^{\mathbb{X}}_{t,A_u})^{-1})(\tilde{x}) \\
&= \int_S \frac{\int_A \ell_{N|\xi=D}(t)(\tilde{x})d(P\xi^{-1})(D)}{\ell_N(t)(\tilde{x})} d\nu_{t,A_u}(\tilde{x}) \\
&= \sum_{m=0}^{\infty}\int_{S_m} \frac{\int_A \ell_{N|\xi=D}(t)(\tilde{x})d(P\xi^{-1})(D)}{\ell_N(t)(\tilde{x})} \ell_N(t)(\tilde{x})dx_1...dx_{k+m} \\
&= \int_A \sum_{m=0}^{\infty}\int_{S_m} \mu_0^{n(\tilde{x})(A_t)}e^{-\mu_0|A_t|}dx_1...dx_{k+m}d(P\xi^{-1})(D) \\
&= \mu_0^k e^{-\mu_0|A_t|}\int_A \sum_{m=0}^{\infty}\frac{1}{k!}\frac{1}{m!}\int_{(A_t\setminus A_u)^m}\int_{(A_u)^k}\mu_0^m dx_1...dx_{k+m}d(P\xi^{-1})(D) \\
&= \mu_0^k e^{-\mu_0|A_t|}\int_A \sum_{m=0}^{\infty}\frac{1}{k!}\frac{1}{m!}\mu_0^m|A_t\setminus A_u|^m|A_u|^k d(P\xi^{-1})(D) \\
&= \frac{(\mu_0|A_u|)^k}{k!}e^{-\mu_0|A_t|+\mu_0|A_t\setminus A_u|}P[\xi\in A] \\
&= \frac{(\mu_0|A_u|)^k}{k!}e^{-\mu_0|A_u|}P[L_t=0].
\end{aligned}
$$

However, it is also true that

$$
\begin{aligned}
\int_G P[L_t=0|\mathcal{F}_t^N]dP &= P[L_t=0,N_u=k] \\
&= P[N_u=k|L_t=0]P[L_t=0] \\
&= \frac{(\mu_0|A_u|)^k}{k!}e^{-\mu_0|A_u|}P[L_t=0].
\end{aligned}
$$

From this it is clear that $\int_G h_t(N)dP = \int_G P[L_t=0|\mathcal{F}_t^N]dP$ for any $G\in\mathcal{F}_t^N$ since $u\le t$ was arbitrary and $N(\omega)$ is a measure for $\omega\in\Omega$. Therefore

$h(N) = P[L_t = 0|\mathcal{F}_t^N]$ holds with probability 1; that is to say

$$P[L_t = 0|\mathcal{F}_t^N] = \frac{\int_{\{D\in\mathcal{U};t\notin D\}} \ell_{N|\xi=D}(t)(N)d(P\xi^{-1})(D)}{\ell_N(t)(N)} = \frac{\ell_{N|L_t=0}(t)(N)P[L_t=0]}{\ell_N(t)(N)}.$$

Note that $\ell_{N|L_t=0}(t)(N) = e^{-\mu_0|A_t|}\mu_0^{N(A_t)}$ and

$$\ell_N(t)(N) = \int_{\mathcal{U}} e^{-\mu_0|A_t|}\mu_0^{N(A_t)}e^{-(\mu_1-\mu_0)|A_t\cap\xi|}\left(\frac{\mu_1}{\mu_0}\right)^{N(A_t\cap\xi)} d(P\xi^{-1})(D).$$

The results of the preceding discussion are listed in the following lemma:

Lemma 2.3.3 *Let L be the first line of a Poisson process, let $\xi = \{t; L_t > 0\}$, and let N be, conditional upon ξ, a Poisson process with intensity μ_0 on ξ^c and μ_1 on ξ. Then*

i) the conditional likelihood of N given ξ is given by

$$\ell_{N|\xi}(t) = e^{-\mu_0|A_t|}\mu_0^{n(A_t)}e^{-(\mu_1-\mu_0)|A_t\cap\xi|}\left(\frac{\mu_1}{\mu_0}\right)^{n(A_t\cap\xi)},$$

ii) $\ell_N(t) = \int_\Omega \ell_{N|\xi}(t)dP$,

iii) $\ell_{N|L_t=0}(t) = \mu_0^{n(A_t)}e^{-\mu_0|A_t|}$, and

iv) $\mathbb{E}[X_t|\mathcal{F}_t^N] = \frac{\ell_{N|L_t=0}(t)(N)\cdot P[L_t=0]}{\ell_N(t)(N)}$ where $X_t = I_{\{L_t=0\}}$.

2.4 The Compensator

In this section the we return to the more general setting of $E = \mathbb{R}_+^d$.

Definition 2.4.1 *An increasing process Λ on E is a compensator for a weak $\mathcal{F}$-submartingale L if it is $\mathcal{F}$-adapted and $M = L - \Lambda$ is a weak martingale.*

Lemma 2.4.2 *Let L be a single line process on E and let $\tilde{\Lambda}$ be the measure defined on rectangles $(s,t]$ via $\tilde{\Lambda}(s,t] := \mathbb{E}(L(s,t])$. If $\tilde{\Lambda}$ is absolutely continuous with respect to Lebesgue measure let $\tilde{\lambda}$ be its Radon-Nikodym derivative and define*

$$\lambda_s := \begin{cases} \frac{\tilde{\lambda}_s}{P(L_s=0)} & \textit{when } P(L_s = 0) > 0 \\ 0 & \textit{otherwise.} \end{cases}$$

Then $\Lambda_t := \int_{[0,t]} \lambda_s I_{\{L_s=0\}} ds$ is an $\mathcal{F}^L$-compensator for L.

Proof:

Let $H_1, ..., H_k \in \mathcal{B}(\mathbb{R}_+)$, $s \leq t \in E$ and $u_1, ..., u_k \leq s$. If $0 \notin \bigcap_{i=1}^k H_i$ then, on $\bigcap_{i=1}^k L_{u_i}^{-1}(H_i) \subseteq \{L_s > 0\}$, $L(s,t] = 0$. Otherwise,

$$\begin{aligned} \int_{\bigcap_{i=1}^k L_{u_i}^{-1}(H_i)} \mathbb{E}[L(s,t]|\mathcal{F}_s^L]dP &= \int_{L_s^{-1}\{0\}} L(s,t]dP + \int_{\bigcap_{i=1}^k L_{u_i}^{-1}(H_i)\setminus\{L_s=0\}} L(s,t]dP \\ &= \int_{\{0\}} \mathbb{E}[L(s,t]|L_s = y]dPL_s^{-1}(y) \\ &= \int_{\{L_s=0\}} \mathbb{E}[L(s,t]|L_s = 0]dP \\ &= \int_{\bigcap_{i=1}^k L_{u_i}^{-1}(H_i)} \mathbb{E}[L(s,t]|L_s = 0]I_{\{L_s=0\}}dP \end{aligned}$$

since $\{L_s = 0\} \subseteq \bigcap_{i=1}^k \{L_{u_i} = 0\} \subseteq \bigcap_{i=1}^k L_{u_i}^{-1}(H_i)$. Since sets of the form $\bigcap_{i=1}^k L_{u_i}^{-1}(H_i)$ represent a π-system generating $\mathcal{F}_s^L$ it follows that

$$\begin{aligned} \mathbb{E}[L(s,t]|\mathcal{F}_s^L] &= \mathbb{E}[L(s,t]|L_s = 0]I_{\{L_s=0\}} \\ &= \begin{cases} \frac{1}{P(L_s=0)} \int_{\{L_s=0\}} L(s,t]dP \cdot I_{\{L_s=0\}} & \text{if } P(L_s = 0) > 0 \\ 0 & \text{else} \end{cases} \\ &= \begin{cases} \frac{\tilde{\Lambda}(s,t]}{P(L_s=0)} I_{\{L_s=0\}} & \text{if } P(L_s = 0) > 0 \\ 0 & \text{else.} \end{cases} \end{aligned}$$

Furthermore, by Fubini's theorem,

$$\mathbb{E}[\Lambda(s,t]|\mathcal{F}_s^L] = \int_{(s,t]} \mathbb{E}[\lambda_u I_{\{L_u=0\}}|\mathcal{F}_s^L]du$$

$$
\begin{aligned}
&= \int_{(s,t]} \frac{\tilde{\lambda}_u}{P(L_u = 0)} \mathbb{E}[I_{\{L_u=0\}}|\mathcal{F}_s^L]du \\
&= \int_{(s,t]} \frac{\tilde{\lambda}_u}{P(L_u = 0)} P(L_u = 0|\mathcal{F}_s^L)du
\end{aligned}
$$

where it is understood that the integrand is 0 when $P(L_u = 0) = 0$. One can similarly show that

$$
P(L_u = 0|\mathcal{F}_s^L) = P(L_u = 0|L_s = 0)I_{\{L_s=0\}} = \frac{P(L_u = 0)}{P(L_s = 0)}I_{\{L_s=0\}}
$$

for $s \leq u$ when $P(L_s = 0) > 0$ (otherwise the value 0 is assumed). Then

$$
\begin{aligned}
\mathbb{E}[\Lambda(s,t]|\mathcal{F}_s^L] &= \begin{cases} \int_{(s,t]} \frac{\tilde{\lambda}_u}{P(L_s=0)} I_{\{L_s=0\}}du & \text{if } P(L_s = 0) > 0 \\ 0 & \text{otherwise} \end{cases} \\
&= \begin{cases} \frac{\tilde{\Lambda}(s,t]}{P(L_s=0)} I_{\{L_s=0\}} & \text{if } P(L_s = 0) > 0 \\ 0 & \text{otherwise} \end{cases}
\end{aligned}
$$

and the conclusion is that $\mathbb{E}[L(s,t]|\mathcal{F}_s^L] = \mathbb{E}[\Lambda(s,t]|\mathcal{F}_s^L]$ w.p.1., i.e. Λ is an $\mathcal{F}^L$-compensator for L.

■

The function λ_s of the lemma 2.4.2 is called the *weak hazard* function of L.

Lemma 2.4.3 *Let L be a single line point process, let $\xi = \{t \in E; L_t > 0\}$, and let N be a conditional Poisson process given ξ, as in definition 2.2.16. If Λ is an $\mathcal{F}^L$-compensator for L then it is also an $\mathcal{F}^{L,N}$-compensator for L.*

Proof:

Fix $s, t \in E$, and let $(L_n(s,t])_{n\geq 1}$ be a sequence of simple, $\mathcal{F}_{(s,t]}^L$-measurable functions converging upward toward $L(s,t]$ almost surely. By the dominated convergence theorem for conditional expectation, $\mathbb{E}[L(s,t]|\mathcal{A}] = \lim_{n\to\infty} \mathbb{E}[L_n(s,t]|\mathcal{A}]$ for $\mathcal{A} = \mathcal{F}_s^L$ or $\mathcal{F}_s^{L,N}$. So if it can be shown that $\mathbb{E}[I_H|\mathcal{F}_s^L] = \mathbb{E}[I_H|\mathcal{F}_s^{L,N}]$ for each $H \in \mathcal{F}_{(s,t]}^L$ it will follow that

$$\mathbb{E}[L(s,t]|\mathcal{F}_s^L] = \mathbb{E}[L(s,t]|\mathcal{F}_s^{L,N}].$$

Of course, this is equivalent to $P[H|\mathcal{F}_s^L] = P[H|\mathcal{F}_s^{L,N}]\ \forall H \in \mathcal{F}_{(s,t]}^L$, and by Proposition 6.6 of [7], this is equivalent to $\mathcal{F}_s^N \perp\!\!\!\perp_{\mathcal{F}_s^L} \mathcal{F}_{(s,t]}^L$ which holds by Lemma 2.2.17. Therefore, the equality holds. Since $\Lambda(s,t]$ is $\mathcal{F}_{(s,t]}^L$-measurable, it follows that $\mathbb{E}[\Lambda(s,t]|\mathcal{F}_s^L] = \mathbb{E}[\Lambda(s,t]|\mathcal{F}_s^{L,N}]$ also holds and hence that

$$\mathbb{E}[L(s,t]|\mathcal{F}_s^{L,N}] = \mathbb{E}[L(s,t]|\mathcal{F}_s^L] = \mathbb{E}[\Lambda(s,t]|\mathcal{F}_s^L] = \mathbb{E}[\Lambda(s,t]|\mathcal{F}_s^{L,N}].$$

Thus Λ is indeed an $\mathcal{F}^{L,N}$-compensator for L.

■

Example 2.4.4 The Single Jump Process

Let Y be an E-valued r.v. with distribution function F and continuous density f, and let $L_t = I_{\{Y\in A_t\}}$. Then L is a single line process; in fact, it is a single jump process. Also $\tilde{\Lambda}_{(s,t]} = \mathbb{E}[L_{(s,t]}] = P(Y \in (s,t]) = \int_{(s,t]} f_u du$, and so according to lemma 2.4.2

$$\lambda_s = \begin{cases} \frac{f_s}{1-F_s} & \text{if } P(L_s = 0) > 0 \\ 0 & \text{otherwise} \end{cases}$$

is the weak hazard of L and $\Lambda_t = \int_{[0,t]} \lambda_s I_{\{L_s=0\}} ds$ is an $\mathcal{F}^L$-compensator for L.

Example 2.4.5 The First Line of A Poisson Process

Consider an inhomogeneous Poisson process J on E with intensity function γ, and let L be the first line of J. Let $t = (t_1, ..., t_d) \in E, n \in \mathbb{N}$ and let $\mathcal{C}_n$ be the n^{th} approximate of $[0,t]$. Since γ is continuous and A_t is compact $\exists K \in \mathbb{R}_+$ such that

$\gamma|_{A_t} \leq K$. Let n be large enough so that $C \in \mathcal{C}_n$ implies $0 \leq \int_C \gamma(u)du \leq 1$. Then $1 - \int_C \gamma(u)du \leq e^{-\int_C \gamma(u)du}$ and hence

$$\begin{aligned}
\int_C \gamma(u)du \cdot e^{-\int_C \gamma(u)du} &\geq \int_C \gamma(u)du - \Big(\int_C \gamma(u)du\Big)^2 \\
&\geq \int_C \gamma(u)du - \big(K\frac{\prod_{i=1}^d t_i}{2^{nd}}\big)^2 \\
&= \int_C \gamma(u)du + o(\frac{1}{2^{(d+1)n}}).
\end{aligned}$$

Then

$$\begin{aligned}
\mathbb{E}[L(C)] &= \mathbb{E}[L(C)I_{\{L_{t_{C_-}}=0\}}] \\
&\geq \mathbb{E}[I_{\{L(C)=1\}}I_{\{L_{t_{C_-}}=0\}}] \\
&\geq \mathbb{E}[I_{\{J(C)=1\}}I_{\{L_{t_{C_-}}=0\}}] \\
&= \mathbb{E}[I_{\{J(C)=1\}}I_{\{J_{t_{C_-}}=0\}}] \\
&= \mathbb{E}[I_{\{J(C)=1\}}]\mathbb{E}[I_{\{J_{t_{C_-}}=0\}}] \\
&= P[J(C)=1]P[J_{t_{C_-}}=0] \\
&= P[J_{t_{C_-}}=0]\int_C \gamma(u)du \cdot e^{-\int_C \gamma(u)du} \\
&\geq P[J_{t_{C_-}}=0]\Big(\int_C \gamma(u)du + o(\frac{1}{2^{(d+1)n}})\Big) \\
&= P[L_{t_{C_-}}=0]\int_C \gamma(u)du + e^{-\int_0^{t_{C-}} \gamma(u)du} o(\frac{1}{2^{(d+1)n}}) \\
&= P[L_{t_{C_-}}=0]\int_C \gamma(u)du + o(\frac{1}{2^{(d+1)n}}),
\end{aligned}$$

as well as

$$\begin{aligned}
\mathbb{E}[L(C)] &= \mathbb{E}[L(C)I_{\{L_{t_{C_-}}=0\}}] \\
&\leq \mathbb{E}[J(C)I_{\{J_{t_{C_-}}=0\}}] \\
&= \mathbb{E}[J(C)]\mathbb{E}[I_{\{J_{t_{C_-}}=0\}}]
\end{aligned}$$

$$
\begin{aligned}
&= \int_C \gamma(u) du P[J_{t_{C-}} = 0] \\
&= P[L_{t_{C-}} = 0] \int_C \gamma(u) du.
\end{aligned}
$$

Define $p_n : A_t \to \mathbb{R}_+$ via $p_n(u) := \sum_{C \in \mathcal{C}_n} P[L_{t_{C-}} = 0] I_C(u)$ so that the above inequalities may be expressed concisely as follows:

$$\int_C \gamma(u) p_n(u) du + o(\frac{1}{2^{n(d+1)}}) \leq \mathbb{E}[L(C)] \leq \int_C \gamma(u) p_n(u) du.$$

Summing over all $C \in \mathcal{C}_n$ then gives

$$\int_{[0,t]} \gamma(u) p_n(u) du + o(\frac{1}{2^n}) \leq \mathbb{E}[L_t] \leq \int_{[0,t]} \gamma(u) p_n(u) du.$$

Finally, observe that for $u \in [0,t]$ there is some sequence of rectangles $(C_n)_{n \geq 1}$ containing u for each n. This sequence will converge to $\{u\}$, and the lower left corners t_{C_n-} will converge to u. Thus $p_n(u) = P[L_{t_{C_n-}} = 0] I_{C_n}(u) \underset{n \to \infty}{\longrightarrow} P[L_u = 0] I_{\{u\}}(u) = P[L_u = 0]$. Therefore, by the dominated convergence theorem, taking the limit of the above inequality as $n \to \infty$ yields

$$\mathbb{E}[L_t] = \int_{[0,t]} \gamma(u) P[L_u = 0] du.$$

The weak hazard function of L is thus

$$\lambda_s = \begin{cases} \gamma(s) & \text{if } P[L_s = 0] > 0 \\ 0 & \text{otherwise.} \end{cases}$$

and so the compensator Λ_t of L (with respect to $\mathcal{F}^L$) is

$$\Lambda_t = \int_{[0,t]} \gamma(u) I_{\{L_u = 0\}} du. \tag{2.4.1}$$

Recall that, by lemma 2.4.3, Λ is also an $\mathcal{F}^{L,N}$-compensator.

Chapter 3

The Model

In this chapter the model alluded to earlier is explained in full detail. To this end the gain function is introduced with some motivation from the 1 dimensional gain function in [1], and is shown to have a smooth semi-martingale representation. The notion of monotonicity of a process is given and its importance is immediately illustrated with an optimality lemma for stopped processes. The conditional expectation of the gain function with respect to $\mathcal{F}^N$ is then manipulated so that sufficient conditions for its monotonicity and relevance to the optimality lemma can be more readily determined. It is then shown that these conditions indeed yield an optimal solution to the sequential detection problem.

3.1 The Gain Function

A stopping point can only be called optimal if there is some criterion according to which the stopping point can be evaluated. In [1] the authors introduce a gain function to serve as this criterion when evaluating stopping times. A reward proportional to the length of time spent before the change point σ is given, and a penalty proportional to the amount of time spent after σ is given. There is additionally a fixed value reward given if and only if stopping has occurred after time σ, and a constant k_0 is added to the gain function to account for costs or rewards that might come independently of the chosen stopping time. The gain function specified by these conditions is

$$Z_t = c_0(t \wedge \sigma) - c_1(t - \sigma)^+ + k_0 + k_1 I_{\{t \geq \sigma\}}$$

where $c_0 \geq 0, c_1 > 0$ and $k_1 \geq 0$. In [2] the authors represent this gain function in a way that has meaning in higher dimensions and indeed agrees with intuition about what a higher dimensional gain function should look like: by defining $L_t = I_{\{\sigma \leq t\}}$ and $\xi = [\sigma, \infty) = \{t; L_t > 0\}$ one sees that L is a single jump point process, ξ is a random upper layer, and

$$\begin{aligned} Z_t &= c_0|A_t \cap \xi^c| - c_1|A_t \cap \xi| + k_0 + k_1 L_t \\ &= k_0 + \int_{A_t} (-c_1 + (c_0 + c_1)X_u)du + k_1 L_t \end{aligned}$$

where $X_t = I_{\{L_t = 0\}}$. As mentioned in the comments preceding section 2.4, the change set ξ of the model will be assumed to be generated by the first line L via $\xi = \{t \in E; L_t > 0\}$. The gain function for points $t \in E = [0, R] \subseteq \mathbb{R}^d$ is simply the generalized representation given above. For lower layers $B \subseteq E$ it is given more generally by

$$\begin{aligned} Z(B) &= c_0|B \cap \xi^c| - c_1|B \cap \xi| + k_0 + k_1 L(B) \\ &= k_0 + \int_B (-c_1 + (c_0 + c_1)X_u)du + k_1 L(B). \end{aligned}$$

It is important to note that the information available to an observer located at the point t (the σ-field known at point t) is the location of the jump points of N in A_t ($\mathcal{F}_t^N$). The jump points of L are *not* known, and it is precisely their influence that must be detected in an optimal way. In the sequential detection model, this is done by deciding to remain at t or push forward based on current information, and Z_t is a measure of how good this decision is.

Naturally, a rule for stopping is optimal if, on average, the values of the gain function that it yields are at least as high as those yielded by other stopping rules. Stopping rules are taken from the collection of stopping sets since they must be lower layers and must be determinable by known information, i.e. $\{t \in \rho\}$ must be $\mathcal{F}_t^N$-measurable for any stopping rule ρ. This notion of optimality is formalized in the following definition.

Definition 3.1.1 *An $\mathcal{F}^N$-stopping set $\hat{\rho}$ is called an optimal solution to the sequential detection problem if it satisfies*

$$\mathbb{E}[Z(\hat{\rho})] = \sup\{\mathbb{E}[Z(\rho); \rho \textit{ is a an } \mathcal{F}^N\textit{-stopping set}\}.$$

As it turns out, finding the optimal solution depends on the existence of a particular represention of the gain function.

Definition 3.1.2 *A process $Z = (Z_t)_{t \in E}$ adapted to the filtration $\mathcal{F} = (\mathcal{F}_t)_{t\in E}$ is a smooth semi-martingale w.r.t $\mathcal{F}$ ($\mathcal{F}$-SSM) if there exists an outer continuous with inner limits (ocil) adapted process U and a weak $\mathcal{F}$-martingale M on E such that*

$$Z_t = Z_0 + \int_{A_t} U_s ds + M_t$$

$\forall t \in E$. The $\mathcal{F}$-SSM is denoted $Z = (U, M)$.

Definition 3.1.3 *A function v on E is monotone if, whenever $v(s) \leq 0$, $v(t) \leq 0$ for all $t \geq s$. A stochastic process V is monotone if $V_.(\omega)$ is monotone for each $\omega \in \Omega$.*

Lemma 3.1.4 *Let V be a monotone stochastic process. Then*

$$\hat{\rho} := \{t \in E; V_s > 0 \ \forall s << t\}$$

is an $\mathcal{F}^V$ stopping set.

Proof:

Let $\omega \in \Omega$, $t \in \hat{\rho}(\omega)$, and let $s \leq t$. If $u << s$ then $u << t$ and so $V_u(\omega) > 0$ since $t \in \hat{\rho}(\omega)$. Therefore $s \in \hat{\rho}(\omega)$, and it follows that $\hat{\rho}(\omega)$ is a lower layer.

Now let $t \in \overline{\hat{\rho}(\omega)}$ and $u << t$. Let $s \in (u, t)$ be arbitrary. If $s \notin \hat{\rho}(\omega)$ then $[s, R] \cap \hat{\rho}(\omega) = \emptyset$; since $[s, R]$ is a neighborhood of t, a contradiction arises. Thus it must be that $s \in \hat{\rho}(\omega)$. Then, since $u << s$, it must be that $V_u(\omega) > 0$. Since $u << t$ was arbitrary, it follows that $t \in \hat{\rho}(\omega)$, and since $t \in \overline{\hat{\rho}(\omega)}$ was arbitrary, it follows that $\hat{\rho}(\omega)$ is closed.

Now let $t \in E$ be arbitrary, let $(t_n)_{n\geq 1}$ be a sequence in E converging to t from below and satisfying $t_n << t$. Let $s << t$ and $\omega \in \cap_{n\geq 1}\{V_{t_n} > 0\}$. Then $\exists\, n \in \mathbb{N}$ such that $s \leq t_n$. Since $V_s(\omega) \leq 0$ would necessitate $V_{t_n}(\omega) \leq 0$ by monotonicity of V, which is forbidden by the choice of ω, it must be that $V_s(\omega) > 0$. Since $s << t$ was arbitrary, it follows that $t \in \hat{\rho}(\omega)$, and since ω was arbitrary it follows that

$$\{t \in \hat{\rho}\} = \cap_{n\geq 1}\{V_{t_n} > 0\} \in \sigma(\cup_{n\geq 1}\mathcal{F}^V_{t_n}) \subseteq \mathcal{F}^V_t.$$

Therefore $\hat{\rho}$ is a stopping set.

■

The preceding lemma and definitions provide the requisites for the following theorem, which serves as a key tool in determining an optimal solution to the sequential detection problem. It is lemma 3.10 of [2].

Lemma 3.1.5 *Let U be a bounded stochastic process on E adapted to a filtration $\mathcal{G}$ such that U is outer-continuous with inner limits. If $\mathcal{F}$ is a subfiltration of $\mathcal{G}$ and if a version of $V_t := \mathbb{E}[U_t|\mathcal{F}_t]$ exists that is outer-continuous with inner limits, then for any $\mathcal{F}$-stopping set ρ on E,*

$$\mathbb{E}[\int_\rho U_t dt] = \mathbb{E}[\int_\rho V_t dt].$$

In addition, if V is monotone then the $\mathcal{F}$-stopping set $\hat{\rho}$ defined by

$$\hat{\rho} = \{t \in E; V_s > 0 \ \forall \ s << t\}$$

satisfies

$$\mathbb{E}[\int_{\hat{\rho}} U_t dt] = \sup\{\mathbb{E}[\int_{\rho} U_t dt]; \rho \textit{ is an } \mathcal{F}\textit{-stopping set}\}.$$

Proof:

Let ρ be a stopping set, let $\mathcal{C}_n$ be the n^{th} rectangular partition of E, and let $\rho_n = \cup_{\{C \in \mathcal{C}_n; t_{C\_} \in \rho\}} C$ be the n^{th}-approximate of ρ. Since U is outer-continuous, it is jointly-measurable by lemma 2.2.2. For $C \in \mathcal{C}_n$ the map $I_{\{t_{C\_} \in \rho\}} : \Omega \times E \to \mathbb{R}_+$ is also jointly measurable, hence $U \cdot I_{\{t_{C\_} \in \rho\}}$ is. Then, by Fubini's theorem, $\int_C U_t \cdot I_{\{t_{C\_} \in \rho\}} dt : \Omega \to \mathbb{R}$ is $\mathcal{F}$-measurable, and so too must be

$$\int_{\rho_n} U_t dt = \sum_{C \in \mathcal{C}_n} \int_C U_t \cdot I_{\{t_{C\_} \in \rho\}} dt.$$

Now, for fixed $\omega \in \Omega$, $\rho(\omega) = \cap_{n \geq 1} \rho_n(\omega)$ and so

$$\int_{\rho(\omega)} U_t(\omega) dt = \lim_{n \to \infty} \int_{\rho_n(\omega)} U_t(\omega) dt.$$

It follows that $\int_\rho U_t dt$ is measurable (the integrals are well defined, and finite in fact, since U is bounded). Similar arguments apply to V. Now, since V is jointly-measurable, Fubini's theorem guaranties that, for $A \in \mathcal{F}_{t_{C\_}}$,

$$\begin{aligned}
\int_A \mathbb{E}[\int_C V_t dt | \mathcal{F}_{t_{C\_}}] dP &= \int_A \int_C V_t dt dP \\
&= \int_C \int_A V_t dP dt \\
&= \int_C \int_A \mathbb{E}[U_t | \mathcal{F}_t] dP dt \\
&= \int_C \int_A U_t dP dt \\
&= \int_A \int_C U_t dt dP
\end{aligned}$$

$$= \int_A \mathbb{E}[\int_C U_t dt | \mathcal{F}_{t_{C-}}] dP$$

since $A \in \mathcal{F}_{t_{C-}} \subseteq \mathcal{F}_t$ for all $t \in C$. Thus, with probability 1, $\mathbb{E}[\int_C V_t dt | \mathcal{F}_{t_{C-}}] = \mathbb{E}[\int_C U_t dt | \mathcal{F}_{t_{C-}}]$. Then

$$\begin{aligned}
\mathbb{E}[\int_{\rho_n} U_t dt] &= \mathbb{E}[\sum_{C \in \mathcal{C}_n} I_{\{t_{C-} \in \rho\}} \int_C U_t dt] \\
&= \mathbb{E}[\sum_{C \in \mathcal{C}_n} I_{\{t_{C-} \in \rho\}} \mathbb{E}[\int_C U_t dt | \mathcal{F}_{t_{C-}}]] \\
&= \mathbb{E}[\sum_{C \in \mathcal{C}_n} I_{\{t_{C-} \in \rho\}} \mathbb{E}[\int_C V_t dt | \mathcal{F}_{t_{C-}}]] \\
&= \mathbb{E}[\sum_{C \in \mathcal{C}_n} I_{\{t_{C-} \in \rho\}} \int_C V_t dt] \\
&= \mathbb{E}[\int_{\rho_n} V_t dt].
\end{aligned}$$

Taking the limit as $n \to \infty$ then yields $\mathbb{E}[\int_\rho U_t dt] = \mathbb{E}[\int_\rho V_t dt]$ by the dominated convergence theorem.

Now suppose that V is monotone and let ρ be an $\mathcal{F}$-stopping set on E. Then for $\omega \in \Omega$, $V_\cdot(\omega) \geq 0$ on $\hat{\rho}(\omega)$, and if $u \in \hat{\rho}(\omega)^c$ $\exists s \leq u$ such that $V_s(\omega) \leq 0$. Then, by monotonicity of V,$V_u(\omega) \leq 0$, and since $u \in \hat{\rho}(\omega)^c$ was arbitrary this means that $V_\cdot(\omega) \leq 0$ on $\hat{\rho}(\omega)^c$. Therefore

$$\begin{aligned}
\mathbb{E}[\int_{\hat{\rho}} U_t dt - \int_\rho U_t dt] &= \mathbb{E}[\int_{\hat{\rho}} V_t dt - \int_\rho V_t dt] \\
&= \mathbb{E}[\int_{\hat{\rho} \setminus \rho} V_t dt + \int_{\hat{\rho} \cap \rho} V_t dt - \int_{\rho \setminus \hat{\rho}} V_t dt - \int_{\rho \cap \hat{\rho}} V_t dt] \\
&= \mathbb{E}[\int_{\hat{\rho} \setminus \rho} V_t dt - \int_{\rho \setminus \hat{\rho}} V_t dt] \\
&\geq 0.
\end{aligned}$$

Since ρ was selected arbitrarily it follows that

$$\mathbb{E}[\int_{\hat{\rho}} U_t dt] = \sup\{\mathbb{E}[\int_{\rho} U_t dt]; \rho \text{ is an } \mathcal{F}\text{-stopping set}\}.$$

■

Assume that the single line process L, forming the boundary between the territories of differing intensities, has continuous weak hazard λ and compensator $\Lambda_t = \int_{A_t} \lambda(u) I_{\{L_u=0\}} du$. If M is the weak $\mathcal{F}^L$-martingale $L - \Lambda$, the gain function Z_t may be expressed as

$$Z_t = k_0 + \int_{A_t} (-c_1 + (c_0 + c_1 + k_1\lambda(u))X_u) du + k_1 M_t.$$

Now $X_u = I_{\{L_u=0\}}$ is outer continuous with inner limits since L is, and λ is continuous, so $U_u := -c_1 + (c_0 + c_1 + k_1\lambda(u))X_u$ is outer continuous with inner limits. This means that $Z = (U, M)$ is a smooth semi-martingale with respect to $\mathcal{F}^L$, and in fact with respect to $\mathcal{F}^{L,N}$ since Λ is also a $\mathcal{F}^{L,N}$-compensator.

In order to apply lemma 3.1.5 to U with $\mathcal{G} = \mathcal{F}^{L,N}$ and $\mathcal{F} = \mathcal{F}^N$ one must first establish that U_t is $\mathcal{F}_t^{L,N}$-measurable and bounded, and that

$$V_t := \mathbb{E}[U_t | \mathcal{F}_t^N] = -c_1 + (c_0 + c_1 + k_1\lambda(t)) P[L_t = 0 | \mathcal{F}_t^N]$$

is monotone. $X_u = I_{\{L_u=0\}}$ is measurable $\mathcal{F}^L \subseteq \mathcal{F}^{L,N}$, so U is adapted to $\mathcal{F}^{L,N}$, and since λ is continuous and E is compact, U is bounded. Thus finding a solution to the sequential detection problem reduces to showing that V is monotone, and determining what conditions yield this property is the goal of the next section.

3.2 Conditions Sufficient for Monotonicity

Let N, L and ξ be as in definition 2.2.16, and let $\ell_{N|\xi}(t)$ be the conditional likelihood of N given ξ. Recall that $\mathcal{U}$ denotes the space of upper layers on E and that $E_t =$

$\{s \in E; s \geq t\}$, so that $L_t > 0 \Leftrightarrow E_t \subseteq \xi$. Then, by lemma 2.3.3 and a change of variable,

$$\begin{aligned}
\ell_N(t)(N) &= \int \ell_{N|\xi(\omega)}(t)(N)dP(\omega) \\
&= \int_{\{L_t=0\}} \ell_{N|\xi(\omega)}(t)(N)dP(\omega) + \int_{\{L_t\neq 0\}} \ell_{N|\xi(\omega)}(t)(N)dP(\omega) \\
&= \int_{\{L_t=0\}} e^{-\mu_0|A_t|}\mu_0^{N_t}dP(\omega) \\
&\quad + \int_{\{D\in\mathcal{U};E_t\subseteq D\}} e^{-\mu_0|A_t|}\mu_0^{N_t}e^{-(\mu_1-\mu_0)|A_t\cap D|}\left(\frac{\mu_1}{\mu_0}\right)^{N(A_t\cap D)} d(P\xi^{-1})(D) \\
&= e^{-\mu_0|A_t|}\mu_0^{N_t}\big[P(L_t=0) \\
&\quad + \int_{\{D\in\mathcal{U};E_t\subseteq D\}} e^{-(\mu_1-\mu_0)|A_t\cap D|}\left(\frac{\mu_1}{\mu_0}\right)^{N(A_t\cap D)} dP\xi^{-1}(D)\big] \\
&= e^{-\mu_0|A_t|}\mu_0^{N_t}\left[P(L_t=0) + e^{-(\mu_1-\mu_0)|A_t|}Q_t\right],
\end{aligned}$$

where $Q_t := \int_{\{D\in\mathcal{U};E_t\subseteq D\}} e^{(\mu_1-\mu_0)|A_t\setminus D|}\left(\frac{\mu_1}{\mu_0}\right)^{N(A_t\cap D)}dP\xi^{-1}(D)$. Note that $\ell_{N|\xi}(t) : \Omega \times \Omega \to \mathbb{R}$ is the map $\ell_{N|\xi}(t)(N)(\omega,\omega') = \ell_{N|\xi}(t)(N(\omega))(\omega')$, and that the above integral "removes" only the dependence on Ω through ξ, leaving $\ell_N(t)(N)$ which is random due to its dependence on N. Further application of the lemma yields

$$\begin{aligned}
\mathbb{E}[X_t|\mathcal{F}_t^N] &= \frac{\ell_{N|L_t=0}(t)(N)\cdot P[L_t=0]}{\ell_N(t)(N)} \\
&= \frac{\mu_0^{N_t}e^{-\mu_0|A_t|}P[L_t=0]}{e^{-\mu_0|A_t|}\mu_0^{N_t}\left[P(L_t=0)+e^{-(\mu_1-\mu_0)|A_t|}Q_t\right]} \\
&= \frac{1}{1+q_tQ_t}
\end{aligned}$$

for $P[L_t = 0] \neq 0$ where $q_t := \frac{e^{-(\mu_1-\mu_0)|A_t|}}{P[L_t=0]}$. If $P[L_t = 0] = 0$ then $q_t = \infty$ and $\mathbb{E}[X_t|\mathcal{F}_t^N] = 0$ so that the above equality holds for all $t \in E$. Notice that Q_t is increasing in each coordinate of t since the integrand is (recall that $\mu_1 > \mu_0$) and the set $\{D \in \mathcal{U}; E_t \subseteq D\}$ is. Conditions for monotonicity of V can now be readily sought; they are given in the following theorem taken from [2].

Theorem 3.2.1 *Let L be a single line process with continuous weak hazard λ, and for $t \in E$ define*

$$q_t = \frac{e^{-(\mu_1-\mu_0)|A_t|}}{P[L_t = 0]}.$$

An optimal solution to the sequential detection problem exists if λ and q are decreasing and increasing, respectively, in each coordinate. In this case the function V given by

$$V_t = -c_1 + (c_0 + c_1 + k_1\lambda_t)\frac{1}{1 + q_tQ_t}$$

is monotone and the optimal solution is

$$\hat{\rho} = \{t \in E; V_s > 0 \ \forall s << t\}.$$

Proof:

Recall that the gain function Z is given by $Z(B) = k_0 + \int_B U du + k_1 M(B)$ for any lower layer B where M is a weak $\mathcal{F}^{L,N}$-martingale and $U := -c_1 + (c_0 + c_1 + k_1\lambda(u))X_u$ is outer-continuous with inner limits, as well as bounded, so that $Z = (U, M)$ is an $\mathcal{F}^{L,N}$-SSM. As was shown preceding the theorem,

$$\mathbb{E}[U_t|\mathcal{F}_t^N] = -c_1 + (c_0 + c_1 + k_1\lambda(u))\frac{1}{1 + q_tQ_t} = V_t.$$

Thus letting $\mathcal{G} := \mathcal{F}^{L,N}$ and $\mathcal{F} := \mathcal{F}^N$, it will follow from lemma 3.1.5 that an optimal solution exists and is given by $\hat{\rho}$ so long as it can be verified that V is outer-continuous with inner limits (ocil) and monotone.

Now λ is continuous and by continuity of the measure P, q_t is outer-continuous with inner limits. Also, $E \ni t \mapsto N(A_t \cap D)$ is outer-continuous with inner limits since N is a point process. Now let $(u_n)_{n\geq 1}$ be a sequence in E decreasing to t. If $U \in \mathcal{U}$ satisfies $E_t \subseteq U$ then $E_{u_n} \subseteq U$ for each $n \geq 1$ and hence $I_{\{D\in\mathcal{U}; E_t\subseteq D\}}(U) = 1 = \lim_{n\to\infty} I_{\{D\in\mathcal{U}; E_{u_n}\subseteq D\}}(U)$. Otherwise, $E_t \not\subseteq U$ and there exists some $m \geq 1$ for which $n \geq m$ implies $E_{u_n} \not\subseteq U$ and hence $I_{\{D\in\mathcal{U}; E_t\subseteq D\}}(U) = 0 = \lim_{n\to\infty} I_{\{D\in\mathcal{U}; E_{u_n}\subseteq D\}}(U)$. It follows that $I_{\{D\in\mathcal{U}; E_t\subseteq D\}}$ is outer-continuous as a function of t, and similar arguments show that it has inner limits.

Thus, by the dominated convergence theorem, Q_t is outer-continuous with inner limits. Therefore V is.

Finally, $V_s \leq 0$ if and only if $k_1\lambda_s \leq c_1 q_s Q_s - c_0$. Let $t \geq s$. Since λ is decreasing in each coordinate and both q and Q are increasing in each coordinate (see the comments preceding the theorem),

$$\begin{aligned} k_1\lambda_t &\leq k_1\lambda_s \\ &\leq c_1 q_s Q_s - c_0 \\ &\leq c_1 q_t Q_t - c_0 \end{aligned}$$

which means that $V_t \leq 0$. Therefore V is monotone and an optimal solution to the sequential detection problem exists and is given by $\hat{\rho}$.

■

Chapter 4

Applications and an Alternate Gain Function

It is in this chapter that the applications of the preceding results are presented. Theorem 3.2.1 provides a powerful tool for solving the sequential detection problem, reducing the problem to finding conditions to ensure a weak hazard that is decreasing in each coordinate and that q is increasing in each coordinate. That tool is applied here to the case that L is the first line of an inhomogeneous Poisson process. Afterwards, an alternate gain function is investigated and it is shown that theorem 3.2.1 can be used to find a solution that is optimal with respect to this new gain function as well.

4.1 Optimal Solution for First Line of a Poisson Process

Lemma 4.1.1 *Let L be the first line of an inhomogeneous Poisson process J with intensity function γ. If γ is decreasing in each component and if*

$$\mu_1 - \mu_0 \leq \frac{1}{|A_t \setminus A_s|} \int_{A_t \setminus A_s} \gamma(u)du$$

whenever $s \leq t$ then an optimal solution to the sequential detection problem exists.

Proof:

As shown previously, the weak hazard of L is $\lambda_s = \gamma(s) \cdot I_{\{P[L_s=0]>0\}}$, and so if γ is decreasing in each coordinate λ is. By theorem 3.2.1 it is sufficient to show that $q_t = \frac{e^{-(\mu_1-\mu_0)|A_t|}}{P[L_t=0]}$ is increasing in each component. Now, for $s \leq t$,

$$\frac{e^{\int_{[0,s]} \gamma(u)du}}{e^{(\mu_1-\mu_0)|A_s|}} = \frac{e^{-(\mu_1-\mu_0)|A_s|}}{e^{-\int_{[0,s]} \gamma(u)du}} = q_s \leq q_t = \frac{e^{\int_{[0,t]} \gamma(u)du}}{e^{(\mu_1-\mu_0)|A_t|}} \Leftrightarrow$$

$$e^{(\mu_1-\mu_0)(|A_t|-|A_s|)} \quad \leq \quad e^{\int_{[0,t]\setminus[0,s]} \gamma(u)du}$$

or

$$\mu_1 - \mu_0 \leq \frac{1}{|A_t \setminus A_s|} \int_{A_t \setminus A_s} \gamma(u)du.$$

which holds by hypothesis. Thus $q(s)$ is increasing in each coordinate, and so an optimal solution exists and is given by

$$\hat{\rho} = \{t \in E; V_s > 0 \ \forall \ s << t\}.$$

■

4.2 An Alternative Valuation

For this example we return to 2 dimensions. The goal is to find an alternate gain function than the one used in the preceding sections, and to investigate how well the results obtained in those sections apply to the new gain function. The term $k_1 L$, which required a compensator in order for the gain function to have a smooth semi-martingale representation, will be replaced.

Let L be a single line process and define $Y_t := I_{\{L_t > 0\}}$. Recall that, for rectangles $(s,t]$, $Y(s,t]$ is defined via the inclusion exlusion formula, and for a finite disjoint union $A = \cup_{i=1}^{k} R_k$ of rectangles define $Y(A) = \sum_{i=1}^{k} Y(R_k)$. Extend Y to the collection $\mathcal{L}$ of closed lower layers as follows:

Let $\mathcal{C}_n$ be the n^{th} rectangular partition, and for $B \in \mathcal{L}$ let

$$B_n^+ := \bigcup_{\{C \in \mathcal{C}_n; B \cap C \neq \emptyset\}} C.$$

Clearly $B_n^+ \downarrow B$. Fix $\omega \in \Omega$, and let $u_1, ..., u_\ell$ be the points in E satisfying $L(\omega)(\{u_i\}) > 0$. Let n be large enough so that $|C \cap \{u_1, ..., u_\ell\}| \leq 1$ for all $C \in \mathcal{C}_n$, let $m \geq n$, and for $C = (s,t] \subseteq B_m^+$ note that

$$Y(C) = \begin{cases} 1 & \text{if } \exists u_i \in C \text{ and } \nexists u_j \in A_{(t_1,s_2)} \cup A_{(s_1,t_2)} \\ -1 & \text{if } \exists u_i \in A_{(s_1,t_2)}, u_j \in A_{(t_1,s_2)} \text{ and } \nexists u_h \in A_s \\ 0 & \text{otherwise.} \end{cases}$$

Consider the collection $\{p_1, ..., p_e\}$ of intersection points of any horizontal line containing some u_i and any vertical line containing some u_j, and such that A_{p_k} does not contain any u_h, $h \neq i, j$, for each $k = 1, ..., e$. Define a new point process L^+ by letting $L^+(\omega)$ have precisely $p_1, ..., p_e$ as its jump points (note that $\{t; L_t(\omega) \geq 2\} = \{t; L_t^+(\omega) \geq 1\}$). Clearly $Y(C) = -1$ if and only if $\exists p_i \in C$. Also, $C \in \mathcal{C}_n$ cannot contain a u_i and a p_j when n is large enough, in which case $Y(C) = 1$ if and only if $\exists u_i \in C$. Thus

$$\begin{aligned} Y(B_m^+) &= \sum_{\{\mathcal{C}_m \ni C \subseteq B_m^+; \exists u_i \in C\}} Y(C) + \sum_{\{\mathcal{C}_m \ni C \subseteq B_m^+; \exists p_i \in C\}} Y(C) \\ &= |B_m \cap \{u_1, ..., u_\ell\}| - |B_m \cap \{p_1, ..., p_e\}| \\ &= L(B_m) - L^+(B_m) \end{aligned}$$

for all $m \geq n$, and it follows that $Y(B) = \lim_{n \to \infty} Y(B_n^+) = L(B) - L^+(B)$. Since $Y = L - L^+$ is therefore the difference of two increasing, integrable processes, corollary 2.2.12 implies that Y_ρ is a random variable for any stopping set ρ.

The goal here is to construct a valuation which is an alternative to Z; rather than adding a reward of $k_1 L_t$ to the gain function, so that having more jump points of L in $\hat{\rho}$ is favorable, we consider adding instead a reward of $k_1 Y_t$ to the gain function, so that just having non-empty intersection with ξ is favorable. However, applying theorem 3.2.1 to the gain function to find an optimal solution necessitated a smooth semi-martingale representation of Z, and this representation required a compensator for L. Recall that, by definition, a process must be a weak submartingale in order to have a compensator.

Unfortunately, Y may not be a weak submartingale. To see this, consider the case that L is the first line of an inhomogeneous Poisson process J with intensity function γ, let $\delta > 0$ and let $H_1 = ((0,1),(1,1+\delta)], H_2 = ((1,0),(1+\delta,1)]$. If ω is such that $Y(\omega)(H_1), Y(\omega)(H_2) > 0$ and $Y(\omega)([0,1]^2), L(\omega)((1,1+\delta]^2) = 0$, then $Y(\omega)((1,1+\delta]^2) = -1$. The probability of selecting such an ω is

$$\begin{aligned}
& P[Y(H_1), Y(H_2) > 0, Y([0,1]^2), L((1,1+\delta]^2) = 0] \\
= \; & P[J(H_1), J(H_2) > 0, J([0,1]^2), J((1,1+\delta]^2) = 0] \\
= \; & (1 - e^{-\int_{H_1} \gamma(u)du})(1 - e^{-\int_{H_2} \gamma(u)du}) e^{-\int_{[0,1]^2} \gamma(u)du} e^{-\int_{(1,1+\delta]^2} \gamma(u)du} \\
> \; & 0.
\end{aligned}$$

Thus, if $\tilde{\Omega}$ represents the collection of such ω,

$$\begin{aligned}
\int_{\tilde{\Omega}} \mathbb{E}[Y((1,1+\delta]^2)|\mathcal{F}_1^Y]dP &= \int_{\tilde{\Omega}} Y((1,1+\delta]^2)dP \\
&= -P[\tilde{\Omega}] \\
&< 0
\end{aligned}$$

from which it follows that Y cannot be a weak submartingale.

Suppose, however, that one wishes to find a process Λ_t satisfying $\mathbb{E}[\Lambda_t|\mathcal{F}_t^Y] = \mathbb{E}[Y_t|\mathcal{F}_t^Y]$. Such a process functions to "compensate" Y but is not formally a compensator since $(Y_u)_{u\in T}$ is not a weak submartingale (such a process will be referred to as a *quasi-*$\mathcal{F}^Y$*-compensator*). The following is an analogue to lemma 2.4.2:

Lemma 4.2.1 *Let L be a single line process on E and let $Y = L - L^+$. Let $\tilde{\Lambda}$ be the signed measure defined on rectangles $(s,t]$ via $\tilde{\Lambda}(s,t] := \mathbb{E}[Y(s,t]]$. If $\tilde{\Lambda}$ is absolutely continuous with respect to Lebesgue measure let $\tilde{\lambda}$ be its Radon-Nikodym derivative and define*

$$\lambda_s := \begin{cases} \frac{\tilde{\lambda}_s}{P[Y_s=0]} & \textit{when } P[Y_s = 0] > 0 \\ 0 & \textit{otherwise.} \end{cases}$$

Then $\Lambda_t := \int_{A_t} \lambda_s I_{\{L_s=0\}} ds$ is a quasi-$\mathcal{F}^Y$-compensator for Y.

<u>Proof:</u>

Let $H_1, ..., H_k \in \mathcal{B}(\mathbb{R}_+)$, $s \le t \in E$ and $u_1, ..., u_k \le s$. If $0 \notin \bigcap_{i=1}^k H_i$ then, on $\bigcap_{i=1}^k Y_{u_i}^{-1}(H_i) \subseteq \{Y_s > 0\}$,

$$Y(s,t] = Y_t - Y_{(s_1,t_2)} - Y_{(t_1,s_2)} + Y_s = 1 - 1 - 1 + 1 = 0.$$

From this the remainder of the arguments can follow identically those of Lemma 2.4.2 with L replaced by Y.

■

Define $\tilde{\Lambda}_t := \mathbb{E}[Y_t]$. In the case that L is the first line of the Poisson process J, differentiating $\tilde{\Lambda}_s = P[L_s > 0] = 1 - e^{-\int_0^{s_1}\int_0^{s_2}\gamma(u_1,u_2)du_2du_1}$ first with respect to s_1 yields, by the fundamental theorem of calculus,

$$\begin{aligned} &-\frac{d}{ds_1}(e^{-\int_0^{s_1}\int_0^{s_2}\gamma(u_1,u_2)du_2du_1}) \\ =\ &(\frac{d}{ds_1}\int_0^{s_1}\int_0^{s_2}\gamma(u_1,u_2)du_2du_1)e^{-\int_0^{s_1}\int_0^{s_2}\gamma(u_1,u_2)du_2du_1} \end{aligned}$$

$$= \int_0^{s_2} \gamma(s_1,u_2)du_2 e^{-\int_0^{s_1}\int_0^{s_2}\gamma(u_1,u_2)du_2du_1}$$

and differentiating this expression with respect to s_2 yields, by another application of the fundamental theorem of calculus,

$$\begin{aligned}
&\frac{d}{ds_2}\Big(\int_0^{s_2} \gamma(s_1,u_2)du_2 e^{-\int_0^{s_1}\int_0^{s_2}\gamma(u_1,u_2)du_2du_1}\Big)\\
&= e^{-\int_0^{s_1}\int_0^{s_2}\gamma(u_1,u_2)du_2du_1}\Big(\gamma(s_1,s_2)\\
&\quad -\int_0^{s_2}\gamma(s_1,u_2)du_2\cdot\frac{d}{ds_2}\int_0^{s_1}\int_0^{s_2}\gamma(u_1,u_2)du_2du_1\Big)\\
&= e^{-\int_{[0,s]}\gamma(u)du}\Big(\gamma(s)-\int_0^{s_2}\gamma(s_1,u_2)du_2\int_0^{s_1}\gamma(u_1,s_2)du_1\Big).
\end{aligned}$$

Therefore

$$f(s) := P[Y_s=0]\big(\gamma(s)-\int_0^{s_1}\int_0^{s_2}\gamma(u_1,s_2)\gamma(s_1,u_2)du_2du_1\big)$$

satisfies $\tilde{\Lambda}_t = \int_{[0,t]} f(s)ds$. In other words, f is the density needed to apply Lemma 4.2.1. Define now

$$\begin{aligned}
\lambda_s^Y &= \begin{cases} \frac{f(s)}{P(Y_s=0)} & \text{when } P(Y_s=0)>0\\ 0 & \text{otherwise}\end{cases}\\
&= \begin{cases} \gamma(s)-\int_0^{s_1}\int_0^{s_2}\gamma(u_1,s_2)\gamma(s_1,u_2)du_2du_1 & \text{when } P(Y_s=0)>0\\ 0 & \text{otherwise}\end{cases}
\end{aligned}$$

so that

$$\begin{aligned}
\Lambda_t &= \int_{[0,t]} \lambda_s^Y I_{\{Y_s=0\}}ds\\
&= \int_{[0,t]} \big(\gamma(s)-\int_0^{s_1}\int_0^{s_2}\gamma(u_1,s_2)\gamma(s_1,u_2)du_2du_1\big)I_{\{L_s=0\}}ds
\end{aligned}$$

is a quasi-$\mathcal{F}^Y$-compensator for Y, according to the lemma. Note that Λ_t is also a quasi-$\mathcal{F}^L$-compensator since $\{L_s=0\}=\{Y_s=0\}$ and, on $\{L_s>0\}=\{Y_s>0\}$,

$Y(s,t] = 0$. Also, following the arguments on the compensator for L, it is easy to show that Λ_t is also a quasi-$\mathcal{F}^{L,N}$-compensator.

Now suppose that a quasi-compensator Λ exists for Y, and that the following gain function is used to describe how favorable it is to "stop" at the lower layer B:

$$\begin{aligned} S(B) &= c_0|B \cap \xi| - c_1|B \cap \xi^c| + k_0 + k_1 Y(B) \\ &= k_0 + \int_B (-c_1 + (c_0 + c_1)X_u)du + k_1 Y(B) \end{aligned}$$

where $X_u := I_{\{L_s=0\}} = 1 - Y_u$. Letting M denote the weak $\mathcal{F}^{L,N}$-martingale $Y - \Lambda$ gives

$$S(B) = k_0 + \int_B (-c_1 + (c_0 + c_1 + k_1\lambda^Y(u))X_u)du + M(B).$$

Clearly the integrand $U_t := -c_1 + (c_0 + c_1 + k_1\lambda_t^Y)X_t$ is bounded and outer-continuous with inner limits, and so the above is a smooth semi-martingale representation of S with respect to $\mathcal{F}^{L,N}$. As shown earlier, $\mathbb{E}[X_t|\mathcal{F}_t^N] = \frac{1}{1+q_tQ_t}$, and so

$$V_t := \mathbb{E}[U_t|\mathcal{F}_t^N] = -c_0 + (c_0 + c_1 + k_1\lambda_t^Y)\frac{1}{1+q_tQ_t}.$$

Note that the form of V remains unchanged when the new gain function is used instead of Z, athough λ^Y is no longer necessarily non-negative. Lemma 3.1.5 can be applied to show that an optimal solution to the sequential detection problem (with respect to the new gain function S) exists and is given by

$$\hat{\rho} = \{t \in [0,R]; V_s > 0 \ \forall s \ll t\}$$

so long as it can be shown that V is outer-continuous with inner limits and is monotone. If q_t and λ_t^Y are increasing and decreasing respectively in each coordinate and λ_t^Y is continuous, the argument in Theorem 3.2.1 may be followed exactly to establish these conditions, however. In this case, therefore, an optimal solution to the sequential detection problem, with gain function S, exists and is given by $\hat{\rho}$.

In the case that L is the first line of the inhomogeneous Poisson process J with continuous intensity γ, λ^Y is clearly continuous. Also, if ν is the function defined by

$$\nu(t) := \int_0^{t_1} \int_0^{t_2} \gamma(u_1, t_2)\gamma(t_1, u_2)du_2du_1$$

then $\lambda^Y = \gamma - \nu$, and if $\gamma(x, y) = \gamma_1(x)\gamma_2(y)$ is decreasing in each coordinate then

$$\begin{aligned}
\lambda^Y(t) &= \gamma(t_1, t_2) - \int_0^{t_1} \int_0^{t_2} \gamma(u_1, t_2)\gamma(t_1, u_2)du_2du_1 \\
&= \gamma_1(t_1)\gamma_2(t_2) - \int_0^{t_1} \int_0^{t_2} \gamma_1(u_1)\gamma_2(t_2)\gamma_1(t_1)\gamma_2(u_2)du_2du_1 \\
&= \gamma_1(t_1)\gamma_2(t_2)(1 - \int_0^{t_1} \int_0^{t_2} \gamma_1(u_1)\gamma_2(u_2)du_2du_1)
\end{aligned}$$

is a decreasing function of t so long as both γ_1 and γ_2 are both decreasing functions, and $\nu \leq \gamma$. Also, as shown earlier, q_t is increasing in each coordinate if and only if

$$\mu_1 - \mu_0 \leq \frac{1}{|A_t \setminus A_s|} \int_{A_t \setminus A_s} \gamma(u)du,$$

and for this it is sufficient that $\gamma(R) \geq \mu_1 - \mu_0$, so long as γ is decreasing in each coordinate. When these conditions are met, then, an optimal solution exists. The conclusions of the preceding discussion are listed in the following lemma.

Lemma 4.2.2 *Let L be the first line of an inhomogeneous Poisson process J with continuous intensity γ, let $\nu(t) := \int_0^{t_1} \int_0^{t_2} \gamma(u_1, t_2)\gamma(t_1, u_2)du_2du_1$, let $\xi = \{t \in E; L_t > 0\}$, let N be a conditional Poisson process given ξ with rate μ_0 on ξ^c and μ_1 on ξ, and let $Y_t = I_{\{L_t>0\}}$. Then a solution to the optimal detection problem with gain function*

$$S_t = c_0|A_t \cap \xi| - c_1|A_t \cap \xi^c| + k_0 + k_1Y_t$$

exists if both γ and $\gamma - \nu$ are non-negative and decreasing in each coordinate, and $\gamma(R) \geq \mu_1 - \mu_0$, and is given by

$$\hat{\rho} = \{t \in [0, R]; V_s > 0 \ \forall s \ll t\}.$$

Bibliography

[1] Tina Herberts and Uwe Jensen, **Optimal Detection of a Change Point in a Poisson Process for Different Observation Schemes**, *Scandinavian Journal of Statistics* Vol 31, pp. 347 - 366, 2004.

[2] Gail Ivanoff and Ely Merzbach, **Optimal Detection of a Change-Set in a Spatial Poisson Process**, preprint, 2008.

[3] Gail Ivanoff and Ely Merzbach, **Set-Indexed Markov Processes**, *Canadian Mathematical Society Conference Proceedings* Vol 26, 2000.

[4] Gail Ivanoff and Ely Merzbach, **Set-Indexed Martingales**, *Monographs on Statistics and Applied Probability 85*, 2000.

[5] Patrick Billingsley, **Probability and Measure, Third Edition**, John Wiley & Sons Inc., 1995.

[6] Olav Kallenberg, **Random Measures, 3rd Edition**, Academic Press Inc., 1983.

[7] Olav Kallenberg, **Foundations of Modern Probability, 2nd Edition**, Springer-Verlag, 2002.

[8] D.J. Daley and D. Vere-Jones, **An Introduction to the Theory of Point Processes, Vol. 1, 2nd Edition**, Springer-Verlag, 2003.

[9] James R. Munkres, **Topology, 2nd Edition**, Prentice Hall, 2000.

[10] E. Khmaladze, R. Mnatsakanov and N. Toronjadze, **The Change-Set Problem for Vapnik-Cervonenkis Classes**, *Mathematical Methods in Statistics* Vol 15, No. 2, pp. 224-231, 2006.

[11] Ts. Hahubia and R. Mnatsakanov, **On the Mode-Change Problem for Random Measures**, *Georgian Mathematical Journal* Vol 3, No. 4, pp. 343-362, 1996.