



**HAL**  
open science

## Inférence statistique sur le processus de Mino

Rabih Damaj

► **To cite this version:**

Rabih Damaj. Inférence statistique sur le processus de Mino. Statistiques [math.ST]. Université de Bretagne Sud, 2015. Français. NNT : 2015LORIS369 . tel-01295051

**HAL Id: tel-01295051**

**<https://theses.hal.science/tel-01295051>**

Submitted on 30 Mar 2016

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



**THESE / UNIVERSITE DE BRETAGNE-SUD**

*sous le sceau de l'Université européenne de Bretagne*

pour obtenir le titre de

**DOCTEUR DE L'UNIVERSITE DE BRETAGNE-SUD**

*Mention : Mathématiques*

**Ecole doctorale : SICMA**

présentée par

**Rabih Damaj**

Laboratoire de Mathématiques de Bretagne

Atlantique URM CNRS 6205 et Université de

Bretagne Sud

## Inférence statistique sur le processus de Mino

**Thèse soutenue le 29 mai 2015**

devant le jury composé de :

**Hicham Abdallah**

Professeur, Université libanaise, Beyrouth / Rapporteur

**Laurent Serlet**

Professeur, Université Blaise Pascal / Rapporteur

**Philippe Castagliola**

Professeur, Université de Nantes / Examineur

**Emmanuel Frenod**

Professeur, Université de Bretagne Sud / Examineur

**Gilles Durrieu**

Professeur, Université de Bretagne Sud / Directeur de thèse

**Evans Gouno**

MCF - HDR, Université de Bretagne Sud / Directeur de thèse

---

Laboratoire de Mathématiques de Bretagne Atlantique (LMBA)  
UMR CNRS 6205 et Université de Bretagne-Sud  
Campus de Tohannic  
56000 Vannes, France

---

# Remerciement

Arriver à la fin, c'est-à-dire qu'il y a un ensemble des gens qui ont mis leurs expériences et leurs compétences à votre disposition, tout le long du chemin que tu as traversé, pour cela j'adresse mes plus chers remerciements aux directeurs de ma thèse : Pr. Gilles Durrieu et Dr. Evans Gouno, pour leurs soutiens, leurs patiences, leurs sympathies, et pour le temps qu'ils m'ont accordé.

Je tiens à remercier l'ensemble des membres du jury. Merci à Pr. Laurent Serlet et à Pr. Hicham Abdallah d'avoir accepté d'être les rapporteurs de la thèse. Leur relecture attentive a largement bénéficié à ce manuscrit. Merci à Pr. Philippe Castagliola et à Pr. Emanuel Frenod d'avoir accepté d'être des examinateurs de la thèse.

J'adresse également mes remerciements à tous les membres du laboratoire LMBA (Laboratoire de Mathématiques de Bretagne Atlantique), qui ont su faire régner une bonne ambiance durant ces années. Je remercie par ailleurs les collègues de l'IUT de Vannes (département STID), où j'ai passé 2 ans pour le poste d'ATER, pour leurs sympathies et leurs bonnes humeurs.

A tous mes amis qui m'ont toujours soutenu, aidé et conseillé, je leur suis entièrement reconnaissant, et je ne rentre pas dans la nomination parce que la liste est très longue.

Enfin, il faut que je cherche des mots inhabituels pour remercier les soldats inconnus, pour leurs sacrifices, leurs prières et leurs encouragements infinis, qui ont formé les principales motifs pour continuer, avec tous mes respects et mes salutation j'adresse mes remerciement à : mon père Kamal, ma mère Nazek, ma deuxième mère Nawal, mon frère Ziad, ma soeur Raghida...



---

# Inférence statistique sur le processus de Mino

**Résumé :** Le sujet de cette thèse concerne l'inférence statistique sur le processus de Mino que nous définissons comme un processus auto-excité de mémoire 1 dont l'intensité est de forme particulière.

Nous donnons tout d'abord une description générale des processus auto-excités et des méthodes possibles pour estimer les paramètres de l'intensité de ces processus. Puis, nous considérons le cas particulier d'un processus auto-excité de mémoire 1 que l'on rencontre en traitement du signal et que nous avons dénommé : *processus de Mino*.

Nous montrons que ce processus est un processus de renouvellement dont les interarrivées ont une distribution particulière que nous étudions en détails.

Nous envisageons alors le problème de l'estimation des paramètres de l'intensité du processus de Mino en utilisant la méthode du maximum de vraisemblance. Nous résolvons les équations de vraisemblance en utilisant l'algorithme de Newton-Raphson. La méthode est appliquée à des données simulées. La convergence de l'algorithme de Newton-Raphson est démontrée, de même que l'existence et l'unicité des estimateurs.

Nous terminons par la construction d'un test d'hypothèses qui permet de détecter si un processus ponctuel est auto-excité ou non.

**Mots clés :** Processus stochastique, Processus auto-excité, Maximum vraisemblance, Optimisation, Données manquante, Algorithme, Test d'hypothèse.



---

# Statistical inference on Mino process

**Abstract :** The subject of this PhD thesis is the statistical inference on Mino process that we define as a one-memory self-exciting point process which intensity has a special form.

We begin with a general description of self-exciting point processes and we present methods used to estimate the intensity parameters of these processes.

We consider the special case of a one-memory self-exciting point process, used in signal processing. We call the process : *the Mino process*. This process can be interpreted as a renewal process which interarrival times that follow a special distribution that we study in details.

In order to estimate the parameters of a Mino process intensity, we utilize the maximum likelihood method. We solve the likelihood equations with a Newton-Raphson algorithm. We show the efficiency of the method on simulated data.

The convergence of the Newton-Raphson algorithm and, the existence and uniqueness of the maximum likelihood estimators are proved.

Lastly, we construct a test of hypothesis to assess whether a point process is self-exciting or not.

**Keywords :** Stochastic processes, Self-exciting processes, Maximum likelihood, Optimization, Missing data, Algorithms, Hypothesis testing.



## Table des matières

---

---

# Table des matières

<b>Introduction générale</b>	<b>5</b>
<b>I Processus ponctuels auto-excités</b>	<b>11</b>
1 Introduction . . . . .	11
2 Analyse spectrale des processus ponctuels . . . . .	15
3 Processus de cluster . . . . .	18
4 Processus auto-excité (PAE) . . . . .	20
4.1 Généralités . . . . .	21
4.2 Loi du nombre d'événements . . . . .	23
4.3 Théorème limite pour la somme d'un processus ponctuel auto-excité . . . . .	24
5 Processus ponctuel auto-excité de mémoire limitée . . . . .	26
5.1 Caractérisation d'un processus ponctuels $m$ -mémoire lorsque $m = 0$ . . . . .	26
5.2 Caractérisation d'un processus ponctuels $m$ -mémoire lorsque $m > 0$ . . . . .	27
5.3 Processus auto-excité et renouvellement . . . . .	27
<b>II Estimation pour les processus auto-excités</b>	<b>29</b>
1 Introduction . . . . .	29
2 Vraisemblance pour un processus stochastique . . . . .	30

2.1	Produit intégrale . . . . .	30
2.2	Vraisemblance pour le PAE . . . . .	31
3	Algorithmes pour le calcul des estimateurs . . . . .	34
3.1	Randomized search methods (RSM) . . . . .	34
3.2	Algorithme Newton-Raphson . . . . .	34
3.3	Algorithme EM . . . . .	35
4	Une illustration de l'algorithme EM . . . . .	38
4.1	Vraisemblance du nombre des durées . . . . .	40
4.2	Estimation de $\theta$ en utilisant l'algorithme de Newton-Raphson . . . . .	40
4.3	Vraisemblance pour le modèle complet . . . . .	42
4.4	Estimation de $\theta$ en utilisant l'algorithme EM . . . . .	43
4.5	Comparaison des deux algorithmes . . . . .	47
5	Analyse bayésienne des processus stochastiques . . . . .	48
5.1	Méthode bayésienne . . . . .	49
5.2	Approche bayésienne pour un processus de Weibull . . . . .	50
5.3	Approche bayésienne pour un processus auto-excité . . . . .	54
<b>III Processus de Mino</b>		<b>57</b>
1	Introduction . . . . .	57
2	Définition . . . . .	57
3	Simulation des interarrivées du processus de Mino . . . . .	61
4	Loi de Mino . . . . .	63
4.1	Propriétés . . . . .	63
4.2	Fonction de répartition . . . . .	64
4.3	Taux de défaillance . . . . .	65
4.4	Fonction génératrice des moments . . . . .	66

---

4.5	Calcul de l'espérance mathématique . . . . .	70
4.6	Calcul de la variance . . . . .	74
5	Estimation statistique pour le Processus de Mino . . . . .	77
5.1	Maximum de vraisemblance . . . . .	80
5.2	Algorithme RSM . . . . .	82
5.3	Application de l'algorithme de Newton-Raphson . . . . .	85
5.4	Existence et unicité de l'EMV . . . . .	88
<b>IV Test d'hypothèses sur le Processus de Mino</b>		<b>91</b>
1	Introduction . . . . .	91
2	Tests de processus . . . . .	91
2.1	Processus de Poisson . . . . .	92
2.2	Processus de renouvellement . . . . .	94
3	Test pour le processus auto-excité . . . . .	95
3.1	Test paramétrique . . . . .	96
3.2	Test non paramétrique . . . . .	98
3.3	Test du processus de Mino . . . . .	100
<b>Conclusion générale et perspectives</b>		<b>103</b>
<b>Références bibliographiques</b>		<b>105</b>

## Table des matières

---



---

## Notations

$\{N(t), t \geq 0\}$	processus ponctuel
$T$	longueur de la fenêtre d'observation
$N(T)$	nombre d'événement dans la fenêtre $[0, T]$
$N(\cdot   y_i)$	famille dénombrable de processus ponctuels
$w_i$	date de la $i$ ème événement, $i = 1, \dots, N(T)$ , $w_0 = 0$ et $w_{N(T)+1} \equiv T$
$t_i$	temps qui s'écoule entre l'événement $i$ et l'événement $i + 1$ : $t_i = w_{i+1} - w_i$
$\lambda(t)$	intensité du processus
$h(t)$	fonction absolument intégrable
$H_t$	histoire des processus
$\theta$	vecteur des paramètres en lien avec l'intensité
$W$	ensemble de $N(T)$ occurrence $0 \leq w_1, w_2, \dots, w_{N(T)} \leq T$
$\delta(\tau)$	fonction de Dirac
$f_c(w)$	fonction de densité spectrale complète
$\mu(\tau)$	fonction de covariance
$\mu_c$	fonction de covariance complète
$g(u)$	fonction de réponse
$\mathcal{X}$	espace métrique complet séparé
$\mathcal{Y}$	espace métrique séparable complet
$X$	processus de cluster de Poisson
$Z_i$	les marques
$C_i$	clusters
$y_i$	centre de cluster
$N_c$	processus de centre de cluster sur $\mathcal{Y}$
$N_i$	les immigrés

---

$\mathcal{B}_x$	mesure Borélienne
$\hat{\lambda}(t, N(t))$	l'espérance de $\lambda(t)$ conditionnellement à $N(t)$
$\prod$	Produit intégrale
$P_n(t)$	probabilité que la population est de taille $n$
$\tilde{P}_n(s)$	transformation de Laplace de $P_n(t)$
$S_k(t)$	probabilité que la population est au plus de dimension $K$ à l'instant $t$
$\tilde{S}_K(s)$	transformation de Laplace de $S_k(t)$
$L(N(t) \theta)$	fonction de vraisemblance
$\log L(N(t) \theta)$	log-vraisemblance
$\log L_{dc}(N(t) \theta)$	log-vraisemblance pour les données complet
$N^1(t)$	processus de comptage observable
$N^0(t)$	processus de comptage inobservable
$\lambda^1(t)$	intensité du processus de comptage observable
$\lambda^0(t)$	intensité du processus de comptage inobservable
$\Lambda(t)$	taux de défaillance
$R(t)$	fiabilité à l'instant $t$
$H(\theta)$	matrice hessienne
$f(X \theta)$	distribution d'échantillonnage
$\pi(\theta)$	distribution a priori
$\pi(\theta X)$	distribution a posteriori
$l(\beta)$	fonction d'information
$\rho_k$	coefficient d'autocorrélation d'ordre $k$
$\phi_T(W^T)$	fonction du test
$\mathcal{K}_\varepsilon$	classe de la fonction du test
$\beta_T(u, \phi_T)$	fonction du puissance
$\phi_T^*(\cdot)$	plus puissant localement asymptotiquement uniformément dans la classe $\mathcal{K}_\varepsilon$
$I_h^*$	information de Fisher
$\Delta_T(W^T)$	limite de la fonction du puissance
$Z_\varepsilon$	$1 - \varepsilon$ quantile de la loi normale $N(0, 1)$
$\propto$	proportionnelle



## Acronymes

PAE	Processus auto-excité
PPNH	Processus ponctuel de Poisson non homogène
EMV	Estimateur du maximum de vraisemblance
i.i.d	Variables indépendantes et identiquement distribuées

---

# Introduction générale

En général, les processus stochastiques permettent de décrire l'évolution au cours du temps d'un phénomène statistique selon des lois de probabilités. Ces processus permettent de modéliser des systèmes dont le comportement n'est que partiellement prévisible.

Les processus ponctuels sont des processus stochastiques qui sont des modèles probabilistes pour les phénomènes d'arrivées que l'on observe au cours du temps : clients arrivants devant des guichets à la poste, requêtes reçues par un routeur d'internet, instants où se produisent les répliques à la suite d'un tremblement de terre, etc.

En 1971, Alan G. Hawkes [1] a introduit la notion de processus ponctuel auto-excité (self-exciting process). Ce processus ne dépend pas seulement de l'instant de réalisation de l'événement mais de l'ensemble des événements arrivant précédemment. Parfois ce processus dépend seulement d'un nombre précis d'événements - noté  $m$  - et on parle alors de processus auto-excité de mémoire  $m$ .

Comme le font remarquer Daley & Vere-jones [2], le processus d'Hawkes admet à la fois une représentation par processus de cluster et une représentation classique par intensité conditionnelle, intensité qui, de plus, est linéaire. Ceci autorise de nombreux développements théoriques. De plus, le processus d'Hawkes est très polyvalent. Il permet de modéliser de nombreuses si-

tuations et possède un champ d'application très large. Le processus auto-excité (PAE) est un processus qui dépend d'une intensité et qu'on peut appliquer dans des domaines variés :

Par exemple, en sismologie, le calcul du risque sismique dans une région donnée s'appuie sur des modèles stochastiques décrivant au mieux l'activité sismique. Il s'agit d'étudier les probabilités d'occurrences de tremblements de terre et d'analyser l'évolution spatio-temporelle de ces probabilités. En statistique sismique, l'outil principal de modélisation est la théorie des processus ponctuels qui permet, pour une date donnée, d'intégrer l'information jusqu'à cette date (exclue) dans l'expression du risque conditionnel. Les modèles de processus ponctuel auto-excité (PAE) permettent d'étudier les distributions spatiales et temporelles des événements sismiques. Les techniques d'inférence statistique sont basées sur la vraisemblance des observations. Afin d'analyser des données de plus en plus nombreuses, des méthodes statistiques appropriées sont requises [3–7].

Dans le domaine médicale et biologique, les PAE sont utilisés pour modéliser des données par exemple en épidémiologie [8] ou encore en génétique [9]. On rencontre dans ces domaines, une extension multidimensionnelle des PAE qui consiste à considérer un processus multivarié dont les composantes s'auto-excitent et/ou s'excitent mutuellement.

En neurophysiologie, les techniques actuellement les plus avancées permettent d'enregistrer simultanément l'activité de plusieurs neurones (par exemple chez le primate vigile). Cependant, ces enregistrements sont effectués en aveugle. Aucune technique ne permet à la fois d'enregistrer précisément l'activité des neurones individuels et de connaître le graphe précis des connexions qui lient des milliers de neurones sur le plan structural et fonctionnel. L'analyse statistique de ce type de données met en évidence des relations statistiques entre activités enregistrées, notamment en termes de synchronisation et de reconstruction de graphes fonctionnels. Les processus ponctuels auto-excité peuvent être utilisés pour modéliser ce type de données [10–

14].

En finance, les processus ponctuels sont appliqués principalement pour expliquer des problèmes relatifs à la microstructure des marchés financiers. Plus précisément, les processus ponctuels peuvent être utilisés pour modéliser la durée entre certains événements sur les marchés et pour modéliser par exemple, le taux de devis et de prix dans les différents marchés.

En fiabilité, le développement de l'informatique et son utilisation dans les domaines les plus variés ont rendu notre société de plus en plus dépendante du bon fonctionnement des systèmes logiciels. Leur fiabilité est donc devenue un enjeu majeur, rendu d'autant plus crucial que leur taille et leur complexité ne cessent de croître exponentiellement. La suite des instants de défaillance dans de tels systèmes peut être modélisée par un processus aléatoire ponctuel. Ainsi la fiabilité d'un système réparable s'exprime comme une probabilité de bon fonctionnement conditionnellement à une historique de vie du système. Les PAE sont donc des outils particulièrement adaptés pour modéliser la fiabilité des logiciels [15–18].

La simulation des processus auto-excités est également une question importante. Nous l'utiliserons pour étudier le processus qui est au centre de ce travail. Il s'agit d'un processus auto-excité de mémoire 1 particulier décrit par Hiroyuki Mino [19]. Nous mènerons l'étude de ce processus par les étapes suivantes :

- identifier la loi des interarrivées du processus, si elle existe, pour proposer des méthodes d'inférence alternatives,
- simuler les interarrivées associées, ce qui nous permettra une meilleure compréhension du modèle,
- estimer les paramètres de l'intensité,
- construire un test d'hypothèses pour détecter si un processus quelconque est auto-excité ou non.

Notre thèse porte donc sur l'inférence statistique pour le processus de Mino. Nous nous intéressons à l'estimation des paramètres d'un processus auto-excité de mémoire 1. Nous utiliserons la méthode du maximum de vraisemblance ce qui nous conduira à des problèmes d'optimisation où on nous utiliserons des algorithmes stochastiques. Nous traiterons également la questions des tests d'hypothèses.

Cette thèse est composée de quatre chapitres.

Le chapitre 1 offre une vue globale sur le processus auto-excité : travaux récents sur les processus auto-excités, propriétés, théorèmes, et autres notions dont nous aurons besoin dans la suite des travaux.

Le chapitre 2 est consacré au problème d'estimation des paramètres d'un processus auto-excité. Nous nous intéressons à l'écriture de la fonction de vraisemblance d'un PAE et nous décrivons les méthodes et les algorithmes qui permettent de résoudre les équations de vraisemblance. Nous illustrons l'usage d'un de ces algorithmes par une application sur des données groupées en survie.

Dans le troisième chapitre, nous considérons un processus auto-excité de mémoire 1, utilisé en traitement du signal que nous appelons processus de Mino. Nous étudions en détails son intensité. Ce travail nous conduit à introduire une loi de distribution spécifique que nous avons appelée loi de Mino. Cette loi se trouve être la loi des interarrivées du processus qui nous intéresse. Nous traitons alors la question de l'estimation des paramètres de l'intensité du processus de Mino par la méthode du maximum de vraisemblance, en utilisant des algorithmes stochastiques. La qualité des méthodes est évaluée sur des données simulées.

Dans le chapitre 4, nous nous intéressons aux tests d'hypothèses pour les processus et nous construisons un test d'hypothèse qui permet de détecter l'appartenance à la famille des processus auto-excités d'un processus ponctuel quelconque.

Enfin, nous concluons en décrivant quelques perspectives de recherche ouvertes par les travaux présentés dans notre thèse.



---

---

# Chapitre I

---

## Processus ponctuels auto-excités

Dans ce chapitre, nous nous intéressons au processus ponctuel auto-excité (PAE) introduit par Alan G. Hawkes en 1971. Nous présentons les propriétés et principaux théorèmes concernant cette famille de processus.

### 1 Introduction

Un processus ponctuel est un modèle mathématique décrivant un phénomène physique caractérisé par des événements très localisés répartis de façon aléatoire dans un espace. Chaque événement est représenté dans le modèle par un point qui identifie la position de l'événement [20]. Les processus ponctuels auto-excités (PAE) (self-exciting point processes (SEPP) dans la littérature anglo-saxonne) sont des processus de comptage, dont l'intensité à l'instant  $t$ , ne dépend pas seulement de  $t$  mais également du nombre d'événements  $N(t)$  survenu avant  $t$ . En cela, l'intensité  $\lambda(t)$  du processus devient un processus stochastique et on parlera du *processus intensité*. Plus précisément, l'intensité s'exprime à l'instant  $t$  comme une fonction de l'ensemble ou d'une partie de l'ensemble des dates et du nombre d'événements antérieurs à cette date. Ainsi, l'évolution du processus dépend de toute son histoire, de sa trajectoire.



Comme nous l'avons déjà mentionné, les PAE trouvent des applications dans de nombreux domaines : sismologie [5], neurophysiologie [11], génétique [9], épidémiologie [8], fiabilité [16] et économie [21].

Un processus ponctuel auto-excité peut être vu comme un processus de Poisson non homogène dont l'intensité est une fonction du processus lui-même. Mais nous verrons dans la suite qu'un PAE n'est pas nécessaire poissonnien comme par exemple le processus de renouvellement considéré dans la section 5.3.

Un exemple élémentaire de processus ponctuel auto-excité, est le processus de naissance [22], où l'occurrence de l'événement d'intérêt dépend du nombre d'événements qui a déjà eu lieu. Par définition, l'intensité d'un processus de naissance s'exprime directement en fonction du nombre d'événements qui ont déjà eu lieu ; c'est donc par essence un processus auto-excité.

Le processus ponctuel auto-excité (PAE) a été introduit en 1971 par Hawkes [1]. De nombreux travaux ont été réalisés sur ce processus [19, 23–28] et dans ce qui suit nous présentons les travaux les plus récents qui concernent la modélisation, l'estimation et la simulation d'un PAE.

Au niveau **modélisation**, les PAE sont de plus en plus utilisés pour modéliser des problèmes dans différents domaines. Dans le but de détecter les distances entre des événements génomiques le long d'une séquence d'ADN et du fait de la complexité du problème biologique, Reynaud-Bouret & Schath [29] en 2010 se sont appuyés sur un modèle de processus auto-excité.

De même, Molher et al. [30] en 2011 ont modélisés par des PAE, les occurrences de crimes en utilisant de données fournies par les départements de police de Los Angeles aux États-Unis. Dans le même esprit, Lewis et al. [31] ont modélisé la mortalité des civils en Irak, en utilisant des données fournis par le projet Iraq Body Count entre 2003 et 2007 en considérant quatre régions irakiennes : Karakh, Najaf, Mosul et Fallujah. Pour chaque région, les auteurs définissent des paramètres qui sont nécessaires pour mettre en oeuvre des modèles de type

processus ponctuels avec des taux non stationnaires. En 2013, Monk & Leib [14] ont utilisé des PAE pour modéliser l'activité neuronale. En 2014, Cho et al. [32] ont utilisé un modèle des PAE latent pour les réseaux spatio-temporelles. En finance, Grothe et al. [33] ont proposé des PAE pour décrire les caractéristiques typiques d'événements extrêmes multivariées observées dans des séries chronologiques.

Au niveau **estimation**, il est très souvent fait appel à l'algorithme EM pour déterminer les paramètres inconnus des modèles. Olson & Carley (2014)[34] utilisent cet algorithme pour estimer les paramètres d'un processus mutuellement excité en soulignant que ce modèle s'avère être réaliste pour des séquences d'événements continus décrivant le comportement social dans une population. Guo & Luk (2014) [35] proposent une solution pour réduire les temps de calcul des estimateurs d'un PAE dans le cas d'un grand nombre d'observations. Pour cela, ces auteurs ont utilisé une stratégie éliminant la dépendance au sein de données complexes réduisant ainsi le nombre d'opérations. Comme nous l'avons remarqué précédemment dans l'introduction, le processus d'Hawkes s'interprète comme un processus de cluster où les centres des clusters sont modélisés par une processus Poisson homogène. Wheatley et al. [36] considèrent le cas où le processus associé aux centres est un processus de renouvellement quelconque (non nécessairement poissonnien) et utilise un algorithme EM pour obtenir des estimateurs des paramètres de l'intensité.

La **simulation** des PAE est nécessaire pour évaluer les procédures d'estimation et de test statistique. Les résultats obtenus par les simulations permettent une meilleure compréhension des modèles. La littérature concernant la simulation de processus auto-excités est relativement abondante. Par exemple, Moller & Rasmussen (2005) [37] s'intéressent à la simulation de PAE de type processus de cluster (*cf.* 3) poissonnien marqué,  $X = \{(w_i, Z_i)\}$  avec  $w_i$  les dates d'événements,  $Z_i$  les marques et  $C_i$  les clusters pour  $i = 1, \dots, n$ ,  $n$  désignant le nombre d'observations. Ainsi, pour  $w_- \in ]-\infty, 0]$  et  $w_+ \in ]0, \infty[$ , Moller & Rasmussen [37] proposent

deux algorithmes de simulation :

**Algorithme 1 :** Il est constitué de deux étapes pour simuler les événements marqués  $(w_i, Z_i) \in X$  avec  $0 \leq w_i \leq w_+$  :

1. Simuler les immigrants sur  $[w_-, w_+[$ ,
2. Pour chaque immigrant  $w_i$ , on simule  $Z_i$  et  $(w_j, Z_j) \in C_i$  avec  $w_i < w_j \leq w_+$ .

En général, cet algorithme souffre d'un effet de bord.

Par contre l'algorithme suivant est sans l'effet de bord ([38]).

**Algorithme 2 :** On suppose que la fonction de répartition pour la longueur de cluster  $F$  est connue. Soit  $N_1$  un processus ponctuel des immigrants sur  $[0, w_+[$ ,  $N_2$  un processus ponctuel des immigrants pour  $w_i \leq 0$ , où  $\{(w_j, Z_j) \in C_i : w_j \in [0, \infty[ \} \neq \emptyset$ .

1. Simuler un processus de Poisson  $N_1$  avec une intensité  $\lambda_1(t)$  sur  $[0, w_+[$ .
2. Pour chaque  $w_i \in N_1$ , simuler  $Z_i$  et  $(w_j, Z_j) \in C_i$  avec  $w_i < w_j < w_+$ .
3. Simuler un processus de Poisson  $N_2$  avec une intensité  $\lambda_2(t) = (1 - F(-t))\lambda_1(t)$  sur  $] - \infty, 0[$ .
4. Pour chaque  $w_i \in N_2$  simuler  $Z_i$  et  $\{(w_j, Z_j) \in C_i : w_j \in [0, w_+[ \}$  conditionnellement à l'événement  $\{(w_j, Z_j) \in C_i : w_j \in [0, \infty[ \} \neq \emptyset$ .
5. La sortie est tous les points marqués à partir des étapes 1, 2 et 4.

Dans les étapes 1 et 2 de l'algorithme 2, on utilise l'algorithme 1 avec  $w_- = 0$ .

En 2006, dans le prolongement de ces travaux, les mêmes auteurs [39] utilisent l'algorithme 1 pour quantifier l'effet de bord et l'améliorent en proposant un algorithme plus rapide et plus pratique.

En 2013, Dassios & Zhao [40] introduisent un algorithme de simulation, également sans effet de bord, pour le PAE avec une intensité exponentielle, qui présente les avantages suivants :

- possibilité de simuler de manière exacte les interarrivées d'un PAE à partir de leur loi sans utiliser la méthode d'inversion,
- mise en oeuvre et généralisation possible dans des dimensions plus élevées,
- flexibilité pour simuler le PAE à partir de n'importe quel temps,
- ne nécessite pas la condition de stationnarité de l'intensité, et donc possibilité de simuler un PAE avec des intensités non bornées dans un intervalle de temps fini,
- facilement modifiable pour simuler des PAE avec des intensités stationnaires.

## 2 Analyse spectrale des processus ponctuels

L'analyse spectrale des processus ponctuels est bien connue comme un outil statistique dans le cas de l'estimation des fonctions de densité spectrale. À partir de ces analyses, il est possible d'accumuler des informations complémentaires sur les différents types des processus ponctuels aléatoires ou processus de Poisson.

Bartlett [41] a considéré un processus ponctuel  $N(t)$  tel que :

$$\frac{E(dN(t))}{dt} = \lambda(t), \quad \frac{E(dN(t)dN(t+\tau))}{(dt)^2} = \mu'(t, t+\tau),$$

où  $E$  désigne l'espérance mathématique,  $dN(t) = N(t+dt) - N(t)$ , et  $\tau > 0$ .

Pour un processus ponctuel stationnaire,  $\lambda(t)$  est constant, on note par  $\lambda(t) = \lambda$  et  $\mu'(t, t+\tau)$  est une fonction qui dépend seulement du  $\tau$ .

On définit la fonction de covariance par :

$$\mu(\tau) = \frac{E(dN(t)dN(t+\tau))}{(dt)^2} - \lambda^2, \quad \tau > 0. \quad (\text{I.1})$$

Pour  $\tau < 0$ ,  $\mu(\tau) = \mu(-\tau)$ , mais pour  $\tau = 0$  on a :

$$E([dN(t)]^2) = E(dN(t)).$$

## Chapitre I. Processus ponctuels auto-excités

---

Ainsi pour  $\tau \in \mathbb{R}$ , la fonction de covariance complète est donnée par :

$$\mu_c = \lambda\delta(\tau) + \mu(\tau), \quad (\text{I.2})$$

où  $\delta(\tau)$  est une fonction de Dirac et  $\mu(\tau)$  est continue pour  $\tau = 0$ .

La fonction de densité spectrale complète est définie par :

$$f_c(w) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp(-i\tau w) \mu_c(\tau) d\tau. \quad (\text{I.3})$$

Dans l'article fondateur du processus ponctuel auto-excité (processus d'Hawkes) [1], le processus d'Hawkes est présenté comme un développement du processus ponctuel  $\{N(t), t \geq 0\}$  tel que  $E(dN(t))/t$  soit égal à  $\lambda$ , une constante et tel que la fonction de covariance des accroissements infinitésimaux  $\mu(\tau) = E(dN(t+\tau)dN(t))/dt^2 - \lambda^2$  ne dépend pas de  $t$ .

On a la définition suivante :

**Définition I.1.** *Un processus d'Hawkes est un processus aléatoire ponctuel  $\{N(t); t \geq 0\}$  tel que :*

$$\begin{aligned} P(N(t+h) - N(t) = 1 \mid N(s), s \leq t) &= \lambda(t)h + o(h), \\ P(N(t+h) - N(t) > 1 \mid N(s), s \leq t) &= o(h), \end{aligned}$$

où  $\{\lambda(t), t > 0\}$ , l'intensité du processus, est définie par :

$$\lambda(t) = \mu + \int_{-\infty}^t g(t-u) dN(u),$$

avec  $\mu \in \mathbb{R}^+$ ,  $g(u) \geq 0$  pour tout  $u \geq 0$  et  $\int_0^{+\infty} g(u) du < 1$ .

**Remarques :**

- Le processus d'Hawkes peut s'interpréter comme un processus de Poisson non homogène.
- Si on note  $w_1, w_2, \dots$ , la suite des dates d'événement du processus, on a :

$$\lambda(t) = \mu + \sum_{\{w_i < t\}} g(t - w_i).$$

Lorsque l'événement d'intérêt se produit, l'intensité du processus se trouve modifiée par la fonction  $g$ . Cette fonction s'interprète en quelque sorte comme une réponse au saut du

## I.2 Analyse spectrale des processus ponctuels

---

processus. On l'appelle pour cette raison **fonction réponse** du processus. Son introduction dans l'expression de l'intensité permet d'élargir à un grand nombre de phénomènes aléatoires, la possibilité de modélisation par les processus ponctuels. La fonction  $g$  peut être croissante ou décroissante. Ce dernier cas apparaît en sismologie par exemple, où juste après une secousse, la probabilité de survenue d'une réplique est élevée pour décroître ensuite avec le temps.

Hawkes considère une décroissance exponentielle en s'intéressant à des fonctions  $g$  de la forme :

$$g(t) = \sum_{j=1}^k \alpha_j e^{-\beta_j t}, \quad t \in \mathbb{R}^+ \quad \text{avec} \quad \sum_{j=1}^k \alpha_j / \beta_j < 1. \quad (\text{I.4})$$

Dans ce cas particulier, une expression analytique de la fonction de covariance des accroissements infinitésimaux du processus, ainsi que de la densité spectrale  $f(\omega)$  peuvent être obtenues (*cf.* [1]) :

$$f(\omega) = \frac{\lambda}{2\pi} \left\{ 1 + \frac{\alpha\lambda(2\beta - \alpha)}{(\beta - \alpha)^2 + \omega^2} \right\}.$$

Si on note  $w_i$ , la date du  $i^{\text{ème}}$  saut du processus, l'intensité sera de la forme :

$$\lambda(t) = \mu + \sum_{\{w_i < t\}} \sum_{j=1}^k \alpha_j e^{-\beta_j(t-w_i)}.$$

Considérons  $w_1, w_2, \dots, w_n$ , dates successives d'événements, on a :

- pour  $t \leq w_1$ ,  $\lambda(t) = \mu$ .

Avant que le premier événement ne se produise, le processus se comporte comme un processus de Poisson homogène et le temps qui s'écoule avant le premier événement, suit donc une loi exponentielle de paramètre  $\mu$ .

- pour  $w_1 \leq t \leq w_2$ ,  $\lambda(t) = \mu + \alpha e^{-\beta(t-w_1)}$ .

Ainsi, après le premier événement, l'intensité n'est plus constante et va dépendre du temps, on a un processus de Poisson non homogène.

- D'une manière générale, on a pour  $w_{i-1} \leq t \leq w_i, i > 1$  :

$$\lambda(t) = \mu + \alpha \sum_{j=1}^{i-1} e^{-\beta(t-w_j)}.$$

L'intensité du processus s'exprime donc en fonction des dates d'occurrence successives des événements. Nous reviendrons dans le chapitre suivant sur cette fonction réponse en évoquant les travaux de Ozaki [24] qui propose des estimateurs par la méthode du maximum de vraisemblance des paramètres  $(\mu, \alpha, \beta)$ .

Une autre forme de fonction réponse de type polynôme de Laguerre est proposée par Ogata & Akaike [25] :

$$g(t) = \sum_{k=0}^K a_k t^k e^{-ct},$$

où ils considèrent dans leur article un processus auto-excité d'intensité  $\lambda(t | H_t)$  qui dépend des paramètres  $\theta \in \Theta$ , et  $H_t$  correspond à l'histoire des processus  $(X_s)$  et  $(N_s)$  avant le temps  $t$ , et ils s'intéressent au modèle linéaire qui est défini par :

$$\lambda(t | H_t) = \mu + \int_0^t g(t-s) dN_s + \int_0^t h(t-s) dX_s,$$

avec

$$g(t) = \sum_{k=0}^K a_k t^k e^{-ct}, \quad h(t) = \sum_{k=0}^K b_k t^k e^{-ct}.$$

Dans sa thèse, Lorenzen [42] s'intéresse à un processus d'Hawkes avec une fonction réponse de type Weibull définie par :

$$g(t) = \left(\frac{k}{w}\right) \left(\frac{t}{w}\right)^{k-1} \exp\left[-\left(\frac{t}{w}\right)^k\right].$$

### 3 Processus de cluster

Le problème de regroupement (clustering) est considéré, pour le cas où chaque point des données est un échantillon généré par un processus ergodique stationnaire *i.e* un processus

aléatoire dont tous les moments statistiques sont égaux aux moments temporels.

Le processus de cluster constitue l'un des modèles les plus importants des processus aléatoires. Il est largement utilisé pour l'étude des processus ponctuels. Ce processus apparaît naturellement pour modéliser la position de point dans le plan ou dans l'espace. Il s'applique dans de nombreux domaines. Les points peuvent être des plantes, des molécules, des protozoaires, des êtres humains, des étoiles, des galaxies ou encore des épices des séismes. On rencontre également le processus de cluster en physique pour modéliser les émissions photoélectriques par exemple, en volcanologie pour les fréquences d'éruptions volcaniques, en logistique, systèmes de files d'attente, failles dans les systèmes informatiques, en neurologie dans la modélisation du système nerveux. Le mécanisme de cluster est aussi un moyen naturel pour décrire la localisation des individus de générations consécutives du processus de branchement *cf.* ([2] P. 175 - 176).

Ce processus comporte deux aspects : les emplacements des clusters et les emplacements des éléments au sein d'un cluster. La superposition de ces deux aspects constitue le processus observé. Pour modéliser les éléments du cluster, nous spécifions une famille dénombrable de processus ponctuels  $N(\cdot | y_i)$  indexé par les centres de cluster  $\{y_i\}$ . Les emplacements des clusters sont modélisés par un processus donné  $N_c$ .

**Définition I.2.** – *On dit que le processus  $N$  à valeurs dans un espace d'états  $\mathcal{X}$  (espace métrique complet séparable), est un processus de clusters caractérisé par un processus des centres  $N_c$  à valeurs dans un espace d'états  $\mathcal{Y}$  (espace métrique complet séparable) et par une famille des processus ponctuels mesurables  $\{N(\cdot | y) : y \in \mathcal{Y}\}$ , si, pour tout ensemble borné  $A$  appartenant à  $\mathcal{B}_{\mathcal{X}}$ , la tribu borélienne, on a :*

$$N(A) = \int_{\mathcal{Y}} N(A | y) dN_c(y) = \sum_{y_i \in N_c(\cdot)} N(A | y_i) < \infty, p.s$$

Lorsque le processus des centres de cluster est un processus de Poisson homogène, on parle de processus de cluster poissonnien dont on peut donner la définition formelle suivante utilisant le vocabulaire des processus de branchement :



**Définition I.3.** *Un processus ponctuel  $X \subset \mathbb{R}$  caractérisé par la superposition de points particuliers appelés “immigrants” et de points particuliers appelés “descendants” est un processus de cluster poissonnien si et seulement si :*

1. *les immigrants  $w_i \in I \subset \mathbb{R}$  constituent un processus de Poisson homogène,*
2. *chaque immigrant  $w_i$  génère un ensemble de descendants constituant un processus  $Y$ , pour former un groupe fini de points contenant  $w_i$  appelé “cluster” et noté  $C_i$ ,*
3. *étant donnés les immigrants, les clusters centrés*

$$C_i - w_i = \{Y - w_i : Y \in C_i\}, \quad w_i \in I,$$

*sont i.i.d et ne dépendent pas de  $I$ .*

4.  *$X$  est l'union de tous les clusters.*

Le processus d'Hawkes est un cas particulier de processus de cluster poissonnien où dans le point 2 de la définition précédente, les descendants d'un immigrant génèrent également des points dans ce même cluster suivant un processus de Poisson d'intensité  $\lambda(\cdot - Y)$  (cette intensité est appelée : “taux de fertilité”).

D'après Hawkes & Oakes [23], tout processus auto-excités avec une intensité finie peuvent être représentés comme un processus de cluster poissonnien. Ainsi, tous les résultats (propriétés, théorèmes, etc.) dont on dispose concernant le processus de cluster poissonnien, s'appliquent au PAE.

## 4 Processus auto-excité (PAE)

Dans cette partie, nous étudions plus en détails le processus auto-excité. Nous donnons quelques propriétés sur la loi du nombre d'événements associé à un PAE et sur la loi des interarrivées. Cette présentation s'inspire largement du livre de Snyder et Miller [20].

## 4.1 Généralités

Nous adoptons ici la notation de Snyder et Miller qui désigne par  $w_1, w_2, \dots, w_n$  les v.a dates de saut d'un processus ponctuels auto-excité.

Soit  $\{N(t); t \geq t_0\}$  un processus de comptage vérifiant les conditions suivantes :

1.  $\{N(t); t \geq t_0\}$  est conditionnellement localement fortement continu (orderly en anglais).
2. Soit la fonction  $a(\Delta t, N(t))$  définie par :

$$a(\Delta t, N(t)) = \begin{cases} \frac{1}{\Delta t} Pr[N(t, t + \Delta t) = 1 \mid N(t)], & N(t) = 0, \\ \frac{1}{\Delta t} Pr[N(t, t + \Delta t) = 1 \mid N(t); w_1, w_2, \dots, w_{N(t)}], & N(t) \geq 1. \end{cases} \quad (I.5)$$

On suppose que la limite de  $a(\Delta t, N(t))$  quand  $\Delta t$  tend vers zéro, existe quelque soit  $w_1, w_2, \dots, w_{N(t)}$ .

3.  $Pr[N(t_0) = 0] = 1$ .

La condition 1. signifie grossièrement que la probabilité qu'il y ait plus de 2 sauts du processus dans un intervalle de longueur infinitésimale  $\Delta t$  est bien plus petite que la probabilité qu'il ait exactement 1 saut. Plus formellement, un processus de comptage est conditionnellement localement fortement continu si et seulement si pour tout  $t \geq t_0$ , pour tout  $\varepsilon$  un réel donné, il existe  $\delta \equiv \delta(t, \varepsilon) > 0$  tel que :

$$Pr[N(t, t + \delta') > 1] \leq \varepsilon Pr[N(t, t + \delta') = 1], \text{ pour tout } \delta' \in (0, \delta).$$

Pour les limites de la condition 2., on pose :

$$\mu(t, 0) = \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} Pr[N(t, t + \Delta t) = 1 \mid N(t)], \quad \text{pour } N(t) = 0, \quad (I.6)$$

et pour  $N(t) \geq 1$  :

$$\mu(t, N(t); w_1, w_2, \dots, w_{N(t)}) = \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} Pr[N(t, t + \Delta t) = 1 \mid N(t); w_1, w_2, \dots, w_{N(t)}]. \quad (I.7)$$

## Chapitre I. Processus ponctuels auto-excités

---

On peut alors définir le processus intensité  $\{\lambda(t); t \geq t_0\}$  d'un processus  $\{N(t); t \geq t_0\}$ , en considérant :

$$\lambda(t) = \begin{cases} \mu(t, 0), & t_0 \leq t \leq w_1, \\ \mu(t, N(t); w_1, w_2, \dots, w_{N(t)}), & w_{N(t)} < t \leq w_{N(t)+1}. \end{cases} \quad (\text{I.8})$$

On dit qu'un processus de comptage  $\{N(t), t \geq t_0\}$ , est auto-excité si et seulement si son intensité  $\{\lambda(t), t > t_0\}$  est un processus aléatoire dont les sauts dépendent des réalisations du processus  $N(t)$  lui-même [20].

Andersen et al. [43] page 73, ont proposé une définition du processus auto-excité en utilisant la notion de filtration. Supposons un fenêtrage d'observation  $[0, T]$ ,  $T > 0$ . Une histoire (également appelée filtration), notée  $(\mathcal{F}_t, t \in [0, T])$  est une famille croissante continue de sous  $\sigma$ -algèbre de  $\mathcal{F}$ . On fait l'hypothèse que cette famille est complète i.e. que pour tout  $t$ , la  $\sigma$ -algèbre  $\mathcal{F}_t$  contient tous les ensembles de mesures nulles de  $\mathcal{F}$ .

$\mathcal{F}_t$  satisfait les conditions suivantes :

- (i)  $\mathcal{F}_s \subseteq \mathcal{F}_t \subseteq \mathcal{F}$ ,  $\forall s < t$  (croissante)
- (ii)  $\mathcal{F}_s = \bigcap_{t>s} \mathcal{F}_t$ ,  $\forall s$  (continue à droite)
- (iii)  $A \subset B \in \mathcal{F}, Pr(B) = 0 \implies A \in \mathcal{F}_0$  (complète)

On peut interpréter l'histoire de la manière suivante :  $\mathcal{F}_t$  est la  $\sigma$ -algèbre qui contient tous les événements dont l'occurrence ou non est fixée à la date  $t$ . C'est la  $\sigma$ -algèbre générée par  $N(s)$ , ( $s \leq t$ ). Soit  $H_t$  un élément de  $\mathcal{F}_t$ .  $H_t$  représente les données disponibles à la date  $t$ . Avec cette notion de filtration, on peut proposer la définition suivante du processus auto-excité :

**Définition I.4.** – *On dit qu'un processus de comptage est auto-excité si sa filtration est générée par le processus lui-même.*

Remarquons que (I.6) et (I.7) permettent d'exprimer les probabilités qu'aucun saut du processus ne se produise dans un intervalle quelconque  $(s, t)$ ,  $(s < t)$  :

$$Pr[N(s, t) = 0 \mid N(s) = 0] = \exp \left[ - \int_s^t \mu(u, 0) du \right], \quad (\text{I.9})$$

$$Pr[N(s, t) = 0 \mid N(s); w_1, w_2, \dots, w_{N(s)}] = \exp \left[ - \int_s^t \mu(u, N(s); w_1, w_2, \dots, w_{N(s)}) du \right]. \quad (\text{I.10})$$

Ceci nous conduit à nous intéresser à la loi du nombre d'événements d'une manière générale.

## 4.2 Loi du nombre d'événements

On montre que la loi du nombre de sauts d'un PAE dans un intervalle de temps donné est identique à celle d'un processus de naissance dont l'intensité serait définie par :

$$\hat{\lambda}(t, N(t)) = E(\lambda(t) \mid N(t)), \quad (\text{I.11})$$

$\lambda(t)$  étant l'intensité du PAE. L'intensité (I.11), qu'on appellera *intensité conditionnelle au comptage*, s'interprète bien comme une fonction intensité d'après le lemme suivant :

**Lemme I.1.** – Soit  $\{N(t); t \geq t_0\}$  un processus ponctuel auto-excité d'intensité  $\{\lambda(t), t > t_0\}$ . Supposons que  $\{N(t); t \geq t_0\}$  est conditionnellement localement fortement continu (*orderly*) et que  $E(\lambda(t)) < \infty$  pour  $t \geq t_0$ .

Alors :

$$E(\lambda(t) \mid N(t)) = \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} Pr[N(t, t + \Delta t) \geq 1 \mid N(t)], \quad (\text{I.12})$$

et

$$E(\lambda(t) \mid N(t)) = \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} Pr[N(t, t + \Delta t) = 1 \mid N(t)]. \quad (\text{I.13})$$

Ce lemme permet d'établir les équations différentielles classiquement obtenues pour les processus de Markov dans le particulier de l'intensité conditionnelle au comptage et on a le théorème suivant :

**Théorème I.2.** – (Snyder & Miller [20]) – Soit  $\{N(t); t \geq t_0\}$  un processus ponctuel auto-excité conditionnellement localement fortement continu, d'intensité  $\{\lambda(t), t > t_0\}$ , tel que  $E(\lambda(t)) < \infty$  et  $Pr[N(t_0) = 0] = 1$ .

Alors, la probabilité  $Pr[N(t) = n]$  qu'il y ait  $n$  sauts du processus dans l'intervalle  $[t_0, t[$ , est solution du système des équations différentielles :

$$\frac{\partial Pr[N(t) = 0]}{\partial t} = -\hat{\lambda}(t, 0)Pr[N(t) = 0], \quad \text{pour } n = 0, \quad (\text{I.14})$$

$$\frac{\partial Pr[N(t) = n]}{\partial t} = -\hat{\lambda}(t, n)Pr[N(t) = n] + \hat{\lambda}(t, n-1)Pr[N(t) = n-1], \quad (\text{I.15})$$

pour  $n \geq 1$ , avec la condition initiale  $Pr[N(t_0) = 0] = 1$ ,  $\hat{\lambda}(t, N(t))$  étant l'espérance mathématique de  $\lambda(t)$  sachant  $N(t)$ .

La résolution des équations différentielles conduit aux expressions suivantes des solutions :

$$Pr[N(t) = 0] = \exp \left[ - \int_{t_0}^t \mu(u, 0) du \right], \quad (\text{I.16})$$

$$Pr[N(t) = n] = \int_{t_0}^t \hat{\lambda}(\tau, n-1) Pr[N(\tau) = n-1] \exp \left[ - \int_{\tau}^t \hat{\lambda}(u, n) du \right] d\tau. \quad (\text{I.17})$$

### 4.3 Théorème limite pour la somme d'un processus ponctuel auto-excité

La figure (I.1) représente la superposition de plusieurs processus ponctuels indépendants. En pratique, il est classique de n'observer que le processus résultant. On envisage ici le cas où les processus observés sont auto-excités. Nous rappelons dans cette section, un théorème qui vaut pour tout processus de comptage et qui est pour la superposition de PAE analogue au théorème central limite pour une somme de variables aléatoires.

On considère  $k$  processus de comptage  $\{N_{i,k}(t) : t \geq t_0\}$ ,  $i = 1, 2, \dots, k$ . On note  $\{M_k(t) : t \geq t_0\}$ , le processus résultant de la superposition de ces  $k$  processus.

$$M_k(t) = \sum_{i=1}^k N_{i,k}(t). \quad (\text{I.18})$$

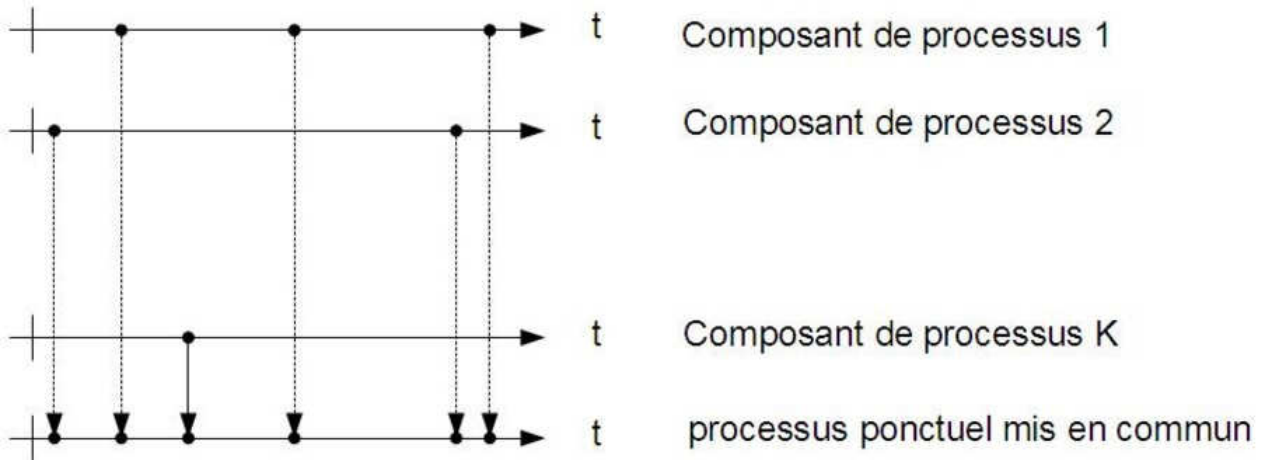


Figure I.1 – Processus ponctuels superposés

On dit que les composantes  $(N_{1,k}(t), \dots, N_{k,k}(t))$  du processus superposé sont *uniformément espacées* si pour tout  $t$  fini, on a :

$$\lim_{k \rightarrow \infty} \sup_{1 \leq i \leq k} Pr[N_{i,k}(t) \geq 1] = 0.$$

L'interprétation de cette notion est la suivante : la probabilité qu'une des composantes génère un saut dans le processus résultant tend vers 0 lorsque le nombre de composantes tend vers l'infini.

On a le théorème suivant :

**Théorème I.3.** (Snyder & Miller [20]) – Soit  $\{N_{i,k}(t), t \geq t_0\}$  pour  $i = 1, 2, \dots, k$ ,  $k$  processus de comptage mutuellement indépendants et uniformément distribués. Alors le processus  $\{M_k(t), t \geq t_0\}$ , résultant de la superposition de ces processus, converge en loi vers un processus de Poisson d'intensité  $\{\lambda(t), t \geq t_0\}$  si et seulement si :

$$\lim_{k \rightarrow \infty} \sum_{i=1}^k Pr[N_{i,k}(t) > 1] = 0 \text{ et } \lim_{k \rightarrow \infty} \sum_{i=1}^k Pr[N_{i,k}(t) = 1] = \int_{t_0}^t \lambda(u) du, \text{ pour } t_0 \leq t < \infty.$$

## 5 Processus ponctuel auto-excité de mémoire limitée

Nous considérons dans ce paragraphe le cas où l'évolution du PAE ne dépend pas de **tous** les sauts passés mais seulement d'une partie d'entre eux. Lorsque seul un nombre restreint de sauts influence l'évolution du processus, on parle alors de **processus auto-excité à mémoire limitée** [20]. Plus précisément, on définit les processus ponctuels auto-excités de mémoire  $m$  de la manière suivante :

**Définition I.5.** – *Un processus auto-excité est de mémoire 0, si pour tous  $N(t) > 1$ , son intensité  $\mu(t, N(t); w_1, w_2, \dots, w_{N(t)})$  ne dépend pas des instants de saut  $w_1, w_2, \dots, w_{N(t)}$  qui ont eu lieu avant  $t$ .*

*On dit qu'un processus ponctuel auto-excité est de mémoire  $m$ , si pour tout  $N(t) \geq 1$ , son intensité  $\mu(t, N(t); w_1, w_2, \dots, w_{N(t)})$  dépend uniquement de  $t$ , de  $N(t)$ , des  $m$  plus récents instants de saut  $w_{N(t)-m+1}, \dots, w_{N(t)-1}, w_{N(t)}$  et ne dépend pas des  $N(t) - m$  premiers sauts.*

### 5.1 Caractérisation d'un processus ponctuels $m$ -mémoire lorsque $m = 0$

Le processus ponctuel auto-excité de mémoire zéro est caractérisé dans le théorème suivant :

**Théorème I.4.** – *(Snyder & Miller [20])*

*Soit  $\{N(t); t \geq t_0\}$  un processus de comptage ponctuel auto-excité, satisfaisant les conditions du théorème I.2. Alors le processus ponctuel est de mémoire 0 si et seulement si  $\{N(t); t \geq t_0\}$ , est un processus de Markov.*

**Théorème I.5.** – *(Snyder & Miller [20])*

*Soit  $\{N(t) : t \geq t_0\}$  un processus de naissance de Markov avec un taux  $\{\mu(t, N(t)) : t \geq t_0\}$  si la somme*

$$\sum_{n=0}^{\infty} \left[ \sup_{t_0 \leq u < t} \mu(u, n) \right]^{-1} \quad (\text{I.19})$$

*est infinie, alors pour tous  $t \geq t_0$  on a :*

$$Pr[N(t) < \infty] = \sum_{n=0}^{\infty} Pr[N(t) = n] = 1, \quad (\text{I.20})$$

*et il n'y a pas d'explosion démographique avec probabilité 1.*

## 5.2 Caractérisation d'un processus ponctuels $m$ -mémoire lorsque $m > 0$

Le processus ponctuel auto-excité de mémoire strictement supérieure à zéro est caractérisé par le théorème suivant :

**Théorème I.6.** – (Snyder & Miller [20])

*Un processus ponctuel auto-excité est de mémoire  $m$  pour  $m = 1, 2, \dots$  si et seulement si les suites des dates de saut  $\{w_n, n = 1, 2, \dots\}$  forment une chaîne de Markov d'ordre  $m$ .*

## 5.3 Processus auto-excité et renouvellement

Snyder & Miller [20] ont établi le lien entre les PAE de mémoire 1 et les processus de renouvellement. Rappelons qu'un processus de renouvellement est un processus de comptage tel que les temps qui s'écoulent entre deux sauts forment des v.a. indépendantes et sont identiquement distribuées. Il est caractérisé par la loi de ces temps. Un processus de renouvellement classique est le processus de renouvellement où les interarrivées sont de loi exponentielle. Pour ce processus, si le paramètre de la loi exponentielle est  $\lambda$ , la loi des dates de saut est connue et la date du  $n$ ième saut est un loi Gamma de paramètre  $(n, \lambda)$ . On a également que la loi du nombre de sauts dans un intervalle de temps donné est une loi de Poisson de paramètre  $\lambda$  fois la longueur de cet intervalle.

Une définition formelle du processus de renouvellement est donnée ci-dessous :

**Définition I.6.** [22] – Soit  $\{w_i, i \in \mathbb{N}^*\}$  la suite des dates de saut d'un processus aléatoire telle que, pour tout  $i$ ,  $w_i = \sum_{j=1}^i t_j$  où les v.a  $t_i$  sont positives, indépendantes et identiquement distribuées de même loi  $F$ .

On dit que le processus de comptage  $\{N(t), t \geq 0\}$  défini par :

$$N(t) = \max\{i \mid w_i \leq t\} = \sum_{i \geq 1} \mathbb{1}_{[0,t]}(w_i), \quad \forall t \geq 0,$$

est un processus de renouvellement.

Ces processus permettent de modéliser et de résoudre les problèmes de renouvellement dans des systèmes complexes, dans le cadre par exemple, d'une politique de maintenance [22]. Les



## Chapitre I. Processus ponctuels auto-excités

---

processus de renouvellement permettent d'étudier une classe de phénomènes qui fait intervenir des instants de régénération (instants où un processus oublie son passé) [44].

On a le théorème suivant :

**Théorème I.7.** – (Snyder & Miller [20]) *Un processus ponctuel auto-excité est de mémoire 1, avec*

$$\mu(t, N(t); w_{N(t)}) = f(N(t), t - w_{N(t)}), \quad (\text{I.21})$$

*pour une fonction  $f(\cdot)$  régulière si et seulement si les interarrivées  $t_1 = w_1, t_2 = w_2 - w_1, \dots, t_n = w_n - w_{n-1}$ , sont des variables aléatoires indépendantes.*

Un processus ponctuel auto-excité dont les interarrivées sont i.i.d est appelé *processus de renouvellement ordinaire*. Si toutes les interarrivées sauf la première, sont identiquement distribuées, il est appelé un *processus de renouvellement modifié*.

On déduit du théorème I.7 la représentation en processus de renouvellement d'un processus auto-excité et on peut énoncer le corollaire suivant :

**Corollaire I.1.** – *Un processus ponctuel auto-excité est un processus de renouvellement modifié si et seulement si l'intensité associée  $\mu(\cdot)$  est la forme :*

$$\mu(t, N(t); w_1, \dots, w_{N(t)}) = h(t - w_{N(t)}), \quad (\text{I.22})$$

*où  $h(\cdot)$  est une fonction régulière.*

*C'est un processus de renouvellement ordinaire si et seulement si de plus,  $\mu(t, 0) = h(t)$ .*

Nous utiliserons par la suite ce corollaire pour caractériser le processus de Mino.

---

---

# Chapitre II

---

## Estimation pour les processus auto-excités

Dans ce chapitre, nous nous intéressons à l'écriture de la fonction de vraisemblance d'un processus ponctuel auto-excité ainsi qu'à l'estimation des paramètres de l'intensité de celui-ci par la méthode du maximum de vraisemblance. Nous décrivons les algorithmes permettant le calcul des estimateurs. Une approche bayésienne est également évoquée.

### 1 Introduction

Les premiers travaux concernant l'inférence sur des processus stochastique au sens large remonte à 1948, 1949 avec les travaux d'Abraham Wald qui s'intéresse à l'estimation par maximum de vraisemblance du paramètre d'un processus aléatoire discret. A la même période Kendall publie des travaux sur la modélisation de l'évolution des populations et s'intéresse en particulier de manière détaillée au processus de naissance qui comme nous l'avons déjà mentionné est un exemple élémentaire de processus auto-excité. [45].

## 2 Vraisemblance pour un processus stochastique

La méthode du maximum de vraisemblance [46, 47] consiste à déterminer les paramètres qui maximise la fonction de vraisemblance. Dans le cas d'un processus ponctuel [2, 20, 43] dépendant d'une intensité  $\lambda(t)$ , l'application de la méthode du maximum de vraisemblance pour ce processus a présenté une difficulté dans les années antérieures à 1970. Pour obtenir la log-vraisemblance d'un processus ponctuel qui dépend d'une intensité  $\lambda(t)$ , Andersen et al. [43] ont utilisé le **produit intégrale** noté  $\prod$ .

### 2.1 Produit intégrale

Le produit intégrale est une notion qui a été introduite à la fin du 19<sup>ème</sup> siècle par le mathématicien italien Victor Volterra, comme un outil pour la résolution d'une certaine classe d'équations différentielles [48, 49].

Dans un problème d'estimation paramétrique d'un processus ponctuel qui dépend d'une intensité  $\lambda(t)$ , on a :

$$P(W) \propto \prod_{t < T} [\lambda(t)dt]^{dN(t)} [1 - \lambda(t)dt]^{1-dN(t)},$$

où  $W$  est l'ensemble des  $N(T)$  occurrences.

D'après le développement de Taylor, on a :

$$1 - \lambda(t)dt \approx \exp(-\lambda(t)dt),$$

donc on peut écrire l'équation de vraisemblance :

$$L(N(t)|\theta) \propto \left( \prod_{t=0}^T \lambda(t)^{dN(t)} \right) \exp \left( - \int_0^T \lambda(t)dt \right),$$

et par suite

$$\log L(N(t)|\theta) = - \int_0^T \lambda(t)dt + \int_0^T \log \lambda(t)dN(t).$$

## II.2 Vraisemblance pour un processus stochastique

---

On donne la définition suivante :

**Définition II.1.** – La log-vraisemblance d'un processus ponctuel qui dépend d'une intensité  $\lambda(t)$  est définie par (voir [43] p.59 et [2] p.23) :

$$\log L(N(t)|\theta) = - \int_0^T \lambda(t)dt + \int_0^T \log \lambda(t)dN(t), \quad (\text{II.1})$$

où  $\theta$  représente les paramètres de l'intensité,  $N(T)$  le nombre d'événement dans la fenêtre  $[0, T]$ .

D'après Daley & Vere-jones [2], une relation entre la log-vraisemblance d'un processus ponctuel et la densité de Janossy (cf [2] chapitre 7) est établie. Le maximum de vraisemblance des événements  $w_1, \dots, w_n$  d'un processus ponctuel  $N$  sur un ensemble de Borel borné  $A \subseteq \mathbb{R}^d$  ( $d$  dénote ici la dimension) est définie par :

$$L_A(N(t)|w_1, \dots, w_n) = j_n(w_1, \dots, w_n|A).$$

où  $j_n$  désigne la densité de Janossy.

**Exemple II.1.** – On considère un processus de Poisson inhomogène qui appartient à  $A \subset \mathbb{R}^d$ . Le cardinal de  $A$  suit une loi de Poisson avec une espérance  $E(A)$  et  $w_1, \dots, w_n$  sont une suite i.i.d. La densité de Janossy s'écrit :

$$j_n(w_1, \dots, w_n|A) = \exp\left(- \int_A \lambda(t)dt\right) \prod_{i=1}^n \log \lambda(w_i) = L_A(w_1, \dots, w_n).$$

On en déduit :

$$\log L_A(N(t)|w_1, \dots, w_n) = - \int_A \lambda(t)dt + \sum_{i=1}^n \log \lambda(w_i).$$

## 2.2 Vraisemblance pour le PAE

Un processus auto-excité, est un processus ponctuel qui dépend d'une intensité. Ainsi, la log-vraisemblance d'un PAE est donnée par l'équation (II.1). Ozaki [24] a utilisé la méthode du

## Chapitre II. Estimation pour les processus auto-excités

---

maximum de vraisemblance pour estimer les paramètres d'un PAE qui dépend d'une intensité :

$$\lambda(t) = \mu + \int_{-\infty}^t g(t-u) dN(u),$$

avec une fonction de réponse :

$$g(t) = \alpha e^{-\beta t} \text{ avec } \alpha < \beta,$$

qui vérifie la condition :  $\int_0^{+\infty} g(t) dt < 1$ .

Considérons  $w_1, w_2, \dots, w_n$  les dates successives de  $n$  événements, on a :

$$\lambda(t) = \mu + \sum_{w_i < t} g(t-w_i) = \mu + \sum_{w_i < t} \alpha e^{-\beta(t-w_i)},$$

et on remarque que pour tout  $t \leq w_1$ ,  $\lambda(t) = \mu$ .

Avant que le premier événement ne se produise, le processus se comporte comme un processus de Poisson homogène et le temps qui s'écoule avant le premier événement, suit donc une loi exponentielle de paramètre  $\mu$ .

Pour  $w_1 \leq t \leq w_2$ ,  $\lambda(t) = \mu + \alpha e^{-\beta(t-w_1)}$ . Ainsi, après le premier événement, l'intensité n'est plus constante et dépend donc du temps. On a alors un processus de Poisson non homogène.

D'une manière générale, on a pour  $w_{i-1} \leq t \leq w_i$  :

$$\lambda(t) = \mu + \alpha \sum_{j=1}^{i-1} e^{-\beta(t-w_j)}.$$

L'intensité du processus s'exprime donc en fonction des dates d'occurrence des événements successives. On a bien un processus auto-excité.

En utilisant l'équation (II.1), la fonction de log-vraisemblance s'écrit :

$$\log L(N(t)|\theta) = -\mu w_n + \sum_{i=1}^n \frac{\alpha}{\beta} (e^{-\beta(w_n-w_i)} - 1) + \sum_{i=1}^n \log[\mu + \alpha B(i)]$$

où  $B(i) = \sum_{w_j < w_i} e^{-\beta(w_i-w_j)}$  pour  $i \geq 2$ .

## II.2 Vraisemblance pour un processus stochastique

---

Pour calculer les estimateurs de  $\theta$ , il est possible d'utiliser l'algorithme de Newton-Raphson.

Nous nous intéressons maintenant à l'existence et l'unicité de l'estimateur du maximum de vraisemblance.

Dans un problème d'estimation d'un vecteur des paramètres en utilisant la méthode du maximum de vraisemblance, on dit qu'un estimateur du maximum de vraisemblance existe et est unique, si la fonction de vraisemblance est continûment dérivable jusqu'à l'ordre 2, et si le déterminant de la matrice hessienne est négative.

Le théorème suivant (II.2) [50] donne une condition suffisante d'existence et d'unicité du maximum de vraisemblance.

**Théorème II.2.** – Soit  $L(\boldsymbol{\theta})$  une fonction de vraisemblance 2 fois continûment dérivable, où  $\boldsymbol{\theta}$  est l'ensemble des paramètres à estimer défini dans  $\Theta \subset \mathbb{R}^d$  ( $d$  désigne ici la dimension). On suppose que :

1.

$$\lim_{\boldsymbol{\theta} \rightarrow \partial\Theta} L(\boldsymbol{\theta}) = 0,$$

où  $\partial\Theta$  désigne la frontière du domaine  $\Theta$ .

2. et la matrice hessienne

$$H(\boldsymbol{\theta}) = \frac{\partial^2 L(\boldsymbol{\theta})}{\partial\theta_i \partial\theta_j},$$

est définie négative pour chaque point  $\theta \in \Theta$  pour lequel le vecteur gradient

$$\nabla L = \frac{\partial L(\boldsymbol{\theta})}{\partial\theta_i}$$

est nulle. Alors

- (a) l'estimateur du maximum de vraisemblance est unique,
- (b) la fonction de vraisemblance atteint :
  - aucun autre maximum dans  $\Theta$ ,
  - aucun minimum ou autre point stationnaire dans  $\Theta$ ,
  - sa borne inférieure est 0 sur la frontière de  $\partial\Theta$ .

### 3 Algorithmes pour le calcul des estimateurs

Nous utilisons des méthodes numériques pour le calcul des estimateurs qui n'admettent pas de solutions analytiques. Pour un processus ponctuel auto-excité qui dépend d'une intensité, l'estimateur du maximum de vraisemblance n'admet pas de solution explicite. Nous présentons trois algorithmes qui sont **randomized search methods (RSM)**, **Newton-Raphson** et l'algorithme **EM** (Expérance-Maximisation).

#### 3.1 Randomized search methods (RSM)

Un algorithme **RSM** [51] est une méthode d'optimisation qui consiste à tirer d'une manière aléatoire la variable d'intérêt d'une fonction pour approcher de son extremum.

#### 3.2 Algorithme Newton-Raphson

L'algorithme Newton-Raphson est attribué aux deux mathématiciens britanniques du XVIIe siècle, Isaac Newton et Joseph Raphson. On utilise cet algorithme dans les problèmes d'optimisation, dans les cas où on a un système d'équations qui n'admet pas de solutions explicites. Dans le cas où l'on souhaite résoudre l'équation  $f(\theta) = 0$  qui dépend d'un seul paramètre  $\theta$ , la méthode de Newton-Raphson [52] utilise le chemin itératif suivant :

$$\theta^{(p+1)} = \theta^{(p)} - \frac{f(\theta^{(p)})}{f'(\theta^{(p)})} \quad \text{pour } p \geq 0. \quad (\text{II.2})$$

où on fixe une valeur arbitraire  $\theta^{(0)}$  puis on suit le chemin itératif (II.2) qui converge en quelques itérations vers la solution de l'équation  $f(\theta) = 0$ .

Dans le cas multidimensionnel, [53, 54] l'algorithme suit le chemin itératif suivant :

$$\boldsymbol{\theta}^{(p+1)} = \boldsymbol{\theta}^{(p)} - [H(\boldsymbol{\theta}^{(p)})]^{-1}U(\boldsymbol{\theta}^{(p)}) \quad \text{pour } p \geq 0, \quad (\text{II.3})$$

où  $\boldsymbol{\theta}$  est le vecteur des paramètres à estimer,  $[H(\boldsymbol{\theta}^{(p)})]^{-1}$  est la matrice hessienne inverse de la fonction d'intérêt et  $U(\boldsymbol{\theta}^{(p)})$  est le vecteur des dérivées premières de cette fonction. Partant d'une valeur initiale  $\boldsymbol{\theta}^{(0)}$  on suit le chemin téraatif (II.3), tant que  $|\boldsymbol{\theta}^{(p+1)} - \boldsymbol{\theta}^{(p)}|$  est supérieur à une valeur  $\varepsilon$  que l'on se fixe très petite.

### 3.3 Algorithme EM

L'algorithme espérance-maximisation (Expectation-maximisation algorithm) abrégé EM a été introduit par Dempster, Laird et Rubin en 1977 [55]. C'est un algorithme qui permet de trouver le maximum de vraisemblance des paramètres de modèles probabilistes lorsque le modèle dépend de variables latentes non observables (données incomplètes). On utilise souvent l'algorithme EM pour la classification de données, l'apprentissage automatique, ou la vision artificielle. On peut également citer son utilisation en imagerie médicale dans le cadre de la reconstruction tomographique. L'algorithme d'espérance-maximisation comporte :

- une étape d'évaluation de l'espérance (E), où l'on calcule l'espérance de la vraisemblance en tenant compte des dernières variables observées,
- une étape de maximisation (M), où l'on estime le maximum de vraisemblance des paramètres en maximisant la vraisemblance trouvée à l'étape E.

On utilise ensuite les paramètres trouvés en M comme point de départ d'une nouvelle phase d'évaluation de l'espérance, et l'on itère ainsi.

L'application de l'algorithme EM est réalisée dans les situations où l'on dispose de données incomplètes et par conséquent l'estimation du maximum de vraisemblance est rendue difficile par l'absence d'une partie des données. Bien que l'algorithme EM est souvent appliquée avec succès, il peut dans certaines situations être lent à converger ou ne pas converger [56].

Snyder & Miller [20] ont proposé la description suivante de l'algorithme EM dans le cas de processus stochastique. L'algorithme espérance-maximisation (EM) est basé sur les concepts



d'un espace de données incomplètes  $N^1$ , un espace de données complètes  $N$ , et un chemin  $h : N \rightarrow N^1$ .

L'espace de données incomplètes est l'espace  $N^1$  dans lequel les données mesurées prennent leurs valeurs. L'espace de données complètes  $N$  est un espace hypothétique qui parvient à atteindre deux objectifs : 1) rendre les étapes espérances et maximisation de l'algorithme EM analytiquement flexible ; et 2) faire les calculs nécessaires à l'algorithme EM pour produire des résultats numérique des estimations. L'espace de données complètes pour un problème donné n'est pas unique, et l'algorithme EM qui en résulte peut être plus ou moins complexe en fonction du choix effectué. Souvent, un choix approprié est suggéré par une compréhension des mécanismes de déformation et du bruit relatif aux données mesurées sur l'espace de données incomplètes. Pour le chemin de  $N$  à  $N^1$ , l'espace de données complètes est plus grand que l'espace de données incomplètes c'est à dire que les données complètes doivent déterminer les données incomplètes. Il doit y avoir une fonction connue  $h(\cdot)$  dont le chemin va des données complètes aux données incomplètes. Il s'agit généralement d'une application avec de nombreux points de  $N$  qui donne le même point dans  $N^1$ . Cette fonction des données incomplètes est déterminée par  $h(N) = N^1$ , sur les valeurs que les données incomplètes  $N$  peuvent avoir.

Notons l'ensemble de paramètres qui doivent être estimés par  $\theta$ , et supposons que ce vecteur appartient dans un ensemble de valeurs de paramètres possibles *i.e*  $\Theta \subseteq \mathbb{R}^n$ . Une estimation de  $\theta$  est appelé recevable que si elle est dans  $\Theta$ . Les paramètres  $\theta$  peuvent être non aléatoires, ou ils peuvent être aléatoires. Nous allons utiliser la méthode du maximum de vraisemblance pour l'estimation des paramètres non aléatoires, et pour les paramètres aléatoires, nous supposons qu'il existe une densité de probabilité *a priori* connue  $\lambda(\theta)$ , et nous utilisons l'estimateur du maximum *a posteriori* (MAP).

Deux fonctions de log-vraisemblance sont importantes pour l'estimation des paramètres  $\theta$  à partir des données mesurées qui appartiennent à  $N^1$  lorsque l'algorithme EM est utilisé. La

### II.3 Algorithmes pour le calcul des estimateurs

---

log-vraisemblance des données incomplètes pour l'estimation de paramètres non aléatoires  $L_{id}(N^1(t)|\theta) = \log[f(N^1(t)|\theta)]$ , où  $f$  désigne la densité des données incomplètes en fonction des paramètres, c'est la log-vraisemblance que nous avons utilisé jusqu'à présent dans l'estimation des paramètres. Le maximum de vraisemblance qui estime  $\hat{\theta} \equiv \hat{\theta}(N^1(t))$  est une fonction de vecteur à valeur admissible de données incomplètes  $N^1(t)$  qui maximise la log-vraisemblance des données incomplètes

$$\hat{\theta} = \arg \max_{\theta \in \Theta} [L_{id}(N^1(t)|\theta)].$$

Si les paramètres sont aléatoires, alors la quantité  $\log f(\theta)$  doit être ajoutée à  $\log[f(N^1(t)|\theta)]$  dans la formation de la log-vraisemblance des données incomplètes.

La deuxième log-vraisemblance qui est aussi importante. C'est la log-vraisemblance de données complètes qui est définie par  $L_{cd}(N(t)|\theta) = \log[g(N(t)|\theta)]$ , où  $g$  désigne la densité de données complètes en fonction des paramètres.

L'algorithme EM est itératif. Il part d'une valeur initiale  $\hat{\theta}^{(0)}$ , où il génère une séquence d'estimateurs  $\hat{\theta}^{(1)}, \hat{\theta}^{(2)}, \dots, \hat{\theta}^{(k)}, \dots$  de façons que la séquence correspondante à log-vraisemblance de données incomplets est croissante  $L_{id}(N^1(t)|\hat{\theta}^{(0)}) \leq L_{id}(N^1(t)|\hat{\theta}^{(1)}) \leq L_{id}(N^1(t)|\hat{\theta}^{(2)}) \leq \dots \leq L_{id}(N^1(t)|\hat{\theta}^{(k)}) \leq \dots$ . Deux étapes sont nécessaires à chaque étape de l'itération pour atteindre l'étape suivante, une espérance (E-étape) et maximisations (M-étape).

**E-étape.** Détermination de l'espérance conditionnelle de la log-vraisemblance des données complètes,

$$Q(\theta|\hat{\theta}^{(k)}) = E[L_{cd}(N(t)|\theta)|N^1(t), \hat{\theta}^{(k)}], \quad (\text{II.4})$$

et compte tenu des données incomplètes on suppose que  $\theta = \hat{\theta}^{(k)}$ .

**M-étape.** Détermination du paramètre d'ordre  $k + 1$  qui maximise  $Q(\theta|\hat{\theta}^{(k)})$ ,

$$\hat{\theta}^{(k+1)} = \arg \max_{\theta \in \Theta} [Q(\theta|\hat{\theta}^{(k)})]. \quad (\text{II.5})$$

Il est établi dans le théorème suivant que l'application répétée des étapes E et M produit une suite croissante de log de vraisemblance de données incomplètes.

**Théorème II.3.** – (Snyder & Miller [20])

Soit la suite  $\{\hat{\theta}^{(k)} : k : 1, 2, \dots\}$  définie à partir des étapes E et M (II.4) et (II.5). Alors, la suite correspondante des log-vraisemblances des données incomplètes  $\{L_{id}(N^1(t)|\hat{\theta}^{(k)}) : k = 1, 2, \dots\}$  est croissante.

## 4 Une illustration de l'algorithme EM

Dans cette partie, nous présentons une application de l'algorithme EM pour estimer les paramètres d'une loi exponentielle tronquée à partir de l'observation de données groupées. On compare les résultats obtenus avec ceux obtenus en appliquant un algorithme d'optimisation classique.

On considère une suite de variables aléatoires  $X_1, \dots, X_n$  i.i.d de loi exponentielle de paramètre  $\theta$ . On rappelle l'expression de la densité de la loi de  $X_i$

$$f(x_i) = \frac{1}{\theta} e^{-x_i/\theta} \mathbf{1}_{\mathbb{R}^+}(x_i).$$

On note  $I_j = [t_{j-1}, t_j], j = 1, \dots, M + 1$  un découpage de  $\mathbb{R}^+$  tel que  $t_0 = 0, t_{M+1} = +\infty$ . Pour  $j = 1, \dots, M + 1$ , on note  $K_j$  la v.a. nombre de durées se réalisant dans l'intervalle  $I_j$ .  $K_j = \sum_{i=1}^n \mathbf{1}_{I_j}(x_i)$  et  $\sum_{j=1}^{M+1} k_j = n$ . Calculons la loi du vecteur  $(K_1, \dots, K_M; K_{M+1})$  où  $K_{M+1} = n - \sum_{j=1}^M K_j$ . Nous notons par  $\mathcal{B}(n, p)$  la loi binomiale de paramètres  $n$  et  $p$ .

On note :  $p_j = P(X_i \leq t_j)$ , pour  $j = 1, \dots, M$ .

On a :  $K_1 \sim \mathcal{B}(n, p_1)$ . On peut alors calculer la loi de  $K_2 | K_1 = k_1$ , ; c'est une loi binomiale de paramètres  $(n - k_1, P(X_i \leq t_2 | X_i > t_1))$ .

Or

$$P(X_i \leq t_2 | X_i > t_1) = \frac{P(X_i \leq t_2) - P(X_i \leq t_1)}{P(X_i > t_1)} = \frac{p_2 - p_1}{1 - p_1}.$$

## II.4 Une illustration de l'algorithme EM

D'une manière générale, la loi de  $K_j \mid K_{j-1} = k_{j-1}, \dots, K_1 = k_1$  est une loi binomiale de paramètres  $(n - \sum_{\ell=1}^{j-1} k_\ell, P(X_i \leq t_j \mid X_i > t_{j-1}))$  où  $P(X_i \leq t_j \mid X_i > t_{j-1}) = \frac{p_j - p_{j-1}}{1 - p_{j-1}}$ , pour  $j = 1, \dots, M$ .

On peut donc exprimer la loi  $Q$  du vecteur  $(K_1, \dots, K_M; K_{M+1})$  en effectuant le produit de ces lois conditionnelles. Il vient :

$$\begin{aligned}
 Q &= P(K_1 = k_1) \prod_{i=2}^M P(K_i = k_i \mid K_{i-1} = k_{i-1}, \dots, K_1 = k_1) \\
 &= C_n^{k_1} p_1^{k_1} (1 - p_1)^{n-k_1} \times C_{n-k_1}^{k_2} \left( \frac{p_2 - p_1}{1 - p_1} \right)^{k_2} \left( \frac{1 - p_2}{1 - p_1} \right)^{n-k_1-k_2} \\
 &\quad \vdots \\
 &\times C_{n-\sum_{j=1}^{M-1} k_j}^{k_M} \left( \frac{p_M - p_{M-1}}{1 - p_{M-1}} \right)^{k_M} \left( \frac{1 - p_M}{1 - p_{M-1}} \right)^{n-\sum_{j=1}^{M-1} k_j} \\
 &= \frac{n!}{k_1!(n - k_1)!} p_1^{k_1} (1 - p_1)^{n-k_1} \\
 &\times \frac{(n - k_1)!}{k_2!(n - k_1 - k_2)!} \left( \frac{p_2 - p_1}{1 - p_1} \right)^{k_2} \left( \frac{1 - p_2}{1 - p_1} \right)^{n-k_1-k_2} \\
 &\quad \vdots \\
 &\times \frac{(n - \sum_{j=1}^{M-1} k_j)!}{k_M!(n - \sum_{j=1}^M k_j)!} \left( \frac{p_M - p_{M-1}}{1 - p_{M-1}} \right)^{k_M} \left( \frac{1 - p_M}{1 - p_{M-1}} \right)^{n-\sum_{j=1}^M k_j}
 \end{aligned}$$

En effectuant les simplifications, on obtient l'expression suivante :

$$Q = \frac{n!}{k_1! \dots k_M! (n - \sum_{j=1}^M k_j)!} p_1^{k_1} (p_2 - p_1)^{k_2} \dots (p_M - p_{M-1})^{k_M} \times (1 - p_M)^{n - \sum_{j=1}^M k_j}.$$

Autrement dit,  $(K_1, \dots, K_M; K_{M+1})$  suit une loi multinomiale de paramètre :

$$(n, p_1, p_2 - p_1, \dots, p_M - p_{M-1}, 1 - p_M).$$

On notera :

$$(K_1, \dots, K_M; K_{M+1}) \sim \mathcal{M}(n, p_1, p_2 - p_1, \dots, p_M - p_{M-1}, 1 - p_M).$$

## 4.1 Vraisemblance du nombre des durées

Puisque la variable aléatoire  $X_i$  est distribuée suivant une loi exponentielle de paramètre  $\theta$ , on a :

$$P(X_i \leq t_j) = 1 - e^{-\frac{t_j}{\theta}} = p_j.$$

La vraisemblance s'écrit :

$$\begin{aligned} L(\theta) &= P(K_1 = k_1, \dots, K_M = k_M) \\ &= \frac{n!}{k_1! \dots k_M! (n - \sum_{j=1}^M k_j)!} \prod_{j=1}^M \left( (e^{-\frac{t_{j-1}}{\theta}} - e^{-\frac{t_j}{\theta}})^{k_j} e^{-\frac{t_M}{\theta} (n - \sum_{j=1}^M k_j)} \right), \end{aligned}$$

et la log-vraisemblance est

$$\log L(\theta) = \sum_{j=1}^M k_j \log(e^{-\frac{t_{j-1}}{\theta}} - e^{-\frac{t_j}{\theta}}) - \frac{t_M}{\theta} (n - \sum_{j=1}^M k_j) + C,$$

où

$$C = \log \left( \frac{n!}{k_1! \dots k_M! (n - \sum_{j=1}^M k_j)!} \right).$$

## 4.2 Estimation de $\theta$ en utilisant l'algorithme de Newton-Raphson

Pour estimer  $\theta$  on dérive par rapport à  $\theta$  la log-vraisemblance  $\log L(\theta)$  :

$$\begin{aligned} \frac{\partial \log L(\theta)}{\partial \theta} &= \sum_{j=1}^M k_j \left[ \frac{t_{j-1} e^{-\frac{t_{j-1}}{\theta}} - t_j e^{-\frac{t_j}{\theta}}}{\theta^2 (e^{-\frac{t_{j-1}}{\theta}} - e^{-\frac{t_j}{\theta}})} \right] + \frac{t_M}{\theta^2} (n - \sum_{j=1}^M k_j) \\ &= \frac{k_j}{\theta^2} \sum_{j=1}^M k_j \left[ \frac{t_{j-1} - t_j e^{-\frac{\Delta_j}{\theta}}}{(1 - e^{-\frac{\Delta_j}{\theta}})} \right] + \frac{t_M}{\theta^2} (n - \sum_{j=1}^M k_j), \end{aligned}$$

où  $\Delta_j = t_j - t_{j-1}$ .

On cherche maintenant à résoudre l'équation  $\frac{\partial \log L(\theta)}{\partial \theta} = 0$ . On obtient :

$$\sum_{j=1}^M k_j \left[ \frac{t_{j-1} - t_j e^{-\frac{\Delta_j}{\theta}}}{(1 - e^{-\frac{\Delta_j}{\theta}})} \right] + t_M (n - \sum_{j=1}^M k_j) = 0.$$

On dénote par :

$$\phi(\theta) = \sum_{j=1}^M k_j \left[ \frac{t_{j-1} - t_j e^{-\frac{\Delta_j}{\theta}}}{1 - e^{-\frac{\Delta_j}{\theta}}} \right] + t_M (n - \sum_{j=1}^M k_j),$$

et on cherche à résoudre l'équation

$$\phi(\theta) = 0.$$

Comme cette équation n'admet pas une solution explicite, on utilise l'algorithme de Newton-Raphson II.2 :

$$\theta^{(p+1)} = \theta^{(p)} - \frac{\phi(\theta^{(p)})}{\phi'(\theta^{(p)})}.$$

Nous calculons dans un premier temps  $\phi'(\theta)$  et on obtient :

$$\phi'(\theta) = \sum_{j=1}^M \frac{k_j}{\theta^2} \left[ \frac{t_{j-1}^2 - t_j^2 e^{-\frac{\Delta_j}{\theta}}}{1 - e^{-\frac{\Delta_j}{\theta}}} - \frac{(t_{j-1} - t_j e^{-\frac{\Delta_j}{\theta}})^2}{(1 - e^{-\frac{\Delta_j}{\theta}})^2} \right],$$

après on initialise  $\theta^{(0)}$  et on applique l'algorithme de Newton-Raphson.

En appliquant cet algorithme sur des données simulées on obtient les résultats suivants :

Tableau 1 : Résultats numériques pour  $M = 7$  et différentes valeurs de  $n$  et  $\theta$  sur 1000 échantillons : SD désigne l'écart type (Standard Deviation en anglais) et MSE est l'erreur quadratique moyenne (Mean Squared Error en anglais).

$\theta$	$n$	Moyenne de $\hat{\theta}$	biais	SD	MSE
10	50	9.951	0.049	1.470	2.164
	200	10.029	0.029	0.747	0.559
	500	9.989	0.011	0.480	0.231
100	50	100.546	0.546	14.492	210.333
	200	103.036	3.036	7.084	50.278
	500	99.132	0.867	4.336	18.819
200	50	197.983	2.017	29.121	852.139
	200	201.049	1.049	14.755	218.826
	500	200.474	0.474	9.471	89.935

## Chapitre II. Estimation pour les processus auto-excités

---

Dans ce tableau, on trouve des bons estimateurs pour les différentes tailles d'échantillon. On remarque cependant que si la taille  $n$  augmente la variance diminue.

Maintenant on veut étudier l'effet du nombre de découpage sur la qualité des estimateurs (voir Tableau 2).

Tableau 2 : Résultats numériques pour  $\theta = 100$  et différentes valeurs de  $n$  et de nombre de découpage sur 1000 échantillons SD désigne l'écart type (Standard Deviation en anglais) et MSE est l'erreur quadratique moyenne (Mean Squared Error en anglais).

$\theta$	$n$	Nbr découpage	Moyenne de $\hat{\theta}$	<i>biais</i>	SD	MSE
100	50	3	100.564	0.564	15.141	229.73
		6	99.524	0.475	14.449	209.019
		12	100.502	0.502	14.651	214.92
		30	100.594	0.594	14.335	205.873
100	200	3	100.257	0.257	7.541	56.926
		6	100.143	0.142	7.439	55.351
		12	99.872	0.127	7.069	49.997
		30	99.853	0.146	6.803	46.303
100	500	3	100.215	0.215	4.654	21.711
		6	99.873	0.127	4.547	20.699
		12	99.931	0.069	4.688	21.988
		30	99.921	0.078	4.759	22.663

D'après les résultats du tableau 2, on trouve que le nombre des découpages ne joue pas un rôle significatif sur la qualité d'estimateur, mais on voit que le biais devient plus petit dans le cas  $\theta = 100$  et  $n = 500$ .

### 4.3 Vraisemblance pour le modèle complet

Le modèle complet suit une loi exponentielle de paramètre  $\theta$  :

$$f(x) = \frac{1}{\theta} e^{-\frac{x}{\theta}}, \quad x \geq 0.$$

En notant  $\underline{x} = (x_1, \dots, x_n)$ , on a :

$$\begin{aligned} L(\underline{x}, \theta) &= \frac{1}{\theta^n} e^{-\frac{1}{\theta} \sum_{i=1}^n x_i}, \\ \log L(\underline{x}, \theta) &= -n \log \theta - \frac{1}{\theta} \sum_{i=1}^n x_i, \\ \frac{\partial \log L(\underline{x}, \theta)}{\partial \theta} &= \frac{-n}{\theta} + \frac{1}{\theta^2} \sum_{i=1}^n x_i. \end{aligned} \tag{II.6}$$

Pour obtenir l'estimateur du maximum de vraisemblance de  $\theta$ , on résout l'équation :

$$\frac{-n}{\theta} + \frac{1}{\theta^2} \sum_{i=1}^n x_i = 0.$$

On obtient :

$$\frac{1}{\theta^2} \sum_{i=1}^n x_i = \frac{n}{\theta}.$$

et par conséquent :

$$\hat{\theta} = \bar{X}.$$

#### 4.4 Estimation de $\theta$ en utilisant l'algorithme EM

Pour appliquer l'algorithme EM, on cherche tout d'abord à déterminer la loi de  $X_i/X_i \in I_j$ , c'est à dire la probabilité conditionnelle  $P(X_i \leq x/t_{j-1} \leq X_i \leq t_j)$  où  $X_i \sim \exp(\theta)$ . On a :

$$\begin{aligned} P(X_i \leq x/t_{j-1} \leq X_i \leq t_j) &= \frac{P(t_{j-1} \leq X_i \leq x)}{P(t_{j-1} \leq X_i \leq t_j)} \\ &= \frac{F(x) - F(t_{j-1})}{F(t_j) - F(t_{j-1})}, \text{ pour } X_i \in I_j, \end{aligned}$$

et

$$\begin{aligned} f(x/t_{j-1} \leq X_i \leq t_j) &= \frac{f(x)}{F(t_j) - F(t_{j-1})} \\ &= \frac{e^{-\frac{x}{\theta}}}{\theta(e^{-\frac{t_{j-1}}{\theta}} - e^{-\frac{t_j}{\theta}})}, \text{ pour } X_i \in I_j, \end{aligned}$$



et comme cette expression correspond à la loi de densité d'une loi exponentielle tronquée alors  $X_i$  sachant  $X_i \in I_j$  suit une loi exponentielle tronquée de paramètre  $\theta$ .

L'espérance mathématique de  $X_i/X_i \in I_j$  s'écrit :

$$\begin{aligned} E(X_i/t_{j-1} \leq X_i \leq t_j) &= \int_{t_{j-1}}^{t_j} \frac{x_i e^{-\frac{x_i}{\theta}}}{\theta(e^{-\frac{t_{j-1}}{\theta}} - e^{-\frac{t_j}{\theta}})} dx_i \\ &= \frac{k_j}{\theta(e^{-\frac{t_{j-1}}{\theta}} - e^{-\frac{t_j}{\theta}})} \int_{t_{j-1}}^{t_j} x_i e^{-\frac{x_i}{\theta}} dx_i \\ &= k_j \left[ \theta + \frac{t_{j-1} - t_j e^{-\frac{\Delta_j}{\theta}}}{1 - e^{-\frac{\Delta_j}{\theta}}} \right], \end{aligned}$$

où  $k_j$  désigne le nombre des durées qui arrivant dans l'intervalle  $I_j$ .

Dans le cas où on étudie un modèle appartenant à la famille exponentielle [57, 58], on calculera l'espérance des expressions obtenues conditionnellement au vecteur des données réellement observées. L'estimateur du maximum de vraisemblance de  $\theta$  sera alors une fonction de l'espérance conditionnelle des données complètes (non observées) sachant les données incomplètes (inobservées).

En pratique, cet algorithme se déroulera de façon itérative. Chaque itération est constituée de deux étapes. A l'itération  $p$ , partant de la valeur  $\theta^p$  du paramètre :

- étape E : calcul de l'espérance conditionnelle  $E(X_i/t_{j-1} \leq X_i \leq t_j, \theta^p)$ .
- étape M : calcul de la nouvelle valeur du paramètre  $\theta^{p+1}$  en remplaçant les statistiques exhaustives  $(x_i)$  par l'espérance conditionnelle  $E(X_i/t_{j-1} \leq X_i \leq t_j, \theta^p)$  dans l'estimateur de la loi complet.

Notre modèle appartenant à la famille exponentielle, l'algorithme EM consiste simplement à remplacer les données complètes par l'espérance conditionnelle i.e l'espérance relativement à la loi du complet sachant l'incomplet, où  $x_i/k_j$  suit une loi exponentielle tronquée de paramètre

$\theta$ . On en déduit :

$$E[\log L(x_i; \theta)/k_j] = \sum_{j=1}^{M+1} \log L(x_i; \theta) f(x_i/k_j). \quad (\text{II.7})$$

D'après 4.3, nous avons :

$$\hat{\theta} = \frac{1}{n} \sum_{i=1}^n x_i.$$

Par conséquent, en remplaçant  $\hat{\theta}$  par  $\theta$  dans (II.7), nous avons

$$\begin{aligned} E[\log L(x_i; \theta)/k_j] &= \sum_{j=1}^{M+1} \frac{1}{n} \sum_{i=1}^n x_i f(x_i/k_j) \\ &= \sum_{j=1}^{M+1} \frac{1}{n} E(x_i/k_j) \\ &= \frac{1}{n} \sum_{j=1}^{M+1} k_j \left[ \theta + \frac{t_{j-1} - t_j e^{-\frac{\Delta_j}{\theta}}}{1 - e^{-\frac{\Delta_j}{\theta}}} \right] \\ &= \theta + \frac{1}{n} \sum_{j=1}^{M+1} k_j \frac{t_{j-1} - t_j e^{-\frac{\Delta_j}{\theta}}}{1 - e^{-\frac{\Delta_j}{\theta}}}, \end{aligned}$$

après on initial  $\theta^{(0)}$  puis on suit ce chemin itératif :

$$\begin{aligned} \theta^{(1)} &= \theta^{(0)} + \frac{1}{n} \sum_{j=1}^{M+1} k_j \frac{t_{j-1} e^{-\frac{t_{j-1}}{\theta^{(0)}}} - t_j e^{-\frac{t_j}{\theta^{(0)}}}}{e^{-\frac{t_{j-1}}{\theta^{(0)}}} - e^{-\frac{t_j}{\theta^{(0)}}}} \\ &\vdots \\ \theta^{(n+1)} &= \theta^{(n)} + \frac{1}{n} \sum_{j=1}^{M+1} k_j \frac{t_{j-1} e^{-\frac{t_{j-1}}{\theta^{(n)}}} - t_j e^{-\frac{t_j}{\theta^{(n)}}}}{e^{-\frac{t_{j-1}}{\theta^{(n)}}} - e^{-\frac{t_j}{\theta^{(n)}}}}. \end{aligned}$$

En appliquant cet algorithme sur des données simulées, on obtient dans le tableau suivant les résultats :

Tableau 3 : Résultats numériques pour différentes valeurs de  $N$  et  $\theta$  sur 1000 échantillons : SD désigne l'écart type (Standard Deviation en anglais) et MSE est l'erreur quadratique moyenne (Mean Squared Error en anglais).

## Chapitre II. Estimation pour les processus auto-excités

---

$\theta$	$N$	moyenne de $\hat{\theta}$	biais	SD	MSE
10	50	10.047	0.047	1.484	2.205
	200	9.975	0.024	0.743	0.553
	500	9.982	0.017	0.463	0.214
100	50	100.849	0.849	15.465	239.898
	200	100.207	0.207	7.715	59.574
	500	99.844	0.155	4.681	21.943
200	50	200.921	0.921	29.097	847.488
	200	200.438	0.438	14.261	203.546
	500	200.312	0.312	9.201	84.746

Dans ce tableau on obtient des bons estimateurs quelque soit la taille de l'échantillon, mais on remarque que si  $N$  augmente la variance diminue.

Maintenant on va savoir si le nombre de découpage joue un rôle sur la qualité des estimateurs c.f (Tableau 4).

Tableau 4 : Résultats numériques pour  $\theta = 100$  et différentes valeurs de  $N$  et de nombre de découpage sur 1000 échantillons : SD désigne l'écart type (Standard Deviation en anglais) et MSE est l'erreur quadratique moyenne (Mean Squared Error en anglais).

$\theta$	$N$	Nbr découpage	moyenne de $\hat{\theta}$	biais	SD	MSE
100	50	3	99.314	0.685	15.239	232.712
		6	100.385	0.385	14.612	213.664
		12	99.683	0.316	15.272	233.363
		30	100.599	0.599	14.848	220.835
100	200	3	100.121	0.121	7.681	59.005
		6	99.661	0.338	7.319	53.695
		12	100.111	0.111	7.485	56.044
		30	100.257	0.257	7.195	51.837
100	500	3	100.069	0.069	4.767	22.737
		6	100.117	0.177	4.889	23.923
		12	100.136	0.136	4.642	21.574
		30	99.921	0.079	4.551	20.722

## II.4 Une illustration de l'algorithme EM

---

D'après ces résultats, le nombre des découpages ne semble pas jouer un rôle significatif sur la qualité des estimateurs. En augmentant la taille d'échantillon, on obtient dans le tableau 5 les résultats :

Tableau 5 : Résultats numériques pour différentes valeurs de  $N$  et  $\theta$  sur 1000 échantillons : SD désigne l'écart type (Standard Deviation en anglais) et MSE est l'erreur quadratique moyenne (Mean Squared Error en anglais).

$\theta$	$N$	moyenne de $\hat{\theta}$	biais	SD	MSE
100	5000	100.087	0.087	1.401	1.972
	10000	100.016	0.016	1.035	1.072
	20000	99.996	0.003	0.755	0.571
200	5000	200.034	0.034	2.863	8.201
	10000	200.032	0.032	2.060	4.247
	20000	200.016	0.016	1.459	2.129
500	5000	500.139	0.139	7.214	52.064
	10000	499.958	0.041	4.946	24.471
	20000	499.949	0.051	3.586	12.864

On remarque que quand la taille de l'échantillon augmente la variance diminue, et on obtient toujours des bons estimateurs.

### 4.5 Comparaison des deux algorithmes

Pour comparer les deux algorithmes EM et Newton-Raphson, on prend un exemple avec  $\theta = 100$  et  $N = 200$ , où on regarde plus en détails les résultats des ces deux algorithmes ci-dessous.

$t_j$	0	50	100	150	200	250	300
$k_j$	100	31	22	14	10	11	

Dans cet exemple, on trouve que  $M = 6$  et on a  $k_{M+1} = N - \sum_{j=1}^M k_j = 200 - 188 = 12$ . En appliquant les deux algorithmes vus précédemment, on obtient :

Tableau 6 : Comparaison des algorithmes EM et Newton-Raphson.

	NR	EM
$i$	$\theta^{(i)}$	$\theta^{(i)}$
0	89	89
1	98.44793	97.65111
2	98.42428	98.36234
3	98.42428	98.41932
4	98.42428	98.42388
5	98.42428	98.42425
6	98.42428	98.42427
7	98.42428	98.42428
8	98.42428	98.42428
9	98.42428	98.42428
10	98.42428	98.42428

Dans ce tableau, l'indice  $i$  correspond au nombre d'itérations. En conclusion, les deux algorithmes convergent et les résultats sont identiques.

## 5 Analyse bayésienne des processus stochastiques

La statistique bayésienne est une approche statistique fondée sur l'inférence bayésienne, qui est une méthode permettant de déduire la probabilité d'un événement à partir de celles d'autres événements déjà évalués. L'analyse bayésienne suppose de choisir une loi de probabilité *a priori* pour les paramètres du modèle, d'écrire la distribution des données conditionnelle aux paramètres (*i.e* l'équivalent de la fonction de vraisemblance) et enfin de calculer la distribution *a posteriori* des paramètres à l'aide théorème de Bayes [59]. Aujourd'hui l'approche bayésienne est beaucoup étudiée autant en théorie que dans les applications [60].

Le statisticien bayésien raisonne différemment que le statisticien classique puisqu'il considère que le paramètre du modèle statistique,  $f(X|\theta)$  est incertain. Il va donc chercher à quantifier son incertitude en mobilisant toutes les informations disponibles. C'est ce qui fait toute la différence puisque cela revient à conférer au paramètre  $\theta$  le statut de variable aléatoire, dès lors qu'il y a un sens à lui attribuer une distribution de probabilité *a priori* qui décrit le savoir actuel sur ce paramètre. Cette distribution de probabilité, souvent appelée prior, est notée  $\pi(\theta)$ . Il faut bien comprendre que le prior quantifie l'état de connaissance d'un expert (et donc son incertitude) sur le problème en main. Cela signifie que l'expert parie plus volontiers sur certaines valeurs de  $\theta$  que sur d'autres (ne pas confondre incertitude et ignorance). Cette information a d'autant plus de valeur que les données sont rares. Il doit être clair que le savoir de l'expert encodé dans le prior doit être tout à fait indépendant de l'échantillon en main, sinon la même source d'information interviendrait deux fois, ce qui ne serait pas cohérent [61].

**Définition II.2.** – *Un modèle statistique bayésien [62] est constitué d'un modèle statistique paramétrique,  $f(X|\theta)$ , et d'une distribution a priori pour les paramètres,  $\pi(\theta)$ .*

Après avoir choisi la distribution *a priori*  $\pi(\theta)$ , on peut calculer la distribution *a posteriori* qui est définie par :

$$\pi(\theta|X) = \frac{f(X|\theta)\pi(\theta)}{\int f(X|\theta)\pi(\theta)d\theta}.$$

### 5.1 Méthode bayésienne

La méthode bayésienne comprend brièvement les étapes principales suivantes [60] :

1. **Loi de l'observation** : Obtenir la fonction de vraisemblance à partir de la distribution d'échantillonnage  $f(X|\theta)$ . Cette étape décrit simplement le processus à l'origine des données  $X$  en fonction des paramètres inconnus  $\theta$ .
2. **Loi a priori** : Choisir la distribution *a priori*  $\pi(\theta)$ . La distribution *a priori* représente les informations *a priori* qu'on sait sur les paramètres  $\theta$  avant l'observation des données.

3. **Loi *a posteriori*** : On obtient la distribution *a posteriori* en utilisant le théorème de Bayes.
4. **Inférence** : On cherche à estimer les paramètres  $\theta$  en utilisant la distribution *a posteriori*.

### 5.2 Approche bayésienne pour un processus de Weibull

Un processus de Weibull est un processus de Poisson d'intensité :

$$\lambda(t) = \frac{\beta}{\alpha} \left( \frac{t}{\alpha} \right)^{\beta-1},$$

où les paramètres  $\alpha$  et  $\beta$  sont strictement positifs. Pour une intensité d'un processus de Weibull,  $\alpha$  désigne un paramètre d'échelle et  $\beta$  est paramètre de forme :

- Si  $\beta < 1$ ,  $\lambda(t)$  est croissant,
- si  $\beta = 1$ ,  $\lambda(t)$  est constant,
- si,  $\beta > 1$ ,  $\lambda(t)$  est décroissant.

La figure II.1 donne des représentations de l'intensité d'un processus de Weibull pour différents valeurs de  $\alpha$  et  $\beta$ .

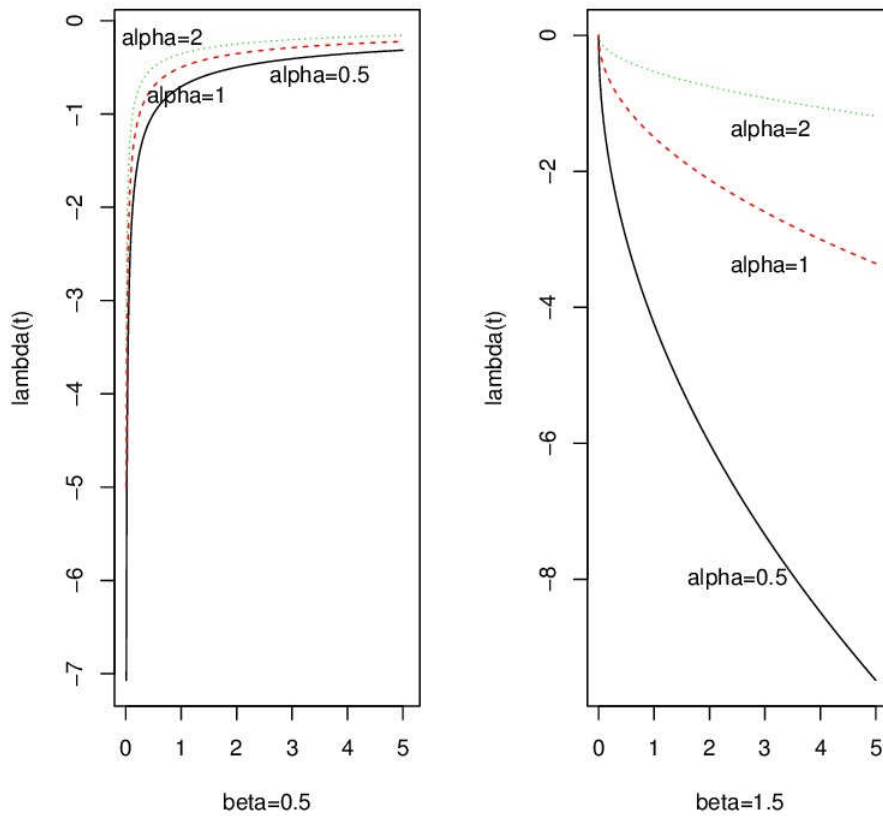
L'inférence bayésienne pour ce processus a été proposée par Guida et al. [63].

Soit  $w_1, w_2, \dots, w_n$  les  $n$  événements d'un processus de Weibull, pour obtenir la densité de la loi conditionnelle de  $w_i$  sachant  $w_{i-1}, \dots, w_1$ , nous déterminons tout d'abord la loi de  $P(w_i > t | w_{i-1})$  puis on obtient :

$$f(w_i | w_{i-1}, \dots, w_1) = \frac{\beta}{\alpha} \left( \frac{w_i}{\alpha} \right)^{\beta-1} \exp \left\{ - \left( \frac{w_i}{\alpha} \right)^\beta + \left( \frac{w_{i-1}}{\alpha} \right)^\beta \right\}.$$

Nous remarquons que la loi conditionnelle de  $w_i$  sachant  $w_{i-1}, \dots, w_1$  suit une loi de Weibull tronquée.

**Figure II.1** – Représentation de l'intensité d'un processus de Weibull pour différentes valeurs de  $\alpha$  et  $\beta$



La fonction de vraisemblance s'écrit :

$$L(N(t)|\beta, \alpha) = \beta^n \alpha^{-n\beta} u^{\beta-1} \exp \left\{ - \left( \frac{w_n}{\alpha} \right)^\beta \right\},$$

où  $u = \prod_{i=1}^n w_i$ .

Après des calculs élémentaires, nous obtenons les estimateurs du maximum de vraisemblance

$\hat{\alpha}, \hat{\beta}$  de  $\alpha$  et  $\beta$  donnés par :

$$\hat{\beta} = \left( \log w_n - \frac{1}{n} \sum_{i=1}^n \log w_i \right)^{-1},$$



et

$$\hat{\alpha} = w_n \exp \left\{ -\frac{1}{\hat{\beta}} \log n \right\}.$$

On considère la loi *a priori* :

$$\pi(\beta, \alpha) \propto (\beta\alpha)^{-1}.$$

Ainsi la loi jointe *a posteriori* est :

$$\pi(\beta, \alpha | w_i) \propto \beta^{n-1} u^\beta \alpha^{-n\beta-1} \exp \left\{ -\left(\frac{w_n}{\alpha}\right)^\beta \right\} \quad \text{pour } \alpha, \beta > 0,$$

et la loi marginale *a posteriori* de  $\beta$  est donnée par

$$\pi(\beta | w_i) \propto \beta^{n-2} \left(\frac{u}{w_n}\right)^\beta,$$

qui a pour

- mode :  $\beta_m = (n - 2)/z$ ,
- espérance mathématique :  $E(\beta | w_i) = (n - 1)/z$ , où  $z = \sum_{i=1}^n \log(w_n/w_i)$ .

De plus comme  $2z\beta$  suit une loi de  $\chi^2$  avec un degré de liberté égale à  $2(n - 1)$ , nous déduisons que  $\beta_m$  est un estimateur sans biais de  $\beta$ .

La loi marginale *a posteriori* de  $\alpha$  est :

$$\pi(\alpha | w_i) \propto \int_0^\infty \beta^{n-1} u^\beta \alpha^{-n\beta-1} \exp \left\{ -\left(\frac{w_n}{\alpha}\right)^\beta \right\} d\beta, \quad \alpha > 0.$$

Son espérance mathématique est donnée par :

$$E(\alpha | w_i) = \frac{t_n z^{n-1}}{\Gamma(n)\Gamma(n-1)} \int_0^\infty \beta^{n-2} \left(\frac{u}{w_n}\right)^\beta \Gamma(n-1/\beta) d\beta.$$

On considère maintenant une loi *a priori* pour  $\beta$  uniformément distribuée sur l'intervalle  $[\beta_1, \beta_2]$ , et une loi non informative pour  $\alpha$ . Les lois *a priori* respectivement pour  $\beta$  et  $\alpha$  sont :

$$\pi(\beta) = \frac{1}{\beta_1 - \beta_2} \quad \beta_1 \leq \beta \leq \beta_2,$$

et

$$\pi(\alpha) = \alpha^{-1} \quad \alpha > 0.$$

La loi jointe *a posteriori* est :

$$\pi(\beta, \alpha | w_i) \propto \beta^n u^\beta \alpha^{-n\beta-1} \exp \left\{ - \left( \frac{w_n}{\alpha} \right)^\beta \right\} \quad \beta_1 \leq \beta \leq \beta_2, \quad \alpha > 0.$$

La loi marginale *a posteriori* de  $\beta$  est :

$$\pi(\beta | w_i) \propto \beta^{n-1} \left( \frac{u}{w_n} \right)^\beta \quad \beta_1 \leq \beta \leq \beta_2.$$

Le mode de cette loi est donnée par :

$$\beta_m = \begin{cases} \beta_1 & \text{si } (n-1)/z \leq \beta_1, \\ \beta_2 & \text{si } (n-1)/z \geq \beta_2, \\ (n-1)/z & \text{sinon.} \end{cases}$$

L'espérance mathématique de cette loi est :

$$E(\beta | w_i) = I(\beta_1, \beta_2; n) / I(\beta_1, \beta_2; n_1),$$

où

$$I(\beta_1, \beta_2, n) = \int_{\beta_1}^{\beta_2} \beta^n \left( \frac{u}{w_n} \right)^\beta d\beta.$$

La loi marginale *a posteriori* de  $\alpha$  est :

$$\pi(\alpha | w_i) \propto \int_{\beta_1}^{\beta_2} \beta^n u^\beta \alpha^{-n\beta-1} \exp \left\{ - \left( \frac{w_n}{\alpha} \right)^\beta \right\} d\beta,$$

et

$$E(\alpha | w_i) = \frac{w_n}{\Gamma(n)} I_1(\beta_1, \beta_2; n-1) / I(\beta_1, \beta_2; n-1),$$

où

$$I_1(\beta_1, \beta_2; n - 1) = \int_{\beta_1} \beta_2 \beta^{n-1} \left( \frac{u}{w_n^n} \right)^\beta \Gamma(n - 1/\beta) d\beta.$$

L'inférence et la prédiction du processus de Weibull est aussi considérée dans les travaux de [64–66] par une approche Bayésienne. L'approche bayésienne a été appliquée sur des données de séismes en Italie [67] et à Taiwan [68]. Ces données ont été modélisées par des processus ponctuels et la loi *a priori* est une loi gamma.

### 5.3 Approche bayésienne pour un processus auto-excité

Un processus auto-excité est un processus qui dépend d'une intensité. Afin d'appliquer l'approche bayésienne sur les PAE, on considère l'intensité du processus pour trouver la densité du processus considéré. On peut considérer cette densité comme étant une distribution d'échantillonnage  $f(X|\theta)$ . En choisissant une distribution *a priori* sur  $\theta$ ,  $\pi(\theta)$ , et en utilisant les travaux de Robert [62], il s'avère que le point le plus critique et le plus critiquable de l'analyse bayésienne est le choix de la loi *a priori*. Une fois cette loi *a priori* connue, l'inférence peut être conduite en minimisant le coût *a posteriori*, en calculant les régions de plus forte densité *a posteriori* ou en intégrant les paramètres pour obtenir la distribution prédictive.

Ainsi, une fois la loi *a priori* définie, on peut déterminer la distribution jointe, la distribution marginale, la distribution *a posteriori* et la distribution prédictive. Les paramètres sont ensuite estimés.

#### Exemple II.4. –

*Insua et al. [69] supposent qu'un nouveau pneumatique de vélo suit un processus de Poisson homogène de paramètres  $\lambda$ , puisqu'il est supposé ne pas se détériorer pendant un intervalle de temps  $[0, T]$ . Nous supposons de plus que les perforations se produisent de façon aléatoire. Alors, à chaque événement,  $w_i, i = 1, \dots, n$ , lorsque une crevaison se produit dans l'intervalle  $[0, T]$ , il est réparé mais celui-ci peut être plus enclins à de nouvelles crevaisons du fait de l'usure. On ajoute alors  $\mu_i$  au paramètre du processus de Poisson homogène  $\lambda$ . On obtient un exemple très*

## II.5 Analyse bayésienne des processus stochastiques

---

simple d'un processus auto-excité, où sa fonction de vraisemblance est donnée par :

$$\prod_{i=1}^n \left( \lambda + \sum_{j=1}^{i-1} \mu_j \right) e^{-\lambda T - \sum_{i=1}^n \mu_i (T - w_i)}.$$

On considère la loi gamma comme une loi a priori pour  $\lambda$  et  $\mu_i$ , et la loi a posteriori correspond à une loi conditionnelle par rapport à tous les paramètres et elle appartient à la famille de la loi gamma.



---

---

# Chapitre III

---

## Processus de Mino

Ce chapitre a pour objet l'étude d'un processus auto-excité, dont l'intensité dépend d'une fonction réponse particulière. On s'intéresse au problème de l'estimation des paramètres. La méthode du maximum de vraisemblance est utilisée et la qualité des estimations est étudiée sur des données simulées.

### 1 Introduction

En 2001, Hiroyuki Mino [19] a considéré une intensité d'un processus ponctuel auto-excité qui dépend seulement du dernier événement, *i.e* un processus auto-excité de mémoire 1. Il a estimé les paramètres de cette intensité en utilisant l'algorithme EM par la méthode du maximum de vraisemblance.

### 2 Définition

Le processus de Mino est le processus ponctuel auto-excité (PAE)  $\{N(t), t \geq 0\}$  de mémoire 1 [19] dont l'intensité est définie par :

$$\lambda(t) = \mu h(t - w_{N(t)}), \quad (\text{III.1})$$

où  $h(t)$  est la fonction :

$$h(t) = 1 + \alpha e^{-\beta t}, \quad t \geq 0. \quad (\text{III.2})$$

On a donc :

$$\lambda(t) = \mu \left( 1 + \alpha e^{-\beta(t-w_{N(t)})} \right), \quad \text{avec } \mu > 0, \alpha \geq -1 \text{ et } \beta \geq 0. \quad (\text{III.3})$$

Supposons que l'on observe les sauts du processus dans une fenêtre de longueur  $T$ .

Soit  $w_{N(t)}$  la date du  $N(t)$ <sup>ème</sup> saut.

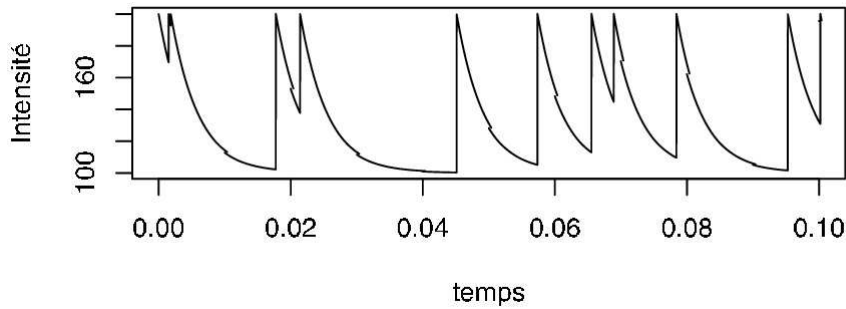
Ainsi, on a :

$$\begin{aligned} &\text{pour } t \in [0, w_1], \lambda(t) = \mu(1 + \alpha e^{-\beta t}), \\ &\text{pour } t \in ]w_1, w_2], \lambda(t) = \mu(1 + \alpha e^{-\beta(t-w_1)}), \\ &\quad \vdots \\ &\text{pour } t \in ]w_{i-1}, w_i], \lambda(t) = \mu(1 + \alpha e^{-\beta(t-w_{i-1})}), \\ &\quad \vdots \\ &\text{et pour } t \in ]w_{N(T)}, T], \lambda(t) = \mu(1 + \alpha e^{-\beta(t-w_{N(T)})}). \end{aligned}$$

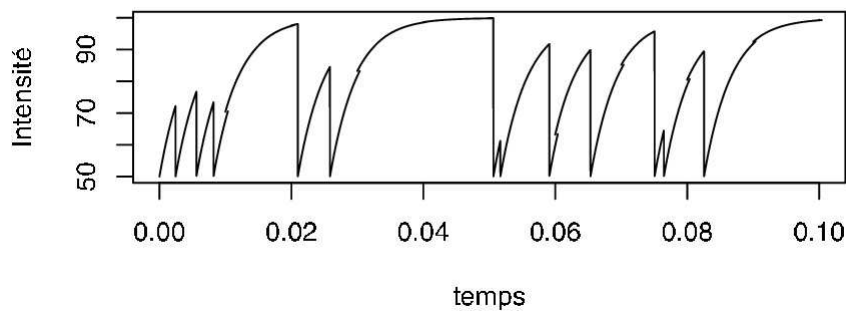
Remarques :

- Si  $\alpha = 0$ ,  $\lambda(t) = \mu$  et on a un processus de Poisson homogène.
- Si  $\alpha > 0$ ,  $\lambda(t)$  est croissant et on dit que le processus est **excité**.
- Si  $-1 \leq \alpha < 0$ ,  $\lambda(t)$  est décroissant et on dit que le processus est **inhibé**.

Les figures III.1 et III.2 donnent des représentations de l'intensité pour différentes valeurs de  $\mu, \alpha, \beta$ . Ces représentations sont obtenues à partir de simulations de  $w_1, \dots, w_{N(T)}$ , dates de sauts d'occurrence du processus.



**Figure III.1** – Représentation de l'intensité d'un processus excité :  $\alpha = 1$ ,  $\mu = 100$ ,  $\beta = 250$  et  $T = 0.1$  ms



**Figure III.2** – Représentation de l'intensité d'un processus inhibé :  $\alpha = -0.5$ ,  $\mu = 100$ ,  $\beta = 250$  et  $T = 0.1$  ms

**Proposition III.1.** – Soit un PAE de mémoire 1, d'intensité (III.3) observée du processus dans une fenêtre  $[0, C]$ . Alors les interarrivées  $T_i$ ,  $i = 1, \dots, N(C)$  sont des variables aléatoires indépendantes identiquement distribuées de loi :

$$Pr(T_i > t) = \exp\left\{-\mu\left[t + \frac{\alpha}{\beta}(1 - e^{-\beta t})\right]\right\}. \quad (\text{III.4})$$

**Preuve :** D'après le théorème 6.3.4, p. 314, de Snyder & Miller [20] on a :

$$\lambda(t) = f(N(t), t - w_{N(t)}). \quad (\text{III.5})$$



La probabilité de survie des interarrivées peut s'écrire :

$$\begin{aligned} Pr[t_n \geq T_n | t_1 = T_1, \dots, t_{n-1} = T_{n-1}] \\ = P_{w_n | w_1, \dots, w_{n-1}}(T_1 + \dots + T_n | T_1, T + T_2, \dots, T_1 + T_2 + \dots + T_{n-1}). \end{aligned}$$

Comme on a un processus de mémoire 1, d'après le théorème 6.3.3, p. 313 de Snyder & Miller [20], on obtient :

$$\begin{aligned} Pr[t_n \geq T_n | t_1 = T_1, \dots, t_{n-1} = T_{n-1}] &= P_{w_n | w_{n-1}}(T_1 + \dots + T_n | T_1 + \dots + T_{n-1}) \\ &= \exp \left[ - \int_{T_1 + \dots + T_{n-1}}^{T_1 + \dots + T_n} \lambda(\mu, n - 1; T_1 + \dots + T_{n-1}) d\mu \right]. \end{aligned}$$

D'après (III.5) et le théorème 6.3.3, p. 313, de Snyder & Miller [20], on obtient :

$$Pr[t_n \geq T_n | t_1 = T_1, \dots, t_{n-1} = T_{n-1}] = \exp \left[ - \int_0^{T_n} f(n - 1, \lambda) d\lambda \right],$$

et donc

$$Pr[t_n \geq T_n | t_1 = T_1, \dots, t_{n-1} = T_{n-1}] = Pr[t_n \geq T_n].$$

On en conclut que les interarrivées sont indépendantes.

On a :

$$Pr(T_i > t) = P(N[w_{i-1}, w_{i-1} + t] = 0).$$

Par définition, on a :

$$\begin{aligned} P(N(w_{i-1}, w_{i-1} + t) = 0) &= \exp \left\{ - \int_{w_i}^{w_i + t} \lambda(u) du \right\} \\ &= \exp \left\{ - \int_{w_{i-1}}^{w_{i-1} + t} \mu(1 + \alpha e^{-\beta(u - w_{i-1})}) du \right\} \\ &= \exp \left\{ -\mu \left[ t + \frac{\alpha}{\beta} (1 - e^{-\beta t}) \right] \right\}. \end{aligned}$$

**Proposition III.2.** – *Un processus de Mino est un processus auto-excité de mémoire 1 dont la densité de la loi de probabilité des durées inter-arrivées est :*

$$f(t) = \mu(1 + \alpha e^{-\beta t}) \exp \left\{ -\mu \left[ t + \frac{\alpha}{\beta} (1 - e^{-\beta t}) \right] \right\}. \quad (\text{III.6})$$

On se propose dans la suite d'étudier en détails cette loi.

## 3 Simulation des interarrivées du processus de Mino

Pour simuler les dates de saut du processus, on propose de simuler les interarrivées. On utilise la méthode d'inversion dont on rappelle le principe ci-dessous.

La méthode d'inversion pour simuler les réalisations d'une variable aléatoire repose sur le résultat suivant :

**Proposition III.3.** – *Soit  $Y$  une variable aléatoire de loi  $F$  alors  $F(Y)$  est une variable aléatoire de loi uniforme sur  $[0, 1]$  et réciproquement.*

**Preuve :**

Supposons  $Y$  de loi  $F$  et calculons la loi de  $F(Y)$ .  $Pr(F(Y) \leq y) = Pr(Y \leq F^{-1}(y)) = F(F^{-1}(y)) = y$  donc  $F(Y)$  est bien une loi uniforme sur  $[0, 1]$ . Réciproquement,  $Pr(Y \leq y) = Pr(F(Y) \leq F(y)) = F(y)$  si  $F(Y)$  suit une loi uniforme sur  $[0, 1]$  donc  $Y$  est de loi  $F$ .

Ainsi, pour obtenir une réalisation de  $Y$  de loi  $F$ , il suffit de tirer  $u$ , un nombre entre  $[0, 1]$  et de calculer  $F^{-1}(u)$ . Dans notre situation, on ne peut calculer  $F^{-1}$  de manière explicite et  $F^{-1}(u)$  est la solution en  $t$  de l'équation :

$$-\frac{\log(1-u)}{\mu} + t + \frac{\alpha}{\beta}(1 - e^{-\beta t}) = 0, \quad (\text{III.7})$$

dont on peut approcher la solution par une méthode itérative.

On pose :

$$\phi(t) = \frac{\log u}{\mu} + t + \frac{\alpha}{\beta}(1 - e^{-\beta t}),$$

alors

$$\phi'(t) = 1 + \alpha e^{-\beta t} \geq 0.$$

Donc  $\phi(t)$  est croissante est strictement croissante pour tout  $\alpha \neq -1$ .

De plus,  $\phi(0) < 0$  et  $\lim_{t \rightarrow +\infty} \phi(t) = +\infty$ , donc l'équation III.7 admet une solution unique. Pour approcher cette solution, on peut appliquer le schéma itératif suivant : Partant d'une valeur initiale  $t^{(0)}$ , tant que  $\|t^{(p+1)} - t^{(p)}\| > \varepsilon$ ,  $\varepsilon$  étant une valeur fixée arbitrairement petite, on détermine :

$$t^{(p+1)} = t^{(p)} - \frac{\phi(t^{(p)})}{\phi'(t^{(p)})}.$$

En général, cet algorithme converge vite.

L'étape d'initialisation peut s'effectuer en approximant  $1 - e^{-\beta t}$  par  $\beta t$ .

On résout l'équation :

$$\frac{\log u}{\mu} + t(1 + \alpha) = 0,$$

on obtient :

$$t^{(0)} = -\frac{\log u}{\mu(1 + \alpha)},$$

Pour  $\alpha = -1$ , on pourra prendre  $-\frac{\log u}{\mu}$ .

Alors pour simuler des réalisations d'une variable aléatoire de loi de Mino on utilise l'algorithme suivant :

1. Génération d'un variable aléatoire uniform  $U$  sur  $[0,1]$ .
2. – si  $\alpha \neq -1$ ,  $t_1 = \frac{-\log U}{\mu(1+\alpha)}$ ,  
– sinon,  $t_1 = \frac{-\log U}{\mu}$ .
3. On calcule :
  - $A_i = \frac{\log U}{\mu} + t_i + \frac{\alpha}{\beta}(1 - e^{-\beta t_i})$ ,
  - $B_i = 1 + \alpha e^{-\beta t_i}$ .

4.  $t_{i+1} = t_i - \frac{A_i}{B_i}$ .

## 4 Loi de Mino

**Définition III.1.** – On dit qu’une variable aléatoire  $Y$  suit une loi de Mino de paramètre  $(\mu, \alpha, \beta)$  si et seulement si la densité de sa loi de probabilité est de la forme (III.6) :

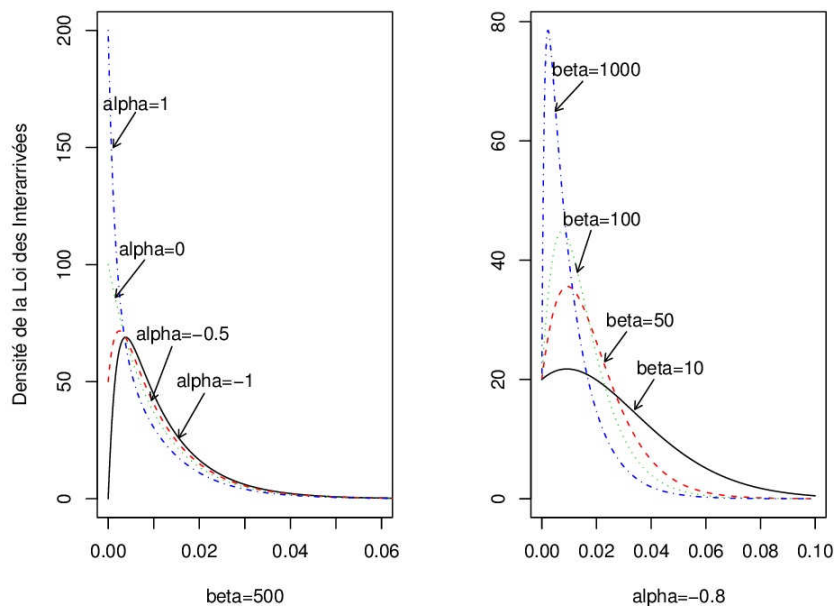
$$f(t) = \mu(1 + \alpha e^{-\beta t}) \exp \left\{ -\mu \left[ t + \frac{\alpha}{\beta} (1 - e^{-\beta t}) \right] \right\} \quad \mu > 0, \quad \beta > 0, \quad \alpha \geq -1.$$

On note :  $Y \sim \text{Mino}(\mu, \alpha, \beta)$ .

### 4.1 Propriétés

1. Si  $\alpha = 0$ , on a  $f(t) = \mu e^{-\mu t}$  alors la loi de Mino est une loi exponentielle de paramètre  $\mu$ .
2. si  $\beta \rightarrow +\infty$ , on retrouve encore la loi exponentielle de paramètre  $\mu$ .

On donne dans la (Figure III.3) des représentations de la densité d’une loi de Mino suivant les valeurs des paramètres.



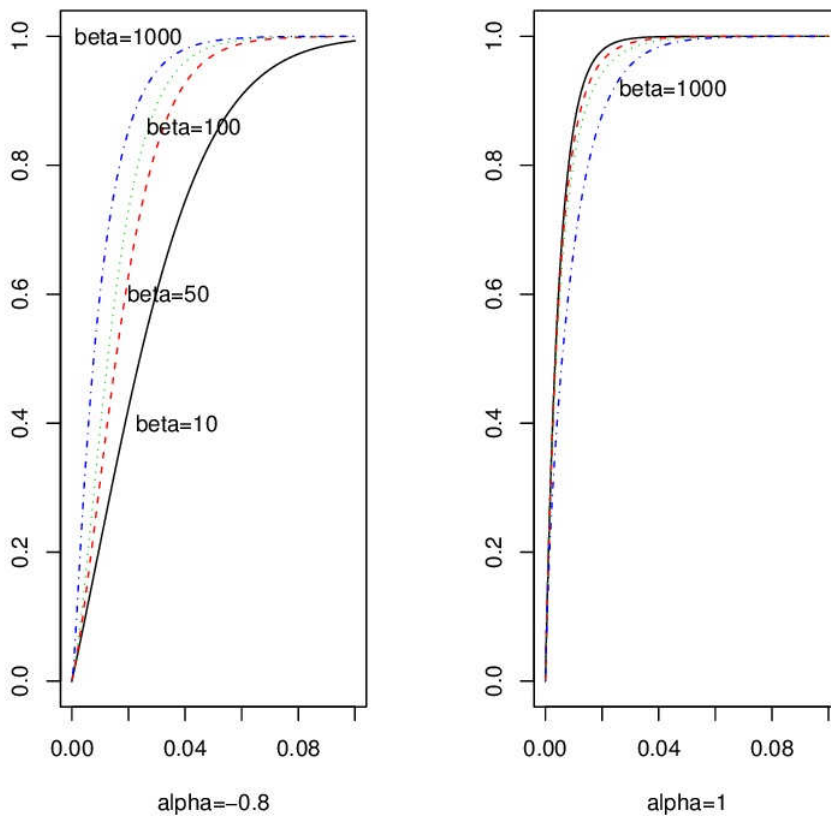
**Figure III.3** – Représentation de la densité de la loi des interarrivées pour  $\mu = 100$  et différentes valeurs de  $\alpha$  et  $\beta$ .

## 4.2 Fonction de répartition

Puisque  $F(t) = 1 - P(T_i > t)$ , on a :

$$F(t) = 1 - \exp \left\{ -\mu \left[ t + \frac{\alpha}{\beta} (1 - e^{-\beta t}) \right] \right\}.$$

Une représentation graphique de la fonction de répartition d'une loi de Mino suivant différentes valeurs des paramètres est donnée dans la Figure III.4.



**Figure III.4** – Représentation de la fonction de répartition de la loi des interarrivées pour  $\mu = 100$  et différentes valeurs de  $\alpha$  et  $\beta$ .

### 4.3 Taux de défaillance

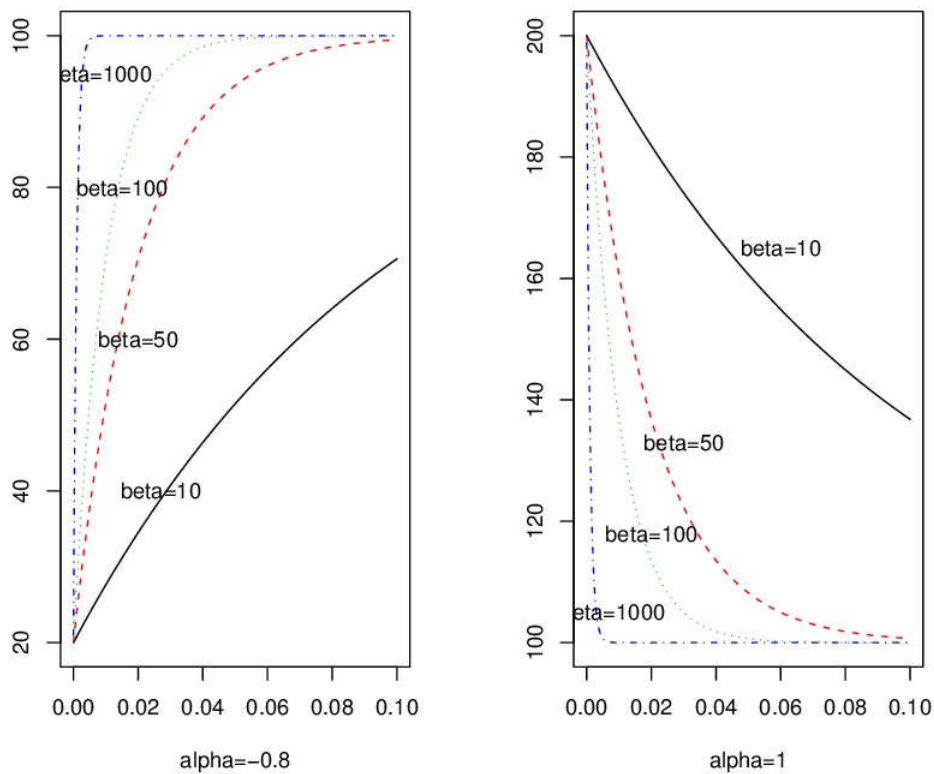
Le taux de défaillance [70], notée  $\Lambda(t)$ , est une expression probabiliste de l'apparition d'une défaillance à un instant  $t$  au cours de la vie d'un produit est a pour expression :

$$\Lambda(t) = \frac{f(t)}{1 - F(t)} = \frac{f(t)}{R(t)},$$

où  $R(t)$  désigne la fiabilité à l'instant  $t$ . Ainsi, le taux de défaillance de la loi de Mino est :

$$\Lambda(t) = \mu(1 + \alpha e^{-\beta t}). \tag{III.8}$$

On donne dans la (Figure III.5) des représentations du taux de défaillance d'une loi de Mino suivant les valeurs des paramètres.



**Figure III.5** – Représentation du taux de défaillance pour  $\mu = 100$  et différentes valeurs de  $\alpha$  et  $\beta$ .

#### 4.4 Fonction génératrice des moments

Afin de calculer la fonction génératrice des moments, on introduit la fonction :

$$\Gamma^*(x; a) = \int_0^x z^{a-1} e^z dz. \quad (\text{III.9})$$

dont la représentation graphique est donnée dans la Figure ci-dessous :

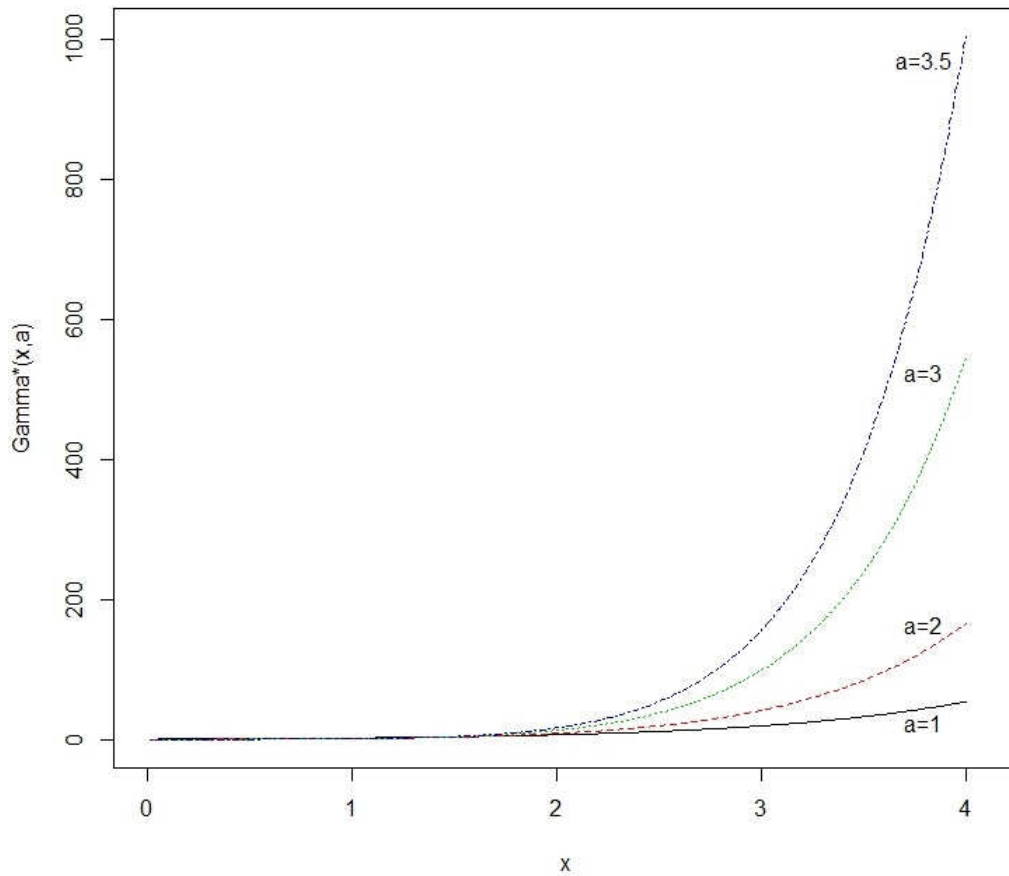


Figure III.6 – Représentation de la fonction  $\Gamma^*(x; a)$ .

**Proposition III.4.** –

1. Si  $0 \leq a < +\infty$ ,  $z^{a-1} e^z$  est intégrable sur  $[0, x[$  et  $\Gamma^*(x; a)$  existe pour  $x$  fini.
2. Si  $a < 0$ ,  $\Gamma^*(x; a)$  est divergente  $\forall x$ .

**Proposition III.5.** –

1. Si  $a = 1$ ,  $\Gamma^*(x; 1) = e^x - 1$ .

2.  $\forall a \geq 1$  on a :

$$\Gamma^*(x; a + 1) + a\Gamma^*(x; a) = x^a e^x.$$

3.  $\forall a \in \mathbb{N}^*$  on a :

$$\Gamma^*(x; a + 1) = (-1)^{a+1} a! + \sum_{k=0}^a \left[ (-1)^k \frac{a!}{(a-k)!} x^{a-k} \right] e^x.$$

**Preuve :**

1.

$$\Gamma^*(x; 1) = \int_0^x e^z dz = [e^z]_0^x = e^x - 1.$$

2. On a  $\Gamma^*(x; a + 1) = \int_0^x z^a e^z dz$ , on intègre par parties on obtient :

$$\Gamma^*(x; a + 1) = x^a e^x - a\Gamma^*(x; a),$$

donc

$$\Gamma^*(x; a + 1) + a\Gamma^*(x; a) = x^a e^x,$$

3. On a :

$$\Gamma^*(x; a + 1) + a\Gamma^*(x; a) = x^a e^x,$$

et

$$\Gamma^*(x; a) = x^{a-1} e^x - (a-1)\Gamma^*(x; a-1).$$



Alors

$$\begin{aligned}
 \Gamma^*(x; a + 1) &= x^a e^x - a x^{a-1} e^x + a(a-1) \Gamma^*(x; a-1) \\
 &= [x^a - a x^{a-1}] e^x + a(a-1) [x^{a-2} e^x + (a-2) \Gamma^*(x; a-2)] \\
 &= [x^a - a x^{a-1} + a(a-1) x^{a-2}] e^x - a(a-1)(a-2) \Gamma^*(x; a-2) \\
 &= \qquad \qquad \qquad \vdots \\
 &= [x^a - a x^{a-1} + a(a-1) x^{a-2} - \dots + \dots + (-1)^a a(a-1) \dots 2] e^x \\
 &+ (-1)^a a! \Gamma^*(x; 1) \\
 &= \sum_{k=0}^{a-1} \left[ (-1)^k \frac{a!}{(a-k)!} x^{a-k} \right] e^x + (-1)^a a! (e^x - 1) \\
 &= (-1)^{a+1} a! + \sum_{k=0}^a \left[ (-1)^k \frac{a!}{(a-k)!} x^{a-k} \right] e^x.
 \end{aligned}$$

**Proposition III.6.** – Soit  $X$  une variable aléatoire dont la loi a une densité de la forme (III.6). La fonction génératrice des moments est :

$$g_X(t) = 1 + t \frac{\beta^{\frac{\mu-t}{\beta}-1}}{(\mu\alpha)^{\frac{\mu-t}{\beta}}} e^{-\frac{\mu\alpha}{\beta}} \Gamma^*\left(\frac{\mu\alpha}{\beta}, \frac{\mu-t}{\beta}\right) \text{ avec } \mu > t \text{ et } \beta \geq 0 \text{ et } \alpha > 0, \tag{III.10}$$

$$g_X(t) = 1 + t \frac{\beta^{\frac{\mu-t}{\beta}-1}}{(\mu|\alpha|)^{\frac{\mu-t}{\beta}}} e^{-\frac{\mu\alpha}{\beta}} \gamma\left(\frac{\mu|\alpha|}{\beta}, \frac{\mu-t}{\beta}\right) \text{ avec } \mu > t \text{ et } \beta \geq 0 \text{ et } \alpha < 0, \tag{III.11}$$

où  $\gamma$  est la fonction gamma incomplète [71] définie par :

$$\gamma(a, x) = \int_0^x t^{a-1} e^{-t} dt.$$

**Preuve :**

- Pour  $\alpha > 0, \frac{\mu}{\beta} \geq 1,$

$$\begin{aligned}
 f(x) &= \mu(1 + \alpha e^{-\beta x}) \exp\left\{-\mu\left[x + \frac{\alpha}{\beta}(1 - e^{-\beta x})\right]\right\} \\
 &= U'(x) e^{-U(x)},
 \end{aligned}$$

avec  $U(x) = \mu[x + \frac{\alpha}{\beta}(1 - e^{-\beta x})]$ .

Par définition  $g_X(t) = E(e^{tx})$  Alors :

$$\begin{aligned}
 g_X(t) &= \int_0^{+\infty} e^{tx} f(x) dx \\
 &= \int_0^{+\infty} e^{tx} U'(x) e^{-U(x)} dx \\
 &= \int_0^{+\infty} U'(x) e^{-[U(x)-tx]} dx \\
 &= \int_0^{+\infty} (U'(x) - t) e^{-[U(x)-tx]} dx + t \int_0^{+\infty} e^{-[U(x)-tx]} dx \\
 &= I + tJ,
 \end{aligned}$$

Calcul de I :

$$\begin{aligned}
 I &= \int_0^{+\infty} (U'(x) - t) e^{-[U(x)-tx]} dx \\
 &= \left[ -e^{-[U(x)-tx]} \right]_0^{+\infty} \\
 &= \left[ -e^{-[(\mu-t)x + \frac{\mu\alpha}{\beta}(1-e^{-\beta x})]} \right]_0^{+\infty} \\
 &= 1 \text{ si } t < \mu.
 \end{aligned}$$

Calcul de J :

$$\begin{aligned}
 J &= \int_0^{+\infty} e^{-[U(x)-tx]} dx \\
 &= \int_0^{+\infty} e^{-[(\mu-t)x + \frac{\mu\alpha}{\beta}(1-e^{-\beta x})]} dx \\
 &= e^{-\frac{\mu\alpha}{\beta}} \int_0^{+\infty} e^{-[(\mu-t)x - \frac{\mu\alpha}{\beta}e^{-\beta x}]} dx,
 \end{aligned}$$

En faisant le changement de variable  $y = e^{-x}$ , on a  $x = -\log y$  et  $dx = -\frac{dy}{y}$ . On obtient :

$$\begin{aligned}
 J &= e^{-\frac{\mu\alpha}{\beta}} \int_0^1 e^{(\mu-t)\log y} e^{\frac{\mu\alpha}{\beta}e^{\log y}} \frac{dy}{y} \\
 &= e^{-\frac{\mu\alpha}{\beta}} \int_0^1 y^{\mu-t} e^{\frac{\mu\alpha}{\beta}y^\beta} \frac{dy}{y} \\
 &= e^{-\frac{\mu\alpha}{\beta}} \int_0^1 y^{\mu-t-1} e^{\frac{\mu\alpha}{\beta}y^\beta} dy.
 \end{aligned}$$

En posant  $u = y^\beta$  alors  $y = u^{\frac{1}{\beta}}$  et  $dy = \frac{1}{\beta}u^{\frac{1}{\beta}-1}du$ , on obtient :

$$\begin{aligned} J &= \frac{1}{\beta} e^{-\frac{\mu\alpha}{\beta}} \int_0^1 u^{\frac{\mu-t-1}{\beta}} e^{\frac{\mu\alpha}{\beta}u} u^{\frac{1}{\beta}-1} du \\ &= \frac{1}{\beta} e^{-\frac{\mu\alpha}{\beta}} \int_0^1 u^{\frac{\mu-t}{\beta}-1} e^{\frac{\mu\alpha}{\beta}u} du. \end{aligned}$$

En posant  $v = \frac{\mu\alpha}{\beta}u$  alors  $u = \frac{\beta v}{\mu\alpha}$  et  $du = \frac{\beta}{\mu\alpha}dv$ , on obtient :

$$\begin{aligned} J &= \frac{e^{-\frac{\mu\alpha}{\beta}}}{\mu\alpha} \int_0^{\frac{\mu\alpha}{\beta}} \left(\frac{\beta v}{\mu\alpha}\right)^{\frac{\mu-t}{\beta}-1} e^v dv \\ &= \frac{e^{-\frac{\mu\alpha}{\beta}}}{\mu\alpha} \left(\frac{\beta}{\mu\alpha}\right)^{\frac{\mu-t}{\beta}-1} \int_0^{\frac{\mu\alpha}{\beta}} v^{\frac{\mu-t}{\beta}-1} e^v dv \\ &= \frac{\beta^{\frac{\mu-t}{\beta}-1}}{(\mu\alpha)^{\frac{\mu-t}{\beta}}} e^{-\frac{\mu\alpha}{\beta}} \Gamma^*\left(\frac{\mu\alpha}{\beta}; \frac{\mu-t}{\beta}\right). \end{aligned}$$

Par conséquent :

$$\begin{aligned} g_X(t) &= I + tJ \\ &= 1 + t \frac{\beta^{\frac{\mu-t}{\beta}-1}}{(\mu\alpha)^{\frac{\mu-t}{\beta}}} e^{-\frac{\mu\alpha}{\beta}} \Gamma^*\left(\frac{\mu\alpha}{\beta}; \frac{\mu-t}{\beta}\right) \text{ avec } \mu > t \text{ et } \frac{\mu-t}{\beta} \geq 0 \text{ et } \alpha > 0. \end{aligned}$$

- Pour  $-1 \leq \alpha < 0$ , on a :

$I = 1$ , et pour  $J$ , on suit la même démarche que pour le cas  $\alpha > 0$  en supposant que  $\alpha = -a$  avec  $a > 0$ , puis on obtient l'équation (III.11).

## 4.5 Calcul de l'espérance mathématique

**Proposition III.7.** – Soit  $X$  une variable aléatoire dont la loi a une densité de la forme (III.6). L'espérance mathématique de  $X$  est :

$$E(X) = \frac{\Gamma^*(\alpha\eta; \eta)}{\beta (\alpha\eta)^\eta} e^{-\eta\alpha} \quad \text{pour } \eta > 0 \text{ et } \alpha > 0, \quad (\text{III.12})$$

et

$$E(X) = \frac{\gamma(|\alpha|\eta; \eta)}{\beta (|\alpha|\eta)^\eta} e^{-\eta\alpha} \quad \text{pour } \eta > 0 \text{ et } \alpha < 0, \quad (\text{III.13})$$

où  $\eta = \frac{\mu}{\beta}$ .

**Preuve :**

- Pour  $\alpha > 0, \frac{\mu}{\beta} \geq 1$ , on a par définition

$$\begin{aligned} E(X) &= \int_0^{+\infty} P(T_i > t) dt \\ &= \int_0^{+\infty} \exp\left\{-\mu\left[t + \frac{\alpha}{\beta}(1 - e^{-\beta t})\right]\right\} dt \\ &= \exp\left\{-\frac{\alpha\mu}{\beta}\right\} \int_0^{+\infty} \exp\left\{-\mu t + \frac{\mu\alpha}{\beta} e^{-\beta t}\right\} dt, \end{aligned}$$

Par le changement de variable,  $y = e^{-t}$  alors  $t = -\log y$  et  $dt = -\frac{dy}{y}$ , on obtient :

$$\begin{aligned} E(X) &= e^{-\frac{\alpha\mu}{\beta}} \int_0^1 \exp\left\{\mu \log y + \frac{\mu\alpha}{\beta} e^{\beta \log y}\right\} \frac{dy}{y} \\ &= e^{-\frac{\alpha\mu}{\beta}} \int_0^1 y^\mu \exp\left\{\frac{\alpha\mu}{\beta} y^\beta\right\} \frac{dy}{y} \\ &= e^{-\frac{\alpha\mu}{\beta}} \int_0^1 y^{\mu-1} \exp\left\{\frac{\alpha\mu}{\beta} y^\beta\right\} dy. \end{aligned}$$

Par le changement de variable  $u = y^\beta$  alors  $y = u^{1/\beta}$  et  $dy = -\frac{1}{\beta} u^{1/\beta-1} du$ , on obtient :

$$\begin{aligned} E(X) &= e^{-\frac{\alpha\mu}{\beta}} \int_0^1 u^{\frac{\mu-1}{\beta}} e^{\frac{\mu\alpha}{\beta} u} \frac{1}{\beta} u^{\frac{1}{\beta}-1} du \\ &= \frac{1}{\beta} e^{-\frac{\alpha\mu}{\beta}} \int_0^1 u^{\frac{\mu-1}{\beta} + \frac{1-\beta}{\beta}} e^{\frac{\mu\alpha}{\beta} u} du \\ &= \frac{1}{\beta} e^{-\frac{\alpha\mu}{\beta}} \int_0^1 u^{\frac{\mu}{\beta}-1} e^{\frac{\mu\alpha}{\beta} u} du. \end{aligned}$$

En posant :  $v = \frac{\mu\alpha}{\beta} u$  alors  $u = \frac{\beta v}{\mu\alpha}$  et  $du = \frac{\beta}{\mu\alpha} dv$ , on obtient :

$$\begin{aligned} E(X) &= \frac{e^{-\frac{\alpha\mu}{\beta}}}{\mu\alpha} \int_0^{\frac{\mu\alpha}{\beta}} \left(\frac{\beta v}{\mu\alpha}\right)^{\frac{\mu}{\beta}-1} e^v dv \\ &= \frac{e^{-\frac{\alpha\mu}{\beta}}}{\mu\alpha} \left(\frac{\beta}{\mu\alpha}\right)^{\frac{\mu}{\beta}-1} \int_0^{\frac{\mu\alpha}{\beta}} v^{\frac{\mu}{\beta}-1} e^v dv. \end{aligned}$$

Comme  $\int_0^{\frac{\mu\alpha}{\beta}} v^{\frac{\mu}{\beta}-1} e^{-v} dv = \Gamma^*\left(\frac{\mu\alpha}{\beta}; \frac{\mu}{\beta}\right)$ , on a :

$$\begin{aligned} E(X) &= \frac{\beta^{\frac{\mu}{\beta}-1}}{(\mu\alpha)^{\frac{\mu}{\beta}}} e^{-\frac{\alpha\mu}{\beta}} \Gamma^*\left(\frac{\mu\alpha}{\beta}; \frac{\mu}{\beta}\right). \\ &= \frac{\Gamma^*(\alpha\eta; \eta)}{\beta (\alpha\eta)^\eta} e^{-\eta\alpha} \quad \text{pour } \eta \geq 1 \text{ et } \alpha \neq 0. \end{aligned}$$

- Pour  $-1 \leq \alpha < 0$ ,

$$E(X) = \int_0^{+\infty} \exp\left\{-\mu\left[t + \frac{\alpha}{\beta}(1 - e^{-\beta t})\right]\right\} dt.$$

En posant  $\alpha = -a$  avec  $a > 0$ , on a :

$$\begin{aligned} E(X) &= \int_0^{+\infty} \exp\left\{-\mu\left[t + \frac{a}{\beta}(e^{-\beta t} - 1)\right]\right\} dt \\ &= \exp\left\{\frac{a\mu}{\beta}\right\} \int_0^{+\infty} \exp\left\{-\mu t - \frac{\mu a}{\beta}e^{-\beta t}\right\} dt. \end{aligned}$$

En posant  $y = e^{-t}$  alors  $t = -\log y$  et  $dt = -\frac{dy}{y}$ , on obtient :

$$\begin{aligned} E(X) &= e^{\frac{a\mu}{\beta}} \int_0^1 \exp\left\{\mu \log y - \frac{a\mu}{\beta}e^{\beta \log y}\right\} \frac{dy}{y} \\ &= e^{\frac{a\mu}{\beta}} \int_0^1 y^\mu \exp\left\{-\frac{a\mu}{\beta}y^\beta\right\} \frac{dy}{y} \\ &= e^{\frac{a\mu}{\beta}} \int_0^1 y^{\mu-1} \exp\left\{-\frac{a\mu}{\beta}y^\beta\right\} dy. \end{aligned}$$

Par le changement de variable  $u = y^\beta$ , alors  $y = u^{1/\beta}$  et  $dy = \frac{1}{\beta}u^{1/\beta-1}du$ . On obtient :

$$\begin{aligned}
 E(X) &= e^{\frac{a\mu}{\beta}} \int_0^1 u^{\frac{\mu-1}{\beta}} e^{-\frac{\mu a}{\beta} u} \frac{1}{\beta} u^{\frac{1}{\beta}-1} du \\
 &= \frac{1}{\beta} e^{\frac{a\mu}{\beta}} \int_0^1 u^{\frac{\mu-1}{\beta} + \frac{1-\beta}{\beta}} e^{-\frac{\mu a}{\beta} u} du \\
 &= \frac{1}{\beta} e^{\frac{a\mu}{\beta}} \int_0^1 u^{\frac{\mu}{\beta}-1} e^{-\frac{\mu a}{\beta} u} du,
 \end{aligned}$$

En posant  $v = \frac{\mu a}{\beta} u$  alors  $u = \frac{\beta v}{\mu a}$  et  $du = \frac{\beta}{\mu a} dv$ , on a :

$$\begin{aligned}
 E(X) &= \frac{e^{\frac{a\mu}{\beta}}}{\mu a} \int_0^{\frac{\mu a}{\beta}} \left( \frac{\beta v}{\mu a} \right)^{\frac{\mu}{\beta}-1} e^{-v} dv \\
 &= \frac{e^{\frac{a\mu}{\beta}}}{\mu a} \left( \frac{\beta}{\mu a} \right)^{\frac{\mu}{\beta}-1} \int_0^{\frac{\mu a}{\beta}} v^{\frac{\mu}{\beta}-1} e^{-v} dv,
 \end{aligned}$$

Comme  $\int_0^{\frac{\mu a}{\beta}} v^{\frac{\mu}{\beta}-1} e^{-v} dv = \gamma\left(\frac{\mu}{\beta}; \frac{\mu a}{\beta}\right)$ , on obtient :

$$E(X) = \frac{\beta^{\frac{\mu}{\beta}-1}}{(\mu a)^{\frac{\mu}{\beta}}} e^{\frac{a\mu}{\beta}} \gamma\left(\frac{\mu a}{\beta}; \frac{\mu}{\beta}\right) = \frac{\gamma(|\alpha|\eta; \eta)}{\beta (|\alpha|\eta)^\eta} e^{-\eta\alpha}.$$

Par simulation, nous comparons les espérances mathématiques données dans (III.12) et (III.13) avec l'estimateur de la moyenne empirique, où les fonctions auxiliaires  $\Gamma^*$  et  $\gamma$  sont calculées par programmation sous le logiciel R. Le tableau suivant résume ces résultats.

Tableau 7 : Comparaison de l'espérance mathématique  $E(X)$  avec l'estimateur de la moyenne empirique pour  $\mu = 100$  et  $n = 2000$  et différentes valeurs de  $\alpha$  et  $\beta$  :

$\alpha$	Moyenne			$\alpha$	Moyenne		
	$\beta$	empirique	$E(X)$		$\beta$	empirique	$E(X)$
-1	10	0.0434	0.0433	1	10	0.0052	0.0051
	50	0.0215	0.0219		50	0.0056	0.0056
	100	0.0172	0.0171		100	0.0065	0.0063
-0.5	10	0.0173	0.0175	3	10	0.0025	0.0025
	50	0.0146	0.0143		50	0.0029	0.0027
	100	0.0127	0.0129		100	0.0033	0.0031
-0.8	10	0.0287	0.0285	5	10	0.0016	0.0016
	50	0.0189	0.0183		50	0.0017	0.0018
	100	0.0154	0.0153		100	0.002	0.0019

D'après ce tableau, nous observons que la moyenne empirique d'un échantillon simulé selon la loi de Mino, est très proche de l'espérance mathématique de la loi de Mino.

#### 4.6 Calcul de la variance

**Proposition III.8.** – Soit  $X$  une variable aléatoire dont la loi a une densité de la forme (III.6). La variance de  $X$  est :

$$\begin{aligned}
 V(X) &= 2e^{-\eta\alpha} \frac{\beta^{\eta-2}}{(\mu\alpha)^\eta} \left[ (\log(\mu\alpha) - \log \beta) \Gamma^*(\eta\alpha; \eta) - \int_0^{\eta\alpha} \log z z^{\eta-1} e^z dz \right] \\
 &\quad - \left( \frac{\Gamma^*(\alpha\eta; \eta)}{\beta (\alpha\eta)^\eta} e^{-\eta\alpha} \right)^2 \quad \text{pour } \eta > 0 \text{ et } \alpha > 0,
 \end{aligned} \tag{III.14}$$

et

$$\begin{aligned}
 V(X) &= 2e^{-\eta\alpha} \frac{\beta^{\eta-2}}{(\mu|\alpha|)^\eta} \left[ (\log(\mu|\alpha|) - \log \beta) \gamma(\eta\alpha; \eta) - \int_0^{\eta|\alpha|} \log z z^{\eta-1} e^{-z} dz \right] \\
 &\quad - \left( \frac{\gamma(|\alpha|\eta; \eta)}{\beta (|\alpha|\eta)^\eta} e^{-\eta\alpha} \right)^2 \quad \text{pour } \eta > 0 \text{ et } \alpha < 0.
 \end{aligned} \tag{III.15}$$

**Preuve :**

On sait que

$$V(X) = E(X^2) - (EX)^2.$$

Pour calculer  $E(X^2)$  on utilise la fonction génératrice des moments où par définition :

$$g_X^{(n)}(0) = E(X^n).$$

On obtient :

- Pour  $\alpha > 0$

$$\begin{aligned} g'_X(t) &= 0 + \frac{\beta^{\frac{\mu-t}{\beta}-1}}{(\mu\alpha)^{\frac{\mu-t}{\beta}}} e^{-\frac{\mu\alpha}{\beta}} \Gamma^*\left(\frac{\mu\alpha}{\beta}; \frac{\mu-t}{\beta}\right) + t \left( \frac{\beta^{\frac{\mu-t}{\beta}-1}}{(\mu\alpha)^{\frac{\mu-t}{\beta}}} e^{-\frac{\mu\alpha}{\beta}} \Gamma^*\left(\frac{\mu\alpha}{\beta}; \frac{\mu-t}{\beta}\right) \right)' \\ g'_X(0) &= \frac{\beta^{\frac{\mu}{\beta}-1}}{(\mu\alpha)^{\frac{\mu}{\beta}}} e^{-\frac{\mu\alpha}{\beta}} \Gamma^*\left(\frac{\mu\alpha}{\beta}; \frac{\mu}{\beta}\right) \\ &= \frac{\Gamma^*(\alpha\eta; \eta)}{\beta (\alpha\eta)^\eta} e^{-\eta\alpha} \quad \text{pour } \eta \geq 1 \\ &= E(X). \end{aligned}$$

$$g''_X(t) = 2 \left( \frac{\beta^{\frac{\mu-t}{\beta}-1}}{(\mu\alpha)^{\frac{\mu-t}{\beta}}} e^{-\frac{\mu\alpha}{\beta}} \Gamma^*\left(\frac{\mu\alpha}{\beta}; \frac{\mu-t}{\beta}\right) \right)' + t \left( \frac{\beta^{\frac{\mu-t}{\beta}-1}}{(\mu\alpha)^{\frac{\mu-t}{\beta}}} e^{-\frac{\mu\alpha}{\beta}} \Gamma^*\left(\frac{\mu\alpha}{\beta}; \frac{\mu-t}{\beta}\right) \right)''.$$

On a :

$$\frac{\partial \Gamma^*\left(\frac{\mu\alpha}{\beta}; \frac{\mu-t}{\beta}\right)}{\partial t} = -\frac{1}{\beta} \int_0^{\frac{\mu\alpha}{\beta}} \log z \, z^{\frac{\mu-t}{\beta}-1} e^{-z} dz,$$

et

$$\left( \frac{\beta^{\frac{\mu-t}{\beta}-1}}{(\mu\alpha)^{\frac{\mu-t}{\beta}}} \right)' = \frac{\beta^{\frac{\mu-t}{\beta}-2} (\log(\mu\alpha) - \log \beta)}{(\mu\alpha)^{\frac{\mu-t}{\beta}}},$$



alors

$$\begin{aligned}
 g_X''(t) &= 2e^{-\frac{\mu\alpha}{\beta}} \left[ \frac{\beta^{\frac{\mu-t}{\beta}-2} (\log(\mu\alpha) - \log \beta)}{(\mu\alpha)^{\frac{\mu-t}{\beta}}} \Gamma^*\left(\frac{\mu\alpha}{\beta}; \frac{\mu-t}{\beta}\right) - \frac{\beta^{\frac{\mu-t}{\beta}-2}}{(\mu\alpha)^{\frac{\mu-t}{\beta}}} \int_0^{\frac{\mu\alpha}{\beta}} \log z z^{\frac{\mu-t}{\beta}-1} e^z dz \right] \\
 &+ t \left( \frac{\beta^{\frac{\mu-t}{\beta}-1}}{(\mu\alpha)^{\frac{\mu-t}{\beta}}} e^{-\frac{\mu\alpha}{\beta}} \Gamma^*\left(\frac{\mu\alpha}{\beta}; \frac{\mu-t}{\beta}\right) \right)'' \\
 &= 2e^{-\frac{\mu\alpha}{\beta}} \frac{\beta^{\frac{\mu-t}{\beta}-2}}{(\mu\alpha)^{\frac{\mu-t}{\beta}}} \left[ (\log(\mu\alpha) - \log \beta) \Gamma^*\left(\frac{\mu\alpha}{\beta}; \frac{\mu-t}{\beta}\right) - \int_0^{\frac{\mu\alpha}{\beta}} \log z z^{\frac{\mu-t}{\beta}-1} e^z dz \right] \\
 &+ t \left( \frac{\beta^{\frac{\mu-t}{\beta}-1}}{(\mu\alpha)^{\frac{\mu-t}{\beta}}} e^{-\frac{\mu\alpha}{\beta}} \Gamma^*\left(\frac{\mu\alpha}{\beta}; \frac{\mu-t}{\beta}\right) \right)'',
 \end{aligned}$$

donc

$$\begin{aligned}
 g_X''(0) &= 2e^{-\frac{\mu\alpha}{\beta}} \frac{\beta^{\frac{\mu}{\beta}-2}}{(\mu\alpha)^{\frac{\mu}{\beta}}} \left[ (\log(\mu\alpha) - \log \beta) \Gamma^*\left(\frac{\mu\alpha}{\beta}; \frac{\mu}{\beta}\right) - \int_0^{\frac{\mu\alpha}{\beta}} \log z z^{\frac{\mu}{\beta}-1} e^z dz \right] \\
 &= 2e^{-\eta\alpha} \frac{\beta^{\eta-2}}{(\mu\alpha)^\eta} \left[ (\log(\mu\alpha) - \log \beta) \Gamma^*(\eta\alpha; \eta) - \int_0^{\eta\alpha} \log z z^{\eta-1} e^z dz \right] \\
 &= E(X^2) \quad \text{pour } \eta \geq 1 \text{ et } \alpha > 0.
 \end{aligned}$$

- Pour  $-1 < \alpha < 0$

$$\begin{aligned}
 g_X'(t) &= 0 + \frac{\beta^{\frac{\mu-t}{\beta}-1}}{(\mu|\alpha|)^{\frac{\mu-t}{\beta}}} e^{-\frac{\mu\alpha}{\beta}} \gamma\left(\frac{\mu|\alpha|}{\beta}; \frac{\mu-t}{\beta}\right) + t \left( \frac{\beta^{\frac{\mu-t}{\beta}-1}}{(\mu|\alpha|)^{\frac{\mu-t}{\beta}}} e^{-\frac{\mu\alpha}{\beta}} \gamma\left(\frac{\mu|\alpha|}{\beta}; \frac{\mu-t}{\beta}\right) \right)', \\
 g_X'(0) &= \frac{\beta^{\frac{\mu}{\beta}-1}}{(\mu|\alpha|)^{\frac{\mu}{\beta}}} e^{-\frac{\mu\alpha}{\beta}} \gamma\left(\frac{\mu|\alpha|}{\beta}; \frac{\mu}{\beta}\right) \\
 &= \frac{\gamma(|\alpha|\eta; \eta)}{\beta (|\alpha|\eta)^\eta} e^{-\eta\alpha} \quad \text{pour } \eta \geq 1 \\
 &= E(X).
 \end{aligned}$$

Pour  $g_X''(0)$  on fait le même calcul que précédemment dans le cas où  $\alpha > 0$  on obtient :

$$\begin{aligned}
 g_X''(0) &= 2e^{-\eta\alpha} \frac{\beta^{\eta-2}}{(\mu|\alpha|)^\eta} \left[ (\log(\mu|\alpha|) - \log \beta) \gamma(\eta\alpha; \eta) - \int_0^{\eta|\alpha|} \log z z^{\eta-1} e^{-z} dz \right] \\
 &= E(X^2) \quad \text{pour } \eta \geq 1 \text{ et } -1 < \alpha < 0.
 \end{aligned}$$

### III.5 Estimation statistique pour le Processus de Mino

Par simulation, nous comparons les variances théoriques données dans (III.14) et (III.15) avec l'estimateur de la variance empirique, où les fonctions auxiliaires  $\Gamma^*$  et  $\gamma$  sont calculées par programmation sous le logiciel R. Le tableau suivant résume ces résultats.

Tableau 8 : Comparaison de la variance théorique  $V(X)$  avec l'estimateur de la variance empirique pour  $\mu = 100$  et  $n = 2000$  et différentes valeurs de  $\alpha$  et  $\beta$  :

$\alpha$	$\beta$	Variance empirique $V(X)$		$\alpha$	$\beta$	Variance empirique $V(X)$	
-1	10	0.00062	0.00059	1	10	$2.98 \cdot 10^{-5}$	$2.76 \cdot 10^{-5}$
	50	0.0002	0.00019		50	$4.24 \cdot 10^{-5}$	$4.09 \cdot 10^{-5}$
	100	0.00014	0.00013		100	$5.22 \cdot 10^{-5}$	$5.7 \cdot 10^{-5}$
-0.5	10	0.00028	0.00025	3	10	$7.63 \cdot 10^{-6}$	$6.75 \cdot 10^{-6}$
	50	0.00014	0.00015		50	$9.9 \cdot 10^{-6}$	$9.83 \cdot 10^{-6}$
	100	0.00013	0.00012		100	$1.81 \cdot 10^{-5}$	$1.73 \cdot 10^{-5}$
-0.8	10	0.00046	0.00047	5	10	$3.34 \cdot 10^{-6}$	$2.94 \cdot 10^{-6}$
	50	0.00017	0.00018		50	$4.1 \cdot 10^{-6}$	$3.85 \cdot 10^{-6}$
	100	0.00011	0.00013		100	$5.82 \cdot 10^{-6}$	$6.29 \cdot 10^{-6}$

D'après ce tableau on trouve que la variance empirique d'un échantillon simulé selon la loi de Mino, est presque égale à la variance mathématique de la loi de Mino.

## 5 Estimation statistique pour le Processus de Mino

Nous allons maintenant établir l'estimation statistique sur le Processus de Mino en nous appuyant sur sa représentation "Processus de renouvellement". Mino [19] a considéré un PAE dont l'intensité s'écrit :

$$\begin{aligned} \lambda(t) &= \lim_{\Delta \rightarrow 0} \frac{Pr[N(t, t + \Delta) = 1 | N(t), W]}{\Delta} \\ &= \mu h(t - w_{N(t)}), \end{aligned}$$

avec

$$h(t) = 1 + \alpha e^{-\beta t}, \quad t \geq 0, \alpha > -1, \beta \geq 0. \quad (\text{III.16})$$

Le problème dans cet article [19] est l'estimation des paramètres de l'intensité considéré  $\theta = [\mu, \alpha, \beta]^T$  en utilisant la méthode du maximum de vraisemblance.

Dans le cas d'un processus aléatoire la log-vraisemblance a pour expression [11, 20, 43] :

$$L(N(t)|\theta) = - \int_0^T \lambda(\sigma) d\sigma + \int_0^T \log[\lambda(\sigma)] dN(\sigma).$$

Mino suggère d'utiliser un algorithme EM [55] pour l'estimation statistique des paramètres. Comme l'algorithme EM est utilisé dans un contexte de données manquantes, Mino dans son article a proposé l'étude d'un processus de comptage observable  $N^1(t)$  représentant les occurrences d'événements ponctuels, et un processus de comptage inobservable  $N^0(t)$  représentant les occurrences ponctuels où il n'y a pas d'événements. Il a associé à chaque processus une intensité, où les intensités des processus  $N^1(t)$  et  $N^0(t)$  sont données respectivement par :

$$\lambda^1(t) = \lim_{\Delta \rightarrow 0} \frac{Pr[N^1(t, t + \Delta) = 1 | N(t), W]}{\Delta} = \mu(1 + \alpha e^{-\beta(t-w_{N(t)})}),$$

et

$$\lambda^0(t) = \lim_{\Delta \rightarrow 0} \frac{Pr[N^0(t, t + \Delta) = 1 | N(t), W]}{\Delta} = \mu(c - \alpha e^{-\beta(t-w_{N(t)})}),$$

où  $c$  est une constante positive et  $\lambda^0(t) > 0$  pour tout  $t$ , et les processus de comptage  $N^1(t)$  et  $N^0(t)$  sont localement fortement continus [72, 73]. Mino a introduit la notion des données manquantes pour utiliser l'algorithme EM. Il a en fait supposé un processus ponctuel discret qui prend les valeurs 0 ou 1 [74], où 0 signifie qu'aucun événement a eu lieu dans un intervalle de temps donné (processus de comptage inobservable), et 1 sinon (processus de comptage observable). Ce processus stochastique est définie comme un processus stochastique discret, dans lequel les variables aléatoires prennent seulement les valeurs 0 ou 1, et par suite le processus de comptage  $N^1(t)$  correspond à l'événement où on a 1 et le processus de comptage  $N^0(t)$

### III.5 Estimation statistique pour le Processus de Mino

correspond à l'événement où on a 0.

Pour appliquer l'algorithme EM, Mino a calculé la log-vraisemblance pour les données complètes qui est définie par :

$$\begin{aligned}
 L_{dc}(N(t)|\theta) &= -\int_0^T \lambda^1(\sigma)d\sigma + \int_0^T \log[\lambda^1(\sigma)]dN^1(\sigma) - \int_0^T \lambda^0(\sigma)d\sigma + \int_0^T \log[\lambda^0(\sigma)]dN^0(\sigma) \\
 &= -(c+1)\mu T + \log[\mu] \left[ N^1(T) + \int_0^T dN^0(\sigma) \right] + \int_0^T \log[1 + \alpha e^{-\beta(\sigma-w_{N(\sigma)})}]dN^1(\sigma) \\
 &+ \int_0^T \log[c - \alpha e^{-\beta(\sigma-w_{N(\sigma)})}]dN^0(\sigma).
 \end{aligned}$$

À la première étape de l'algorithme EM, l'espérance mathématique de la log-vraisemblance est :

$$E[L_{dc}(N(t)|\theta)|W, N(T), \theta^{(old)}] \equiv Q(\mu|\theta^{(old)}) + Q(\alpha e^{-\beta\tau}|\theta^{(old)}),$$

avec

$$\begin{aligned}
 Q(\mu|\theta^{(old)}) &= -(c+1)\mu T + \log[\mu] \left[ N^1(T) + \int_0^T E[dN^0(\sigma)|W, N(T), \theta^{(old)}] \right] \\
 &= -(c+1)\mu T + \log[\mu] \cdot \left[ N^1(T) + \int_0^T \mu^{(old)}(c - \alpha^{(old)} e^{\beta^{(old)}(\sigma-w_{N(\sigma)})}) \right],
 \end{aligned}$$

et

$$\begin{aligned}
 Q(\alpha e^{-\beta\tau}|\theta^{(old)}) &= \int_0^T \log[1 + \alpha e^{-\beta(\sigma-w_{N(\sigma)})}]dN^1(\sigma) \\
 &+ \int_0^T \log[c - \alpha e^{-\beta(\sigma-w_{N(\sigma)})}] \cdot E[dN^0(\sigma), N(T), \theta^{(old)}] \\
 &= \int_0^T \log \left[ 1 + \int_0^{t_{N(\sigma)}} \alpha e^{-\beta\tau} \delta(\sigma - w_{N(\sigma)} - \tau) d\tau \right] dN^1(\sigma) \\
 &+ \int_0^T \log \left[ c - \int_0^{t_{N(\sigma)}} \alpha e^{-\beta\tau} \delta(\sigma - w_{N(\sigma)} - \tau) d\tau \right] \\
 &\times \mu^{(old)}(c - \alpha^{(old)} e^{-\beta^{(old)}(\sigma-w_{N(\sigma)})}) d\sigma,
 \end{aligned}$$

où  $t_{N(\sigma)} = w_{N(\sigma)+1} - w_{N(\sigma)}$  et  $\delta(t)$  est la mesure de dirac.

Après le calcul de l'espérance mathématique de la log-vraisemblance, usuellement nous passons à l'étape de maximisation de l'algorithme EM qui consiste à chercher les estimateurs de

$\mu$ ,  $\alpha$  et  $\beta$  qui maximisent l'espérance du log de vraisemblance  $E[L_{dc}(N(t)|\theta)|W, N(T), \theta^{(old)}]$ . Il s'agit de calculer la dérivée par rapport à chaque paramètre, puis de résoudre les équations obtenues lorsque ces dérivées sont nulles. Dans son article, Mino a pour maximiser  $E[L_{dc}(N(t)|\theta)|W, N(T), \theta^{(old)}]$  résolu l'équation

$$\frac{\partial}{\partial \alpha e^{-\beta \tau}} Q(\alpha e^{-\beta \tau} | \theta^{(old)}) = 0.$$

Selon Mino, si les estimateurs des deux paramètres  $\alpha$  et  $\beta$  sont indépendants, la résolution de l'équation précédente revient à déterminer les estimateurs du maximum de vraisemblance de  $\alpha$  et  $\beta$ .

Ensuite, Mino a cherché les valeurs  $\mu^{(new)}, \theta^{(new)}, \beta^{(new)}$  qui maximisent par simulation  $E[L_{dc}(N(t)|\theta)|W, N(T), \theta^{(old)}]$ . Il a en fait simulé les interarrivées en utilisant la méthode de Monte carlo, pour différentes valeurs des paramètres et pour calculer ces estimateurs.

Dans ce travail, nous nous sommes intéressés à introduire une méthode plus simple pour estimer les paramètres de l'intensité considérée par Mino, où on suggère d'utiliser la méthode du maximum de vraisemblance. On utilise l'algorithme de Newton-Raphson sur des données simulées par la méthode d'inversion [24], et pour vérifier l'existence et l'unicité des estimateurs, nous utilisons la méthode décrite dans [50]. Ce travail nous conduit à introduire une loi de distribution spécifique que nous avons appelée loi de Mino définie ci-dessus.

## 5.1 Maximum de vraisemblance

On propose d'obtenir les estimateurs du maximum de vraisemblance des paramètres  $\mu$ ,  $\alpha$ ,  $\beta$ . La log-vraisemblance est donnée par :

$$\log L(N(T)|\theta) = - \int_0^T \lambda(t) dt + \int_0^T \log \lambda(t) dN(t). \quad (\text{III.17})$$

### III.5 Estimation statistique pour le Processus de Mino

---

Pour le processus de Mino, on a donc :

$$\begin{aligned}
 -\int_0^T \lambda(t)dt &= -\int_0^T \mu + \mu\alpha e^{-\beta(t-w_{N(t)})} dt \\
 &= -[\mu t]_0^T - \sum_{i=0}^{N(T)} \int_{w_i}^{w_{i+1}} \mu\alpha e^{-\beta(t-w_i)} dt \\
 &= -\mu T - \sum_{i=0}^{N(T)} \left[ \frac{-\mu\alpha}{\beta} e^{-\beta(t-w_i)} \right]_{w_i}^{w_{i+1}} \\
 &= -\mu T - \sum_{i=0}^{N(T)} \frac{-\mu\alpha}{\beta} \left( e^{-\beta(w_{i+1}-w_i)} - 1 \right) \\
 &= -\mu T - \sum_{i=0}^{N(T)} \frac{\mu\alpha}{\beta} \left( 1 - e^{-\beta(w_{i+1}-w_i)} \right).
 \end{aligned}$$

On pose  $t_i = w_{i+1} - w_i$ ,  $w_0 = 0$  et  $w_{N(T)+1} \equiv T$ . On a donc :

$$-\int_0^T \lambda(t)dt = -\mu T - \frac{\mu\alpha}{\beta} \sum_{i=0}^{N(T)} \left( 1 - e^{-\beta t_i} \right).$$

et

$$\begin{aligned}
 \int_0^T \log \lambda(t) dN(t) &= \sum_{w_i < T} \log \lambda(w_i) \\
 &= \sum_{w_i < T} \log \mu(1 + \alpha e^{-\beta(w_i - w_{i-1})}) \\
 &= \sum_{i=0}^{N(T)} \log \mu(1 + \alpha e^{-\beta t_i}).
 \end{aligned}$$

La log-vraisemblance est donc :

$$\begin{aligned}
 \log L(N(T)|\theta) &= -\mu T - \frac{\mu\alpha}{\beta} \sum_{i=1}^{N(T)} (1 - e^{-\beta t_i}) + \sum_{i=1}^{N(T)} \log \mu(1 + \alpha e^{-\beta t_i}) \\
 &= -\mu T - \frac{\mu\alpha}{\beta} \sum_{i=1}^{N(T)+1} (1 - e^{-\beta t_i}) + N(T) \log \mu + \sum_{i=1}^{N(T)} \log(1 + \alpha e^{-\beta t_i}).
 \end{aligned}$$

On peut obtenir le même résultat en utilisant l'indépendance des interarrivées, et en calculant alors le produit densité.

On considère la loi de densité (III.6) définie par :

$$f(t) = \mu(1 + \alpha e^{-\beta t}) \exp \left\{ -\mu \left[ t + \frac{\alpha}{\beta} (1 - e^{-\beta t}) \right] \right\}.$$

La fonction de vraisemblance pour cette densité est :

$$L(N(T)|\theta) = \mu^n \prod_{i=1}^n (1 + \alpha e^{-\beta t_i}) \exp \left\{ -\mu \left[ \sum_{i=1}^n t_i + \frac{\alpha}{\beta} \sum_{i=1}^n (1 - e^{-\beta t_i}) \right] \right\}.$$

Donc la log-vraisemblance est :

$$\log L(N(T)|\theta) = n \log \mu + \sum_{i=1}^n \log(1 + \alpha e^{-\beta t_i}) - \mu \sum_{i=1}^n t_i - \frac{\mu \alpha}{\beta} \sum_{i=1}^n (1 - e^{-\beta t_i}). \quad (\text{III.18})$$

### 5.2 Algorithme RSM

Pour estimer les paramètres qui maximisent l'équation (III.18), on génère les interarrivées en utilisant la méthode d'inversion décrite dans (3). On va aussi mettre en oeuvre l'algorithme RSM (Random Stochastic Maximization) décrit dans le paragraphe (3.1). Nous appliquons cet algorithme sur un nombre de  $n = 100$  interarrivées et pour  $\mu = 100$  et différentes valeurs de  $\alpha$  et  $\beta$  et sur 50 jeux de données et on trouve la moyenne des estimateurs dans les tableaux suivants :

### III.5 Estimation statistique pour le Processus de Mino

Tableau 9 : Résultats obtenus par l'algorithme RSM pour  $n = 100$ ,  $\mu = 100$ ,  $\alpha = \{-1, -0.8, -0.5, 0\}$ ,  $\beta = \{10, 50, 100, 250, 500, 1000\}$  : SD désigne l'écart type (Standard Deviation en anglais).

$\alpha$	$\beta$	moyenne de $\hat{\mu}$	SD $\hat{\mu}$	moyenne de $\hat{\alpha}$	SD $\hat{\alpha}$	moyenne de $\hat{\beta}$	SD $\hat{\beta}$
-1	10	104.209	18.960	-0.992	0.171	9.769	1.790
	50	103.882	18.177	-0.980	0.169	46.186	8.177
	100	101.833	17.818	-0.976	0.169	95.521	16.963
	250	101.017	17.644	-0.966	0.167	236.992	42.021
	500	101.352	17.732	-0.967	0.168	470.754	83.495
	1000	101.380	17.682	-0.910	0.163	941.792	167.404
-0.8	10	110.706	19.396	-0.826	0.143	9.783	1.950
	50	104.365	1.917	-0.817	0.013	48.935	1.200
	100	103.672	1.680	-0.823	0.017	97.947	2.533
	250	102.691	1.591	-0.806	0.021	248.627	6.292
	500	102.644	1.509	-0.807	0.306	489.671	12.460
	1000	102.796	1.310	-0.816	0.027	952.976	25.803
-0.5	10	111.139	19.484	-0.560	0.099	11.407	2.520
	50	103.066	18.168	-0.483	0.093	49.051	8.720
	100	101.144	17.791	-0.488	0.094	93.582	16.642
	250	101.654	17.813	-0.490	0.098	249.559	44.439
	500	100.274	17.504	-0.484	0.108	516.079	91.555
	1000	101.162	17.662	-0.467	0.116	988.657	176.230
0	10	104.903	2.575	0.0004	0.963	17.317	0.689
	50	100.072	2.270	0.032	0.036	52.300	1.209
	100	101.515	2.150	0.042	0.045	101.221	2.526
	250	101.855	1.797	0.013	0.060	243.725	6.465
	500	102.143	1.695	-0.002	0.060	519.053	12.031
	1000	103.288	1.506	-0.0395	0.072	1013.997	26.309



### Chapitre III. Processus de Mino

Tableau 10 : Résultats obtenus par l'algorithme RSM pour  $n = 100$ ,  $\mu = 100$ ,  
 $\alpha = \{1, 3, 4, 7, 15\}$ ,  $\beta = \{10, 50, 100, 250, 500, 1000\}$  : SD désigne l'écart type (Standard  
 Deviation en anglais)

$\alpha$	$\beta$	moyenne de $\hat{\mu}$	SD $\hat{\mu}$	moyenne de $\hat{\alpha}$	SD $\hat{\alpha}$	moyenne de $\hat{\beta}$	SD $\hat{\beta}$
1	10	92.646	16.343	1.246	0.233	9.835	2.721
	50	102.207	18.138	1.071	0.210	50.315	8.955
	100	98.992	17.490	1.155	0.229	104.441	18.513
	250	101.136	17.782	1.046	0.208	254.114	45.156
	500	100.240	17.570	1.015	0.215	485.606	86.130
	1000	101.159	17.660	0.988	0.229	1012.642	180.078
3	10	94.918	1.837	3.303	0.080	8.905	1.320
	50	105.747	2.262	2.999	0.553	50.603	1.337
	100	108.975	2.093	2.737	0.096	95.161	2.424
	250	100.521	2.042	3.007	0.110	248.243	5.542
	500	100.333	1.762	3.078	0.107	504.916	12.315
	1000	102.887	2.887	2.978	0.094	982.495	25.739
4	10	92.090	16.160	4.461	0.795	7.735	2.476
	50	106.593	18.806	3.864	0.701	46.773	8.352
	100	107.628	18.947	3.758	0.675	98.922	17.606
	250	97.701	17.212	4.322	0.786	254.438	44.997
	500	99.010	17.394	4.106	0.743	491.658	86.997
	1000	99.517	17.421	4.068	0.740	959.944	170.333
7	10	93.024	1.765	7.771	0.177	8.514	1.348
	50	105.366	2.392	6.826	0.219	47.803	1.339
	100	110.925	1.916	6.404	0.195	95.528	2.630
	250	106.563	2.372	6.731	0.231	250.785	5.786
	500	103.261	2.018	7.019	0.214	497.386	11.710
	1000	100.242	1.716	6.959	0.245	993.501	23.100
15	10	93.93	1.489	16.728	0.302	7.535	1.337
	50	104.728	2.078	14.654	0.442	48.112	1.353
	100	103.555	2.377	14.899	0.463	96.384	2.596
	250	113.626	19.847	13.328	2.367	237.810	42.346
	500	104.613	18.486	14.832	2.659	498.016	88.027
	1000	102.273	17.968	15.509	2.765	1030.133	181.61

Les résultats obtenus en appliquant l'algorithme "RSM" pour différentes valeurs des paramètres, nous confirment que la log-vraisemblance est concave. Nous estimons les paramètres en utilisant l'algorithme Newton-Raphson.

### 5.3 Application de l'algorithme de Newton-Raphson

Pour trouver les valeurs de  $\mu, \alpha, \beta$  qui maximisent l'équation (III.18), on calcule les dérivées premières de la log-vraisemblance par rapport à chaque paramètre, est on obtient le système d'équations suivant :

$$\frac{\partial \log L}{\partial \mu} = -\sum_{i=1}^n t_i - \frac{\alpha}{\beta} \sum_{i=1}^n (1 - e^{-\beta t_i}) + \frac{n}{\mu} = 0, \quad (\text{III.19})$$

$$\frac{\partial \log L}{\partial \alpha} = -\frac{\mu}{\beta} \sum_{i=1}^n (1 - e^{-\beta t_i}) + \sum_{i=1}^n \frac{e^{-\beta t_i}}{1 + \alpha e^{-\beta t_i}} = 0, \quad (\text{III.20})$$

$$\frac{\partial \log L}{\partial \beta} = \frac{\mu \alpha}{\beta^2} \sum_{i=1}^n (1 - e^{-\beta t_i}) - \frac{\mu \alpha}{\beta} \sum_{i=1}^n t_i e^{-\beta t_i} - \alpha \sum_{i=1}^n \frac{t_i e^{-\beta t_i}}{1 + \alpha e^{-\beta t_i}} = 0. \quad (\text{III.21})$$

Ce système n'admet pas des solutions explicites. Nous utilisons alors l'algorithme de Newton-Raphson selon le chemin itératif suivant :

$$\theta^{(p+1)} = \theta^{(p)} - [H(\theta^{(p)})]^{-1} U(\theta^{(p)}), \forall p \geq 0$$

où  $\theta$  est le vecteur des paramètres à estimer, tant que  $\|\theta^{(p+1)} - \theta^{(p)}\|$  est supérieur à  $\varepsilon$  une valeur arbitrairement fixée petite,  $U(\theta^{(p)})$  étant le vecteur des dérivées premières et  $[H(\theta^{(p)})]^{-1}$  la matrice hessienne inverse.

### Chapitre III. Processus de Mino

---

La matrice hessienne s'écrit :

$$\frac{\partial^2 \log L}{\partial \mu^2} = -\frac{n}{\mu^2},$$

$$\frac{\partial^2 \log L}{\partial \alpha^2} = -\sum_{i=1}^n \frac{e^{-2\beta t_i}}{(1 + \alpha e^{-\beta t_i})^2},$$

$$\frac{\partial^2 \log L}{\partial \beta^2} = -\frac{2\mu\alpha}{\beta^3} \sum_{i=1}^n (1 - e^{-\beta t_i}) + \frac{2\mu\alpha}{\beta^2} \sum_{i=0}^n t_i e^{-\beta t_i} + \frac{\mu\alpha}{\beta} \sum_{i=1}^n t_i^2 e^{-\beta t_i} + \alpha \sum_{i=1}^n \frac{t_i^2 e^{-\beta t_i}}{(1 + \alpha e^{-\beta t_i})^2},$$

$$\frac{\partial^2 \log L}{\partial \mu \partial \alpha} = -\frac{1}{\beta} \sum_{i=1}^n (1 - e^{-\beta t_i}),$$

$$\frac{\partial^2 \log L}{\partial \mu \partial \beta} = \frac{\alpha}{\beta^2} \sum_{i=1}^n (1 - e^{-\beta t_i}) - \frac{\alpha}{\beta} \sum_{i=1}^n t_i e^{-\beta t_i},$$

$$\frac{\partial^2 \log L}{\partial \alpha \partial \beta} = \frac{\mu}{\beta^2} \sum_{i=1}^n (1 - e^{-\beta t_i}) - \frac{\mu}{\beta} \sum_{i=1}^n t_i e^{-\beta t_i} - \sum_{i=1}^n \frac{t_i e^{-\beta t_i}}{1 + \alpha e^{-\beta t_i}} + \alpha \sum_{i=1}^n \frac{t_i e^{-2\beta t_i}}{(1 + \alpha e^{-\beta t_i})^2}.$$

On applique la méthode décrite ci-dessus selon le chemin itératif définie dans (II.3), après on prend les valeurs suivant :  $\beta = 500, \mu = (100, 200)$  et  $\alpha = (-1, -0.5, 1)$ , et on calcule les estimateurs suivant le nombre des interarrivées  $N(t) = (20000, 10000, 5000)$ .

On calcule les moyennes des estimateurs par la méthode de Newton-Raphson, à partir de 100 jeux de données simulées on obtient :

### III.5 Estimation statistique pour le Processus de Mino

Tableau 11 : Estimation par rapport à  $N(T)$ . Les moyennes des estimateurs et le SE sont calculées à partir l'exécution de 100 Newton-Raphson sur des données simulés. Pour  $\mu = 100, 200$   $\alpha = -1, -0.5, 1$   $\beta = 500$  : SE désigne l'erreur type (Standard error en anglais)

N(T)	Paramètre	Moyenne de	SE de	Paramètre	Moyenne de	SE de
		$(\hat{\mu}, \hat{\alpha}, \hat{\beta})$	$(\hat{\mu}, \hat{\alpha}, \hat{\beta})$		$(\hat{\mu}, \hat{\alpha}, \hat{\beta})$	$(\hat{\mu}, \hat{\alpha}, \hat{\beta})$
20000	$\mu = 100$	99.753	0.992	$\mu = 200$	198.649	2.061
	$\alpha = -1$	-1.002	0.013	$\alpha = -1$	-1.021	0.008
	$\beta = 500$	521.867	23.097	$\beta = 500$	527.448	19.837
10000	$\mu = 100$	99.820	1.378	$\mu = 200$	199.108	2.795
	$\alpha = -1$	-1.020	0.015	$\alpha = -1$	-1.022	0.014
	$\beta = 500$	517.082	30.471	$\beta = 500$	525.699	24.603
5000	$\mu = 100$	99.595	17.903	$\mu = 200$	198.979	4.644
	$\alpha = -1$	-1.020	0.020	$\alpha = -1$	-1.021	0.016
	$\beta = 500$	520.823	44.628	$\beta = 500$	524.578	36.261
20000	$\mu = 100$	100.261	2.026	$\mu = 200$	200.347	2.715
	$\alpha = -0.5$	-0.502	0.034	$\alpha = -0.5$	-0.501	0.016
	$\beta = 500$	494.628	142.556	$\beta = 500$	501.277	48.943
10000	$\mu = 100$	100.170	2.412	$\mu = 200$	200.524	4.361
	$\alpha = -0.5$	-0.5003	0.024	$\alpha = -0.5$	-0.502	0.023
	$\beta = 500$	488.221	176.037	$\beta = 500$	496.981	81.537
5000	$\mu = 100$	100.481	3.421	$\mu = 200$	200.202	5.151
	$\alpha = -0.5$	-0.498	0.013	$\alpha = -0.5$	-0.499	0.033
	$\beta = 500$	486.176	229.881	$\beta = 500$	502.108	103.064
20000	$\mu = 100$	100.076	4.497	$\mu = 200$	200.065	3.281
	$\alpha = 1$	1.009	0.136	$\alpha = 1$	0.999	0.047
	$\beta = 500$	498.702	131.931	$\beta = 500$	500.495	36.367
10000	$\mu = 100$	100.033	3.376	$\mu = 200$	199.868	5.274
	$\alpha = 1$	1.035	0.148	$\alpha = 1$	1.007	0.061
	$\beta = 500$	523.564	145.251	$\beta = 500$	500.486	54.935
5000	$\mu = 100$	99.808	8.922	$\mu = 200$	199.868	6.507
	$\alpha = 1$	1.010	0.213	$\alpha = 1$	1.007	0.084
	$\beta = 500$	500.517	183.564	$\beta = 500$	503.351	79.664

### 5.4 Existence et unicité de l'EMV

Pour démontrer l'existence et l'unicité de l'EMV, on considère la matrice hessienne calculée en  $\hat{\theta} = (\hat{\mu}, \hat{\alpha}, \hat{\beta})$  désignant l'estimateur du maximum de vraisemblance de  $\mu$ ,  $\alpha$  et  $\beta$ . Nous déterminons maintenant la matrice hessienne  $H|\hat{\theta}$  :

$$H|\hat{\theta} = \begin{pmatrix} \frac{\partial^2 \log L}{\partial \mu^2} & \frac{\partial^2 \log L}{\partial \mu \partial \alpha} & \frac{\partial^2 \log L}{\partial \mu \partial \beta} \\ \frac{\partial^2 \log L}{\partial \mu \partial \alpha} & \frac{\partial^2 \log L}{\partial \alpha^2} & \frac{\partial^2 \log L}{\partial \alpha \partial \beta} \\ \frac{\partial^2 \log L}{\partial \mu \partial \beta} & \frac{\partial^2 \log L}{\partial \alpha \partial \beta} & \frac{\partial^2 \log L}{\partial \beta^2} \end{pmatrix}. \quad (\text{III.22})$$

On pose :

$$A_i = \frac{e^{-\beta t_i}}{1 + \alpha e^{-\beta t_i}},$$

puis on calcule les composantes de la matrice hessienne  $H|\hat{\theta}$  on obtient :

$$\frac{\partial^2 \log L}{\partial \mu^2} = -\frac{n}{\mu^2},$$

$$\begin{aligned} \frac{\partial^2 \log L}{\partial \alpha^2} &= -\sum_{i=1}^n \frac{e^{-2\beta t_i}}{(1 + \alpha e^{-\beta t_i})^2} \\ &= -\sum_{i=1}^n A_i^2, \end{aligned}$$

$$\begin{aligned} \frac{\partial^2 \log L}{\partial \beta^2} &= -\frac{2\mu\alpha}{\beta^3} \sum_{i=1}^n (1 - e^{-\beta t_i}) + \frac{2\mu\alpha}{\beta^2} \sum_{i=0}^n t_i e^{-\beta t_i} + \frac{\mu\alpha}{\beta} \sum_{i=1}^n t_i^2 e^{-\beta t_i} + \alpha \sum_{i=1}^n \frac{t_i^2 e^{-\beta t_i}}{(1 + \alpha e^{-\beta t_i})^2} \\ &= -\frac{2}{\beta} \left( \frac{\mu\alpha}{\beta^2} \sum_{i=1}^n (1 - e^{-\beta t_i}) - \frac{\mu\alpha}{\beta} \sum_{i=0}^n t_i e^{-\beta t_i} \right) + \frac{\mu\alpha}{\beta} \sum_{i=1}^n t_i^2 e^{-\beta t_i} + \alpha \sum_{i=1}^n \frac{t_i^2 e^{-\beta t_i}}{(1 + \alpha e^{-\beta t_i})^2} \\ &= -\frac{2\alpha}{\beta} \sum_{i=1}^n \frac{t_i e^{-\beta t_i}}{1 + \alpha e^{-\beta t_i}} + \frac{\mu\alpha}{\beta} \sum_{i=1}^n t_i^2 e^{-\beta t_i} + \alpha \sum_{i=1}^n \frac{t_i^2 e^{-\beta t_i}}{(1 + \alpha e^{-\beta t_i})^2} \quad c.f.(III.21) \\ &= -\frac{2\alpha}{\beta} \sum_{i=1}^n t_i A_i + \frac{\mu\alpha}{\beta} \sum_{i=1}^n t_i^2 e^{-\beta t_i} + \alpha \sum_{i=1}^n t_i^2 e^{\beta t_i} A_i^2, \end{aligned}$$

$$\begin{aligned}
 \frac{\partial^2 \log L}{\partial \mu \partial \alpha} &= -\frac{1}{\beta} \sum_{i=1}^n (1 - e^{-\beta t_i}) \\
 &= -\frac{1}{\mu} \sum_{i=1}^n \frac{e^{-\beta t_i}}{1 + \alpha e^{-\beta t_i}} \quad c.f(III.20) \\
 &= -\frac{1}{\mu} \sum_{i=1}^n A_i,
 \end{aligned}$$

$$\begin{aligned}
 \frac{\partial^2 \log L}{\partial \mu \partial \beta} &= \frac{\alpha}{\beta^2} \sum_{i=1}^n (1 - e^{-\beta t_i}) - \frac{\alpha}{\beta} \sum_{i=1}^n t_i e^{-\beta t_i} \\
 &= \frac{\alpha}{\beta \mu} \sum_{i=1}^n A_i - \frac{\alpha}{\beta} \sum_{i=1}^n t_i e^{-\beta t_i},
 \end{aligned}$$

$$\begin{aligned}
 \frac{\partial^2 \log L}{\partial \alpha \partial \beta} &= \frac{\mu}{\beta^2} \sum_{i=1}^n (1 - e^{-\beta t_i}) - \frac{\mu}{\beta} \sum_{i=1}^n t_i e^{-\beta t_i} - \sum_{i=1}^n \frac{t_i e^{-\beta t_i}}{1 + \alpha e^{-\beta t_i}} + \alpha \sum_{i=1}^n \frac{t_i e^{-2\beta t_i}}{(1 + \alpha e^{-\beta t_i})^2} \\
 &= \frac{1}{\alpha} \left( \frac{\partial \log L}{\partial \beta} \right) + \alpha \sum_{i=1}^n \frac{t_i e^{-2\beta t_i}}{(1 + \alpha e^{-\beta t_i})^2} \\
 &= \alpha \sum_{i=1}^n t_i A_i^2.
 \end{aligned}$$

En remplaçant dans (III.22), on obtient :

$$H|\hat{\boldsymbol{\theta}} = \begin{pmatrix} -\frac{n}{\hat{\mu}^2} & -\frac{1}{\hat{\mu}} \sum_{i=1}^n A_i & \frac{\hat{\alpha}}{\hat{\beta} \hat{\mu}} \sum_{i=1}^n A_i - \frac{\hat{\alpha}}{\hat{\beta}} \sum_{i=1}^n t_i e^{-\hat{\beta} t_i} \\ -\frac{1}{\hat{\mu}} \sum_{i=1}^n A_i & -\sum_{i=1}^n A_i^2 & \hat{\alpha} \sum_{i=1}^n t_i A_i^2 \\ \frac{\hat{\alpha}}{\hat{\beta} \hat{\mu}} \sum_{i=1}^n A_i - \frac{\hat{\alpha}}{\hat{\beta}} \sum_{i=1}^n t_i e^{-\hat{\beta} t_i} & \hat{\alpha} \sum_{i=1}^n t_i A_i^2 & \frac{\partial^2 \log L}{\partial \hat{\beta}^2} \end{pmatrix}.$$

Nous avons montré numériquement en simulant des échantillons des interarrivées pour différent valeurs des paramètres  $\mu, \alpha, \beta$  que cette matrice est définie négative quand elle est calculée aux estimateurs du maximum de vraisemblance de  $\hat{\mu}, \hat{\alpha}$  et  $\hat{\beta}$ . Ainsi, la matrice  $H|\hat{\boldsymbol{\theta}}$  vérifie les conditions cité par Maklaine et al. [50].



---

---

# Chapitre IV

---

## Test d'hypothèses sur le Processus de Mino

Ce chapitre a pour objet de détecter l'auto-excitation d'un processus ponctuel en construisant un test d'hypothèses. Afin de construire ce test, on s'intéresse aux tests appliqués aux processus de Poisson et de renouvellement.

### 1 Introduction

Un test d'hypothèse est un mécanisme qui permet de trancher entre deux hypothèse antagonistes  $H_0$  (hypothèse nulle) et  $H_1$  (hypothèse alternative).

Les tests sur les processus ont été beaucoup étudiés en statistique. Nous pouvons en particulier citer les travaux de [75–77] sur le processus de Poisson et sur l'homogénéité du processus de Poisson [44]. Dachian & Kutoyants [78] ont aussi construit un test basé sur le processus de Poisson, en utilisant le processus auto-excité. Notre objectif ici est de tester si un processus est auto excité ou non.

### 2 Tests de processus

D'après Yagouti et al. [77], on identifie généralement cinq principales étapes au cours du processus de la modélisation d'un test d'hypothèse :



1. définition du problème et des variables des intérêts,
2. formulation des hypothèses et des équations du modèle qui lient ces variables,
3. estimation des paramètres du modèle et son application aux données observées,
4. validation du modèle (application des tests statistiques),
5. simulation et analyse de sensibilité.

### 2.1 Processus de Poisson

Le processus de Poisson sur la droite est un processus à temps continu et à valeurs entières positives. On dit encore que c'est un processus de comptage, qu'on note  $N(t) : t \geq 0$ . Il s'agit d'étudier le nombre aléatoire  $N(t)$  de certains événements qui se produisent dans un intervalle de temps  $[0, t]$  donné. Sa grande popularité dans les applications vient notamment du fait que beaucoup de calculs le concernant sont explicites [79]. L'objet de ce test est la détection de tendance à partir du taux d'occurrence, où on examine l'existence d'une tendance dans le taux d'occurrence par un modèle de régression linéaire [77].

On note que ce test n'est pas valable que pour un processus ponctuel de Poisson non homogène (PPNH) et on suppose que la taille de la série d'événements est assez grande. La fonction d'intensité doit avoir une forme exponentielle dans le temps.

Les hypothèses du test sont :

- $H_0$  : Il n'existe pas de tendance dans la série d'événements ;
- $H_1$  : Il existe une tendance dans la série d'événements.

Pour appliquer ce test on suppose l'existence d'une certaine tendance dans le taux d'occurrence représentée par un changement dans le temps :

$$\lambda(t) = \exp(\alpha + \beta t). \tag{IV.1}$$

Au voisinage de  $\beta = 0$ , la tendance est linéaire.

Dans le cas d'un PPNH, on fait l'hypothèse que les intervalles de temps sont indépendants, étant donné le dernier événement, et que la probabilité d'occurrence d'un événement dans l'intervalle  $(w_{i+1}, w_{i+1} + \Delta t]$  est donnée par :

$$P[N_{w_{i+1}, w_{i+1} + \Delta t} = 1] = \lambda(w_{i+1}) \exp \left\{ - \int_{w_i}^{w_{i+1}} \lambda(u) du \right\} \Delta t + o(\Delta t). \quad (\text{IV.2})$$

Si la série est observée dans l'intervalle  $(0, w_0]$  et si les événements arrivent aux temps  $w_1, \dots, w_n$ , la fonction de vraisemblance est donnée par :

$$L(N(t)|\theta) = \prod_{i=1}^n P[N_{w_{i+1}, w_{i+1} + \Delta t} = 1] = \left\{ \prod_{i=1}^n \lambda(w_i) \right\} \exp \left\{ - \int_0^{w_0} \lambda(u) du \right\}. \quad (\text{IV.3})$$

En remplaçant  $\lambda(u)$  par sa valeur (équation (IV.1)), et en divisant l'équation (IV.3) par la probabilité marginale de  $n$  événements, on obtient une distribution conditionnelle  $L(N(t)|\beta)$  qui dépend uniquement de paramètre  $\beta$ . La transformée logarithmique de la fonction de vraisemblance conditionnelle est donnée par :

$$LL = \log[L(N(t)|\beta)] = n \log \beta - n \log(e^{\beta w_0} - 1) + \beta \sum_{i=1}^n w_i + \log n!. \quad (\text{IV.4})$$

Afin de tester l'hypothèse nulle  $H_0 : \beta = \beta_0$ , on construit la statistique suivante :

$$S_L = \frac{LL'(\beta_0)}{\sqrt{l(\beta_0)}}. \quad (\text{IV.5})$$

La fonction  $LL'(\beta)$  représente la dérivée première par rapport à  $\beta$  de la fonction  $LL(\beta)$ , et la fonction  $l(\beta)$  dite fonction d'information, est définie par :

$$l(\beta) = E \left\{ - \frac{\partial^2 LL}{\partial^2 \beta^2}(\beta) \right\}. \quad (\text{IV.6})$$

Sous l'hypothèse nulle, cette statistique suit asymptotiquement une loi normale centrée réduite  $N(0, 1)$ .

Nous présentons maintenant le test d'Homogénéité pour un processus de Poisson. L'objectif de ce test et de savoir si un processus de Poisson est homogène ou non. Nous testons :

- l'hypothèse nulle  $H_0$  stipulant que le processus de Poisson observé est homogène;
- contre l'hypothèse alternative  $H_1$  stipulant que le processus de Poisson observé est non homogène.

Nous supposons observer le processus de Poisson sur  $[0, t]$ . Selon, Coccozza-thivent [44] nous divisons l'intervalle  $[0, t]$  en  $d$  intervalles de même longueur notés  $I_1, \dots, I_d$ . Soit  $N_k (1 \leq k \leq d)$ , le nombre d'observations appartenant au  $k^{\text{ème}}$  intervalle :

$$N_k = \sum_{i=1}^{N(t)} \mathbf{1}_{\{T_i \in I_k\}}.$$

D'après la proposition 2.3 dans [44], sous  $H_0$  la loi de  $(N_1, \dots, N_d)$  sachant  $\{N(t) = n\}$  est la loi de  $\sum_{i=1}^n \mathbf{1}_{\{X_i \in I_1\}}, \dots, \sum_{i=1}^n \mathbf{1}_{\{X_i \in I_d\}}$ , les variables aléatoires  $X_i$  étant indépendantes et de loi uniforme sur  $[0, t]$ . Par conséquent, d'après le théorème classique de convergence vers la loi du  $\chi^2$ , sous  $H_0$  quand  $n$  tend vers l'infini, la loi conditionnellement à  $\{N(t) = n\}$ , de la variable aléatoire

$$Z_n = \frac{d}{n} \sum_{k=1}^d \left( N_k - \frac{n}{d} \right)^2,$$

tend sous  $H_0$  vers la loi du  $\chi^2$  à  $d - 1$  degrés de liberté.

Donc en pratique, pour construire un test de niveau approximativement égale à  $\gamma$ , si l'on observe  $N(t)=n$  avec  $n$  grand, on prend comme région critique :

$$D = \{Z_n > \chi_{1-\gamma}^2(d-1)\}. \quad (\text{IV.7})$$

## 2.2 Processus de renouvellement

Un processus de renouvellement est défini comme un processus de comptage dont les inter-arrivées sont des variables aléatoires indépendantes de la loi de probabilité qui correspond, cf.

section 5.3 pour plus de détails.

Pour tester si un processus est un processus de renouvellement Yagouti et al. [77] ont utilisé le coefficient d'autocorrélation entre les intervalles, pour examiner l'indépendance des intervalles séparent deux événements successifs.

Les hypothèses du test sont :

- $H_0$  : Il n'existe aucune dépendance entre les intervalles de temps (on a un processus de renouvellement) ;
- $H_1$  : Il existe une dépendance entre les intervalles de temps.

Le coefficient d'autocorrélation d'ordre  $k$ , noté  $\rho_k$ , est défini par la formule suivante :

$$\rho_k = \frac{n}{n-k} \frac{\sum_{i=1}^{n-k} (w_i - \bar{w})(w_{i+k} - \bar{w})}{\sum_{i=1}^n (w_i - \bar{w})^2},$$

où  $t_i$  représente le temps entre deux événements successifs.

Les hypothèses du test se transforment sous la forme suivante :

- $H_0$  :  $\rho_k = 0$  (indépendance des intervalles) ;
- $H_1$  :  $\rho_k \neq 0$  (dépendance des intervalles).

La statistique du test est donnée par :

$$\hat{\rho}_k \sqrt{n-1},$$

où  $\hat{\rho}_k$  est l'estimation de  $\rho_k$ . Cette statistique suit sous l'hypothèse nulle une loi normale  $N(0, 1)$  pour un nombre d'observations assez grand.

### 3 Test pour le processus auto-excité

Pour tester si un processus est auto-excité ou non, Dachian & Kutoyants [78] ont considéré le problème du test d'hypothèse avec l'hypothèse de base : la séquence des événements observée correspond à un processus de Poisson stationnaire avec une intensité connue, contre l'hypothèse

alternative : cette séquence correspond à un processus auto-excité stationnaire. Dans ce contexte ils ont envisagé une approche paramétrique et une approche non paramétrique.

### 3.1 Test paramétrique

Dachian & Kutoyants [78] ont supposé qu'on observe une séquence des événements d'un processus ponctuel  $W^T = \{W_t, 0 \leq t \leq T\}$  avec une fonction d'intensité  $\lambda(t) = \{\lambda(\gamma, t, z), 0 \leq t \leq T\}$ . Si  $\gamma = 0$ , alors  $\lambda(t) = \lambda_*$ , i.e ce processus ponctuel est un processus de Poisson homogène d'intensité  $\lambda_* > 0$ . Dans le cas alternatif quand  $\gamma > 0$ ,  $\lambda(t)$  est la fonction d'intensité d'un processus auto-excité. Comme souvent dans ces problèmes on fait un changement de variable tel que  $\gamma = u/\sqrt{T}$  et par suite les hypothèses de test sont :

- $H_0 : u = 0$ ;
- $H_1 : u > 0$ .

On note  $E_0$  l'espérance mathématique sous l'hypothèse  $H_0$ , et  $E_u$  celle de l'hypothèse  $H_1$ . Nous fixons  $\varepsilon \in (0, 1)$ , et on note par  $\mathcal{K}_\varepsilon$  la classe de la fonction du test  $\phi_T(W^T)$  pour la taille asymptotique de  $\varepsilon$ , i.e pour  $\phi_T(W^T) \in \mathcal{K}_\varepsilon$  on a :

$$\lim_{T \rightarrow \infty} E_0 \phi_T(W^T) = \varepsilon,$$

$\phi_T(W^T)$  est la probabilité pour accepter l'hypothèse  $H_1$ . La fonction du puissance qui correspond est :

$$\beta_T(u, \phi_T) = E_u \phi_T(W^T), \quad u \geq 0.$$

On introduit l'optimalité asymptotique pour les tests en utilisant la définition de Le Cam [80] :

**Définition IV.1.** – Un test  $\phi_T^*(.)$  est appelé le plus puissant localement asymptotiquement uniformément dans la classe  $\mathcal{K}_\varepsilon$  si pour tout autre test  $\phi_T(.) \in \mathcal{K}_\varepsilon$  et pour tout constante  $K > \varepsilon$  on a :

$$\lim_{t \rightarrow \infty} \inf_{0 \leq u \leq K} [\beta_T(u, \phi_T^*) - \beta_T(u, \phi_T)] \geq 0.$$

Le but est de construire le test le plus puissant localement asymptotiquement uniformément dans la classe  $\mathcal{K}_\varepsilon$ .

L'intensité d'un processus auto-excité est :

$$\lambda(t) = \lambda_* + \int_0^{t-} g(t-s)dW_s = \lambda_* + \sum_{w_i < t} g(t-w_i),$$

où  $\lambda_* > 0$ ,  $w_i$  est les événements du processus, et la fonction  $g(\cdot) \geq 0$  satisfait la condition

$$\int_0^\infty g(t)dt < 1. \tag{IV.8}$$

Nous supposons que le processus observé est un processus de Poisson avec une intensité constante  $\lambda_*$ , ou processus auto-excité avec une fonction d'intensité contiguë.

$$\lambda(t) = \lambda_* + \gamma_T \int_0^t h(t-s)dW_s.$$

Le terme contiguë, signifie que le rapport de vraisemblance est asymptotiquement non-dégénéré.

On suppose que la fonction  $h(\cdot)$  est connue et bornée et

$$h(\cdot) \in \mathcal{L}_+^1(\mathbb{R}_+) = \left\{ f(\cdot) \geq 0 : \int_0^\infty f(t)dt < \infty \right\}.$$

Pour avoir des alternatives contiguës nous choisissons  $\gamma_T = \frac{u}{\sqrt{T}}$  i.e

$$\lambda(t) = \lambda_* + \frac{u}{\sqrt{T}} \int_0^t h(t-s)dW_s, \quad u \geq 0.$$

On note que pour tout  $h(\cdot) \in \mathcal{L}_+^1(\mathbb{R}_+)$  et tout  $u \leq K$ , et pour  $T$  suffisamment grand, la condition (IV.8) est accomplie pour la fonction  $g(\cdot) = uT^{-1/2}h(\cdot)$ . Ceci nous conduit aux hypothèses de test suivants :

- $H_0 : u = 0$  (processus de Poisson) ;
- $H_1 : u > 0$  (processus auto-excité).

L'information de Fisher est définie par :

$$I_h^* = \int_0^\infty h(t)^2 dt + \lambda_* \left( \int_0^\infty h(t) dt \right)^2.$$

La limite de la fonction du puissance est donnée par :

$$\Delta_T(W^T) = \frac{1}{\lambda_* \sqrt{TI_h^*}} \int_0^T \int_0^{t-} h(t-s) dW_s [dW_t - \lambda_* dt],$$

et  $z_\varepsilon$  le  $1 - \varepsilon$  quantile de la loi normale  $N(0, 1)$ , et Dachian & Kutoyants [78] ont donné le théorème suivant :

**Théorème IV.1.** – Soit  $h(\cdot) \in \mathcal{L}_+^1(\mathbb{R}_+)$  et borné, alors le test :

$$\hat{\phi}_T(W^T) = \mathbf{1}_{\{\Delta_T(W^T) > z_\varepsilon\}},$$

est le plus puissant localement asymptotiquement uniformément dans la classe  $\mathcal{K}_\varepsilon$  et pour tout  $u > 0$  on a la fonction du puissance suivante :

$$\beta_T(u, \hat{\phi}_T) \rightarrow \hat{\beta}(u) = P\{\zeta > z_\varepsilon - u\sqrt{I_h^*}\},$$

où  $\zeta \sim N(0, 1)$ .

### 3.2 Test non paramétrique

Comme pour le test paramétrique Dachian & Kutoyants [78] ont supposé que sous l'hypothèse  $H_0$ , le processus ponctuel observé  $W^T = \{W_t, 0 \leq t \leq T\}$  est un processus de Poisson avec une fonction d'intensité  $\lambda(t) = \lambda_* > 0$ , et sous l'hypothèse  $H_1$  le processus ponctuel est un processus auto-excité avec une fonction d'intensité

$$\lambda(t) = \lambda_* + \int_{-\infty}^t g(t-s) dW_s, \quad 0 \leq t \leq T,$$

où  $g(\cdot)$  est inconnue pour l'instant, on suppose que

$$\int_0^T g(t) dt < 1, \tag{IV.9}$$

d'où le processus  $W^T$  est stationnaire. Pour décrire la classe de l'hypothèse  $H_1$  on écrit l'intensité :

$$\lambda(t) = \lambda_* + \frac{1}{\sqrt{T}} \int_{-\infty}^t u(t-s) dW_s, \quad 0 \leq t \leq T,$$

où la fonction  $u(\cdot)$  appartient à l'ensemble  $\mathcal{U}_r$  défini ci-dessous. On note l'ensemble des fonctions positive borné par une constante par  $C_+^b$ , et on introduit l'ensemble :

$$\mathcal{U}_r = \{u(\cdot) \in C_+^b : \int_0^\infty u(t) dt = r, \text{ supp } u(\cdot) \text{ est borné} \}.$$

On note que la condition (IV.9) est accomplie pour tout  $r > 0$  et  $T > r^2$ .

Donc les hypothèses de test sont :

- $H_0 : u(\cdot) \equiv 0$  ;
- $H_1 : u(\cdot) \in \mathcal{U}_r, \quad r > 0$ .

La fonction de puissance pour le test  $\phi_T$  dépend de la fonction  $u(\cdot)$  et on l'écrit :

$$\beta_T(u, \phi_T) = E_u \phi_T(W^T), \tag{IV.10}$$

où  $u = u(\cdot) \in \mathcal{U}_r$  avec  $r > 0$ , et pour tout test  $\phi_T \in \mathcal{K}_\varepsilon$  on a :

$$\inf_{u(\cdot) \in \mathcal{U}_r} \beta_T(u, \phi_T) \leq \varepsilon, \tag{IV.11}$$

puisque pour tout  $T > 0$  on peut prendre une fonction depuis  $\mathcal{U}_r$  égale à 0 dans  $[0, T]$ . D'où on introduit l'ensemble :

$$\mathcal{U}_{r,n} = \{u(\cdot) \in \mathcal{U}_r : \text{supp } u(\cdot) \subset [0, N]\},$$

et

$$B_T(r, N, \phi_T) = \inf_{u(\cdot) \in \mathcal{U}_r} \beta_T(u, \phi_T).$$



et on donne la définition suivant :

**Définition IV.2.** – Un test  $\phi_T^*(.)$  est appelé le plus puissant localement asymptotiquement uniformément dans la classe  $\mathcal{K}_\varepsilon$ , si pour tout autre test  $\phi_T(.) \in \mathcal{K}_\varepsilon$  et tout  $K > 0$  on a :

$$\lim_{N \rightarrow \infty} \lim_{T \rightarrow \infty} \inf_{0 \leq r \leq K} [B_T(r, N, \phi_T^*) - B_T(r, N, \phi_T)] \geq 0. \quad (\text{IV.12})$$

On introduit la fonction :

$$\hat{\phi}_T(W^T) = \mathbf{I}_{\{\delta_T(W^T) > z_\varepsilon\}}, \quad \delta_T = \frac{W^T - \lambda_* T}{\sqrt{\lambda_* T}}.$$

**Théorème IV.2.** – Le test  $\hat{\phi}_T$  est le plus puissant localement asymptotiquement uniformément dans la classe  $\mathcal{K}_\varepsilon$ , et pour tout  $u(.) \in \mathcal{U}_r$  la fonction de puissance est :

$$\beta_T(u, \hat{\phi}_T \rightarrow \hat{\beta}(u)) = P \left\{ \zeta > z_\varepsilon - r \sqrt{\lambda_*} \right\},$$

où  $\zeta \sim N(0, 1)$ .

### 3.3 Test du processus de Mino

L'objet de ce test est de savoir si le processus de Mino est un processus auto-excité ou un processus de Poisson homogène.

Les hypothèses du test sont :

- $H_0$  : Le processus de Mino est auto-excité ;
- $H_1$  : Le processus de Mino est un processus de Poisson homogène.

L'intensité du processus de Mino est donnée par :

$$\lambda(t) = \mu(1 + \alpha e^{-\beta(t-w_{N(t)})})$$

et on a :

- si  $\alpha = 0$ ,  $\lambda(t) = \mu = \lambda_*$  (processus de Poisson homogène) ;
- si  $\alpha > 0$ , (Processus auto-excité).

On pose que  $\alpha = u/\sqrt{T}$ , alors les hypothèses du test sont :

- $H_0$  :  $u > 0$  (processus auto-excité) ;

–  $H_1 : u = 0$  (processus de Poisson homogène).

On remarque que le processus de Mino vérifie les conditions mis par Dachian & Kutoyants [78], donc on peut appliquer leur test paramétrique le plus puissant localement asymptotiquement uniformément dans la classe  $\mathcal{K}_\varepsilon, \hat{\phi}_T(W^T)$  qui est définie par :

$$\hat{\phi}_T(W^T) = \mathbf{1}_{\{\Delta_T(W^T) > z_\varepsilon\}},$$

où

$$\Delta_T(W^T) = \frac{1}{\mu\sqrt{TI_h^*}} \int_0^T \int_0^{t-} h(t-s)dW_s[dW_t - \mu dt],$$

et

$$I_h^* = \int_0^\infty h(t)^2 dt + \mu \left( \int_0^\infty h(t) dt \right)^2,$$

et  $Z_\varepsilon$  le  $1 - \varepsilon$  quantile de la loi normale  $N(0,1)$ .

Donc  $\Delta_T(W^T) > z_\varepsilon$  est la zone d'acceptation de l'hypothèse  $H_0$ .

Nous effectuons maintenant une étude par simulation sur un cas particulier décrit dans l'article de Dachian & Kutoyants [78]. Les interarrivées du processus de Mino sont simulés par la méthode d'inversion. Pour appliquer le théorème (IV.1), on prend  $\lambda_* = 1$  et  $h(t) = e^{-\beta t}$  alors :

$$\lambda(t) = 1 + \frac{u}{\sqrt{T}} \sum_{w_i \leq t} e^{-\beta(t-w_i)}, \quad u \geq 0, \quad 0 \leq t \leq T.$$

Dans ce cas on a :

$$\Delta_T(W^T) = \frac{1}{\sqrt{5T}} \sum_{0 \leq w_j \leq T} \sum_{w_i < w_j} e^{-\beta(w_j-w_i)} - \frac{2}{\sqrt{5T}} \left( w_T - \sum_{0 \leq w_j \leq T} e^{-\beta(T-w_j)} \right),$$

## Chapitre IV. Test d'hypothèses sur le Processus de Mino

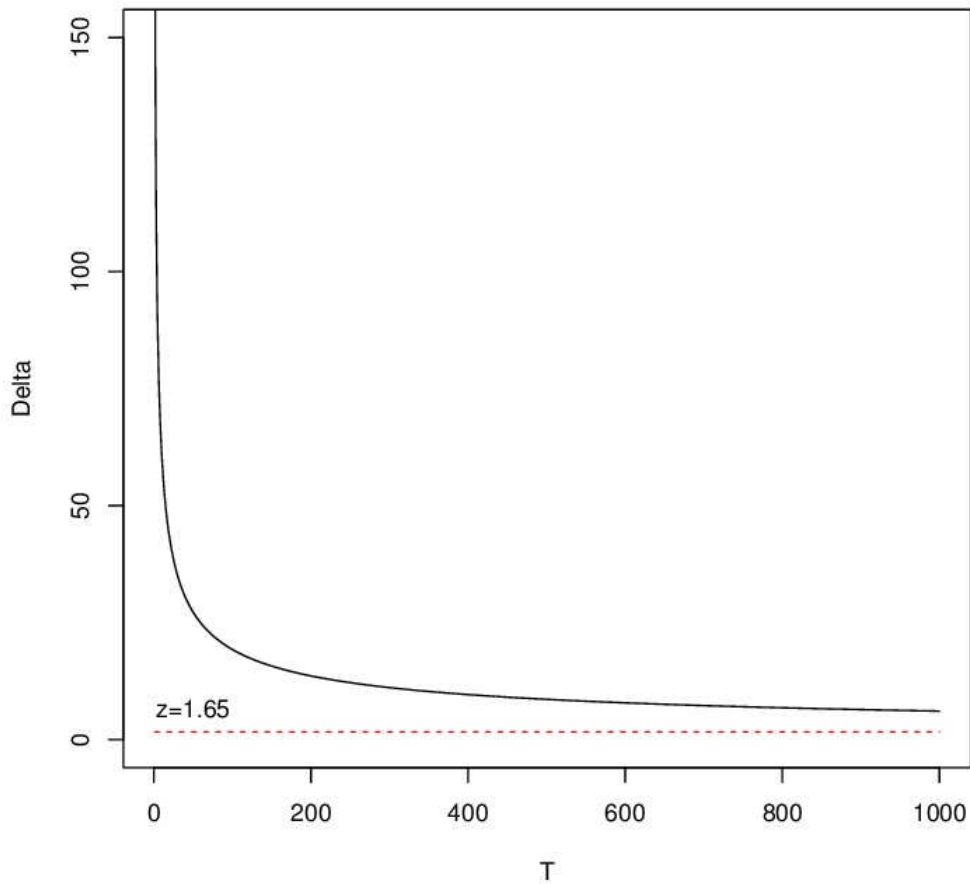
---

où les  $w_i$  sont les interarrivées du processus de Mino, et le test  $\hat{\phi}_T^\varepsilon$  est donné par :

$$\hat{\phi}_T^\varepsilon = \hat{\phi}_T(W^T) = \mathbf{I}_{\{\Delta_T(W^T) > z_\varepsilon\}},$$

qui est le plus puissant localement asymptotiquement uniformément dans la classe  $\mathcal{K}_\varepsilon$ .

La figure (IV.1) représente la fonction  $\Delta_T(W^T)$  en fonction de  $T \in [0, 1000]$ . D'après cette



**Figure IV.1** – Représentation de la fonction  $\Delta_T(W^T)$  en fonction de  $T \in [0, 1000]$ .

figure on remarque  $\Delta_T(W^T) > z_{0.05} \approx 1.65$  pour tout  $T \in [0, 1000]$ , et donc nous sommes dans la zone d'acceptation de  $H_0$ .

---

## Conclusion générale et perspectives

Nous avons dans cette thèse, défini et étudié de manière détaillée, le processus de Mino. Ce processus appartient à la famille des processus auto-excités pour lesquels nous avons fait un état de l'art. Nous avons présenté différentes méthodes et algorithmes pour estimer les paramètres de l'intensité.

L'étude du processus de Mino, qui n'est autre qu'un processus auto-excité de mémoire 1 particulier, nous a conduit à introduire la loi de Mino et a proposé son étude. Nous avons alors défini la fonction  $\Gamma^*$  qui apparaît dans le calcul de la fonction génératrice des moments. Ainsi, l'espérance mathématique et la variance ont été obtenues.

Le problème de l'inférence sur processus de Mino a été envisagé par la méthode de maximum de vraisemblance et un algorithme de Newton-Raphson a été développé pour obtenir les estimateurs des paramètres de l'intensité du processus. La qualité des estimateurs a été étudiée à partir de données simulées. L'existence et l'unicité de ces estimateurs ont été prouvées. Un test d'hypothèse pour détecter l'auto-excitation d'un processus quelconque a aussi été construit.

De nombreuses perspectives ont été dégagées à travers ce travail. L'étude des propriétés asymptotiques des estimateurs restent à faire. Il serait également intéressant d'envisager des méthodes bayésiennes pour analyser les processus auto-excités. La construction de structures de

lois a priori particulières, adaptées à cette famille de processus pourrait être très intéressantes. De plus, le calcul des lois *a posteriori* et autres indicateurs bayésiens devraient conduire à imaginer et mettre en oeuvre des algorithmes MCMC originaux.

D'un point de vue des applications, nous prévoyons d'appliquer ces différents résultats sur des données de surveillance de la qualité des eaux par l'étude de bioindicateurs de pollution, de tremblements de terre et des données d'activités neuronales.

Ainsi, l'ensemble de ce travail permet une meilleure compréhension du comportement des processus auto-excités et devrait trouver de riches applications.

---

## Références bibliographiques

- [1] A. HAWKES. *Spectra of some self-exciting and mutually exciting point process*. *Biometrika* **58**, 83–90 (1971). [5](#), [12](#), [16](#), [17](#)
- [2] D.J. DALEY AND D. VERE-JONES. *An introduction to the theory of point processes*. Springer (2002). [5](#), [19](#), [30](#), [31](#)
- [3] A.G. HAWKES AND L. ADAMOPOULOS. *Cluster models for earthquakes—regional comparisons*. *Bull. Int. Statist. Inst* **45**(3), 454–461 (1973). [6](#)
- [4] L. ADAMOPOULOS. *Cluster models for earthquakes : Regional comparisons*. *Journal of the International Association for Mathematical Geology* **8**(4), 463–475 (1976).
- [5] Y. OGATA. *Statistical models for earthquake occurrences and residual analysis for point processes*. *Ann. Inst. Statist. Math.* **83**(401), 9–27 (1988). [12](#)
- [6] Y. OGATA. *Space-time point-process models for earthquake occurrences*. *Annals of the Institute of Statistical Mathematics* **50**(2), 379–402 (1998).
- [7] D. VERE-JONES. *Forecasting earthquakes and earthquake risk*. *International Journal of Forecasting* **11**(4), 503–538 (1995). [6](#)

- [8] J. BOUYER. *Epidémiologie : principes et méthodes quantitatives*. Lavoisier (2009). 6, 12
- [9] B.C. SEAN, A.J. F. GRIFFITHS, S. WESSLER, AND R.C. LEWONTIN. *Introduction à l'analyse génétique*. De Boeck Supérieur (2010). 6, 12
- [10] E.S. CHORNOBOY, L.P. SCHRAMM, AND A.F. KARR. *Maximum likelihood identification of neural point process systems*. *Biological cybernetics* **59**(4-5), 265–275 (1988). 6
- [11] D.H. JOHNSON. *Point process models of single-neuron discharges*. *Computational Neuroscience* **3**, 275–299 (1996). 12, 78
- [12] W. TRUCCOLO, U.T. EDEN, M.R. FELLOWS, J.P. DONOGHUE, AND E.N. BROWN. *A point process framework for relating neural spiking activity to spiking history, neural ensemble, and extrinsic covariate effects*. *Journal of neurophysiology* **93**(2), 1074–1089 (2005).
- [13] V. BENEŠ AND B. FRCALOVÁ. *Modelling and simulation of a neurophysiological experiment by spatio-temporal point processes*. *Image Analysis & Stereology* **27**(1), 47–52 (2011).
- [14] M. SCOTT AND L. HARRY. *A self exciting point process model for neural spike sequences, and its rate estimation*. 26th IEEE Canadian Conference Of Electrical And Computer Engineering (CCECE) (2013). 7, 13
- [15] Y. CHEN AND N. D. SINGPURWALLA. *Unification of software reliability models by self-exciting point processes*. *Advances in Applied Probability* pages 337–352 (1997). 7
- [16] O. GAUDOIN AND J. LEDOUX. *Modélisation aléatoire en fiabilité des logiciels*. Lavoisier (2007). 12

- 
- [17] J. LEDOUX. *Point processes in software reliability*. Mathematical Methods in Survival Analysis, Reliability and Quality of Life pages 219–236 (2008).
- [18] F. RUGGERI, R. SOYER, AND C.N.R. IMATI. *Advances in bayesian software reliability modeling*. Advances in Mathematical Modeling for Reliability page 165 (2008). [7](#)
- [19] H. MINO. *Parameter estimation of the intensity process of self-exciting point processes using the em algorithm*. IEEE **50**, 658–664 (2001). [7](#), [12](#), [57](#), [77](#), [78](#)
- [20] L. SNYDER AND I. MILLER. *Random Point processes in Time and Space*. Springer (1991). [11](#), [20](#), [22](#), [24](#), [25](#), [26](#), [27](#), [28](#), [30](#), [35](#), [38](#), [59](#), [60](#), [78](#)
- [21] C.G. BOWSER. *Modelling security market events in continuous time : Intensity based, multivariate point process models*. Journal of Econometrics **141**(2), 876–912 (2007). [12](#)
- [22] Y. CAUMEL. *Probabilités et processus stochastique*. Springer (2011). [12](#), [27](#)
- [23] A. HAWKES AND D. OAKES. *A cluster process representation of a self-exciting process*. J. Appl. Prob **11**, 493–503 (1974). [12](#), [20](#)
- [24] T. OZAKI. *Maximum likelihood estimation of hawkes' self-exciting point processes*. Annals of the Institute of Statistical Mathematics **31**(1), 145–155 (1979). [18](#), [31](#), [80](#)
- [25] Y. OGATA AND H. AKAIKE. *On linear intensity models for mixed doubly stochastic poisson and self-exciting point processes*. J. R. Statist. Soc. B **44**, 102–107 (1982). [18](#)
- [26] H. VAN DEN BOOGARD. *Maximum likelihood estimations in a nonlinear self-exciting point process model*. Biological Cybernetics, Springer-Verlag **55**, 219–225 (1986).
- [27] M. A. INGRAM AND A. H. HADAD. *A sequential detection approach to state estimation of linear systems driven by self-excited point process*. Conference on Decision and Control Austin, Texas (1988).



- [28] A. KWIECINSKI AND R. SZEKLI. *Some monotonicity and dependence properties of self-exciting point processes*. The Annals of Applied Probability **6**(4), 1211–1231 (1996). [12](#)
- [29] P. REYNAUD-BOURET AND S. SCHBATH. *Adaptive estimation for hawkes processes; application to genome analysis*. The Annals of Statistics **38**(5), 2781–2822 (2010). [12](#)
- [30] G.O. MOLHER, M.B. SHORT, P.J. BRANTINGHAM, F.P. SCHOENBERG, AND G.E. TITA. *Self-exciting point process modeling of crime*. Journal of the American Statistical Association pages 100–108 (2011). [12](#)
- [31] E. LEWIS, G. MOHLER, P.J. BRANTINGHAM, AND A.L. BERTOZZI. *Self-exciting point process models of civilian deaths in iraq*. Macmillan Publishers Ltd (2011). [12](#)
- [32] Y.S. CHO, A. GALSTYAN, P.J. BRANTINGHAM, AND G. TITA. *Latent self-exciting point process model for spatial-temporal networks*. Discrete and Continuous Dynamical Systems - Series B **19**(5), 1335–1354 (2014). [13](#)
- [33] O. GROTHE, V. KORNIICHUK, AND H. MANNER. *Modeling multivariate extreme events using self-exciting point processes*. Journal of Econometrics (2014). [13](#)
- [34] J.F. OLSON AND K.M. CARLEY. *Exact and approximate em estimation of mutually exciting hawkes processes*. Statistical Inference for Stochastic Processes **16**(1), 63–80 (2013). [13](#)
- [35] C. GUO AND W. LUK. *Accelerating parameter estimation for multivariate self-exciting point processes*. In *Proceedings of the 2014 ACM/SIGDA international symposium on Field-programmable gate arrays*, pages 181–184. ACM (2014). [13](#)

- [36] S. WHEATLEY, V. FILIMONOV, AND D. SORNETTE. *Estimation of the hawkes process with renewal immigration using the em algorithm*. Swiss Finance Institute Research Paper (14-53) (2014). [13](#)
- [37] J. MOLLER AND J.G. RASMUSSEN. *Perfect simulation of hawkes process*. Adv. Appl. Prob. (SGSA) **37**, 629–646 (2005). [13](#)
- [38] A. BRIX AND W.S. KENDALL. *Simulation of cluster point processes without edge effects*. Advances in Applied Probability pages 267–280 (2002). [14](#)
- [39] J. MOLLER AND J.G. RASMUSSEN. *Approximate simulation of hawkes process*. Methodol Comput Appl Probab **8**, 53–64 (2006). [14](#)
- [40] A. DASSIOS AND H. ZHAO. *Exact simulation of hawkes process with exponentially decaying intensity*. Electron. Commun Probab. **18**(62), 1–13 (2013). [14](#)
- [41] M.S. BARTLETT. *The spectral analysis of point process*. J. R. Statist. Soc. B **25**, 264–296 (1963). [15](#)
- [42] F. LORENZEN. *Analysis of Order Clustering Using High Frequency Data : A Point Process Approach*. (2012). [18](#)
- [43] P.K. ANDERSEN, O. BORGAN, R.D. GILL, AND N. KEIDING. *Statistical Models Based on Counting Processes*. Springer (1992). [22](#), [30](#), [31](#), [78](#)
- [44] C. COCOZZA-THIVENT. *Processus stochastiques et fiabilité des systèmes*. Springer (1997). [28](#), [91](#), [94](#)
- [45] M. MOORE. *Inférence statistique dans les processus stochastiques : Aperçu historique*. The Canadian Journal of Statistics **15**(3), 185–207 (1987). [29](#)

- [46] S. MORGENTHALER. *Introduction à la statistique*. Presses polytechniques et universitaires romandes (2007). [30](#)
- [47] P.P.B EGGERMONT AND V.N. LARICCIA. *Maximum penalized likelihood estimation Volume I : Density estimation*. Springer (2001). [30](#)
- [48] A. SLAVIK. *Product integration, its history and application*. History of mathematics **20** (2007). [30](#)
- [49] R.D. GILL AND S. JOHANSEN. *A survey of product-integration with a view toward application in survival analysis*. The annals of statistics pages 1501–1555 (1990). [30](#)
- [50] T. MAKELAINEN, K. SCHMIDT, AND G.P.H. STYAN. *On the existence and uniqueness of the maximum likelihood estimate of a vector-valued parameter in fixed-size samples*. The Annals of Statistics **9**, 758–767 (1981). [33](#), [80](#), [89](#)
- [51] J.C. SPALL. *Introduction to stochastic search and optimization : estimation, simulation, and control*. John Wiley & Sons (2005). [34](#)
- [52] M. MINOUX. *Programation mathématique théorie et algorithmes*. Lavoisier (2008). [34](#)
- [53] B.S. EVERITT. *Introduction to optimization methods and their application in statistics*. Chapan & Hall (1987). [34](#)
- [54] R. MICHEL. *A multidimensional newton-raphson method and its applications to the existence of asymptotic  $f_n$ -estimators and their stochastic expansions*. Journal of Multivariate Analysis **7**, 235–248 (1977). [34](#)
- [55] A. P. DEMPEMSTER, N. M. LAIRD, AND D. B. RUBIN. *Maximum likelihood from incomplete data via the em algorithm*. Journal of the Royal Statistical Society. Series B (Methodological) **39**, 1–38 (1977). [35](#), [78](#)

- 
- [56] G.J. MCLACHLAN AND T. KRISHNAN. *The EM algorithm and extensions*. Wiley Inter-science (1997). 35
- [57] O. CAPPÉ. *Online sequential monte-carlo em algorithm*. IEEE/SP 15th Workshop on statistical signal processing pages 37–40 (2009). 44
- [58] O. CAPPÉ. *Online em algorithm for hidden markov models*. Journal of computational and graphical statistics **20**(3), 728–749 (2011). 44
- [59] L. WASSERMAN. *All of statistics a concise course in statistical inference*. Springer (2004). 48
- [60] A. O’HAGAN AND J. FORSTER. *Bayesian inference*. Arnold (2004). 48, 49
- [61] J. BOREUX, E. PARENT, AND J. BERNIER. *Pratique du calcul bayésien*. Springer (2010). 49
- [62] C. ROBERT. *Le choix bayésien Principes et pratique*. Springer (2006). 49, 54
- [63] M. GUIDA, R. CALABRIA, AND G. PULCINI. *Bayes inference for a non-homogeneous poisson process with power intensity law*. IEEE Transactions on reliabilty **38**(5), 603–609 (1989). 50
- [64] K.B.L. SHAUL, L. LAVI, AND B. REISER. *Bayesian inference for the power law process*. Ann. Inst. Statist. Math. **44**(4), 623–639 (1992). 54
- [65] A. SEN. *Bayesian estimation and prediction of the intensity of the power law process*. J. Statist. Comput. Simul.. **78**(8), 613–631 (2002).
- [66] J.W. YU, G.L. TIAN, AND M.L. TANG. *Predictive analysis for nonhomogeneous poisson process with power law using bayesian approach*. Computational Statistic and Data Analysis **51**, 4254–4268 (2007). 54

- [67] M. PERUGGIA AND T. SANTNER. *Bayesian analysis of time evolution of earthquakes*. Journal of the American Statistical Association **91**(435), 1209–1218 (1996). [54](#)
- [68] T.H. FAN AND E.N. KUO. *A bayesian analysis for the seismic data on taiwan*. Annals of the Institute of Statistical Mathematics **56**(4), 599–609 (2004). [54](#)
- [69] D.R. INSUA, F. RUGGERI, AND M.P. WIPER. *Bayesian Analysis of Stochastic Process Models*. WileyJohn Wiley & Sons, Ltd (2012). [54](#)
- [70] M. FINKELSTEIN. *Failure Rate Modelling for Reliability and Risk*. Springer (2008). [65](#)
- [71] M. ABRAMOWITZ AND I. A. STEGUN. *Handbook of mathematical functions*. National Bureau of Standards (1972). [68](#)
- [72] D. J. DALEY. *Various concepts of orderliness for point processes*. Stochastic Geometry a Tribute to the Memory of Rollo Davidson, John Wiley, New York pages 148–161 (1974). [78](#)
- [73] D.J. DALEY AND D. VERE-JONES. *A summary of the theory of point processes*. Stochastic Point Process : Statistical Analysis, Theory, an Application, Wiley, New York pages 299–283 (1972). [78](#)
- [74] Z. A. LOMNICKI AND S.K. ZAREMABA. *Some applications of zero-one processes*. Journal of the Royal Statistical Society. Series B (Methodological) **17**(2), 243–255 (1955). [78](#)
- [75] L.D. BROWN AND L.H. ZHAO. *A test for the poisson distribution*. The Indian Journal of Statistics. Special issue in memory of D. Basu **64**, 611–625 (2002). [91](#)
- [76] A.W.L. PETER. *Some results on tests for poisson processes*. Biometrika **52**, 67–77 (1965).

- [77] A. YAGOUTI, I. ABI-ZEIDAND, B.M.J. TAHA, AND B.B. OUARDA. *Revue de processus ponctuels et synthèse de tests statistiques pour le choix d'un type de processus*. Revue des sciences de l'eau pages 323–361 (2001). [91](#), [92](#), [95](#)
- [78] S. DACHIAN AND Y.A. KUTOYANTS. *Hypotheses testing : Poisson versus self-exciting*. Board of the Foundation of the Scandinavian Journal of Statistics 2005 **33**, 391–408 (2006). [91](#), [95](#), [96](#), [98](#), [101](#)
- [79] D. FOATA AND A. FUCHS. *Processus stochastiques, Processus de Poisson, chaînes de Markov et martingales*. Dunod (2002). [92](#)
- [80] L. LE CAM. *Locally asymptotically normal families of distributions*. Univ. Calif. Publ. Statist. **3**, 27–98 (1956). [96](#)