FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTOCTORALES

uOttawa

L'Université canadienne
Canada's university

FACULTY OF GRADUATE AND POSDOCTORAL STUDIES

Farid Elktaibi

AUTEUR DE LA THÈSE / AUTHOR OF THESIS

M.Sc. (Mathematics)

GRADE / DEGREE

Department of Mathematics and Statistics

FACULTÉ, ÉCOLE, DÉPARTEMENT / FACULTY, SCHOOL, DEPARTMENT

Optimal Detection of the Support of a Poisson Process

TITRE DE LA THÈSE / TITLE OF THESIS

G. Ivanoff

DIRECTEUR (DIRECTRICE) DE LA THÈSE / THESIS SUPERVISOR

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

R. Kulik

B. Szyszkowicz

Gary W. Slater

Le Doyen de la Faculte des etudes superieures et postdoctorales / Dean of the Faculty of Graduate and Postdoctoral Studies

Optimal detection of the support of a Poisson process

Farid Elktaibi

Thesis submitted to the Faculty of Graduate and Postdoctoral Studies
in partial fulfillment 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

Your file Votre référence
ISBN: 978-0-494-73862-7
Our file Notre référence
ISBN: 978-0-494-73862-7

Abstract

We address the problem of the optimal detection of the support of a Poisson process in one and two dimensions. The point process N is assumed to be distributed according to a Poisson process with intensity μ on an unobservable random set ξ. In one dimension, $\xi = [\sigma, \infty)$ where σ is a change-point. In two dimensions, the boundary of ξ is determined by one or more incomparable points. The goal is to detect ξ in an optimal way by maximizing the expected value of a reward function. Optimal solutions are found in the one dimensional case for two different information schemes. Analytical and practical approaches are provided in the two dimensional case.

Acknowledgements

First and foremost I am heartily thankful to my supervisor Professor Gail Ivanoff, whose encouragement, guidance and support from the initial to the final level enabled me to develop an understanding of the subject. Her way of removing misunderstanding gave me the opportunity to pursue this work in the best conditions. One simply could not wish for a better or friendlier supervisor. Thank you deeply.

I convey a special acknowledgment to Dr Rafal Kulik for his help and his valuable advices.

Many thanks go in particular to Dr Mahmoud Zarepour for his precious time to answer my questions and give his comments.

I also offer my regards to all of those who supported me in any respect throughout the process of writing this thesis.

Lastly, I wish to express my love and gratitude to my beloved family for their understanding and endless love throughout the duration of my studies.

Dedication

This work is dedicated to my little one Yasmine Jannat, to my lovely wife Aouatif, to my mother Nadia and to the spirit of my grand-parents.

Contents

List of Figures

Chapter 1

Introduction

Statistical methods dealing with the detection of changes in the characteristics of a random process can be of great use in many problems. These methods have accompanied the rapid growth in data during the last few decades. According to a tradition of more than thirty years, we call this sphere of statistical analysis the "theory of change-point detection".

During the last twenty years, we have witnessed many exciting developments in this theory. Traditional trends have flourished anew and new promising directions of research have emerged such as the detection of a change-set which can be viewed as the multiparameter version of the classical change-point detection problem.

Medicine, biology, technology, archaeology and forestry are many potential areas of application. The broad applicability of change-point or change-set methods in various areas makes this an attractive field of research, with still many open and unsettled questions, particularly in the relatively new area of change-set detection.

In this thesis, we address the problem of *optimal* detection. There is a vast literature on the general change-point problem in one dimension but very little work has been done on change-set detection:

The optimal detection of an exponential change time in a Poisson process, in one dimension, was studied in [2] and [6] using different techniques . For the two dimensional case, we found only a few papers addressing the problem of change-set (cf.[1], [3], [4], [5]). Since we will be focusing on *optimal support-detection*, we restrict our attention to the two relevant papers [2] and [3].

In paper [2] the authors (Herberts and Jensen) consider a point process N on the line which is, given an unobservable random time σ, Poisson with intensity $\mu_0(s)$ up to σ and $\mu_1(s)$ after σ. They supposed that $0 < \mu_0(s) \le \mu_1(s)$ for all $s \in \mathbb{R}_+$, and that σ is exponentially distributed. Different observation schemes were discussed, namely sequential observation where the available information at time t is the collection of jump times of N prior to t; and the ex-post decision after observing the point process up to a fixed time T. Despite the fact that these information schemes differ in their structure, the detection problem can be viewed as an optimal stopping problem. This problem can be solved by maximizing the expected value of a reward function that describes the gain associated with stopping at time t:

$$Z_t = c_0(t \wedge \sigma) - c_1(t - \sigma)^+.$$

Inspired by the work done in [2], Gail Ivanoff and Ely Merzbach proposed an extension of the one-dimensional case to a two-dimensional case. In their paper [3], Ivanoff and Merzbach used the theory of set-indexed martingales to prove the existence of an optimal solution to the detection problem. The goal was to detect the occurrence of an unobservable random set ξ on which an observable Poisson process N changes its intensity from μ_0 to μ_1.

The goal in this research paper is to discuss the special case when $\mu_0 = 0$ that relates to the so-called support detection problem. This is the primary concern of this thesis. In this case, the interval $[\sigma, \infty)$ in one dimension or the random set ξ in two dimensions denotes the support of Poisson process with intensity μ_1. We note that this question was not addressed in [2] and was only discussed briefly in [3] in comment 4.2. Moreover, the formula given cannot go beyond the theoretical view since the level of abstraction and complexity does not allow its use in practical situations.

From a graduate student perspective, we will deal with the problem of detecting the support despite of all the issues and the difficulties that complicate our analysis and that might not be resolved at this stage. To motivate the two dimensional problem, we will first analyze the one dimensional case in two different schemes: the sequential observation and the ex-post analysis. To this end, an introduction to the theory of the support detection problem is given in the next chapter. Lemma 2.1.2 which is the work's main tool for the one dimensional case is given in Section 2.1. Following this, Section 2.2 presents all the details necessary to solve the optimal support-detection problem for the sequential structure using a Bayes-type formula. Chapter 2 ends with an analysis of the optimal solution presented previously. We will show that the reward function can be chosen to yield unbiased solutions in certain situations.

In Chapter 3, the ex-post model is stated in relation to Section 2.1. It is important to note here that the optimal solution of the ex-post analysis is entirely developed in Section 3.2. Finally, the last section is dedicated to analyzing the optimal solution of the ex-post model as was done in Chapter 2.

Chapter 4 introduces the optimal support-detection problem in two dimensions. The model and the necessary background including the powerful lemma 4.1.8 are introduced in Section 4.1. The sequential observation structure will be the only one of relevance here. The solution of the support problem, restricted to certain conditions, terminates this Section. Section 4.2 develops an analytical approach to theoretically simplify the optimal solution found before in order to be applied to practical situations. We will see that the analytic approach unfortunately does not yield a closed-form solution, and so the remainder of this Chapter is devoted to developing a simulation based approach to constructing the optimal solution to the detection problem. Detailed applications will be presented for some arbitrary parameters that play a role in our assumptions.

Finally, in Chapter 5 we discuss the estimation of some of the parameters related to the detection-support problem in both the one and the two dimensional cases. We conclude with a discussion of some directions for further research.

Chapter 2

Sequential observation in one dimension

This chapter aims to provide an overview of the optimal detection of an exponential change time σ introducing a Poisson Process N for the sequential observation scheme. The first section is devoted to describing the model and presenting sufficient conditions leading to an optimal solution to the detection problem. In section 1.2, solutions of this problem are presented for equal and different rates of the exponential change point σ and the Poisson process N. Finally, analysis of the solution presented previously and its bias constitutes the last section.

2.1 The Model

We are given a non-explosive point process $N = \{N_t, t \in \mathbb{R}_+\}$ defined on $\mathbb{R}_+$ and an unobservable random time $\sigma \geq 0$. Given σ, we assume that N is a Poisson process with intensity μ on $\xi = [\sigma, \infty)$. Our goal is to detect, as well as possible, the occurrence of the support ξ or more precisely the change point σ. Since the change point σ is unobservable, we would like to find, if it exists, a random time $\hat{\sigma}$ adapted to the available information structure that maximizes the expected value of a specified valuation: if the information available at $t \in \mathbb{R}_+$ is represented by the σ-field $\mathcal{F}_t$, then we must have that $\hat{\sigma}$ is an $\mathcal{F}_t$-stopping time: i.e. $\{\hat{\sigma} \leq t\} \in \mathcal{F}_t$ for all t.

For the sequential model described in this chapter, the information is described by the filtration $\mathbb{F}^N = (\mathcal{F}_t^N, t \in \mathbb{R}_+)$ where $\mathcal{F}_t^N = \sigma(N_s : s \leq t)$.

The gain function at t is defined by:

$$Z_t = c_0(t \wedge \sigma) - c_1(t - \sigma)^+ \text{ where } c_0 > 0,\ c_1 > 0.$$

The gain function is piecewise linear, increasing at rate c_0 before the change point σ and decreasing at rate c_1 after. It is clearly maximized when $t = \sigma$. The parameters c_0 and c_1 can be interpreted respectively as a reward before the change time point σ and a penalty after.

If $A_t = [0, t]$, $X_u = I_{(u \in \xi^c)} = 1 - I_{(u \in \xi)} = I_{(\sigma > u)}$ and $|\ .\ |$ denotes Lebesgue measure, we can rewrite the gain function as :

$$\begin{aligned}
Z_t &= c_0|A_t \cap \xi^c| - c_1|A_t \cap \xi| \\
&= (c_0 + c_1)|A_t \cap \xi^c| - c_1|A_t| \text{ since } A_t = (A_t \cap \xi) \cup (A_t \cap \xi^c) \\
&= \int_{A_t} \left(-c_1 + (c_0 + c_1) I_{(u \in \xi^c)}\right) du \\
&= \int_{A_t} (-c_1 + (c_0 + c_1) X_u)\, du \\
&= \int_{A_t} U_u \, du \quad \text{where } U_u = -c_1 + (c_0 + c_1) X_u. \qquad (2.1.1)
\end{aligned}$$

The detection problem described above can be viewed as an optimal stopping problem. Detecting the change point 'as well as possible' means that we will determine an observable stopping time $\hat{\sigma}$ that is optimal in the following sense :

Definition 2.1.1 *Let $\mathbb{F}$ be an arbitrary filtration. An $\mathbb{F}$-stopping time $\hat{\sigma}$ is called an optimal solution for the sequential detection problem if:*

$$E[Z_{\hat{\sigma}}] = \sup\{E[Z_\sigma], \sigma \text{ is an } \mathbb{F}\text{-stopping time}\}.$$

Lemma 2.1.2 *(cf.[2]; Theorem 2) Let $\mathbb{F} = (\mathcal{F}_u, u \in \mathbb{R}_+)$ be an arbitrary filtration and let V_u a version of $E[U_u|\mathcal{F}_u]$ that is right continuous with left limits. Then for any $\mathbb{F}$-stopping time $\sigma > 0$,*

$$E\left[\int_0^\sigma U_u\,du\right] = E\left[\int_0^\sigma V_u\,du\right].$$

Moreover, if V is decreasing on $\mathbb{R}_+$, then the $\mathbb{F}$-stopping time $\hat{\sigma}$ defined by

$$\hat{\sigma} = \inf\{u \in \mathbb{R}_+ : V_u \leq 0\} \tag{2.1.2}$$

is an optimal solution for the sequential detection problem.

This means that the solution is exact in the sense that it is explicitly defined by the observed data.

■

2.2 Optimal Solution for the Sequential Detection problem

In order to give an explicit formula of (2.1.2), we will use some Bayesian arguments combined with the following fact:

$$\begin{aligned} N_u > 0 \ &\Rightarrow u \geq \sigma \\ &\Rightarrow u \in \xi. \end{aligned}$$

Now, let F_σ and f_σ to be respectively the distribution and the density of σ. Recalling that $\mathcal{F}_u = \mathcal{F}_u^N$, we have:

$$\begin{aligned} E[X_u|\mathcal{F}_u] &= E\left[I_{u\in\xi^c}|\mathcal{F}_u\right] \\ &= E\left[I_{u\in\xi^c}|\mathcal{F}_u\right] I_{(N_u=0)} + E\left[I_{u\in\xi^c}|\mathcal{F}_u\right] I_{(N_u>0)} \end{aligned}$$

$$\begin{aligned}
&= E\left[I_{(\sigma>u)}|N_u=0\right] I_{(N_u=0)} \\
&= P\left(\sigma>u|N_u=0\right) I_{(N_u=0)} \\
&= \frac{P\left[(\sigma>u)\cap(N_u=0)\right]}{P(N_u=0)} I_{(N_u=0)} \\
&= \frac{P(\sigma>u)}{P(N_u=0)} I_{(N_u=0)} \\
&= \frac{1-F_\sigma(u)}{P(N_u=0)} I_{(N_u=0)}.
\end{aligned}$$

On the other hand,

$$\begin{aligned}
P(N_u=0) &= E[P(N_u=0|\sigma)] \\
&= E[P(N_u=0|\sigma)I_{(\sigma\le u)} + P(N_u=0|\sigma)I_{(\sigma>u)}] \\
&= E[e^{-\mu(u-\sigma)}I_{(\sigma\le u)} + I_{(\sigma>u)}] \\
&= 1-F_\sigma(u) + \int_0^u e^{-\mu(u-s)} f_\sigma(s)ds.
\end{aligned}$$

Therefore,

$$E[X_u|\mathcal{F}_u] = \frac{I_{(N_u=0)}}{1+\frac{1}{1-F_\sigma(u)}\int_0^u e^{-\mu(u-s)} f_\sigma(s)\, ds}$$

$$\text{and} \quad V_u = -c_1 + \frac{(c_0+c_1)I_{(N_u=0)}}{1+\frac{1}{1-F_\sigma(u)}\int_0^u e^{-\mu(u-s)} f_\sigma(s)\, ds}.$$

Let $\{\tau_i\}$ be the ordered jump points of the Poisson process N. We notice that $\{N_u>0\} = \{\tau_1 \le u\}$. If we assume now that σ is exponentially distributed with parameter λ and we use the formula above for V_u, we get:

$$\begin{aligned}
V_u &= -c_1 + \frac{(c_0+c_1)}{1+e^{\lambda u}\int_0^u e^{-\mu(u-s)}\lambda e^{-\lambda s}\, ds} I_{(N_u=0)} \\
&= -c_1 + \frac{(c_0+c_1)}{1+\lambda e^{-u(\mu-\lambda)}\int_0^u e^{(\mu-\lambda)s}\, ds} I_{(N_u=0)}. \qquad (2.2.1)
\end{aligned}$$

According to the expression above, two cases are to be discussed: $\mu = \lambda$ and $\mu \neq \lambda$.

i) $\mu = \lambda:$

From (2.2.1) we get:

$$V_u = -c_1 + \frac{(c_0 + c_1)}{1 + \lambda u} I_{(N_u=0)}.$$

It is obvious that V_u is decreasing in u.

$$\begin{aligned} V_u \leq 0 \quad &\Leftrightarrow \quad -c_1 + \frac{(c_0 + c_1)}{1 + \lambda u} I_{(N_u=0)} \leq 0 \\ &\Leftrightarrow \quad \{N_u > 0\} \text{ or } \{N_u = 0 \text{ and } \frac{c_0}{c_1} + 1 \leq 1 + \lambda u\} \\ &\Leftrightarrow \quad \{N_u > 0\} \text{ or } \{N_u = 0 \text{ and } \frac{c_0}{\lambda c_1} \leq u\}, \end{aligned}$$

and

$$\hat{\sigma} = \frac{c_0}{\lambda c_1} \wedge \tau_1.$$

ii) $\mu \neq \lambda:$

Using (2.2.1) again,

$$\begin{aligned} V_u &= -c_1 + \frac{c_0 + c_1}{1 + \frac{\lambda}{\mu-\lambda} e^{-u(\mu-\lambda)} (e^{(\mu-\lambda)u} - 1)} I_{(N_u=0)} \\ &= -c_1 + \frac{c_0 + c_1}{1 + \frac{\lambda}{\mu-\lambda} (1 - e^{-(\mu-\lambda)u})} I_{(N_u=0)}. \end{aligned}$$

We observe clearly that V_u is decreasing in u.

$$\begin{aligned} V_u \leq 0 \quad &\Leftrightarrow \quad -c_1 + \frac{c_0 + c_1}{1 + \frac{\lambda}{\mu-\lambda} (1 - e^{-(\mu-\lambda)u})} I_{(N_u=0)} \leq 0 \\ &\Leftrightarrow \quad \{N_u > 0\} \text{ or } \{N_u = 0 \text{ and } \frac{c_0}{c_1} + 1 \leq 1 + \frac{\lambda}{\mu - \lambda} (1 - e^{-(\mu-\lambda)u})\} \\ &\Leftrightarrow \quad \{N_u > 0\} \text{ or } \{N_u = 0 \text{ and } \frac{c_0}{c_1} \leq \frac{\lambda}{\lambda - \mu} (e^{(\lambda-\mu)u} - 1)\}. \end{aligned} \tag{2.2.2}$$

Two subcases are to be discussed here according to the sign of $\lambda - \mu$.

- $\mu < \lambda:$

$$(2.2.2) \quad \Leftrightarrow \quad \{N_u > 0\} \text{ or } \{N_u = 0 \text{ and } \frac{(\lambda - \mu) c_0}{\lambda c_1} + 1 \leq e^{(\lambda-\mu)u}\}$$

$$\Leftrightarrow \quad \{N_u > 0\} \text{ or } \{N_u = 0 \text{ and } \frac{1}{\lambda-\mu}\ln\left[\frac{(\lambda-\mu)c_0}{\lambda c_1}+1\right] \leq u\},$$

so

$$\hat{\sigma} = \frac{1}{\lambda-\mu}\ln\left[\frac{(\lambda-\mu)c_0}{\lambda c_1}+1\right] \wedge \tau_1.$$

- $\mu > \lambda$:

$$(2.2.2) \Leftrightarrow \{N_u > 0\} \text{ or } \{N_u = 0 \text{ and } \frac{(\lambda-\mu)c_0}{\lambda c_1}+1 \geq e^{(\lambda-\mu)u}\}.$$

Remark that:

$$\begin{aligned}
\frac{(\lambda-\mu)c_0}{\lambda c_1}+1 > 0 \quad &\Leftrightarrow \quad \frac{(\mu-\lambda)c_0}{\lambda c_1} < 1 \\
&\Leftrightarrow \quad \frac{\mu}{\lambda} - 1 < \frac{c_1}{c_0} \\
&\Leftrightarrow \quad \mu < \lambda\left(1+\frac{c_1}{c_0}\right),
\end{aligned}$$

So

$$\text{if } \mu \geq \lambda\left(1+\frac{c_1}{c_0}\right), \qquad V_u \leq 0 \Leftrightarrow N_u > 0.$$

In this case,

$$\hat{\sigma} = \tau_1.$$

In the second situation when $\mu < \lambda\left(1+\frac{c_1}{c_0}\right)$

$$\begin{aligned}
(2.2.2) \quad &\Leftrightarrow \quad \{N_u > 0\} \text{ or } \{N_u = 0 \text{ and } \ln\left[\frac{(\lambda-\mu)c_0}{\lambda c_1}+1\right] \geq (\lambda-\mu)u\} \\
&\Leftrightarrow \quad \{N_u > 0\} \text{ or } \{N_u = 0 \text{ and } \frac{1}{\lambda-\mu}\ln\left[\frac{(\lambda-\mu)c_0}{\lambda c_1}+1\right] \leq u\},
\end{aligned}$$

and

$$\hat{\sigma} = \frac{1}{\lambda-\mu}\ln\left[\frac{(\lambda-\mu)c_0}{\lambda c_1}+1\right] \wedge \tau_1.$$

Summarizing the formulas found for $\hat{\sigma}$, we provide the following expression for the optimal solution of the detection problem for the sequential observation scheme:

$$\hat{\sigma} = \begin{cases} \tau_1 & \text{if } \mu \geq \lambda(1 + \frac{c_1}{c_0}) \\ \\ \frac{c_0}{\lambda c_1} \wedge \tau_1 & \text{if } \lambda = \mu \\ \\ \frac{1}{\lambda - \mu} \ln\left[\frac{(\lambda-\mu)c_0}{\lambda c_1} + 1\right] \wedge \tau_1 & \text{elsewhere} \end{cases} \tag{2.2.3}$$

■

2.3 Analysis of the optimal solution

Now, we are going to evaluate $E[\hat{\sigma}]$ for the above found different cases of $\hat{\sigma}$ according to μ and λ. We will also try to find, if it is possible, an "unbiased estimator" for σ in each case: i.e. $E[\hat{\sigma}] = \dfrac{1}{\lambda}$. To begin this task, we write first:

$$\hat{\sigma} = M \wedge \tau_1 = M\ I_{(M \leq \tau_1)} + \tau_1 I_{(\tau_1 < M)}.$$

Thus,

$$\begin{aligned} E[\hat{\sigma}] &= MP(M \leq \tau_1) + E[\tau_1 I_{(\tau_1 < M)}] \\ &= M(1 - F_{\tau_1}(M)) + \int_0^M t f_{\tau_1}(t)\, dt. \end{aligned} \tag{2.3.1}$$

Therefore, in order to simplify (2.3.1) we need to find the distribution of τ_1.

$$\begin{aligned} F_{\tau_1}(t) &= P(\tau_1 \leq t) \\ &= E\left[P(\tau_1 \leq t|\sigma)\right] \\ &= E\left[P(\tau_1 \leq t|\sigma) I_{(\sigma \leq t)} + P(\tau_1 \leq t|\sigma) I_{(\sigma > t)}\right] \\ &= E\left[P(\tau_1 \leq t|\sigma) I_{(\sigma \leq t)}\right] \\ &= E\left[\left(1 - e^{-\mu(t-\sigma)}\right) I_{(\sigma \leq t)}\right] \end{aligned}$$

$$\begin{aligned} &= \int_0^t \left(1 - e^{-\mu(t-s)}\right) f_\sigma(s)\, ds \\ &= \int_0^t \left(1 - e^{-\mu(t-s)}\right) \lambda e^{-\lambda s}\, ds \\ &= \int_0^t \lambda e^{-\lambda s}\, ds - \lambda e^{-\mu t} \int_0^t e^{(\mu-\lambda)s}\, ds. \end{aligned} \tag{2.3.2}$$

We remark that the value of $F_{\tau_1}(t)$ depends on λ and μ. In the following, two cases are to be discussed.

i) $\mu = \lambda$:

In this case, we have:

$$\begin{aligned} F_{\tau_1}(t) &= 1 - e^{-\lambda t} - \lambda t e^{-\lambda t}, t > 0 \\ f_{\tau_1}(t) &= \lambda^2 t e^{-\lambda t}, t > 0 \\ \text{and} \quad M &= \frac{c_0}{\lambda c_1}. \end{aligned}$$

Substituting now in (2.3.1) , we get:

$$\begin{aligned} E[\hat{\sigma}] &= M(e^{-\lambda M} + \lambda M e^{-\lambda M}) + \int_0^M \lambda^2 t^2 e^{-\lambda t}\, ds \\ &= \frac{2}{\lambda} - Me^{-\lambda M} - \frac{2}{\lambda} e^{-\lambda M} \\ &= \frac{1}{\lambda} + \left[\frac{1}{\lambda} - \frac{c_0}{\lambda c_1} e^{-\frac{c_0}{c_1}} - \frac{2}{\lambda} e^{-\frac{c_0}{c_1}}\right] \\ &= \frac{1}{\lambda} + \frac{1}{\lambda}\left[1 - \frac{c_0}{c_1} e^{-\frac{c_0}{c_1}} - 2e^{-\frac{c_0}{c_1}}\right]. \end{aligned}$$

To get an unbiased estimator, we need $1 - \frac{c_0}{c_1} e^{-\frac{c_0}{c_1}} - 2e^{-\frac{c_0}{c_1}} = 0$.

Define for $x > 0$

$$f(x) = 1 - (x+2)e^{-x};$$

f is continuous and strictly increasing on $(0, \infty)$, thus f is a bijection from $(0, \infty)$ to $(-1, 1)$. So $\exists!\ x_1 > 0$ such that $f(x_1) = 0$, in fact $x_1 \simeq 1.1474025$.

Therefore, if $c_0 = x_1 c_1$ then $E[\hat{\sigma}] = \frac{1}{\lambda}$.

ii) $\mu \neq \lambda:$

Using (2.3.2) again, we have:

$$\begin{aligned} F_{\tau_1}(t) &= 1 - e^{-\lambda t} - \frac{\lambda}{\mu - \lambda} e^{-\mu t}\left(e^{(\mu-\lambda)t} - 1\right) \\ &= 1 - \frac{\mu e^{-\lambda t} - \lambda e^{-\mu t}}{\mu - \lambda}, t > 0 \\ f_{\tau_1}(t) &= \frac{\mu\lambda}{\mu - \lambda}\left(e^{-\lambda t} - e^{-\mu t}\right), t > 0. \end{aligned}$$

Thus,

$$\begin{aligned} E[\hat{\sigma}] &= M\left(\frac{\mu e^{-\lambda M} - \lambda e^{-\mu M}}{\mu - \lambda}\right) + \frac{\mu\lambda}{\mu - \lambda}\int_0^M t(e^{-\lambda t} - e^{-\mu t})\, dt \\ &= \frac{\mu^2 - \lambda^2}{\mu\lambda(\mu - \lambda)} + \frac{1}{\mu - \lambda}\left(\frac{\lambda}{\mu}e^{-\mu M} - \frac{\mu}{\lambda}e^{-\lambda M}\right) \\ &= \frac{\mu + \lambda}{\mu\lambda} + \frac{1}{\mu - \lambda}\left(\frac{\lambda}{\mu}e^{-\mu M} - \frac{\mu}{\lambda}e^{-\lambda M}\right) \\ &= \frac{1}{\lambda} + \left[\frac{1}{\mu} + \frac{1}{\mu - \lambda}\left(\frac{\lambda}{\mu}e^{-\mu M} - \frac{\mu}{\lambda}e^{-\lambda M}\right)\right]. \end{aligned} \tag{2.3.3}$$

Once again, it is necessary to separate our analysis into three subcases:

- $\mu < \lambda:$

We have seen in (2.2.3) that $M = \frac{1}{\lambda - \mu}\ln[\frac{(\lambda-\mu)c_0}{\lambda c_1} + 1]$, so (2.3.3) becomes:

$$E[\hat{\sigma}] = \frac{1}{\lambda} + \left\{\frac{1}{\mu} + \frac{1}{\mu - \lambda}\left[\frac{\lambda}{\mu}\left(\frac{(\lambda - \mu)c_0}{\lambda c_1} + 1\right)^{\frac{-\mu}{\lambda - \mu}} - \frac{\mu}{\lambda}\left(\frac{(\lambda - \mu)c_0}{\lambda c_1} + 1\right)^{\frac{-\lambda}{\lambda - \mu}}\right]\right\}.$$

Define for $x > 0$

$$g(x) = \frac{1}{\mu} + \frac{1}{\mu - \lambda}\left[\frac{\lambda}{\mu}\left(\frac{(\lambda - \mu)x}{\lambda} + 1\right)^{\frac{-\mu}{\lambda - \mu}} - \frac{\mu}{\lambda}\left(\frac{(\lambda - \mu)x}{\lambda} + 1\right)^{\frac{-\lambda}{\lambda - \mu}}\right];$$

g is continuous and strictly increasing on $(0, \infty)$, thus g is a bijection from $(0, \infty)$ to $(-\frac{1}{\lambda}, \frac{1}{\mu})$. Therefore $\exists!\ x_2 > 0$ such that $g(x_2) = 0$.

Thus, if $c_0 = x_2 c_1$ then $E[\hat{\sigma}] = \dfrac{1}{\lambda}$.

- $\lambda < \mu < \lambda(1 + \frac{c_1}{c_0})$:

In this case $M = \frac{1}{\lambda - \mu} \ln[\frac{(\lambda - \mu)c_0}{\lambda c_1} + 1]$, so (2.3.3) becomes:

$$E[\hat{\sigma}] = \frac{1}{\lambda} + \left\{ \frac{1}{\mu} + \frac{1}{\mu - \lambda} \left[\frac{\lambda}{\mu} \left(\frac{(\lambda - \mu)c_0}{\lambda c_1} + 1 \right)^{\frac{-\mu}{\lambda - \mu}} - \frac{\mu}{\lambda} \left(\frac{(\lambda - \mu)c_0}{\lambda c_1} + 1 \right)^{\frac{-\lambda}{\lambda - \mu}} \right] \right\}.$$

Define for $x \in (0, \frac{\lambda}{\mu - \lambda})$

$$\begin{aligned} h(x) &= \frac{1}{\mu} + \frac{1}{\mu - \lambda} \left[\frac{\lambda}{\mu} \left(\frac{(\lambda - \mu)x}{\lambda} + 1 \right)^{\frac{-\mu}{\lambda - \mu}} - \frac{\mu}{\lambda} \left(\frac{(\lambda - \mu)x}{\lambda} + 1 \right)^{\frac{-\lambda}{\lambda - \mu}} \right] \\ &= \frac{1}{\mu} + \frac{1}{\mu - \lambda} \left[\frac{\lambda}{\mu} \left(\frac{(\lambda - \mu)x}{\lambda} + 1 \right)^{\frac{\mu}{\mu - \lambda}} - \frac{\mu}{\lambda} \left(\frac{(\lambda - \mu)x}{\lambda} + 1 \right)^{\frac{\lambda}{\mu - \lambda}} \right]; \end{aligned}$$

h is continuous and strictly increasing on $(0, \frac{\lambda}{\mu - \lambda})$, thus h is a bijection from $(0, \frac{\lambda}{\mu - \lambda})$ to $(-\frac{1}{\lambda}, \frac{1}{\mu})$. So $\exists!\ x_3 \in (0, \frac{\lambda}{\mu - \lambda})$ such that $h(x_3) = 0$.

Thus, if $c_0 = x_3 c_1$ then $E[\hat{\sigma}] = \frac{1}{\lambda}$.

- $\mu \geq \lambda(1 + \frac{c_1}{c_0})$:

In this case $\hat{\sigma} = \tau_1$.

So,

$$\begin{aligned} E[\hat{\sigma}] &= \int_0^\infty \frac{\lambda\mu}{\mu - \lambda} t(e^{-\lambda t} - e^{-\mu t})\, dt \\ &= \frac{1}{\mu - \lambda} \left[\mu \int_0^\infty \lambda t e^{-\lambda t}\, dt \; - \lambda \int_0^\infty \mu t e^{-\mu t}\, dt \right] \\ &= \frac{1}{\mu - \lambda} \left[\frac{\mu}{\lambda} - \frac{\lambda}{\mu} \right] \\ &= \frac{1}{\lambda} + \frac{1}{\mu}. \end{aligned}$$

We see that the bias is constant and equal to $\frac{1}{\mu}$.

Comment 2.3.1 *In the cases when $\mu < \lambda$ and $\lambda < \mu < \lambda(1 + \frac{c_1}{c_0})$, we proved the existence of an "unbiased estimator" of σ. Meanwhile, we need the values of the parameters μ and λ to be able to produce numerically the approximation of the numbers*

x_2 and x_3 in order to give an explicit expression for the "unbiased estimator" $\hat{\sigma}$.

■

Chapter 3

Ex-post analysis in one dimension

In this chapter, we will focus on a different information scheme, namely the ex-post structure. The model is expounded in relation to definition 2.1.1 and lemma 2.1.2 and is presented in the first section. The optimal solution for the ex-post detection problem is examined in detail in the second section. Different cases are to be discussed corresponding to the parameters which determine the optimal solution. Finally, expectation of the solution and its bias compose the last section.

3.1 The Model

Here we assume that the counting process N has been observed up to the fixed time T and we aim to estimate the position of the change point (in particular, it has to be checked if $\sigma(\omega) \leq T$ or $\sigma(\omega) > T$). The corresponding filtration is given by $\mathbb{A} = (\mathcal{A}_t, t \in \mathbb{R}_+)$, where $\mathcal{A}_t \equiv \mathcal{F}_T^N = \sigma(N_s, 0 \leq s \leq T)$ for all $t \in \mathbb{R}_+$.
This is a generalization of the ex-post analysis as here we also consider $t > T$, i.e.at time T one is interested in making inference about a change point not only in the past but also in the future.

Based on the definition 2.1.1 and lemma 2.1.2, the task is to determine a stopping time with respect to the corresponding filtration such that the expected gain is maximized. As assumed in the first chapter, σ is exponentially distributed with parameter λ and given σ, N is a Poisson process with intensity μ on $[\sigma, \infty)$.

If we denote by τ_1 the first jump point of the Poisson process N, we can see that $\tau_1 = \sigma + \upsilon$, where υ is exponentially distributed with rate μ and independent of the change point σ. The distribution of τ_1 will be given later.

The gain function, whose expectation is to be maximized, remains the same as seen in (2.1.1):

$$Z_t = \int_{A_t} U_u \, du \quad \text{where} \quad U_u = -c_1 + (c_0 + c_1)X_u, \;\; c_0 > 0, \;\; c_1 > 0,$$

and we recall that $A_t = [0, t]$ and $X_u = I_{(u \in \xi^c)} = I_{(\sigma > u)}$.

As mentioned before in lemma 2.1.2, if V_u is decreasing on $\mathbb{R}_+$, then the optimal solution is given by:

$$\hat{\sigma} = \inf\{u \in \mathbb{R}_+ : V_u \leq 0\},$$

where now $V_u = E[U_u|\mathcal{A}_u] = E[U_u|\mathcal{A}_T]$.

■

3.2 Optimal Solution for the Ex-post Analysis

In order to provide the optimal solution for the ex-post analysis, we need to evaluate V_u or more precisely $E[X_u|\mathcal{A}_T]$.

$$\begin{aligned}
E[X_u|\mathcal{A}_T] &= P(\sigma > u|\mathcal{A}_T) \\
&= P(\sigma > u|\mathcal{A}_T)I_{(\tau_1 \leq u)} + P(\sigma > u|\mathcal{A}_T)I_{(u < \tau_1 \leq T)} + P(\sigma > u|\mathcal{A}_T)I_{(u \leq T < \tau_1)} \\
&\quad + P(\sigma > u|\mathcal{A}_T)I_{(u > T, \tau_1 > T)} \\
&= P(\sigma > u|\tau_1)I_{(u < \tau_1 \leq T)} + P(\sigma > u|\tau_1 > T)I_{(u \leq T < \tau_1)} \\
&\quad + P(\sigma > u|\tau_1 > T)I_{(u > T, \tau_1 > T)}. \qquad (3.2.1)
\end{aligned}$$

Now, we will compute separately each term in the right side of the formula above to get an explicit form for $E[X_u|\mathcal{A}_T]$.

We first note that:

$$
\begin{aligned}
F_{\tau_1|\sigma}(t) &= P(\tau_1 \leq t|\sigma) \\
&= \left[1 - e^{-\mu(t-\sigma)}\right] I_{(\sigma<t)} \\
f\tau_1|\sigma(t) &= \mu e^{-\mu(t-\sigma)} I_{(\sigma<t)}.
\end{aligned}
$$

Using Bayes' formula, we get:

$$
\begin{aligned}
f_{\sigma|\tau_1}(s) &= \frac{\mu e^{-\mu(\tau_1-s)} I_{(s<\tau_1)} \lambda e^{-\lambda s}}{\int_0^\infty \mu e^{-\mu(\tau_1-y)} I_{(y<\tau_1)} \lambda e^{-\lambda y}\, dy} \\
&= \frac{e^{(\mu-\lambda)s} I_{(s<\tau_1)}}{\int_0^{\tau_1} e^{(\mu-\lambda)y}\, dy}
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
F_{\sigma|\tau_1}(u) &= \int_0^u f_{\sigma|\tau_1}(s)\, ds \\
&= \frac{\int_0^u e^{(\mu-\lambda)s} I_{(s<\tau_1)}\, ds}{\int_0^{\tau_1} e^{(\mu-\lambda)y}\, dy} \\
&= \begin{cases}
0 & \text{if } u \leq 0 \\
\frac{u}{\tau_1} & \text{if } 0 < u < \tau_1 \text{ and } \lambda = \mu \\
\frac{e^{(\mu-\lambda)u}-1}{e^{(\mu-\lambda)\tau_1}-1} & \text{if } 0 < u < \tau_1 \text{ and } \lambda \neq \mu \\
1 & \text{if } u \geq \tau_1.
\end{cases}
\end{aligned}
\tag{3.2.2}
$$

To determine the second and the third terms in the right side of (3.2.1), we observe that if $u \leq T < \tau_1$ then

$$\begin{aligned}\{\sigma > u\} &= [\{\sigma > u\} \cap \{\sigma > T\}] \cup [\{\sigma > u\} \cap \{\sigma \leq T\}] \\ &= \{\sigma > T\} \cup \{u < \sigma \leq T\}.\end{aligned}$$

Hence,

$$\begin{aligned}P(\sigma > u|\tau_1 > T) &= \frac{P\left[\{\sigma > u\} \cap \{\tau_1 > T\}\right]}{P(\tau_1 > T)} \\ &= \frac{P\left[(\{\sigma > T\} \cap \{\tau_1 > T\}) \cup (\{u < \sigma \leq T\} \cap \{\tau_1 > T\})\right]}{P(\tau_1 > T)} \\ &= \frac{P(\sigma > T) + P\left[\{u < \sigma \leq T\} \cap \{\tau_1 > T\}\right]}{P(\tau_1 > T)}.\end{aligned}$$

On the other hand, we can also note that:

$$\begin{aligned}P\left[\{u < \sigma \leq T\} \cap \{\tau_1 > T\}\right] &= E\left[E[I_{\{u<\sigma\leq T\}} I_{(\tau_1 > T)} | \sigma]\right] \\ &= E\left[I_{\{u<\sigma\leq T\}} P(\tau_1 > T|\sigma)\right] \\ &= E\left[e^{-\mu(T-\sigma)} I_{\{u<\sigma\leq T\}}\right] \\ &= \int_u^T e^{-\mu(T-y)} \lambda e^{-\lambda y}\, dy \\ &= \lambda e^{-\mu T} \int_u^T e^{(\mu-\lambda)y}\, dy.\end{aligned}$$

Therefore,

$$\begin{aligned}P(\sigma > u|\tau_1 > T) &= \frac{P(\sigma > T) + \lambda e^{-\mu T} \int_u^T e^{(\mu-\lambda)y} dy}{P(\tau_1 > T)} \\ &= \frac{e^{-\lambda T} + \lambda e^{-\mu T} \int_u^T e^{(\mu-\lambda)y}\, dy}{P(\tau_1 > T)}.\end{aligned} \tag{3.2.3}$$

If we suppose now that $\tau_1 > T$ and $u > T$, then

$$\begin{aligned}P(\sigma > u|\tau_1 > T) &= \frac{P\left[\{\sigma > u\} \cap \{\tau_1 > T\}\right]}{P(\tau_1 > T)} \\ &= \frac{P(\sigma > u)}{P(\tau_1 > T)}\end{aligned}$$

$$= \frac{e^{-\lambda u}}{P(\tau_1 > T)}. \tag{3.2.4}$$

From what has been presented previously and using (3.2.3) and (3.2.4), we can conclude that:

$$\begin{aligned} E[X_u|\mathcal{A}_T] &= \left[1 - F_{\sigma|\tau_1}(u)\right] I_{\{u<\tau_1\leq T\}} + \frac{e^{-\lambda T} + \lambda e^{-\mu T}\int_u^T e^{(\mu-\lambda)y}\,dy}{P(\tau_1 > T)} I_{\{u\leq T<\tau_1\}} \\ &+ \frac{e^{-\lambda u}}{P(\tau_1 > T)} I_{\{u>T,\tau_1>T\}}. \end{aligned} \tag{3.2.5}$$

At this stage, we will use (2.3.2) to compute $P(\tau_1 > T)$ with the purpose of being able to get a final expression for $E[X_u|\mathcal{A}_T]$.

$$P(\tau_1 > T) = \begin{cases} (\lambda T + 1)e^{-\lambda T} & \text{if } \lambda = \mu \\ \\ \frac{\mu e^{-\lambda T} - \lambda e^{-\mu T}}{\mu - \lambda} & \text{otherwise.} \end{cases} \tag{3.2.6}$$

Given the necessary tools presented in (3.2.2), (3.2.5) and (3.2.6), we will establish expressions for $E[X_u|\mathcal{A}_T]$ and consequently for V_u. We will treat separately three different cases according to the relative position of T, τ_1 and $\hat{\sigma}$. Two subcases are to be discussed in each case according to $\lambda = \mu$ and $\lambda \neq \mu$.

<u>i) $\tau_1 \leq T$:</u>

In this case, we note that $\sigma \leq T$ since $\sigma \leq \tau_1$. Three subcases are to be discussed here according to λ and μ.

i1) $\lambda = \mu$:

Using (3.2.2) and (3.2.5), we get

$$E[X_u|\mathcal{A}_T] = 1 - \frac{u}{\tau_1}.$$

Then

$$V_u = -c_1 + (c_0 + c_1)(1 - \frac{u}{\tau_1}).$$

It is obvious that V_u is decreasing in u.

$$V_u \leq 0 \;\Leftrightarrow\; -c_1 + (c_0 + c_1)\left(1 - \frac{u}{\tau_1}\right) \leq 0$$

$$\begin{aligned}
&\Leftrightarrow\quad 1-\frac{u}{\tau_1}\leq\frac{c_1}{c_0+c_1}\\
&\Leftrightarrow\quad 1-\frac{c_1}{c_0+c_1}\leq\frac{u}{\tau_1}\\
&\Leftrightarrow\quad \frac{c_0}{c_0+c_1}\tau_1\leq u
\end{aligned}$$

Since $0<\dfrac{c_0}{c_0+c_1}\tau_1<\tau_1$, the optimal solution is found to be:

$$\hat{\sigma}=\frac{c_0}{c_0+c_1}\tau_1.$$

i2) $\lambda\neq\mu$:

Using again the formulas (3.2.2) and (3.2.5), we have:

$$\begin{aligned}
E[X_u|\mathcal{A}_T] &= 1-\frac{e^{(\mu-\lambda)u}-1}{e^{(\mu-\lambda)\tau_1}-1}\\
&= \frac{e^{(\mu-\lambda)\tau_1}-e^{(\mu-\lambda)u}}{e^{(\mu-\lambda)\tau_1}-1},
\end{aligned}$$

hence,

$$V_u=-c_1+(c_0+c_1)\left[\frac{e^{(\mu-\lambda)\tau_1}-e^{(\mu-\lambda)u}}{e^{(\mu-\lambda)\tau_1}-1}\right].$$

Again, V_u is clearly decreasing in u.

$$V_u\leq 0\Leftrightarrow\frac{e^{(\mu-\lambda)\tau_1}-e^{(\mu-\lambda)u}}{e^{(\mu-\lambda)\tau_1}-1}\leq\frac{c_1}{c_0+c_1}.$$

Here, we have to distinguish between the two situations $\lambda<\mu$ and $\lambda>\mu$.

- $\lambda<\mu$:

$$\begin{aligned}
V_u\leq 0 &\Leftrightarrow\quad e^{(\mu-\lambda)\tau_1}-e^{(\mu-\lambda)u}\leq\frac{c_1}{c_0+c_1}\left[e^{(\mu-\lambda)\tau_1}-1\right]\\
&\Leftrightarrow\quad e^{(\mu-\lambda)\tau_1}-\frac{c_1}{c_0+c_1}\left[e^{(\mu-\lambda)\tau_1}-1\right]\leq e^{(\mu-\lambda)u}\\
&\Leftrightarrow\quad \frac{c_0e^{(\mu-\lambda)\tau_1}+c_1}{c_0+c_1}\leq e^{(\mu-\lambda)u}\\
&\Leftrightarrow\quad \frac{1}{\mu-\lambda}\ln\left[\frac{c_0e^{(\mu-\lambda)\tau_1}+c_1}{c_0+c_1}\right]\leq u.
\end{aligned}$$

We remark that

$$\mu-\lambda>0\quad\Longrightarrow\quad c_1<c_1e^{(\mu-\lambda)\tau_1}$$

$$\begin{aligned}
&\Longrightarrow \quad c_0 e^{(\mu-\lambda)\tau_1} + c_1 < (c_0 + c_1) e^{(\mu-\lambda)\tau_1} \\
&\Longrightarrow \quad \frac{c_0 e^{(\mu-\lambda)\tau_1} + c_1}{c_0 + c_1} < e^{(\mu-\lambda)\tau_1} \\
&\Longrightarrow \quad \frac{1}{\mu-\lambda} \ln\left[\frac{c_0 e^{(\mu-\lambda)\tau_1} + c_1}{c_0 + c_1}\right] < \tau_1.
\end{aligned}$$

Furthermore, we can easily see that

$$\frac{c_0 e^{(\mu-\lambda)\tau_1} + c_1}{c_0 + c_1} > 1.$$

Therefore

$$0 < \hat{\sigma} < \tau_1,$$

where

$$\hat{\sigma} = \frac{1}{\mu-\lambda} \ln\left[\frac{c_0 e^{(\mu-\lambda)\tau_1} + c_1}{c_0 + c_1}\right].$$

- $\lambda > \mu$:

$$\begin{aligned}
V_u \leq 0 \quad &\Leftrightarrow \quad e^{(\mu-\lambda)\tau_1} - e^{(\mu-\lambda)u} \geq \frac{c_1}{c_0 + c_1}\left[e^{(\mu-\lambda)\tau_1} - 1\right] \\
&\Leftrightarrow \quad e^{(\mu-\lambda)\tau_1} - \frac{c_1}{c_0 + c_1}\left[e^{(\mu-\lambda)\tau_1} - 1\right] \geq e^{(\mu-\lambda)u} \\
&\Leftrightarrow \quad \frac{c_0 e^{(\mu-\lambda)\tau_1} + c_1}{c_0 + c_1} \geq e^{(\mu-\lambda)u} \\
&\Leftrightarrow \quad \ln\left[\frac{c_0 e^{(\mu-\lambda)\tau_1} + c_1}{c_0 + c_1}\right] \geq (\mu-\lambda)u \\
&\Leftrightarrow \quad \frac{1}{\mu-\lambda} \ln\left[\frac{c_0 e^{(\mu-\lambda)\tau_1} + c_1}{c_0 + c_1}\right] \leq u.
\end{aligned}$$

We use the same argument seen in the previous case when $\lambda < \mu$ to prove that (in this case $c_0 e^{(\mu-\lambda)\tau_1} + c_1 < c_0 + c_1$):

$$\frac{1}{\mu-\lambda} \ln\left[\frac{c_0 e^{(\mu-\lambda)\tau_1} + c_1}{c_0 + c_1}\right] < \tau_1.$$

Therefore

$$0 < \hat{\sigma} < \tau_1,$$

where

$$\hat{\sigma} = \frac{1}{\mu-\lambda} \ln\left[\frac{c_0 e^{(\mu-\lambda)\tau_1} + c_1}{c_0 + c_1}\right].$$

ii) $u < T < \tau_1$:

In order to achieve an expression for the optimal solution in this case, we will make use of (3.2.5) and (3.2.6) for the usual situations $\lambda = \mu$ and $\lambda \neq \mu$.

ii1) $\lambda = \mu$

Substituting (3.2.6) in (3.2.5) for the case when $\lambda = \mu$, we get:

$$E[X_u|\mathcal{A}_T] = \frac{1+\lambda(T-u)}{1+\lambda T},$$

hence

$$V_u = -c_1 + (c_0 + c_1)\left[\frac{1+\lambda(T-u)}{1+\lambda T}\right].$$

Again, V_u is clearly decreasing in u.

$$\begin{aligned}
V_u \leq 0 \quad &\Leftrightarrow \quad -c_1 + (c_0 + c_1)\left[\frac{1+\lambda(T-u)}{1+\lambda T}\right] \leq 0 \\
&\Leftrightarrow \quad \frac{1+\lambda(T-u)}{1+\lambda T} \leq \frac{c_1}{c_0+c_1} \\
&\Leftrightarrow \quad 1+\lambda(T-u) \leq \frac{c_1}{c_0+c_1}(1+\lambda T) \\
&\Leftrightarrow \quad 1+\lambda T - \frac{c_1}{c_0+c_1}(1+\lambda T) \leq \lambda u \\
&\Leftrightarrow \quad (1+\lambda T)(1-\frac{c_1}{c_0+c_1}) \leq \lambda u \\
&\Leftrightarrow \quad \frac{c_0(1+\lambda T)}{\lambda(c_0+c_1)} \leq u.
\end{aligned}$$

This will occur in $[0, T]$ if and only if:

$$\begin{aligned}
\frac{c_0(1+\lambda T)}{\lambda(c_0+c_1)} \leq T \quad &\Leftrightarrow \quad \frac{c_0}{\lambda(c_0+c_1)} \leq T(1-\frac{c_0}{c_0+c_1}) \\
&\Leftrightarrow \quad \frac{c_0}{c_1} \leq \lambda T.
\end{aligned}$$

Therefore, we can conclude that

$$\hat{\sigma} = \frac{c_0}{c_0+c_1}\left(\frac{1}{\lambda}+T\right) \quad \text{if} \quad \frac{c_0}{c_1} \leq \lambda T.$$

ii2) $\lambda \neq \mu$:

From (3.2.5) and (3.2.6) provided $\lambda \neq \mu$, we have:

$$E[X_u|\mathcal{A}_T] \quad = \quad \frac{e^{-\lambda T} + \lambda e^{-\mu T} \times \frac{e^{(\mu-\lambda)T} - e^{(\mu-\lambda)u}}{\mu-\lambda}}{\frac{\mu e^{-\lambda T} - \lambda e^{-\mu T}}{\mu - \lambda}}$$

$$
\begin{aligned}
&= \frac{\mu e^{-\lambda T} - \lambda e^{-\lambda T} + \lambda e^{-\lambda T} - \lambda e^{(\mu-\lambda)u-\mu T}}{\mu e^{-\lambda T} - \lambda e^{-\mu T}} \\
&= \frac{\mu e^{(\mu-\lambda)T} - \lambda e^{(\mu-\lambda)u}}{\mu e^{(\mu-\lambda)T} - \lambda}.
\end{aligned}
$$

Thus,

$$
V_u = -c_1 + (c_0 + c_1)\left[\frac{\mu e^{(\mu-\lambda)T} - \lambda e^{(\mu-\lambda)u}}{\mu e^{(\mu-\lambda)T} - \lambda}\right].
$$

We can easily prove that V_u is decreasing in u for every $\lambda > 0, \mu > 0$ such that $\lambda \neq \mu$.

$$
V_u \leq 0 \Leftrightarrow \frac{\mu e^{(\mu-\lambda)T} - \lambda e^{(\mu-\lambda)u}}{\mu e^{(\mu-\lambda)T} - \lambda} \leq \frac{c_1}{c_0 + c_1}.
$$

As we have done before, we need to separate our analysis into two subcases.

- $\lambda < \mu$:

$$
\begin{aligned}
V_u \leq 0 \quad &\Leftrightarrow \quad \mu e^{(\mu-\lambda)T} - \lambda e^{(\mu-\lambda)u} \leq \frac{c_1}{c_0 + c_1}\left[\mu e^{(\mu-\lambda)T} - \lambda\right] \\
&\Leftrightarrow \quad \mu c_0 e^{(\mu-\lambda)T} + \mu c_1 e^{(\mu-\lambda)T} - \lambda(c_0 + c_1)e^{(\mu-\lambda)u} \leq \mu c_1 e^{(\mu-\lambda)T} - \lambda c_1 \\
&\Leftrightarrow \quad \mu c_0 e^{(\mu-\lambda)T} + \lambda c_1 \leq \lambda(c_0 + c_1)e^{(\mu-\lambda)u} \\
&\Leftrightarrow \quad \frac{1}{\mu - \lambda}\ln\left[\frac{\mu c_0 e^{(\mu-\lambda)T} + \lambda c_1}{\lambda(c_0 + c_1)}\right] \leq u.
\end{aligned}
$$

On the other hand, we observe that

$$
\begin{aligned}
\frac{1}{\mu - \lambda}\ln\left[\frac{\mu c_0 e^{(\mu-\lambda)T} + \lambda c_1}{\lambda(c_0 + c_1)}\right] \leq T \quad &\Leftrightarrow \quad \frac{\mu c_0 e^{(\mu-\lambda)T} + \lambda c_1}{\lambda(c_0 + c_1)} \leq e^{(\mu-\lambda)T} \\
&\Leftrightarrow \quad \mu c_0 e^{(\mu-\lambda)T} + \lambda c_1 \leq \lambda c_0 e^{(\mu-\lambda)T} + \lambda c_1 e^{(\mu-\lambda)T} \\
&\Leftrightarrow \quad c_0(\mu - \lambda)e^{(\mu-\lambda)T} \leq \lambda c_1\left[e^{(\mu-\lambda)T} - 1\right] \\
&\Leftrightarrow \quad \frac{c_0}{c_1} \leq \frac{\lambda}{\mu - \lambda}\left[\frac{e^{(\mu-\lambda)T} - 1}{e^{(\mu-\lambda)T}}\right] \\
&\Leftrightarrow \quad \frac{c_0}{c_1} \leq \frac{\lambda}{\mu - \lambda}\left[1 - e^{(\lambda-\mu)T}\right].
\end{aligned}
$$

Of course, it is easy to see that

$$
\frac{1}{\mu - \lambda}\ln\left[\frac{\mu c_0 e^{(\mu-\lambda)T} + \lambda c_1}{\lambda(c_0 + c_1)}\right] > 0.
$$

Therefore,

$$\hat{\sigma} = \frac{1}{\mu - \lambda} \ln \left[\frac{\mu c_0 e^{(\mu-\lambda)T} + \lambda c_1}{\lambda(c_0 + c_1)} \right] \quad \text{if} \quad \frac{c_0}{c_1} \leq \frac{\lambda}{\mu - \lambda} \left[1 - e^{(\lambda-\mu)T} \right].$$

- $\lambda > \mu$:

$$\begin{aligned} V_u \leq 0 \quad &\Leftrightarrow \quad \mu e^{(\mu-\lambda)T} - \lambda e^{(\mu-\lambda)u} \geq \frac{c_1}{c_0 + c_1} \left[\mu e^{(\mu-\lambda)T} - \lambda \right] \\ &\Leftrightarrow \quad \mu c_0 e^{(\mu-\lambda)T} + \mu c_1 e^{(\mu-\lambda)T} - \lambda(c_0 + c_1) e^{(\mu-\lambda)u} \geq \mu c_1 e^{(\mu-\lambda)T} - \lambda c_1 \\ &\Leftrightarrow \quad \mu c_0 e^{(\mu-\lambda)T} + \lambda c_1 \geq \lambda(c_0 + c_1) e^{(\mu-\lambda)u} \\ &\Leftrightarrow \quad \frac{1}{\mu - \lambda} \ln \left[\frac{\mu c_0 e^{(\mu-\lambda)T} + \lambda c_1}{\lambda(c_0 + c_1)} \right] \leq u. \end{aligned}$$

Using the same argument as before and using now the fact that $\lambda > \mu$, we can show that:

$$\hat{\sigma} = \frac{1}{\mu - \lambda} \ln \left[\frac{\mu c_0 e^{(\mu-\lambda)T} + \lambda c_1}{\lambda(c_0 + c_1)} \right] \quad \text{if} \quad \frac{c_0}{c_1} \leq \frac{\lambda}{\mu - \lambda} \left[1 - e^{(\lambda-\mu)T} \right].$$

iii) $T < u < \tau_1$:

Following again the same methodology utilized previously and using (3.2.5) and (3.2.6), we will discuss separately the cases $\lambda = \mu$ and $\lambda \neq \mu$ as a means to get the optimal solution for the detection problem.

iii1) $\lambda = \mu$:

We have, using (3.2.5) and (3.2.6)

$$E[X_u | \mathcal{A}_T] = \frac{e^{\lambda(T-u)}}{1 + \lambda T},$$

thus,

$$V_u = -c_1 + (c_0 + c_1) \frac{e^{\lambda(T-u)}}{1 + \lambda T}.$$

Again, it is clear that V_u is decreasing in u.

$$\begin{aligned} V_u \leq 0 \quad &\Leftrightarrow \quad \frac{e^{\lambda(T-u)}}{1 + \lambda T} \leq \frac{c_1}{c_0 + c_1} \\ &\Leftrightarrow \quad e^{\lambda(T-u)} \leq \frac{c_1(1 + \lambda T)}{c_0 + c_1} \end{aligned}$$

$$\Leftrightarrow \quad \lambda(T-u) \leq \ln\left[\frac{c_1(1+\lambda T)}{c_0+c_1}\right]$$
$$\Leftrightarrow \quad T - \frac{1}{\lambda}\ln\left[\frac{c_1(1+\lambda T)}{c_0+c_1}\right] \leq u.$$

This will occur after T if and only if:

$$\begin{aligned} T - \frac{1}{\lambda}\ln\left[\frac{c_1(1+\lambda T)}{c_0+c_1}\right] > T \quad &\Leftrightarrow \quad \ln\left[\frac{c_1(1+\lambda T)}{c_0+c_1}\right] < 0 \\ &\Leftrightarrow \quad \frac{c_1(1+\lambda T)}{c_0+c_1} < 1 \\ &\Leftrightarrow \quad \frac{c_0}{c_1} > \lambda T \end{aligned}$$

We conclude that:

$$\hat{\sigma} = T - \frac{1}{\lambda}\ln\left[\frac{c_1(1+\lambda T)}{c_0+c_1}\right] \quad \text{if} \quad \frac{c_0}{c_1} > \lambda T.$$

iii2) $\lambda \neq \mu$:

Based on the formulas (3.2.5) and (3.2.6) for the case when $\lambda \neq \mu$, we get the following results for $E[X_u|\mathcal{A}_T]$ and V_u.

$$E[X_u|\mathcal{A}_T] = \frac{(\mu-\lambda)e^{-\lambda u}}{\mu e^{-\lambda T} - \lambda e^{-\mu T}},$$

and

$$V_u = -c_1 + (c_0+c_1)\frac{(\mu-\lambda)e^{-\lambda u}}{\mu e^{-\lambda T} - \lambda e^{-\mu T}}.$$

Remark that:

$$\frac{\mu-\lambda}{\mu e^{-\lambda T} - \lambda e^{-\mu T}} > 0 \quad \forall \lambda > 0, \forall \mu > 0 \ \text{ such that } \ \lambda \neq \mu.$$

Hence, V_u is decreasing in u for every $\lambda \neq \mu$.

$$V_u \leq 0 \quad \Leftrightarrow \quad \frac{(\mu-\lambda)e^{-\lambda u}}{\mu e^{-\lambda T} - \lambda e^{-\mu T}} \leq \frac{c_1}{c_0+c_1}$$

$$\Leftrightarrow \quad e^{-\lambda u} \leq \frac{c_1(\mu e^{-\lambda T} - \lambda e^{-\mu T})}{(c_0+c_1)(\mu-\lambda)}$$

$$\Leftrightarrow \quad -\lambda u \leq \ln\left[\frac{c_1(\mu e^{-\lambda T} - \lambda e^{-\mu T})}{(c_0 + c_1)(\mu - \lambda)}\right]$$

$$\Leftrightarrow \quad -\frac{1}{\lambda}\ln\left[\frac{c_1(\mu e^{-\lambda T} - \lambda e^{-\mu T})}{(c_0 + c_1)(\mu - \lambda)}\right] \leq u.$$

This occurs after T if and only if:

$$-\frac{1}{\lambda}\ln\left[\frac{c_1(\mu e^{-\lambda T} - \lambda e^{-\mu T})}{(c_0 + c_1)(\mu - \lambda)}\right] > T \quad \Leftrightarrow \quad \frac{c_1(\mu e^{-\lambda T} - \lambda e^{-\mu T})}{(c_0 + c_1)(\mu - \lambda)} < e^{-\lambda T}$$

$$\Leftrightarrow \quad \frac{\mu e^{-\lambda T} - \lambda e^{-\mu T}}{(\mu - \lambda)e^{-\lambda T}} < \frac{c_0}{c_1} + 1$$

$$\Leftrightarrow \quad \frac{\mu e^{-\lambda T} - \lambda e^{-\mu T} - \mu e^{-\lambda T} + \lambda e^{-\lambda T}}{(\mu - \lambda)e^{-\lambda T}} < \frac{c_0}{c_1}$$

$$\Leftrightarrow \quad \frac{\lambda(e^{-\lambda T} - e^{-\mu T})}{(\mu - \lambda)e^{-\lambda T}} < \frac{c_0}{c_1}$$

$$\Leftrightarrow \quad \frac{\lambda\left[1 - e^{(\lambda-\mu)T}\right]}{\mu - \lambda} < \frac{c_0}{c_1}.$$

Therefore, we have the following expression for the optimal solution in this case:

$$\hat{\sigma} = -\frac{1}{\lambda}\ln\left[\frac{c_1(\mu e^{-\lambda T} - \lambda e^{-\mu T})}{(c_0 + c_1)(\mu - \lambda)}\right] \quad \text{if} \quad \frac{c_0}{c_1} > \frac{\lambda\left[1 - e^{(\lambda-\mu)T}\right]}{\mu - \lambda}.$$

To conclude, the previously achieved results of the optimal solution for the ex-post detection problem can be summarized as follows:

$\star$ $\underline{\lambda = \mu:}$

$$\hat{\sigma} = \begin{cases} \frac{c_0}{c_0+c_1}\tau_1 & \text{if } \tau_1 \leq T \\ \frac{c_0}{c_0+c_1}\left(\frac{1}{\lambda} + T\right) & \text{if } \tau_1 > T,\ \frac{c_0}{c_1} \leq \lambda T \\ T - \frac{1}{\lambda}\ln\left[\frac{c_1(1+\lambda T)}{c_0+c_1}\right] & \text{if } \tau_1 > T,\ \frac{c_0}{c_1} > \lambda T \end{cases} \tag{3.2.7}$$

$\star$ $\underline{\lambda \neq \mu:}$

$$\hat{\sigma} = \begin{cases} \frac{1}{\mu-\lambda} \ln \left[\frac{c_0 e^{(\mu-\lambda)\tau_1} + c_1}{c_0 + c_1} \right] & \text{if } \tau_1 \leq T \\ \\ \frac{1}{\mu-\lambda} \ln \left[\frac{\mu c_0 e^{(\mu-\lambda)T} + \lambda c_1}{\lambda(c_0 + c_1)} \right] & \text{if } \tau_1 > T, \ \frac{c_0}{c_1} \leq \frac{\lambda}{\mu-\lambda} \left[1 - e^{(\lambda-\mu)T} \right] \\ \\ -\frac{1}{\lambda} \ln \left[\frac{c_1(\mu e^{-\lambda T} - \lambda e^{-\mu T})}{(c_0 + c_1)(\mu-\lambda)} \right] & \text{if } \tau_1 > T, \ \frac{c_0}{c_1} > \frac{\lambda}{\mu-\lambda} \left[1 - e^{(\lambda-\mu)T} \right]. \end{cases} \tag{3.2.8}$$

$\blacksquare$

3.3 Analysis of the optimal solution

Based on the expressions above of $\hat{\sigma}$, we are going to calculate $E[\hat{\sigma}]$ for the different cases $\lambda = \mu$ and $\lambda \neq \mu$ and try to find an "unbiased estimator"for σ $(E[\hat{\sigma}] = \frac{1}{\lambda})$. To do that, we need to rewrite the expression of $\hat{\sigma}$ as follows:

$$\hat{\sigma} = g(\tau_1) \ I_{(\tau_1 \leq T)} + M \ I_{(\tau_1 > T)},$$

where the function $g(\tau_1)$ and the constant M are deduced from (3.2.6) and (3.2.7). Therefore,

$$E[\hat{\sigma}] = \int_0^T g(t) dF_{\tau_1}(t) + MP(\tau_1 > T).$$

$\underline{\text{i) } \mu = \lambda:}$

From (2.3.2), we know that:

$$F_{\tau_1}(t) = 1 - (\lambda t + 1)e^{-\lambda t} \quad \forall t > 0,$$

hence,

$$f_{\tau_1}(t) = \lambda^2 t e^{-\lambda t} \quad \forall t > 0.$$

Using now (3.2.7), we get:

$$g(\tau_1) = \frac{c_0}{c_0 + c_1} \tau_1,$$

therefore,

$$\begin{aligned} E[\hat{\sigma}] &= \int_0^T \frac{c_0}{c_0+c_1}\lambda^2 t^2 e^{-\lambda t} dt + M(\lambda T+1)e^{-\lambda T} \\ &= \frac{c_0}{c_0+c_1}\left[-\lambda T^2 e^{-\lambda T} - 2Te^{-\lambda T} - \frac{2}{\lambda}e^{-\lambda T} + \frac{2}{\lambda}\right] + M(\lambda T+1)e^{-\lambda T} \quad (3.3.1) \end{aligned}$$

Since the constant M depends on $\frac{c_0}{c_1}$, λ and T, two possible situations are to be treated.

- $\frac{c_0}{c_1} \leq \lambda T$:

In this case, using (14),

$$M = \frac{c_0}{c_0+c_1}\left[\frac{1+\lambda T}{\lambda}\right],$$

so,

$$\begin{aligned} E[\hat{\sigma}] &= \frac{c_0}{c_0+c_1}\left[-\frac{e^{-\lambda T}}{\lambda} + \frac{2}{\lambda}\right] \\ &= \frac{1}{\lambda} + \frac{c_0}{c_0+c_1}\left[-\frac{e^{-\lambda T}}{\lambda} + \frac{2}{\lambda} - \frac{c_0+c_1}{\lambda c_0}\right] \\ &= \frac{1}{\lambda} + \frac{c_0}{\lambda(c_0+c_1)}\left[-e^{-\lambda T} + 1 - \frac{c_1}{c_0}\right]. \end{aligned}$$

In order to get an "unbiased estimator" for σ, we should have:

$$\begin{aligned} \frac{c_0}{c_1} \leq \lambda T \text{ and } -e^{-\lambda T} + 1 - \frac{c_1}{c_0} = 0 &\Leftrightarrow \frac{c_0}{c_1} \leq \lambda T \text{ and } \frac{c_1}{c_0} = 1 - e^{-\lambda T} \\ &\Leftrightarrow \frac{c_0}{c_1} = \frac{1}{1-e^{-\lambda T}} \leq \lambda T. \end{aligned}$$

Define for $x > 0$

$$f(x) = x - xe^{-x} - 1;$$

f is continuous and strictly increasing on $(0, \infty)$, hence f is a bijection from $(0, \infty)$ to $(-1, \infty)$. Therefore $\exists!\ a > 0$ such that $f(a) = 0$, in fact $a \simeq 1.349976485$.

Consequently, we get:

$$\forall\, \lambda \geq \frac{a}{T}, \qquad \frac{1}{1-e^{-\lambda T}} \leq \lambda T.$$

We conclude that if $\lambda \geq \dfrac{a}{T}$ and $\dfrac{c_0}{c_1} = \dfrac{1}{1-e^{-\lambda T}}$ then $E[\hat{\sigma}] = \dfrac{1}{\lambda}$.

- $\dfrac{c_0}{c_1} > \lambda T$:

From (3.2.7), we have

$$M = T - \frac{1}{\lambda} \ln \left[\frac{c_1(1+\lambda T)}{c_0 + c_1} \right],$$

hence, from (3.3.1) we obtain:

$$\begin{aligned} E[\hat{\sigma}] &= \frac{c_0}{c_0+c_1}\left[(-\lambda T^2 - 2T - \frac{2}{\lambda})e^{-\lambda T} + \frac{2}{\lambda}\right] + \frac{1}{\lambda}\left[\lambda T - \ln\left(\frac{c_1(1+\lambda T)}{c_0+c_1}\right)\right] \\ &\quad \times (\lambda T + 1)e^{-\lambda T} \\ &= \frac{1}{\lambda} + \frac{1}{\lambda}\left[(\lambda T+1)\left(\lambda T - \ln\left[\frac{c_1(\lambda T+1)}{c_0+c_1}\right]\right) - \frac{c_0(\lambda^2T^2+2\lambda T+2)}{c_0+c_1}\right]e^{-\lambda T} \\ &\quad + \frac{c_0 - c_1}{\lambda(c_0+c_1)}. \end{aligned}$$

To get an "unbiased estimator" for σ, c_0 and c_1 must verify: $\dfrac{c_0}{c_1} > \lambda T$ and

$$\left[(\lambda T+1)\left(\lambda T - \ln\left[\frac{c_1(\lambda T+1)}{c_0+c_1}\right]\right) - \frac{c_0(\lambda^2T^2+2\lambda T+2)}{c_0+c_1}\right]e^{-\lambda T} = \frac{c_1-c_0}{c_0+c_1}. \quad (*)$$

If we put $\alpha = \frac{c_0}{c_1}$ and $\beta = \lambda T$, then $\alpha > \beta > 0$ and $(*)$ is equivalent to solve for α:

$$(1-\alpha)(\beta - e^{\beta}) + (1+\alpha)(1+\beta)\ln\left(\frac{1+\alpha}{1+\beta}\right) + \beta^2 - 2\alpha = 0. \quad (**)$$

Define for $x > \beta$

$$g(x) = (1-x)(\beta - e^{\beta}) + (1+x)(1+\beta)\ln\left(\frac{1+x}{1+\beta}\right) + \beta^2 - 2x;$$

g is continuous and strictly increasing from (β, ∞) to $(\beta e^{\beta} - e^{\beta} - \beta, \infty)$.

If we define

$$h(y) = ye^y - e^y - y \quad \forall y > 0,$$

we can show that h has a unique minimum, say $a > 0$. h is strictly decreasing from $(0, a)$ to $(-1, h(a))$ and strictly increasing from (a, ∞) to $(h(a), \infty)$. Using Newton's Method, we found the approximation $a \simeq 0.5671432905$.

Note that:

$$\begin{aligned} h'(a) = 0 \ &\Rightarrow \ ae^a = 1 \\ &\Rightarrow \ h(a) = 1 - e^a - a < 0, \end{aligned}$$

hence, $\exists!\ b \in (a, \infty)$ such that $h(b) = 0$. Therefore, if $\beta < b$ then $g(\beta) < 0$. Newton's method gives us the approximation $b \simeq 1.349976485$.

We conclude that if $\lambda T < b$ then $\exists! \alpha^* > \lambda T$ satisfying $(**)$ and we proved the existence of an "unbiased estimator" for σ when $\lambda T < b$ and $\frac{c_0}{c_1} = \alpha^*$. No approximation for $\alpha*$ can be given at this stage since we need to know the chosen value of β.

ii) $\mu \neq \lambda$:

Using (2.3.2), we have:

$$F_{\tau_1}(t) = 1 - \frac{\mu e^{-\lambda t} - \lambda e^{-\mu t}}{\mu - \lambda} \quad \forall t > 0,$$

hence,

$$f_{\tau_1}(t) = \mu\lambda \frac{e^{-\lambda t} - e^{-\mu t}}{\mu - \lambda} \quad \forall t > 0.$$

From (3.2.8), we can see that:

$$g(\tau_1) = \frac{1}{\mu - \lambda} \ln \left[\frac{c_0 e^{(\mu-\lambda)\tau_1} + c_1}{c_0 + c_1} \right].$$

Letting now $c = \frac{c_0}{c_1}$, we have:

$$\begin{aligned} E[\hat{\sigma}] &= M(\frac{\mu e^{-\lambda T} - \lambda e^{-\mu T}}{\mu - \lambda}) + \frac{\mu\lambda}{(\mu-\lambda)^2} \int_0^T (e^{-\lambda t} - e^{-\mu t}) \ln \left[\frac{c_0 e^{(\mu-\lambda)t} + c_1}{c_0 + c_1} \right] dt \\ &= M(\frac{\mu e^{-\lambda T} - \lambda e^{-\mu T}}{\mu - \lambda}) + \frac{\mu\lambda}{(\mu-\lambda)^2} \int_0^T (e^{-\lambda t} - e^{-\mu t}) \ln \left[\frac{c e^{(\mu-\lambda)t} + 1}{c + 1} \right] dt. \end{aligned}$$

Since the above integrand cannot be written in closed form [1] for general values of μ, λ, T and $\frac{c_0}{c_1}$; we will try to find the "unbiased estimator" in the following way: we

[1]Under some conditions, the integrand can be written in the form of the hypergeometric function ${}_2F_1$.

first choose some values for μ, λ and T; then we substitute them in the expression of $E[\hat{\sigma}]$. Finally, we will solve numerically the equation $E[\hat{\sigma}] = \frac{1}{\lambda}$ for $\frac{c_0}{c_1}$.

- $\frac{c_0}{c_1} \leq \frac{\lambda}{\mu-\lambda}\left[1-e^{(\lambda-\mu)T}\right]$:

From (3.2.8), we have:

$$\begin{aligned} M &= \frac{1}{\mu-\lambda}\ln\left[\frac{\mu c_0 e^{(\mu-\lambda)T}+\lambda c_1}{\lambda(c_0+c_1)}\right] \\ &= \frac{1}{\mu-\lambda}\ln\left[\frac{\mu c e^{(\mu-\lambda)T}+\lambda}{\lambda(c+1)}\right]. \end{aligned}$$

In this case, a solution for the equation $E[\hat{\sigma}] = \frac{1}{\lambda}$ appears to exist for only large values of T.

λ	μ	T	$\frac{\lambda}{\mu-\lambda}\left[1-e^{(\lambda-\mu)T}\right]$	$\frac{c_0}{c_1}$
1	2	10	0.99995464	0.7563149964
1	2	20	0.9999999	0.7562435742
1	2	30	0.9999999	0.7562435709
2	1	10	44050.93	1.322325291
2	1	20	970330389	1.322325291
2	1	30	$2.14\ 10^{13}$	1.322325291
3	4	10	2.999864	0.8725186358
3	4	20	2.9999999	0.8725186358
3	4	30	2.9999999	0.8725186358
4	3	10	88101.86	1.146107325
4	3	20	1940660778	1.146107325
4	3	30	$4.27\ 10^{13}$	1.146107325

- $\frac{c_0}{c_1} > \frac{\lambda}{\mu-\lambda}\left[1-e^{(\lambda-\mu)T}\right]$:

From (3.2.8) again, we have:

$$M = -\frac{1}{\lambda}\ln\left[\frac{c_1(\mu e^{-\lambda T}-\lambda e^{-\mu T})}{(c_0+c_1)(\mu-\lambda)}\right]$$

$$= \quad -\frac{1}{\lambda}\ln\left[\frac{\mu e^{-\lambda T}-\lambda e^{-\mu T}}{(c+1)(\mu-\lambda)}\right].$$

In contrast to the preceding case, a solution for the equation $E[\hat{\sigma}] = 1/\lambda$ appears to exist for only small values of T.

λ	μ	T	$\frac{\lambda}{\mu-\lambda}\left[1-e^{(\lambda-\mu)T}\right]$	$\frac{c_0}{c_1}$
1	2	0.1	0.09516258	1.717347281
1	2	0.2	0.1812692	1.712043312
1	2	0.3	0.2591818	1.700413662
2	1	0.1	0.2103418	1.716357322
2	1	0.2	0.4428055	1.705135757
2	1	0.3	0.6997176	1.680041573
3	4	0.1	0.2854877	1.70471515
3	4	0.2	0.5438077	1.641380992
3	4	0.3	0.7775453	1.528077952
4	3	0.1	0.4206837	1.699775264
4	3	0.2	0.885611	1.612591394
4	3	0.3	1.399435	1.461495523

■

Chapter 4

Sequential observation in two dimensions

In this chapter, we consider the two-dimensional version of the optimal detection problem described in the first chapter. The goal now is to detect the occurrence of an unobservable random set ξ that is the support of an observable Poisson process N. In the first section, we will define the model and present some tools leading to the optimal solution of the detection problem. In section 3.2, we are going to evaluate theoretically the solution found in the previous section. The last section is dedicated to developing a practical approach to determine the optimal solution.

4.1 The model

Before describing the two-dimensional model, it is required to introduce some definitions and properties that play an important role in our assumptions.

Definition 4.1.1 *A point process L is called a single line point process if each of its jump points are all incomparable:*

$$s, t \in \mathbb{R}_+^2 \textit{ are incomparable if both } s \nleq t \textit{ and } t \nleq s.$$

In the following, " $\leq$ " denotes the usual order on $\mathbb{R}_+^2$:

$$s = (s_1, s_2) \leq t = (t_1, t_2) \Leftrightarrow s_1 \leq t_1, s_2 \leq t_2.$$

Definition 4.1.2 *A set ξ is an upper layer if $t \in \xi \Rightarrow s \in \xi \quad \forall s \geq t$. A set B is a lower layer if its complement is an upper layer.*

Definition 4.1.3 *The single line point process L is called the first line of a point process J if:*

$$\Delta_L = \min(\Delta_J) = \{\tau \in \Delta_J : \tau' \nleq \tau \quad \forall \tau' \in \Delta_J \text{ such that } \tau' \neq \tau\}$$

where Δ_J, Δ_L denote respectively the set of jump points of J and L.

We are given a non-explosive point process $N = \{N_t, t \in \mathbb{R}_+^2\}$ on the positive quadrant of the plane and some random Borel set $\xi \subset \mathbb{R}_+^2$. Given the set ξ, N is a Poisson process with intensity μ on ξ. Our objective now is to detect the occurrence of the random set ξ.

Since the change-set ξ is unobservable, we would like to find, if it exists, a random set $\hat{\xi}$ adapted to the underlying information structure $\mathbb{F}$ that maximizes the expected value of a specified valuation or gain function.

Using notation similar to that used for the one-dimensional detection problem, for $t = (t_1, t_2) \in \mathbb{R}_+^2$ let $A_t = [0, t_1] \times [0, t_2]$ and $X_t = 1 - I_{\{t \in \xi\}}$. Then the valuation (gain) function at $t \in \mathbb{R}_+^2$ is defined exactly as before:

$$\begin{aligned} Z_t &= c_0|A_t \cap \xi^c| - c_1|A_t \cap \xi| \quad \text{where } c_0 \geq 0, c_1 > 0 \\ &= \int_{A_t} \left(-c_1 + (c_0 + c_1) I_{(u \in \xi^c)}\right) du \\ &= \int_{A_t} (-c_1 + (c_0 + c_1) X_u) \, du \\ &= \int_{A_t} U_u \, du \quad \text{where } U_u = -c_1 + (c_0 + c_1) X_u. \end{aligned}$$

The gain function is increasing at rate c_0 outside the change-set ξ and decreasing at rate c_1 inside ξ.

We will assume that the set ξ is generated by the single line process L. Of particular interest is the case that L is the first line of Poisson process J because it is analogous to the case considered in the first chapter where the exponential change point σ can be interpreted as the "first line" of a Poisson process on $\mathbb{R}_+$. Therefore the set ξ can be defined as: $\xi = \{t : L_t > 0\} = \{t : J_t > 0\}$. We observe that ξ is an upper layer, it consists of all the points to the northeast of one or more jump points of L. Neither L nor J are observed, we only see the jump points of N. This situation is illustrated in Figure 4.1.

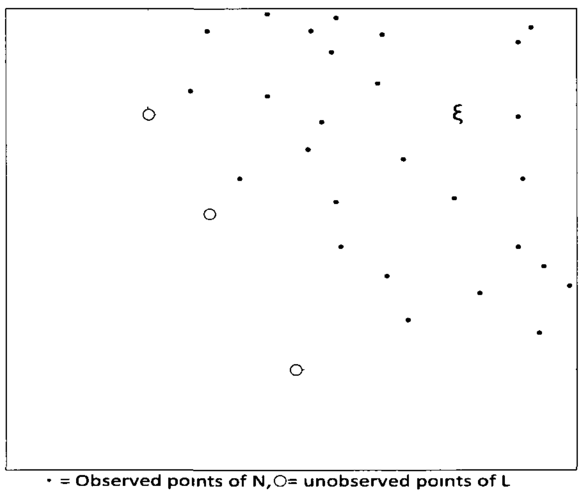

Figure 4.1: *A change-set ξ generated by a single line process L*

We can also define our gain function more generally over the class of *lower layers* as:

$$\begin{aligned} Z_B &= c_0|B \cap \xi^c| - c_1|B \cap \xi| \\ &= \int_B (-c_1 + (c_0 + c_1)X_u)\, du \\ &= \int_B U_u\, du. \end{aligned}$$

A lower layer B and the change set ξ are illustrated in Figure 4.2; we observe that $L(B) = 1$ in this case.

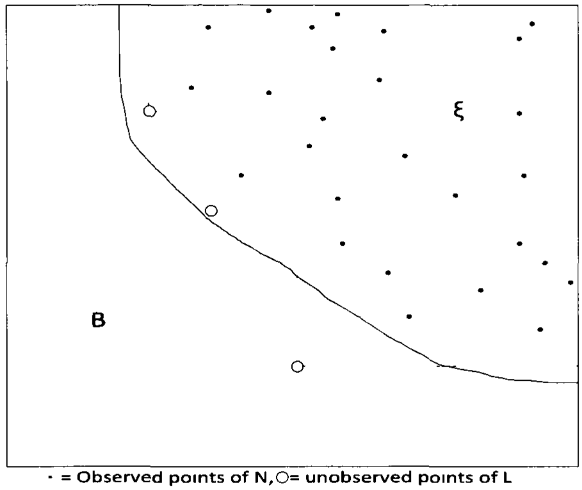

Figure 4.2: *A lower layer B and the change-set ξ*

The gain evaluated at B increases in proportion to the area of B outside the change-set ξ, and decreases in proportion to the area inside of ξ. The gain is maximized when $B = \overline{\xi^c}$.

We are aiming at finding a random lower layer that maximizes the expected value of the gain function and depends on the available information, or more precisely, the underlying *filtration*.

Definition 4.1.4 *A class of σ-fields $\mathbb{F} = \{\mathcal{F}_t, t \in \mathbb{R}_+^2\}$ is a filtration if*

- *$\mathbb{F}$ is increasing: $s \leq t \Rightarrow \mathcal{F}_s \subseteq \mathcal{F}_t$*
- *$\mathbb{F}$ is outer continuous: $\mathcal{F}_t = \cap_n \mathcal{F}_{t_n}$ for every decreasing sequence $(t_n \subset \mathbb{R}_+^2)$ with $t_n \downarrow t$*

Definition 4.1.5 *A closed random lower layer ρ is an $\mathbb{F}$-stopping set if*

$$\{t \in \rho\} \in \mathcal{F}_t, \quad \forall t \in \mathbb{R}_+^2.$$

The optimal set-detection problem can now be stated as follows: for a given filtration $\mathbb{F}$, our goal is to maximize $E[Z_\rho]$ where ρ is an $\mathbb{F}$-stopping set.

In this chapter, we will be focusing on the *sequential* estimation problem: i.e. we assume that the information is described by the filtration $\mathbb{F}^N = (\mathcal{F}_t^N, t \in \mathbb{R}_+^2)$ where $\mathcal{F}_t^N = \sigma\{N_s : s \leq t\}$. Again, the detection problem can be viewed as an optimal stopping problem in the following sense:

Definition 4.1.6 *An $\mathbb{F}^N$-stopping set $\hat{\rho}$ is called an optimal solution to the sequential detection problem on R if :*

$$E[Z_{\hat{\rho}}] = \sup\{E[Z_\rho], \rho \subseteq R, \rho \textit{ is an } \mathbb{F}^N\textit{-stopping set}\}.$$

Here, we restrict the detection problem to a bounded rectangle $R = [0, r]^2$ to ensure that $\hat{\rho}$ is bounded and so $E[Z_{\hat{\rho}}]$ is well-defined. In this case, we have an optimal estimate $\hat{\xi_R}$ of $\xi \cap R$ defined by $\hat{\xi_R} = \overline{R \backslash \hat{\rho}}$.

In what follows, T denotes either $\mathbb{R}_+^2$ or R and $(\Omega, \mathbb{F}, P)$ is a complete probability space where $\mathbb{F} = \{\mathcal{F}_t, t \in T\}$. A T-indexed process $X = \{X_t, t \in T\}$ is adapted to $\mathbb{F}$ if X_t is $\mathcal{F}_t$-measurable, for all $t \in T$.

Definition 4.1.7 *A function $v = \{v_t : t \in T\}$ is monotone on T if $v_s \leq 0 \Rightarrow v_t \leq 0 \;\; \forall\, t \geq s \in T$ A process V is monotone if $V(\omega)$ is monotone for each $\omega \in \Omega$.*

We can see that if a process V is decreasing in each component separately on T, then V is monotone on T.

In [3], Ivanoff and Merzbach proved the following useful lemma which ensures the existence of an optimal solution to the sequential detection problem (under certain conditions). It will be the key tool, in this chapter, leading to the optimal solution.

Lemma 4.1.8 *(cf.[3]; Lemma 3.10) Let U be a bounded T-indexed process adapted to a filtration $\mathbb{G}$ such that U is outer-continuous with inner limits. If $\mathbb{F} = \{\mathcal{F}_t : t \in T\}$ is a subfiltration of $\mathbb{G}$ and if a version of $V_t = E[U_t|\mathcal{F}_t]$ exists that is outer-continuous with inner limits, then for any $\mathbb{F}$-stopping set $\rho \subseteq R = [0, r]^2$,*

$$E\left[\int_\rho U_t dt\right] = E\left[\int_\rho V_t dt\right]$$

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

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

is optimal in the sense that:

$$E\left[\int_{\hat{\rho}} U_t dt\right] = \sup\left\{E\left[\int_\rho U_t dt\right] : \rho \subseteq R, \rho \text{ is an } \mathbb{F}\text{-stopping set}\right\}.$$

In this case, the optimal estimate of $\xi \cap R$ is given by $\hat{\xi_R} = \{t \in R : V_t \leq 0\}$.

Note that:
$s = (s_1, s_2) \ll t = (t_1, t_2) \Leftrightarrow s_i < t_i$ if $t_i > 0$, and $s_i = 0$ if $t_i = 0$, $i = 1, 2$. Also, a function is outer-continuous with inner limits if it is continuous from above with limits from the other three quadrants.

We are now ready to develop a general formula for the optimal solution of the sequential detection problem using the previous lemma.
Recall that $\xi = \{t : L_t > 0\}$ and $N_t > 0 \Rightarrow t \in \xi$, then

$$\begin{aligned} E[X_t|\mathcal{F}_t^N] &= E[I_{(t \in \xi^c)}|\mathcal{F}_t^N] \\ &= E[I_{(t \in \xi^c)}|\mathcal{F}_t^N] I_{(N_t = 0)} \\ &= P(L_t = 0|N_t = 0) I_{(N_t = 0)} \\ &= \frac{P(L_t = 0)}{P(N_t = 0)} I_{(N_t = 0)}. \end{aligned} \tag{4.1.1}$$

On the other hand,

$$\begin{aligned} P(N_t = 0) &= E[I_{(N_t=0)}] \\ &= E\left[E[I_{(N_t=0)}|\xi]\right] \\ &= E\left[P(N(A_t \cap \xi) = 0|\xi)\right] \\ &= E\left[e^{-\mu|A_t \cap \xi|}\right]. \end{aligned} \tag{4.1.2}$$

Hence from (4.1.1) and (4.1.2), we get:

$$E[X_t|\mathcal{F}_t^N] = \frac{P(L_t = 0)}{E\left[e^{-\mu|A_t \cap \xi|}\right]} I_{(N_t=0)}. \tag{4.1.3}$$

If L is the first line of Poisson process with rate λ, we can see that:

$$\begin{aligned} P(L_t = 0) &= P(A_t \cap \xi = \emptyset) \\ &= P(\text{no } L\text{-jumps occurred before } t) \\ &= e^{-\lambda|A_t|}. \end{aligned} \tag{4.1.4}$$

Substituting (4.1.4) in (4.1.3), we obtain:

$$\begin{aligned} E[X_t|\mathcal{F}_t^N] &= \frac{e^{-\lambda|A_t|}}{E\left[e^{-\mu|A_t \cap \xi|}\right]} I_{(N_t=0)} \\ &= \frac{e^{-(\lambda-\mu)|A_t|}}{e^{\mu|A_t|} E\left[e^{-\mu|A_t \cap \xi|}\right]} I_{(N_t=0)} \\ &= \frac{e^{-(\lambda-\mu)|A_t|}}{E\left[e^{\mu|A_t \setminus \xi|}\right]} I_{(N_t=0)}. \end{aligned}$$

Therefore

$$\begin{aligned} V_t &= E[U_t|\mathcal{F}_t^N] \\ &= -c_0 + (c_1 + c_1) E[X_t|\mathcal{F}_t^N] \\ &= -c_0 + \frac{(c_1 + c_1) e^{-(\lambda-\mu)|A_t|}}{E\left[e^{\mu|A_t \setminus \xi|}\right]} I_{(N_t=0)}. \end{aligned} \tag{4.1.5}$$

We observe that the denominator of the second term in the right side is increasing in t and the numerator is decreasing in t provided that $\lambda \geq \mu$, in which case V is monotone and an optimal solution exists.

Now, it is easy to see that for every $t = (t_1, t_2) \in R$:

$$\begin{aligned} V_t > 0 \;&\Leftrightarrow\; -c_1 + (c_0 + c_1)\frac{e^{-(\lambda-\mu)|A_t|}}{E\left[e^{\mu|A_t\backslash\xi|}\right]} > 0 \;\text{ and }\; N_t = 0 \\ &\Leftrightarrow\; \frac{e^{-(\lambda-\mu)|A_t|}}{E\left[e^{\mu|A_t\backslash\xi|}\right]} > \frac{c_1}{c_0 + c_1} \;\text{ and }\; N_t = 0 \\ &\Leftrightarrow\; e^{(\lambda-\mu)|A_t|}E\left[e^{\mu|A_t\backslash\xi|}\right] < 1 + \frac{c_0}{c_1} \;\text{ and }\; N_t = 0 \\ &\Leftrightarrow\; e^{(\lambda-\mu)t_1t_2}E\left[e^{\mu|A_t\backslash\xi|}\right] < 1 + \frac{c_0}{c_1} \;\text{ and }\; N_t = 0. \end{aligned}$$

Therefore

$$\hat{\rho} \;=\; D \cap \Gamma \text{ where } \begin{cases} D = \left\{t = (t_1, t_2) \in R : e^{(\lambda-\mu)t_1t_2}E\left[e^{\mu|A_t\backslash\xi|}\right] \leq 1 + \frac{c_0}{c_1}\right\} \\ \text{and} \\ \Gamma = \{t = (t_1, t_2) \in R : N_t = 0\} \end{cases} \tag{4.1.6}$$

■

4.2 Analytic approach

In this section, we will try to develop an explicit formula for the optimal solution of the two-dimensional sequential problem. In fact, we will make use of (4.1.6) with the hope of getting a closed form of the deterministic set D, or more precisely $E\left[e^{\mu|A_t\backslash\xi|}\right]$. Since the value of $|A_t\backslash\xi|$ is determined by the locations of the jump points of L inside A_t, the task now is to find their joint distribution.

Let $t = (t_1, t_2) \in \mathbf{R}^2_+$ and we denote respectively by $\tau_j = (x_j, y_j)$ for $j = 1, 2, \cdots, i$ and $\tau_{i+1} = (x_{i+1}, y_{i+1})$ the i jumps of L inside the set A_t and the first jump outside A_t such that:

$$\begin{cases} 0 = x_0 < x_1 < x_2 < \cdots < x_i < t_1 < x_{i+1} \\ t_2 = y_0 > y_1 > y_2 > \cdots > y_i > y_{i+1} > 0. \end{cases}$$

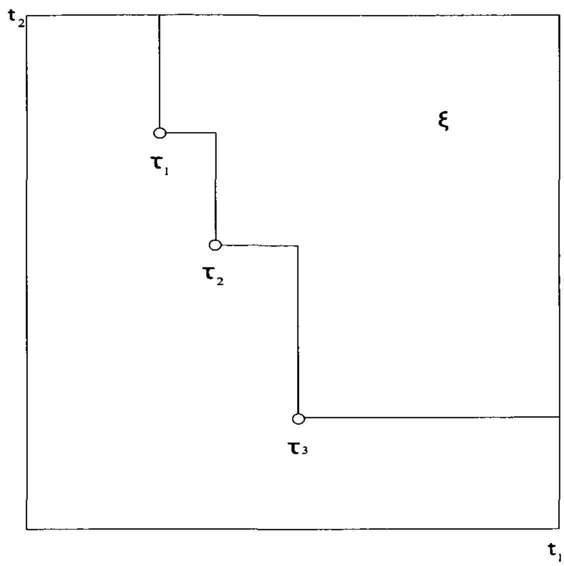

Figure 4.3: *The jump points of L inside* A_t

From Figure 4.3, we see that:

$$
\begin{aligned}
|A_t \backslash \xi| &= \sum_{i=0}^{\infty} \left[x_1 y_0 + (x_2 - x_1) y_1 + \cdots + (x_i - x_{i-1}) y_{i-1} + (t_1 - x_i) y_i\right] I_{(x_i < t_1 < x_{i+1})} \\
&= \sum_{i=0}^{\infty} \left[\sum_{j=1}^{i} (x_j - x_{j-1}) y_{j-1} + (t_1 - x_i) y_i\right] I_{(x_i < t_1 < x_{i+1})}. \qquad (4.2.1)
\end{aligned}
$$

In order to evaluate the expected value of $e^{\mu|A_t \backslash \xi|}$, we need to determine the joint density of $(\tau_1, \tau_2, \cdots, \tau_{i+1})$. To do that, we will use Figure 4.3 as follows:

If we start from the northwest corner $(0, t_2)$ and we follow the horizontal line to the east, we can see that: $X_1 \sim \exp(\lambda t_2)$, $Y_1 \sim U(0, t_2)$, and X_1 and Y_1 are independent. As a result,

$$
\begin{aligned}
f_{\tau_1}(x_1, y_1) &= \lambda t_2 e^{-\lambda t_2 x_1} \frac{1}{t_2} \qquad && 0 < x_1,\ y_1 < t_2 \\
&= \lambda e^{-\lambda t_2 x_1} && 0 < x_1,\ y_1 < t_2.
\end{aligned}
$$

Now, if we start from the point $\tau_1 = (x_1, y_1)$ and we take the second horizontal line to the east, we observe that given (x_1, y_1), $X_2 - x_1 \sim \exp(\lambda y_1)$ and $Y_2 \sim U(0, y_1)$.

X_2 and Y_2 are conditionally independent given τ_1.

Therefore,

$$\begin{aligned} f_{\tau_2|\tau_1}(x_2, y_2) &= \lambda y_1 e^{-\lambda y_1 (x_2 - x_1)} \frac{1}{y_1} && x_1 < x_2,\ y_2 < y_1 \\ &= \lambda e^{-\lambda y_1 (x_2 - x_1)} && x_1 < x_2,\ y_2 < y_1. \end{aligned}$$

Following the same argument, we get:

$$\begin{aligned} f_{\tau_{i+1}|\tau_i}(x_{i+1}, y_{i+1}) &= \lambda y_i e^{-\lambda y_i (x_{i+1} - x_i)} \frac{1}{y_i} && x_i < x_{i+1}, y_{i+1} < y_i \\ &= \lambda e^{-\lambda y_i (x_{i+1} - x_i)} && x_i < x_{i+1}, y_{i+1} < y_i. \end{aligned}$$

Before giving the joint density of $(\tau_1, \tau_2, \cdots, \tau_{i+1})$, we remark that given $\tau_1, \tau_2, \ldots, \tau_k$ for $k \leq i$

$$f_{\tau_{k+1}|\tau_1,\ldots,\tau_k}(x_{k+1}, y_{k+1}) = f_{\tau_{k+1}|\tau_k}(x_{k+1}, y_{k+1}); x_k < x_{k+1}, y_{k+1} < y_k.$$

Thus,

$$\begin{aligned} f_{\tau_1,\ldots,\tau_{i+1}}((x_1, y_1), \ldots, (x_{i+1}, y_{i+1})) &= f_{\tau_{i+1}|\tau_i}(x_{i+1}, y_{i+1}) f_{\tau_1,\ldots,\tau_i}((x_1, y_1), \ldots, (x_i, y_i)) \\ f_{\tau_1,\ldots,\tau_i}((x_1, y_1), \ldots, (x_i, y_i)) &= f_{\tau_i|\tau_{i-1}}(x_i, y_i) f_{\tau_1,\ldots,\tau_{i-1}}((x_1, y_1), \ldots, (x_{i-1}, y_{i-1})) \\ &\vdots \\ f_{\tau_1,\ldots,\tau_3}((x_1, y_1), \ldots, (x_3, y_3)) &= f_{\tau_3|\tau_2}(x_3, y_3) f_{\tau_1,\tau_2}((x_1, y_1), (x_2, y_2)) \\ f_{\tau_1,\tau_2}((x_1, y_1), (x_2, y_2)) &= f_{\tau_2|\tau_1}(x_2, y_2) f_{\tau_1}(x_1, y_1). \end{aligned}$$

Multiplying both sides of the equalities, we get:

for $0 = x_0 < x_1 < x_2 < \cdots < x_{i+1}$ and $t_2 = y_0 > y_1 > y_2 > \cdots > y_{i+1} > 0$

$$\begin{aligned} f_{\tau_1,\ldots,\tau_{i+1}}((x_1, y_1), \ldots, (x_{i+1}, y_{i+1})) &= f_{\tau_{i+1}|\tau_i}(x_{i+1}, y_{i+1}) \cdots f_{\tau_2|\tau_1}(x_2, y_2) f_{\tau_1}(x_1, y_1) \\ &= \lambda^{i+1} e^{-\lambda[t_2 x_1 + y_1(x_2 - x_1) + \cdots + y_i(x_{i+1} - x_i)]} \\ &= \lambda^{i+1} e^{-\lambda \sum_{j=1}^{i+1} (x_j - x_{j-1}) y_{j-1}}. \end{aligned} \tag{4.2.2}$$

Define $Q_t = E\left[e^{\mu|A_t\setminus\xi|} I_{(L_t>0)}\right]$.

Hence,

$$
\begin{aligned}
E\left[e^{\mu|A_t\setminus\xi|}\right] &= E\left[e^{\mu|A_t\setminus\xi|} I_{(L_t>0)} + e^{\mu|A_t\setminus\xi|} I_{(L_t=0)}\right] \\
&= Q_t + E\left[e^{\mu|A_t\setminus\xi|} I_{(L_t=0)}\right] \\
&= Q_t + E\left[e^{\mu|A_t|} I_{(L_t=0)}\right] \\
&= Q_t + e^{\mu|A_t|} e^{-\lambda|A_t|} \\
&= Q_t + e^{(\mu-\lambda)|A_t|}. \qquad (4.2.3)
\end{aligned}
$$

Finally, using (4.1.5) and (4.2.3) we have:

$$
\begin{aligned}
\forall\ t=(t_1,t_2) \in R \qquad V_t &= -c_1 + \frac{(c_0+c_1)e^{(\mu-\lambda)|A_t|}}{Q_t + e^{(\mu-\lambda)|A_t|}} I_{(N_t=0)} \\
&= -c_1 + \frac{c_0+c_1}{1+e^{(\lambda-\mu)t_1t_2}Q_t} I_{(N_t=0)}. \qquad (4.2.4)
\end{aligned}
$$

Using (4.2.1) and (4.2.2), we will try to find a closed formula for Q_t which is as simple as possible:

$$
\begin{aligned}
Q_t &= E\left[e^{\mu|A_t\setminus\xi|} I_{(L_t>0)}\right] \\
&= \sum_{i=1}^{\infty} \int_0^{t_1}\int_{x_1}^{t_1}\cdots\int_{x_{i-1}}^{t_1}\int_{t_1}^{\infty}\int_0^{t_2}\int_0^{y_1}\cdots\int_0^{y_{i-1}}\int_0^{y_i} e^{\mu\left[\sum_{j=1}^{i}(x_j-x_{j-1})y_{j-1}+(t_1-x_i)y_i\right]} \\
&\qquad \lambda^{i+1} e^{-\lambda\left[\sum_{j=1}^{i+1}(x_j-x_{j-1})y_{j-1}\right]}\, dy_{i+1}\cdots dy_1\, dx_{i+1}\cdots dx_1 \\
&= \sum_{i=1}^{\infty} \int_0^{t_1}\int_{x_1}^{t_1}\cdots\int_{x_{i-1}}^{t_1}\int_0^{t_2}\int_0^{y_1}\cdots\int_0^{y_{i-1}} \lambda^{i+1} e^{(\mu-\lambda)\left[\sum_{j=1}^{i}(x_j-x_{j-1})y_{j-1}\right]+\mu(t_1-x_i)y_i} \\
&\qquad I(x_i,y_i)\, dy_i\cdots dy_1\, dx_i\cdots dx_1.
\end{aligned}
$$

where

$$
\begin{aligned}
I(x_i,y_i) &= \int_{t_1}^{\infty}\int_0^{y_i} e^{-\lambda(x_{i+1}-x_i)y_i}\, dx_{i+1}\, dy_{i+1} \\
&= \frac{1}{\lambda} e^{-\lambda(t_1-x_i)y_i}.
\end{aligned}
$$

So Q_t is given by:

$$Q_t = \sum_{i=1}^{\infty} \lambda^i \int_0^{t_1} \int_{x_1}^{t_1} \cdots \int_{x_{i-1}}^{t_1} \int_0^{t_2} \int_0^{y_1} \cdots \int_0^{y_{i-1}} e^{(\mu-\lambda)\left[\sum_{j=1}^{i}(x_j - x_{j-1}) y_{j-1}\right]}$$
$$e^{(\mu-\lambda)(t_1 - x_i) y_i}\, dy_i \cdots dy_1\, dx_i \cdots dx_1. \tag{4.2.5}$$

We note that if $\lambda = \mu$, $Q_t = \sum_{i=1}^{\infty} \frac{(\lambda t_1 t_2)^i}{(i!)^2}$.

It turns out that the formula reached in (4.2.5) doesn't lead to a simplified theoretical form for either V_t or $\hat{\rho}$. Consequently, we will use the formula of the stopping set $\hat{\rho}$ stated in (4.1.6), the area $|A_t \backslash \xi|$ mentioned in (4.2.1) and the distribution of the L-jump points to compute $E\left[e^{\mu |A_t \backslash \xi|}\right]$ using Monte Carlo simulation in order to find the boundary of the sets D and $\hat{\rho}$.

■

4.3 Practical approach

As we failed to produce an "explicit" formula of the deterministic set D theoretically, we are going to approach this problem in different manner using simulations. The following section will describe the steps that form the main algorithm, programmed in R [1], which provides the upper-boundaries of D and $\hat{\rho}$. To achieve that, we will explain in detail the arguments used in each step, produce its outcome and delineate it by figures when it is necessary. To complete this analysis, we will also simulate the random set Γ in order to provide an outline of the possible optimal solution to the sequential detection problem. The parameters considered throughout this analysis are: $r = 1$, $\lambda = 30$, $\mu = 28$, $n = 10$, $m = 10000$, and $c_0/c_1 = 10$.

Step 1: Simulation of the first line L with rate λ.

This step is to simulate one sample of the first line L with rate λ on the square $R = [0, r]^2$. To this end, we recall what we have seen earlier:

[1] The detailed algorithm is provided in the appendix.

Let $\tau_i = (x_i, y_i)$ be the jump points of L inside R. Then given τ_i, X_{i+1} and Y_{i+1} are conditionally independent and $X_{i+1} - x_i \sim \exp(\lambda y_i)$ and $Y_{i+1} \sim U(0, y_i)$.

The outcome will be accumulated in a matrix P for future use. An example of the matrix P is found to be:

$$P = \begin{array}{c} \begin{array}{cc} \quad x_L & \qquad\quad y_L \end{array} \\ \begin{pmatrix} 0.01751037 & 0.87742568 \\ 0.04879942 & 0.56310752 \\ 0.06090159 & 0.22923729 \\ 0.45731178 & 0.10944978 \\ 0.51801568 & 0.09286938 \\ 0.87400566 & 0.08937724 \end{pmatrix} \end{array}$$

Furthermore, we can use the matrix P to plot the obtained sample of the first line L in two dimensions as it is illustrated in Figure 4.4.

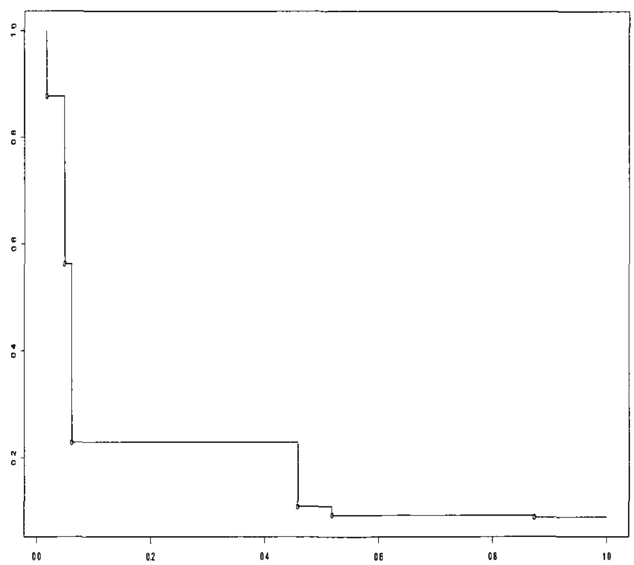

Figure 4.4: *First line of Poisson process on* $[0, 1]^2$ *with* $\lambda = 30$

<u>Step 2</u>: Detection of the jump points of L inside A_t for $t \in R = [0, r]^2$.

Now, we will introduce the function "Detectpoint"that makes use of the matrix P

obtained in step (1). It searches for the L-jump points, if they exist, which take place inside A_t for every $t \in R$. The jump points detected are then gathered in a new matrix P_t.

If $t = (0.6, 0.4)$ for instance, the outcome of Detectpoint$(0.6, 0.4, P)$ will be:

$$P_t = \begin{array}{c} \begin{array}{cc} \quad x_t \quad & \quad y_t \quad \end{array} \\ \begin{pmatrix} 0.06090159 & 0.22923729 \\ 0.45731178 & 0.10944978 \\ 0.51801568 & 0.09286938 \end{pmatrix} \end{array}$$

Step 3: Calculation of the area $A_t \backslash \xi$.

Using the expression stated in (4.2.1) and the result produced from the previous step, we calculate the area of $A_t \backslash \xi$ for any chosen point $t = (t_1, t_2) \in R$.

For the same point $t = (0.6, 0.4)$ we get:

$$\text{Area}(0.6, 0.4, P) = 0.1294905$$

Step 4: Tabulation of areas $A_t \backslash \xi$ for all t on a grid in $[0, r]^2$.

We are going to impose an $n \times n$ grid on R and execute step (3) for every $t \in R$ which falls among the points that compose the grid. The purpose of the function, that is now to follow, is to assemble the results obtained for the $(n+1)^2$ points created in a new matrix.

Step 5: Monte Carlo simulation.

The objective of this step is to repeat the procedures (1) and (4) m times and estimate $E\left[e^{\mu|A_t \backslash \xi|}\right]$ using Monte Carlo simulation [2] .

Step 6a: Detection of the boundary of the lower layer D.

First, we make use of the outcome from the preceding step and the inequality (4.1.6)

[2] The outcomes of steps (4) and (5) will be omitted since n takes large values in practice.

presented earlier which is satisfied by the points belonging to the deterministic set D. As a result, we obtain a matrix called D_1 that is composed of either 0's or 1's. Consequently, the deterministic set D is totally determined by the 1's.

In our example where $r = 1$, $\lambda = 30$, $\mu = 28$, $n = 10$, $m = 10000$ and $c_0/c_1 = 10$, the matrix D_1 is found to be :

$$D_1 = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}$$

Step 6b: Detection of the boundary of the lower layer D.

At this stage, we transform the 1's provided in the step 6a to coordinates (x, y) of the points shaping the curve of the desired upper-bound of the deterministic set D. This curve is totally determined by the 1's which end each row or the first 1 below the 0's in each column.

Taking in consideration the values chosen for all the parameters which play a role in determining the upper-bound of the lower layer D, the coordinates of the obtained points are found to be:

x_D	0.1	0.1	0.1	0.1	0.2	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
y_D	1.0	0.9	0.8	0.7	0.6	0.5	0.4	0.3	0.2	0.2	0.1	0.1	0.1	0.1

The upper-bound of the deterministic set D for the same values is illustrated in Figure 4.5.

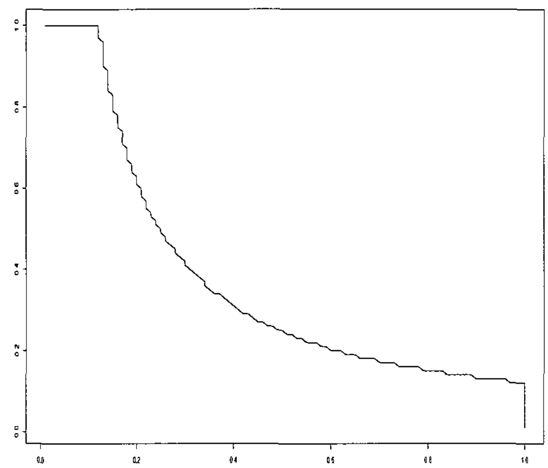

Figure 4.5: *The upper-bound of the deterministic set D for $c_0/c_1 = 10$*

Examples of the upper-bound of the set D, for $\lambda = 30, \mu = 28, n = 100, m = 10000$ and different values of the ratio c_0/c_1, are illustrated in Figures 4.6 and 4.7.

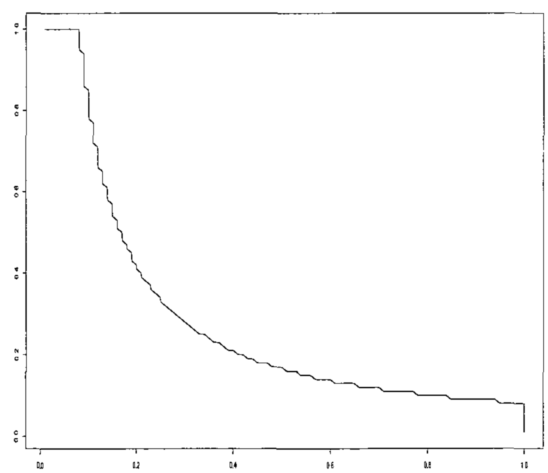

Figure 4.6: *The upper-bound of the deterministic set D for $c_0/c_1 = 5$*

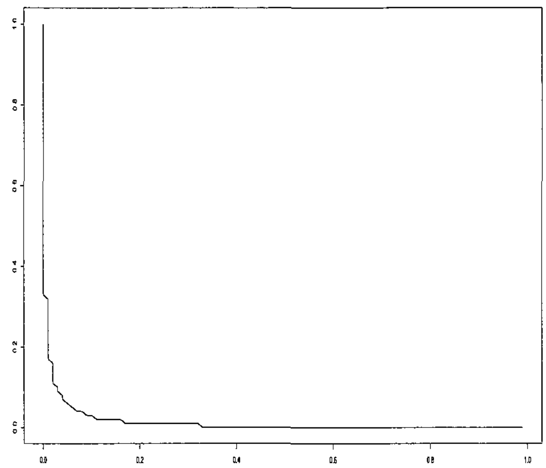

Figure 4.7: *The upper-bound of the deterministic set D for $c_0/c_1 = 0.1$*

We observe that the larger c_0/c_1 gets, the bigger the area of the deterministic set D becomes. In fact, we are not surprised by this result since it is compatible with the definition of D itself.

Step 7: Simulation of Poisson process N on ξ.

Now, we are going to determine Γ, the random part of the stopping set $\hat{\rho}$ shown earlier in (4.1.6). To do that, we will simulate the first line L which generates the set ξ. Given that we obtain i jump points of L inside R, say $\{\tau_j = (x_j, y_j)\}_{j=1}^{i}$, we decompose the set ξ to i rectangles $a_j = [x_j, x_0] \times [y_j, y_{j-1}], j = 1 \cdots i$, provided that $x_0 = y_0 = r$. Then, we simulate i independent Poisson processes inside each rectangle with the appropriate rate $\mu|a_j|$.

For the same values considered before $r = 1$ and $\mu = 28$, we obtain [3] :

[3]See appendix for the obtained matrix P_N

$$P_N = \begin{array}{c} \begin{array}{cc} x_N & y_N \end{array} \\ \begin{pmatrix} 0.2612112 & 0.9060352 \\ 0.2955656 & 0.9696374 \\ \vdots & \vdots \\ 0.8946426 & 0.0696154 \\ 0.9918674 & 0.3541605 \end{pmatrix} \end{array}$$

Since $\hat{\rho} = D \cap \Gamma$, one has to stop if he reaches the upper-bound of the deterministic lower layer or even before that if he meets any jump point of the Poisson process N. This situation is illustrated in Figures 4.8 and 4.9 for two different values of c_0/c_1.

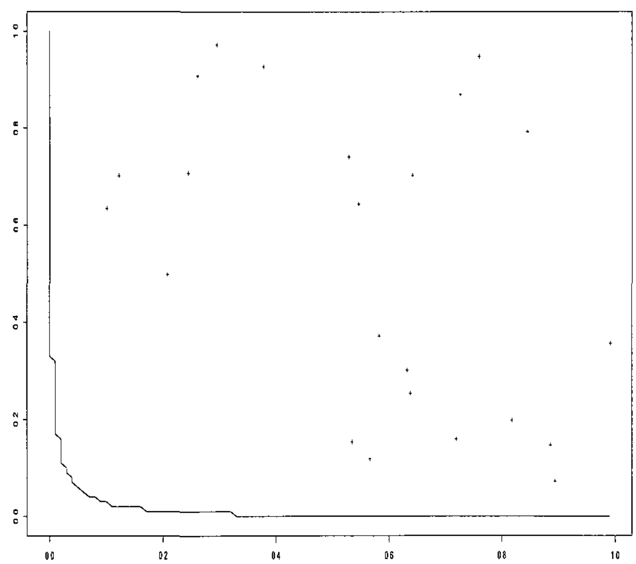

Figure 4.8: *Detection of the optimal solution for* $c_0/c_1 = 0.1$

We can see that the optimal solution $\hat{\rho}$ coincides with the deterministic set D for $c_0/c_1 = 0.1$ since all the jump points of N are outside D.

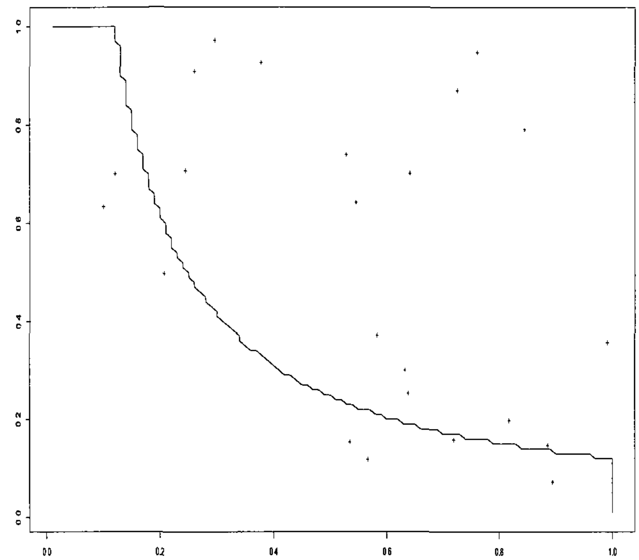

Figure 4.9: *Detection of the optimal solution for* $c_0/c_1 = 10$

In this case, we see some jump points of N inside the deterministic lower-layer D. Therefore, $\hat{\rho}$ consists of the points of D that do not lie to the northeast of any jump point of N.

■

Chapter 5

Estimation of the parameters

In practical situations, the parameters related to the change point (set) and the Poisson process need to be estimated to be able to use all the optimal solutions described in the previous chapters. To this end, the first section will be devoted to discussing the one dimensional case. The two dimensional case is examined in the second section.

5.1 One dimensional case

Recall that given the exponential change point σ with rate λ, N is a Poisson process with rate μ on $[\sigma, \infty)$. We denote by τ_i the jump points of the Poisson process N. As seen previously, we can write that $\tau_1 = \sigma + v$ where v is exponentially distributed with rate μ and independent of the change point σ.
For $k = 1, 2$, let $\tau_{k,1}, \ldots, \tau_{k,n}$ be independent and identically distributed with the same distribution as τ_k. We will mention, for the one dimensional case, two different methods to estimate λ and μ.
We denote by m_1 the first sample moment:

$$m_1 = \frac{\sum_{i=1}^{n} \tau_{1,i}}{n}.$$

Therefore,

$$E[m_1] = \frac{1}{\lambda} + \frac{1}{\mu}.$$

Both methods require n replicates of τ_1 to estimate $\frac{1}{\lambda} + \frac{1}{\mu}$. Method one uses n replicates of τ_2, while method two requires for a fixed $T > 0$ a single replicate of $\tau_k I_{(\tau_k \leq T)}$, $k = 1, \ldots, \infty$.

i) In the first method, we get the following estimates for μ and λ:

$$\hat{\mu} = \frac{n}{\sum_{i=1}^{n}(\tau_{2,i} - \tau_{1,i})} \quad \text{and} \quad \hat{\lambda} = \frac{1}{m_1 - \frac{1}{\hat{\mu}}}.$$

ii) Fix $T > 0$ (T large enough). We will count the jump points of N which fall inside the interval $(\tau_1, T]$.

Since

$$E\left[N(\tau_1, T]\right] = \mu(T - \tau_1),$$

hence,

$$E\left[N(T) - 1\right] = \mu(T - \tau_1).$$

Therefore, our estimates are given by:

$$\hat{\mu} = \frac{N(T) - 1}{T - \tau_1} \quad \text{and} \quad \hat{\lambda} = \frac{1}{m_1 - \frac{1}{\hat{\mu}}}.$$

5.2 Two dimensional case

To estimate the intensity of the Poisson process N in two dimensions, we can use a similar estimator to the one used in one dimension.

Fix $r > 0$, and keeping the same notation used in the previous chapter, we write $R = [0, r]^2$. We observe the jump points of N on R ($\Delta_N \cap R$). Recalling definition 4.1.3, let N_1 denote the first line of N and let $A = \cup_{\tau \in \Delta_{N_1} \cap R}\{t \in R : \tau \ll t\}$.

Then $E[N(A)] = \mu|A|$.

Therefore, we can estimate the intensity μ by:

$$\hat{\mu} = \frac{N(A)}{|A|} = \frac{\text{number of jump points inside } A}{\text{area of } A}.$$

Contrary to the one dimensional case, we are not able to give an estimator for the intensity of the first line process L. The complexity of this task resides in the evaluation of the distribution of the area between the first line L and the first jump points of N.

■

Chapter 6

Conclusion

As indicated in the introduction, this research was built from a graduate student's perspective to attack the problem of the detection-support in two dimensions. We were able to solve this problem in one dimension for two different observation schemes and to give an explicit solution in each case. Also, we solved the case in two dimensions for sequential observation under some conditions using both a theoretical and a practical approach. However, there are still many open questions which need to be answered that may require a different methodology.

- The restrictive condition of monotonicity, which was of great use in solving the detection-support in the one dimensional case, has not been proved for two dimensions when $\lambda < \mu$. This case might need to be solved using other tools than the ones used here.

- Another open question is how to proceed in estimation of the parameters in the two dimensional case, especially the intensity of the first line.

- As done in one dimension, the ex-post scheme in two dimensions is another problem which can be studied . We note that other information structures can be considered such as the combination of the sequential and the ex-post schemes for instance.

- More general reward functions can be considered. Indeed, in papers [2] and [3], the authors considered a slightly more general valuation function.

- The example of Poisson process is so far reasonable in many practical situations. However, one can be interested in solving the detection-support problem for more general processes. Moreover, it will be interesting to deal with the detection-support problem using nonparametric methods which can be of great use in practice because of the lack of a priori information about the process observed in general.

- The approach taken in this thesis can be applied to find a practical solution to the more general change-set problem studied in [3]

■

Appendix

Step 1: Simulation of the first line L with rate λ.

```
# lambda = the rate of the first line process L
Simfirst< −function(r,lambda){
   X< −NULL ; Y< −NULL ; y< −r
   test< −0
   while(test<r)
     {e< −rexp(1,lambda*y) ; u< −runif(1,0,y)
     X< −c(X,e) ; Y< −c(Y,u)
     y< −u
     test< −test+e}
   k< −length(X)-1
   T1< −X[1:k] ; T2< −Y[1:k]
   X< −cumsum(T1) ; Y< −T2
   P=cbind(X,Y)}
```

The following $R-$code is used to plot the result of the previous function in two dimensions as it is illustrated in figure 3, but it is not a necessary piece in our main algorithm.

```
Plot2fl< −function(r,P){
   X< −P[,1] ; Y< −P[,2] ;x< −X ; y< −Y
   X< −c(X,X) ; Y< −c(Y,Y)
```

```
    X< -sort(X) ; Y< -sort(Y,decreasing=TRUE)
    X< -c(X,r) ; Y< -c(r,Y)
    P1< -cbind(X,Y)
    plot(P1,xlab=" ",ylab=" ",type="l")
    points(x,y,xlab=" ",ylab=" ",type="p")}
```

Step 2: Detection of the jump points of L inside A_t for $t \in R = [0, r]^2$.

```
Detectpoint< -function(t1,t2,P){
    X< -P[,1] ; Y< -P[,2] ; xt< -NULL ; yt< -NULL
    k< -length(X)
    for(i in 1:k) {if (X[i]<t1 & Y[i]<t2){xt< -c(xt,X[i]); yt< -c(yt,Y[i])}}
    Pt< -cbind(xt,yt)}
```

Step 3: Calculation of the area $A_t \backslash \xi$.

```
Area< -function(t1,t2,P){
    Pt< -Detectpoint(t1,t2,P)
    xt< -Pt[,1] ; yt< -Pt[,2] ; kt< -length(xt)
    if(kt==0){At< -t1*t2}
    else{VA=c(rep(0,kt+1)) ; xt< -c(0,xt) ; yt< -c(t2,yt)
        for(i in 1:kt) {VA[i]=(xt[i+1]-xt[i])*yt[i]}
        VA[kt+1]< -(t1-xt[kt+1])*yt[kt+1]
        At=sum(VA)}
    At}
```

Step 4: Tabulation of areas $A_t \backslash \xi$ for all t on a grid in $[0, r]^2$.

```
# n = grid
Totalarea< -function(r,n,P){
    areat< -matrix(0,n+1,n+1)
```

```
    a< −seq(0,r,length.out=n+1) ; b< −sort(a,decreasing=TRUE)
    for (i in 1:(n+1)){for(j in 1:(n+1)){areat[i,j]< −Area(a[j],b[i],P)} }
    areat}
```

Step 5: Monte carlo simulation.

```
# mu = the rate of Poisson process N; m = number of iterations
Montsim< −function(r,lambda,mu,n,m){
    A< −matrix(0,n+1,n+1)
    for(i in 1:m){
        P< −Simfirst(r,lambda)
        B< −Totalarea(r,n,P)
        A< −A+exp(mu*B)}
    A< −A/m}
```

Step 6a: Detection of the boundary of the lower layer D.

```
# d = c0/c1
Detectbound1< −function(r,lambda,mu,n,m,d,A){
    D1< −matrix(0,n+1,n+1) ; G< −matrix(0,n+1,n+1)
    D2< −matrix(0,n+1,n+1)
    a< −seq(0,r,length.out=n+1) ; b< −sort(a,decreasing=TRUE)
    for (i in 1:(n+1)){for (j in 1:(n+1)){G[i,j]=(lambda-mu)*a[j]*b[i]}}
    G< −exp(G)
    D2< −G*A-1-d
    for(i in 1:(n+1)){for(j in 1:(n+1)){if(D2[i,j]>0){D1[i,j]=0}else{D1[i,j]=1}}}
    D1}
```

Step 6b: Detection of the boundary of the lower layer D.

```
Detectbound2< −function(n,D1){
    X< −NULL ; Y< −NULL
    for(i in 1:n){
        a< −sum(D1[i,])
        if(a>0){X< −c(X,(a-1)/n) ; Y< −c(Y,(n+1-i)/n)}}
    i< −n+1 ; j< −n
    while(i>2){
        b< −sum(D1[,i]) ; c< −sum(D1[,j])
        if(b==c){X< −c(X,(j-1)/n) ; Y< −c(Y,(b-1)/n)}
        i< −j ; j< −j-1}
    X< −sort(X) ; Y< −sort(Y,decreasing=TRUE)
    D2< −cbind(X,Y)
    if(length(X)==0){print(”there are no jump points inside  R”)}
    else{return(D2)}}
```

Once the previous functions are stated, one can use the following algorithm with the appropriate parameters to get the deterministic upper-bound of the lower layer D:

```
    A< −Montsim(r,lambda,mu,n,m)
    D1< −Detectbound1(r,lambda,mu,n,m,d,A)
    D2< −Detectbound2(n,D1)
    plot(D2,xlab=””,ylab=””,type=”l”)
```

Step 7: Simulation of Poisson process N on ξ.

```
    Simpoisson< −function(r,mu,P){
    XN< −NULL ; YN< −NULL
    XL< −P[,1] ; YL1< −P[,2]
    n< −length(XL) ; YL< −c(r,YL1)
    for(i in 1:n){
```

```
    xn< −NULL ; yn< −NULL
    T1< −NULL ; T2< −NULL ; T3< −NULL ; T4< −NULL
    test< −XL[i]
    while(test<r)
      {e< −rexp(1,mu*(YL[i]-YL[i+1])) ; u< −runif(1,YL[i+1],YL[i])
      xn< −c(xn,e) ; yn< −c(yn,u)
      test< −test+e}
    k< −length(xn)-1
    if(k>0){
      T1< −xn[1:k] ; T2< −cumsum(T1)
      T3< −T2+XL[i] ; T4< −yn[1:k]}
   XN< −c(XN,T3) ; YN< −c(YN,T4)}
  PN=cbind(XN,YN)}
```

$$
\begin{array}{cc}
 & \begin{array}{cc} \quad x_N \quad & \quad y_N \quad \end{array} \\
P_N = & \begin{pmatrix}
0.2612112 & 0.9060352 \\
0.2955656 & 0.9696374 \\
0.3776079 & 0.9245080 \\
0.7261425 & 0.8678994 \\
0.7600570 & 0.9454704 \\
0.1005544 & 0.6318021 \\
0.1215025 & 0.6987850 \\
0.2450285 & 0.7048870 \\
0.5285783 & 0.7378753 \\
0.5458719 & 0.6409081 \\
0.6410205 & 0.6994561 \\
0.8453175 & 0.7893326 \\
0.2072118 & 0.4955613 \\
0.5345843 & 0.1517567 \\
0.5665885 & 0.1159166 \\
0.5830720 & 0.3689551 \\
0.6322667 & 0.2984585 \\
0.6378314 & 0.2513216 \\
0.7187938 & 0.1565363 \\
0.8178011 & 0.1954032 \\
0.8862095 & 0.1444145 \\
0.8946426 & 0.0696154 \\
0.9918674 & 0.3541605
\end{pmatrix}
\end{array}
$$

Bibliography

[1] Hahubia, Ts.and Mnatsakanov, R., **On the mode-change problem for random measures**, *Georg.Math.J.* Vol 3, pp. 343 - 362, 1996.

[2] Herberts, T. and Jensen, U., **Optimal Detection of a Change Point in a Poisson Process for Different Observation Schemes**, *Scand. J. Stat.* Vol 31, pp. 347 - 366, 2004.

[3] Ivanoff, B.G. and Merzbach, E., **Optimal detection of a change-set in a spatial Poisson process**, *The Annals of Applied Probability.* Vol 20, pp. 640 - 659, 2010.

[4] Khmaladze, E., Mnatsakanov, R. and Toronjadze, N., **The change-set problem for Vapnik-Cervonenkis classes** , *Mathematical Methods in Stat.* Vol 15, pp. 224 - 231, 2006.

[5] Khmaladze, E., Mnatsakanov, R. and Toronjadze, N., **The change-set problem and local covering numbers** , *Mathematical Methods in Stat.* Vol 15, pp. 289 - 308, 2006.

[6] Peskir, G. and Shiryaev, A.N., **Solving the Poisson disorder problem, in: Advances in Finance and Statistics, Essays in Honour of Dieter Sondermann and P. Schonbucher, eds**, pp. 295 - 312, Springer, Berlin, 2002.