



HAL
open science

Construction d'estimateurs pour les processus markoviens déterministes par morceaux

Romain Azaïs

► **To cite this version:**

Romain Azaïs. Construction d'estimateurs pour les processus markoviens déterministes par morceaux. Doctorat. Édition 2024 de l'École de Jeunes Chercheur×se×s en Informatique Mathématique, Nantes, France. 2024. hal-04567166

HAL Id: hal-04567166

<https://hal.science/hal-04567166>

Submitted on 3 May 2024

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.

Construction d'estimateurs pour les processus markoviens déterministes par morceaux

Romain Azais

Ces notes ont été rédigées dans le cadre d'un cours qui sera dispensé lors de l'édition 2024 de l'École de Jeunes Chercheur-se-s en Informatique Mathématique, organisée à Nantes en juin 2024. Elles seront publiées dans l'ouvrage collectif *Informatique Mathématique : une photographie en 2024* (éds Benoît Delahaye, Didier Lime) chez CNRS éditions.

Résumé

Les processus de Markov déterministes par morceaux forment une classe générale de processus stochastiques à temps continu non-diffusifs. Ils sont constitués de trajectoires déterministes ponctuées, à des instants aléatoires, de perturbations aléatoires. Leur dynamique est donc régie par trois caractéristiques locales : le flot déterministe, le taux de saut qui gouverne l'apparition des perturbations aléatoires, et le noyau de transition qui donne la loi des perturbations. Ces processus permettent de modéliser une grande variété de phénomènes dans des champs d'application variés, notamment en biologie. Pour mettre en œuvre ce type de modélisation, il est nécessaire d'estimer, à partir d'observations, les paramètres de la dynamique. C'est ce qu'on se propose de faire dans ces notes, en s'intéressant particulièrement au taux de saut et à des fonctionnelles de dépassement d'un seuil.

1 Introduction

Les processus stochastiques à temps continu sont principalement le résultat de trois ingrédients :

- de la diffusion ;
- des sauts aléatoires ;
- des mouvements déterministes.

Les processus de Markov déterministes par morceaux sont une classe générale de modèles non-diffusifs : ils font intervenir des dynamiques déterministes (régies par des équations différentielles) perturbées par des sauts aléatoires. L'aléa opère donc de manière ponctuelle et non à chaque instant. Il est de deux ordres : les durées des trajectoires déterministes et les perturbations sont aléatoires. On distingue généralement deux types de sauts aléatoires : ceux qui affectent directement le mouvement déterministe (en changeant instantanément sa position) et ceux qui changent les paramètres de l'équation différentielle sous-jacente. Cette classe de processus a été rigoureusement introduite dans [6, 7] (voir également [14]).

De nombreux phénomènes peuvent être modélisés par les processus de Markov déterministes par morceaux, notamment en fiabilité et en assurance mais aussi, comme on va le voir dans la suite de ces notes, en biologie en commençant avec [17]. On dirige le lecteur vers [5, 22] pour des références récentes sur l'emploi de ces modèles pour des considérations biologiques. Ces processus ont été et sont largement étudiés, dans des cas généraux comme pour des instances particulières, dans le domaine des probabilités appliquées : comportement asymptotique, limite d'échelle, contrôle... Mais lorsqu'on souhaite tirer des conclusions qualitatives ou quantitatives pour l'application considérée, il est crucial de connaître les paramètres qui gouvernent les trajectoires observées. C'est une des questions traitées par la statistique des processus.

Si la composante déterministe est bien souvent donnée par des lois physiques ou de la biologie, ce n'est généralement pas le cas des deux aléas qui interviennent dans la dynamique. Il nous faut donc estimer la loi des durées de parcours des équations différentielles (on parle de taux de saut) et la loi des perturbations aléatoires (on parle de noyau de transition). Dans ces notes, on montre, sans trop de technicité, comment construire des estimateurs du taux de saut d'un processus déterministe par morceaux. On s'intéresse aussi d'un point de vue statistique à des événements critiques, comme le dépassement d'un seuil, qui sont les phénomènes d'intérêt typiques dans de nombreuses applications.

2 Notions utiles de théorie des probabilités

L'objet de cette partie est de présenter succinctement les concepts probabilistes dont nous aurons besoin dans la suite de ces notes. Évidemment, elle ne se substitue en aucun cas à un véritable cours sur ces thèmes. De nombreuses références sont disponibles dans la littérature, par exemple [11, 8] (et [18] pour les théorèmes limites des chaînes de Markov sur un espace d'état général).

2.1 Variables aléatoires à densité

Lorsque la loi d'une variable aléatoire X à valeur dans \mathbf{R}^d est diffuse, i.e. $\mathbf{P}(X = x) = 0$ pour tout $x \in \mathbf{R}^d$, alors il existe une fonction (suffisamment régulière) $f_X : \mathbf{R}^d \rightarrow \mathbf{R}_+$, appelée densité de X , telle que

$$\mathbf{P}(X \in A) = \int_A f_X(x) dx.$$

Deux lois à densité joueront un rôle important dans la suite :

- la loi exponentielle $\mathcal{E}(\lambda)$ sur \mathbf{R}_+ dont la densité est donnée par $x \mapsto \lambda \exp(-\lambda x)$;
- la loi normale $\mathcal{N}(m, \sigma^2)$ sur \mathbf{R} dont la densité est donnée par $x \mapsto \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-m)^2}{\sigma^2}\right)$.

Pour toute fonction g , la valeur moyenne théorique de $g(X)$, appelée espérance de $g(X)$, est donnée par

$$\mathbf{E}[g(X)] = \int_{\mathbf{R}^d} g(x) f_X(x) dx$$

(tant que l'intégrale est bien définie). En particulier, la moyenne $\mathbf{E}[X]$ d'une variable aléatoire réelle X vérifie $\mathbf{E}[X] = \int_{\mathbf{R}} x f_X(x) dx$ alors que sa variance $\mathbf{Var}(X)$ est donnée par $\mathbf{Var}(X) = \mathbf{E}[(X - \mathbf{E}[X])^2]$. Notons que les paramètres m et σ^2 de la loi normale désignent sa moyenne et sa variance.

On considère maintenant deux variables aléatoires X et Y : on dit qu'elles sont indépendantes si la densité du couple est le produit des densités, i.e.

$$f_{X,Y}(x, y) = f_X(x) f_Y(y).$$

Dans le cas où X et Y ne sont pas indépendantes, on peut considérer la loi conditionnelle de Y sachant X ,

$$\mathbf{P}(Y \in A | X = x) = \int_A f_{Y|X}(y|x) dy,$$

où, avec la convention $\frac{0}{0} = 0$,

$$f_{Y|X}(y|x) = \frac{f_{X,Y}(x, y)}{f_X(x)}.$$

Il est intéressant de noter que, dans le cas indépendant, on retrouve $f_{Y|X} = f_Y$. La covariance de X et Y (mesure de leur dépendance linéaire) est donnée par

$$\mathbf{Cov}(X, Y) = \mathbf{E}[(X - \mathbf{E}[X])(Y - \mathbf{E}[Y])]$$

de telle sorte que $\mathbf{Var}(X) = \mathbf{Cov}(X, X)$.

2.2 Suites de variables aléatoires

Soit $(X_n)_{n \in \mathbf{N}}$ une suite de variables aléatoires. Il existe plusieurs façons de donner du sens à la convergence d'une telle suite. On en donne deux ci-après. S'il existe une variable aléatoire X_∞ telle que $\mathbf{P}(X_n \rightarrow X_\infty) = 1$, on dit que $(X_n)_{n \in \mathbf{N}}$ converge presque sûrement vers X_∞ , ce qu'on note $X_n \xrightarrow{p.s.} X_\infty$. Si pour toute fonction continue bornée g , on a $\mathbf{E}[g(X_n)] \rightarrow \mathbf{E}[g(X_\infty)]$ alors la suite $(X_n)_{n \in \mathbf{N}}$ converge en loi vers X_∞ , ce qu'on note $X_n \xrightarrow{\mathcal{L}} X_\infty$.

Lorsqu'on considère une suite $(X_n)_{n \in \mathbf{N}}$ de variables aléatoires indépendantes et de même loi f_X (dont on note m la moyenne et σ^2 la variance), la suite des moyennes empiriques $(n^{-1}S_n)_{n \in \mathbf{N}^*}$, où $S_n = \sum_{i=0}^{n-1} X_i$, vérifie deux théorèmes importants de convergence :

— la loi forte des grands nombres,

$$\frac{S_n}{n} \xrightarrow{p.s.} m ;$$

— le théorème central limite,

$$\frac{S_n - nm}{\sqrt{n}\sigma} \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1).$$

2.3 Chaînes de Markov

Une chaîne de Markov sur \mathbf{R}^d est une suite de variables aléatoires $(X_n)_{n \in \mathbf{N}}$ satisfaisant la propriété suivante,

$$\mathbf{P}(X_{n+1} \in A | X_n = x_n, \dots, X_0 = x_0) = \mathbf{P}(X_{n+1} \in A | X_n = x_n),$$

i.e. le futur ne dépend du passé qu'à travers le présent. Lorsque cette quantité ne dépend pas de n , la chaîne est dite homogène : elle est alors caractérisée par sa loi initiale μ_0 et son noyau de transition P (supposé à densité ici) qui est la loi conditionnelle de X_{n+1} sachant X_n ,

$$\mathbf{P}(X_{n+1} \in A | X_n = x) = \int_A P(y|x) dy.$$

En itérant P , on peut calculer la loi f_{X_n} de X_n pour n'importe quel n ,

$$f_{X_n}(x) = \int_{\mathbf{R}^d} P(x|u) f_{X_{n-1}}(u) du,$$

avec $f_{X_0} = \mu_0$.

On suppose désormais qu'il existe une unique loi μ_∞ telle que

$$\int_{\mathbf{R}^d} \mu_\infty(x) P(y|x) dx = \mu_\infty(y).$$

Cette loi est appelée stationnaire (ou invariante) pour la raison suivante : si $\mu_0 = \mu_\infty$ alors, pour tout n , $f_{X_n} = \mu_\infty$. Sous certaines conditions techniques qu'on ne détaillera pas ici, quelle que soit sa loi initiale μ_0 , la chaîne converge vers sa loi stationnaire μ_∞ ,

$$X_n \xrightarrow{\mathcal{L}} \mu_\infty,$$

et on a un équivalent des théorèmes de convergence énoncés ci-dessus pour la somme de variables aléatoires indépendantes et de même loi :

— le théorème ergodique presque sûr,

$$\frac{1}{n} \sum_{i=0}^{n-1} g(X_i) \xrightarrow{p.s.} \int_{\mathbf{R}^d} g(x) \mu_\infty(x) dx ;$$

— un théorème central limite,

$$\sqrt{n} \left(\frac{1}{n} \sum_{i=0}^{n-1} g(X_i) - \int_{\mathbf{R}^d} g(x) \mu_\infty(x) dx \right) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma^2),$$

où la variance limite σ^2 vérifie

$$\sigma^2 = \mathbf{Var}(g(Y_0)) + 2 \sum_{k=1}^{+\infty} \mathbf{Cov}(g(Y_0), g(Y_k)),$$

pour une chaîne de Markov $(Y_n)_{n \in \mathbf{N}}$ de loi initiale μ_∞ et de noyau P indépendante de $(X_n)_{n \in \mathbf{N}}$.

3 Processus de Markov déterministes par morceaux

Dans cette partie, on introduit sans ambages techniques les processus de Markov déterministes par morceaux comme une généralisation des processus de renouvellement, eux-mêmes vus comme extension du processus de Poisson. Là aussi, de nombreuses références peuvent être consultées à profit, notamment [8, 7].

3.1 Processus de Poisson

Un processus de Poisson $(X_t)_{t \in \mathbf{R}_+}$ d'intensité $\lambda > 0$ est un processus à temps continu et à valeur entière défini de la façon suivante :

- ses trajectoires sont continues à droite et telles que $X_t - X_{t-} \in \{0, 1\}$;
- ses accroissements $X_{t+s} - X_t$ sont distribués selon la loi de Poisson de paramètre λs ,

$$\mathbf{P}(X_{t+s} - X_t = k) = \frac{(\lambda s)^k}{k!} \exp(-\lambda s);$$

- ses accroissements $(X_{t_i} - X_{t_{i-1}})_{1 \leq i \leq n}$ sont indépendants.

On peut montrer qu'un tel processus partant de $X_0 = 0$ admet une représentation de la forme

$$X_t = \sum_{n \geq 1} \mathbf{1}\{T_n \leq t\}, \tag{1}$$

où la suite $(T_n)_{n \geq 1}$ (suite des instants où le processus voit sa valeur augmenter de 1) vérifie la condition suivante : en notant $T_0 = 0$, les durées $S_n = T_n - T_{n-1}$ sont indépendantes et de loi exponentielle de paramètre λ .

Sans entrer dans les détails, il nous faut remarquer que le processus $(X_t)_{t \in \mathbf{R}_+}$ est un processus de Markov. Cette caractéristique vient de l'absence de mémoire de la loi exponentielle qu'on peut écrire ainsi,

$$\forall 0 \leq s \leq t, \mathbf{P}(S_n > t | S_n > s) = \mathbf{P}(S_n > t - s).$$

Le processus qu'on vient de décrire est dit homogène : toutes les durées inter-sauts S_n sont distribués sous la même loi. On peut rendre le processus inhomogène en considérant une intensité dépendante du temps $\lambda : \mathbf{R}_+ \rightarrow \mathbf{R}_+$ et en supposant que les accroissements $X_{t+s} - X_t$ suivent la loi de Poisson de paramètre $\int_t^{t+s} \lambda(u) du$. On perd alors l'indépendance et le caractère identiquement distribué des durées inter-sauts. Mais il existe une autre généralisation du processus de Poisson qui permet de conserver ces propriétés.

3.2 Processus de renouvellement

Un processus de renouvellement $(X_t)_{t \in \mathbf{R}_+}$ est un processus de la forme (1) pour lequel les durées inter-sauts $S_n = T_n - T_{n-1}$ sont indépendantes et identiquement distribuées. D'après ce qui précède, on retrouve le processus de Poisson homogène lorsqu'elles sont exponentielles.

On peut écrire l'équation d'un processus de renouvellement ainsi,

$$X_t = \sum_{n \geq 0} Z_n \mathbf{1}\{T_n \leq t < T_{n+1}\},$$

où $Z_n = n$. Lors d'un saut, comme c'est le cas du processus de Poisson, un processus de renouvellement voit sa valeur augmenter d'une unité. On peut généraliser cet aspect en introduisant la notion de marque, c'est-à-dire en choisissant une autre suite $(Z_n)_{n \in \mathbf{N}}$ que la suite des entiers. On peut supposer par exemple qu'elle suit une dynamique markovienne sur \mathbf{R}^d . Enfin, et c'est tout l'intérêt d'introduire des marques, on peut faire dépendre la loi des durées inter-sauts de la marque : la loi de S_{n+1} , durée qui sépare T_{n+1} de T_n , dépend de la marque courante Z_n . Finalement, si on note P le noyau des marques et $f_{S|Z}$ la densité conditionnelle des durées inter-sauts, la dynamique peut être décrite itérativement ainsi. À l'instant T_n , le processus est en $X_{T_n} = Z_n$, alors :

- la durée inter-sauts $S_{n+1} = T_{n+1} - T_n$ suit $f_{S|Z}(\cdot|Z_n)$;
- entre T_n et T_{n+1} (exclu), le processus est constant de valeur Z_n ;
- la nouvelle marque Z_{n+1} est tirée selon $P(\cdot|Z_n)$ et $X_{T_{n+1}} = Z_{n+1}$.

On obtient un processus de renouvellement marqué dont la suite des marques $(Z_n)_{n \in \mathbf{N}}$ est markovienne : elle est appelée chaîne immergée du processus.

Si la suite des marques est markovienne, le processus de renouvellement lui-même n'est markovien que lorsque les durées inter-sauts sont exponentielles (dont le paramètre peut dépendre de la marque courante), à nouveau par la propriété d'absence de mémoire de la loi exponentielle. Dans le cas général, on peut toujours écrire

$$\mathbf{P}(S_{n+1} > t|Z_n) = \exp\left(-\int_0^t \lambda(s|Z_n) ds\right),$$

où la fonction λ est donnée par

$$\lambda(t|z) = \frac{f_{S|Z}(t|z)}{\int_t^\infty f_{S|Z}(s|z) ds}. \quad (2)$$

Pour la loi exponentielle, on peut voir aisément que la fonction $\lambda(\cdot|z)$ est constante et on note $\lambda(z)$ sa valeur. Par conséquent, les durées inter-sauts d'un processus de renouvellement markovien sont données par

$$\mathbf{P}(S_{n+1} > t|Z_n) = \exp(-t\lambda(Z_n)),$$

et la fonction $z \mapsto \lambda(z)$ est appelée taux de saut du processus.

3.3 Processus de Markov déterministe par morceaux

Entre les sauts, la dynamique d'un processus de renouvellement est constante. On propose de généraliser cet aspect tout en conservant le caractère déterministe de cette partie de la dynamique. Dans ce but, on se donne un flot $\varphi(\cdot|\cdot)$ sur $\mathbf{R} \times \mathbf{R}^d$ satisfaisant

$$\varphi(s|\varphi(t|x)) = \varphi(s+t|x).$$

En clair, $\varphi(t|x)$ désigne la solution à l'instant t d'une équation différentielle dont la condition au temps 0 est x . Entre les instants T_n et T_{n+1} , au lieu d'être constant égal à Z_n , le processus va suivre le flot $\varphi(\cdot|Z_n)$. Mais on souhaite aussi préserver le caractère markovien de la dynamique à temps continu, ce qui a deux incidences.

- D'après ce qui précède sur les processus de renouvellement, toutes les lois de durées inter-sauts ne sont pas admises : il nous faut conserver l'idée d'un taux de saut dépendant de la valeur du processus et non du temps. Pour jouer ce rôle, on se donne une fonction $\lambda : \mathbf{R}^d \rightarrow \mathbf{R}_+$.
- Pour un processus de renouvellement, la position $X_{T_{n+1}} = Z_{n+1}$ à l'instant T_{n+1} est tirée en appliquant le noyau P à la marque précédente Z_n . Le caractère markovien du processus à temps continu vient aussi du fait que cela revient au même que d'appliquer P à la position juste avant le saut $X_{T_{n+1}^-}$. Cette propriété est évidemment perdue si le processus varie entre les instants de saut.

Un processus de Markov déterministe par morceaux est défini à partir de trois caractéristiques : son noyau de transition P , son taux de saut λ et son flot φ . Sa dynamique peut être décrite itérativement ainsi. À l'instant T_n , le processus est en $X_{T_n} = Z_n$, alors :

- la durée inter-sauts $S_{n+1} = T_{n+1} - T_n$ est tirée selon

$$\mathbf{P}(S_{n+1} > t | Z_n) = \exp\left(-\int_0^t \lambda(\Phi(s|Z_n)) ds\right) ;$$

- entre les instants T_n et T_{n+1} (exclu), le processus évolue selon le flot,

$$\forall 0 \leq t < S_{n+1}, X_{T_n+t} = \varphi(t|Z_n),$$

et on note $Z_{n+1}^- = X_{T_{n+1}^-}$;

- la position Z_{n+1} à l'instant T_{n+1} est tirée selon P appliqué à la position juste avant le saut,

$$\mathbf{P}(Z_{n+1} \in A | Z_{n+1}^-) = \int_A P(x | Z_{n+1}^-) dx.$$

Par construction, la dynamique est à temps continu mais l'aléa intervient de manière discrète : il est notamment contenu dans le processus $(Z_n, S_{n+1})_{n \in \mathbf{N}}$ qui est lui-même une chaîne de Markov. Si on connaît le flot, toute la trajectoire peut être reconstruite à partir de la connaissance de cette chaîne immergée.

3.4 Exemples en biologie

On s'intéresse d'abord au phénomène de croissance d'une bactérie suivie de division en bactéries filles [9]. À partir de sa naissance, on suppose qu'une bactérie croît de manière déterministe mais pendant une durée aléatoire avant de se diviser en deux bactéries filles qui suivront chacune une dynamique analogue. Si on regarde une seule branche de l'arbre formé par la descendance de la bactérie initiale, on obtient une succession de croissances déterministes suivies de fragmentations qui ont lieu à des instants aléatoires. Quant à la loi des fragmentations elles-mêmes, on peut supposer qu'elles sont déterministes (par exemple, la bactérie fille a exactement la moitié de la taille de sa mère) ou aléatoires (même exemple avec un bruit autour de la taille cible).

Cette évolution peut être décrite par un processus de Markov déterministe par morceaux sur \mathbf{R}_+ : on note X_t la taille de la bactérie présente le long de la branche suivie à l'instant t . Le flot déterministe est donné par $\varphi(t|x) = x + rt$ dans le cas de croissances linéaires ou $\varphi(t|x) = x \exp(rt)$ dans le cas de croissances exponentielles. Lorsque les fragmentations divisent la bactérie mère en parts égales, on a $P(\cdot|x) = \delta_{\{x/2\}}$. Enfin, le taux de saut n'est pas spécifié ici : il s'agit typiquement de la quantité d'intérêt d'un point de vue statistique [9, 15, 16].

Partant du même modèle, on peut imaginer que X_t représente la taille d'une population qui subit à des instants aléatoires des catastrophes (épidémies, famines, guerres) qui déciment en un temps très court une partie de sa population. Si on note γ la loi sur $[0, 1]$ de la proportion de la population toujours en vie après l'événement redouté, on peut écrire le noyau de transition de la façon suivante,

$$P(y|x) = \frac{1}{x} \gamma\left(\frac{y}{x}\right). \quad (3)$$

Si on suppose de plus que la taille de la population n'a pas d'effet sur l'occurrence des catastrophes, on est amené à prendre λ constant. Enfin, on peut supposer que la croissance de la population n'a lieu que si le nombre d'individus est suffisant. Dans le cas contraire, la population finit par s'éteindre à cause des catastrophes successives. On obtient finalement le modèle décrit dans [2] où le flot s'écrit

$$\varphi(t|x) = \begin{cases} (x-1)\exp(rt) + 1 & \text{si } x > 1, \\ x & \text{sinon.} \end{cases} \quad (4)$$

D'un point de vue statistique, c'est la probabilité d'extinction et son instant qui sont les quantités d'intérêt.

On présente maintenant un modèle de pharmacocinétique [4] qui est en quelque sorte le modèle de croissance-fragmentation précédent (sans le seuil) pris à rebours. Ici X_t modélise la quantité d'un contaminant alimentaire présent dans l'organisme à l'instant t . Le métabolisme l'élimine au cours du temps selon le flot $\varphi(t|x) = x \exp(-rt)$. Mais à des instants aléatoires, l'individu en ingère une quantité aléatoire distribuée selon une densité γ , de telle sorte que le noyau P est donné par

$$P(y|x) = \gamma(y-x).$$

S'il n'y a pas de raison de supposer que les ingestions de poison sont en lien avec la quantité de contaminant dans l'organisme, on peut admettre que le taux de saut est constant. Ici aussi, les quantités d'intérêt sont typiquement la probabilité de dépassement d'un seuil de dangerosité et la loi du temps passé au-dessus de celui-ci.

4 Cadre statistique

4.1 Schémas d'observation

En statistique des processus, on cherche généralement à estimer les paramètres du modèle ou des fonctionnelles des trajectoires, c'est-à-dire à construire, à partir d'observations, des approximations de ces quantités. On souhaite de plus obtenir des garanties sur ces estimateurs, comme leur convergence vers la quantité cible quand le nombre d'observations tend vers l'infini.

Dans le cas des processus de Markov déterministes par morceaux, et comme illustré dans les exemples ci-dessus, on peut notamment s'intéresser à l'estimation du taux de saut λ et à la loi de dépassement d'un seuil critique. La suite de ces notes abordera ces deux problèmes. Mais avant d'aller plus loin, il nous faut décrire le schéma d'observation du processus choisi.

Tout d'abord, on peut se demander si on observe une seule trajectoire ou plusieurs trajectoires du processus. Dans le second cas, le comportement asymptotique des estimateurs peut être regardé en le nombre de trajectoires observées. Dans le premier, c'est plutôt l'horizon d'observation de la trajectoire qu'on va faire tendre vers l'infini pour étudier les estimateurs. Dans les deux cas, il faut aussi s'interroger sur le mode d'observation des trajectoires : sont-elles inspectées à temps continu (on connaît X_t pour tout t) ou sur une grille temporelle (on connaît X_{t_i} pour certains t_i fixés)? Lorsqu'on regarde le processus sur une grille, les instants de saut sont inconnus même si, en tout cas dans des cas simples, on a une bonne idée de là où ils se situent. Il est important de noter que cette incertitude entache nécessairement les estimateurs d'une erreur supplémentaire par rapport à l'observation à temps continu.

Dans toute la suite, on suppose qu'on observe une seule trajectoire à temps continu. L'estimation des paramètres du flot ne présente aucune difficulté (en tout cas si sa forme est connue) par des techniques d'interpolation. On peut toutefois remarquer que le flot est plus difficile à estimer lorsque la trajectoire est observée à travers un bruit, mais ce n'est pas l'hypothèse qu'on fait dans ces notes. On suppose donc la dynamique déterministe connue. Observer la trajectoire à temps continu revient alors à observer les durées inter-sauts et les positions lors des sauts, c'est-à-dire la chaîne immergée $(Z_n, S_{n+1})_{n \in \mathbf{N}}$. Finalement, il reste à fixer l'horizon d'observation : celui-ci peut être sur le temps du processus à temps continu (on observe

$(X_t)_{t \in [0, T]}$) ou sur le temps du processus à temps discret (on observe $(Z_n, S_{n+1})_{0 \leq n \leq N-1}$). C'est la seconde option qui est privilégiée dans la suite de ce document.

En somme, on observe les N premiers pas $(Z_n, S_{n+1})_{0 \leq n \leq N-1}$ de la chaîne immergée d'une trajectoire d'un processus de Markov déterministe par morceaux. De plus, la seule quantité supposée connue est le flot déterministe.

4.2 Estimation à noyau

Avant de passer aux processus qui nous intéressent, on regarde le cas plus facile de l'estimation de la densité de données indépendantes et identiquement distribuées. On se donne donc N variables aléatoires (X_1, \dots, X_N) indépendantes et de même densité f sur \mathbf{R}^d et on se demande comment l'estimer, c'est-à-dire comment construire une fonction des X_i qui approche f , avec des garanties lorsque N tend vers l'infini. Ce type d'approche est non-paramétrique car ne fait aucune hypothèse de forme sur la quantité inconnue.

Une solution à ce problème est l'estimateur à noyau de Parzen-Rosenblatt [19, 20], qu'on peut voir comme une généralisation de l'histogramme. Étant donné un noyau K (c'est-à-dire une fonction d'intégrale 1), il s'écrit

$$\hat{f}_N(x) = \frac{1}{Nh_N^d} \sum_{i=1}^N K\left(\frac{X_i - x}{h_N}\right),$$

où la fenêtre h_N vérifie $h_N \rightarrow 0$ et $Nh_N \rightarrow \infty$. Sans entrer dans des considérations techniques, cet estimateur converge ponctuellement vers sa cible, i.e. $\hat{f}_N(x) \xrightarrow{p.s.} f(x)$, et vérifie un théorème central limite du type

$$\sqrt{Nh_N^d} \left(\hat{f}_N(x) - f(x) \right) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma^2(x)),$$

pour une certaine variance limite $\sigma^2(x)$.

Lorsqu'on souhaite calculer l'estimateur au fur et à mesure qu'on obtient les données, il peut être avantageux de construire une version récursive de cet estimateur, comme notamment l'estimateur de Wolverton-Wagner [21] qui s'écrit

$$\hat{f}_N(x) = \frac{1}{N} \sum_{i=1}^N \frac{1}{h_i^d} K\left(\frac{X_i - x}{h_i}\right),$$

et qui vérifie la formule de récursion

$$\hat{f}_N(x) = \frac{N-1}{N} \hat{f}_{N-1}(x) + \frac{1}{Nh_N^d} K\left(\frac{X_N - x}{h_N}\right).$$

Lorsque la fenêtre est de la forme $h_N = N^{-\alpha}$ avec $0 < \alpha < 1$, cet estimateur aussi converge ponctuellement vers sa cible et vérifie un théorème central limite à la même vitesse $\sqrt{Nh_N^d} = N^{(1-\alpha d)/2}$ lorsque αd n'est pas trop proche de 0.

Supposons maintenant que les données dont on dispose ne sont plus indépendantes mais sont markoviennes et convergent vers leur loi invariante. Il n'est pas question ici d'entrer dans les hypothèses techniques (nécessairement plus restrictives) qui rendent valides les résultats, mais on peut retenir que les deux estimateurs ci-dessus vérifient des résultats analogues (avec notamment les mêmes vitesses dans les normalités asymptotiques) pour l'estimation de la loi invariante [10].

Lorsqu'on sera amené à estimer la densité de données indépendantes ou la loi invariante de données markoviennes, on utilisera l'un ou l'autre des estimateurs ci-dessus en ayant en tête qu'on dispose de certaines garanties sur leur comportement asymptotique.

5 Estimation non-paramétrique du taux de saut

Cette partie est dédiée à l'estimation non-paramétrique du taux de saut d'un processus de Markov déterministe par morceaux. La clé des approches ci-dessous réside dans la façon d'écrire la quantité d'intérêt en fonction de la loi invariante de l'une ou l'autre des chaînes immergées du processus.

5.1 En dimension 1

On commence ici avec l'approche suivie dans [15, 16] où on suppose que le processus observé évolue sur \mathbf{R}_+ . De plus, on fait l'hypothèse que le flot définit un changement de variable. L'application typique est le modèle de croissance-fragmentation de bactérie décrit auparavant.

Dans un premier temps [15], on fait l'hypothèse que les sauts sont purement déterministes, i.e. $P(\cdot|x) = \delta_{\{f(x)\}}$, où la fonction f , supposée connue, est croissante sur \mathbf{R}_+ et définit également un changement de variable. Alors, la composée $f \circ \varphi(\cdot|x)$ définit elle aussi un changement de variable et on peut exprimer le taux de saut ainsi,

$$\lambda(y) = \frac{\mu_\infty(f(y))}{\mathbf{E}[g_{Z'_0}(f(y))\mathbf{1}\{f(Z'_0) \leq f(y)\}\mathbf{1}\{Z'_1 \geq f(y)\}]}, \quad (5)$$

où μ_∞ désigne la loi invariante de $(Z_n)_{n \in \mathbf{N}}$,

$$g_x(y) = [(f \circ \varphi(\cdot|x))'((f \circ \varphi(\cdot|x))^{-1}(y))]^{-1},$$

et $(Z'_n)_{n \in \mathbf{N}}$ est une chaîne de Markov de même noyau de transition que $(Z_n)_{n \in \mathbf{N}}$ mais débutant sous la loi stationnaire μ_∞ .

La forme (5) permet d'obtenir aisément un estimateur de λ : le numérateur est la loi invariante d'une chaîne de Markov qu'on sait estimer par des méthodes à noyau, alors que le dénominateur est une valeur moyenne sous la loi invariante qu'on sait estimer par le théorème ergodique. L'autrice montre notamment des résultats de convergence en norme \mathbf{L}^2 lorsque l'estimateur de Parzen-Rosenblatt est employé pour l'estimation de μ_∞ .

Désormais on relâche l'hypothèse sur les sauts du processus. Les calculs ne permettent plus d'obtenir une forme aussi simple que (5). En particulier, on doit maintenant conserver les positions juste avant les sauts. D'après [16], on peut exprimer le taux de saut ainsi,

$$\lambda(y) = \frac{\mu_\infty^-(y)}{\mathbf{E}[(\varphi(\cdot|Z'_0)^{-1})'(y)\mathbf{1}\{Z'_0 \leq y < Z'_1\}]}, \quad (6)$$

où μ_∞^- désigne la loi invariante de la chaîne $(Z_n^-)_{n \in \mathbf{N}^*}$. Pour les mêmes raisons que précédemment, on peut déduire de cette expression un estimateur de λ même si, cette fois, le dénominateur fait intervenir la loi invariante d'un processus de dimension 2. Dans [16], les autrices ont montré pour cet estimateur des résultats de convergence en norme \mathbf{L}^2 .

On considère maintenant la méthode développée dans [12] sous des hypothèses semblables mais lorsque l'observation est faite sur une fenêtre à temps continu fixée, i.e. lorsqu'on regarde la trajectoire continue entre les instants 0 et T , et non un certain nombre de pas de la chaîne immergée (le nombre de sauts observés est donc aléatoire). L'auteur suppose que le flot s'écrit

$$\varphi(t|x) = x + \int_0^t \Delta(\varphi(s|x))ds,$$

ce qui permet d'écrire, pour faire le lien avec les estimateurs précédents, $\Delta(y) = \varphi(\varphi(\cdot|x)^{-1}(y)|x)'$ ou encore $(\varphi(\cdot|x)^{-1})' = 1/\Delta$ (quantités indépendantes de x). L'auteur considère alors le temps local du processus,

$$r(x) = \frac{\#\{t \in [0, T] : X_t = x\}}{\Delta(x)},$$

qui fournit, une fois normalisé, un estimateur de la loi invariante du processus à temps continu $(X_t)_{t \in \mathbf{R}_+} : r(x)/T$. À partir de l'estimateur de Parzen-Rosenblatt $\widehat{\mu}_{\infty, T}^-$ de la loi invariante de $(Z_n^-)_{n \in \mathbf{N}^*}$ (défini sur une fenêtre à temps continu contrairement à ce qui précède),

$$\widehat{\mu}_{\infty, T}^-(y) = \frac{1}{Th_T} \sum_{\{n \in \mathbf{N}^* : T_n \leq t\}} K\left(\frac{Z_n^- - y}{h_T}\right),$$

il construit l'estimateur suivant de $\lambda(y)$,

$$\widehat{\lambda}_T(y) = \frac{T\widehat{\mu}_{\infty, T}^-(y)}{r(y)}.$$

On peut premièrement remarquer que $\widehat{\mu}_{\infty, T}^-(y)$ estime le numérateur de l'expression (6) de $\lambda(y)$, alors qu'avec un jeu de réécriture on voit que

$$\frac{r(y)}{T} = \frac{(\varphi(\cdot|x)^{-1})'(y)}{T} \sum_{\{n \in \mathbf{N}^* : T_n \leq t\}} \mathbf{1}\{X_{T_{n-1}} \leq y\} \mathbf{1}\{X_{T_n} \geq y\}$$

estime son dénominateur. Par conséquent, par des approches très différentes, [16] et [12] aboutissent au même estimateur du taux de saut, le premier pour une observation sur une fenêtre à temps discret, l'autre pour une fenêtre à temps continu. Dans [12], l'auteur prouve de plus, dans le cadre de son schéma d'observation, un résultat de convergence en norme infinie.

5.2 En dimension $d \geq 2$

On présente désormais l'approche très différente suivie dans [3], qui consiste à passer par les lois conditionnelles des durées inter-sauts. Cette méthode est moins directe mais a l'avantage de fonctionner quelque soit la dimension de l'espace. Si on note $f_{S|Z}$ la densité conditionnelle de la durée S_{n+1} sachant Z_n , on peut toujours écrire $f_{S|Z}(t|z) = \frac{\mu_{\infty}(z)f_{S|Z}(t|z)}{\mu_{\infty}(z)}$. Cette tautologie fournit une stratégie d'estimation de $f_{S|Z}$ puisque le dénominateur désigne la loi invariante de $(Z_n)_{n \in \mathbf{N}}$ (qu'on sait estimer) et le numérateur correspond à la loi invariante du couple $(Z_n, S_{n+1})_{n \in \mathbf{N}}$ (qu'on sait donc aussi estimer même si c'est en dimension $d + 1$). Toutefois, il reste ensuite à revenir au taux de saut d'intérêt λ .

Dans ce but, on remarque que la composée $\lambda \circ \varphi$ s'écrit

$$\lambda \circ \varphi(t|z) = \frac{f_{S|Z}(t|z)}{G_{S|Z}(t|z)},$$

où $G_{S|Z}(t|z) = \mathbf{P}(S_{n+1} > t | Z_n = z)$ peut s'estimer par une technique similaire à celle décrite pour $f_{S|Z}$. On peut remarquer au passage qu'il s'agit en fait de la même formule que (2). Cette expression nous fournit par conséquent une technique d'estimation de $\lambda \circ \varphi$ pour laquelle sont montrés dans [3] la convergence presque sûre et un théorème central limite dont la variance asymptotique vérifie

$$\sigma^2(t|z) \propto \frac{\lambda \circ \varphi(t|z)}{\mu_{\infty}(z)G_{S|Z}(t|z)}.$$

Afin d'estimer $\lambda(x)$, on peut dès lors remarquer que, pour tout couple (z, t) tel que $x = \varphi(t|z)$, on a un estimateur candidat puisque $\lambda \circ \varphi(t|z) = \lambda(x)$. De plus, on sait que chacun converge vers la quantité cible. Afin de dégager le meilleur estimateur parmi cette infinité de candidats, on propose de choisir celui de variance asymptotique minimale, donc celui qui maximise $\mu_{\infty}(z)G_{S|Z}(t|z)$ puisque le numérateur de la variance ne dépend pas du couple (z, t) . Cette quantité est inconnue mais a déjà été estimée précédemment. Finalement, cette technique permet d'extraire numériquement le meilleur estimateur de $\lambda(x)$ au sens de la variance limite.

Remarquons en passant que le noyau de transition P désigne la loi conditionnelle de Z_n sachant Z_n^- et peut donc s'estimer en écrivant $P(y|x) = \frac{\mu_\infty^-(x)P(y|x)}{\mu_\infty^-(x)}$, où le dénominateur désigne la loi invariante de $(Z_n^-)_{n \in \mathbf{N}^*}$ et le numérateur correspond à la loi invariante du couple $(Z_n^-, Z_n)_{n \in \mathbf{N}^*}$, qu'on sait toutes deux estimer par des méthodes à noyau [1].

6 Dépassement d'un seuil critique

6.1 Approche numérique

On regarde le modèle de population sujette à des catastrophes ponctuelles déjà décrit ci-dessus. Il s'agit d'un processus de Markov déterministe par morceaux dont le taux de saut λ est constant, le flot donné par (4) et le noyau par (3). Son extinction réfère au passage sous le seuil critique 1. Les quantités d'intérêt ici sont la probabilité d'extinction

$$\rho(x) = \mathbf{P}(\exists n \in \mathbf{N}, Z_n < 1 | Z_0 = x)$$

et la loi de l'instant d'extinction

$$\tau = \min\{n \in \mathbf{N}^* : Z_{n-1} \geq 1 \text{ et } Z_n < 1\}$$

conditionnellement à la position initiale du processus, i.e.

$$\eta_n(x) = \mathbf{P}(\tau = n | Z_0 = x).$$

Il est important de noter que l'extinction du processus à temps continu ne peut avoir lieu que lors d'un saut puisque les parties déterministes sont croissantes. À proprement parler, τ est en fait le nombre de catastrophes qui mènent à l'extinction, c'est-à-dire l'instant d'extinction de la chaîne $(Z_n)_{n \in \mathbf{N}}$, alors que T_τ est l'instant d'extinction du processus à temps continu $(X_t)_{t \in \mathbf{R}_+}$. A contrario, qu'elle soit regardée à temps continu ou à temps discret, la probabilité d'extinction est la même.

Pour accéder à ces quantités, on peut chercher à les exprimer comme fonctions des paramètres du modèle. La clé est de remarquer que la probabilité d'extinction est l'unique solution de l'équation intégrale

$$\rho(x) = \int_0^1 R(y|x)dy + \int_1^\infty \rho(y)R(y|x)dy,$$

sous les conditions $\lim_{x \rightarrow 1} \rho(x) = 1$ et $\lim_{x \rightarrow \infty} \rho(x) = 0$, où R désigne le noyau de transition de la chaîne de Markov $(Z_n)_{n \in \mathbf{N}}$, qu'on peut aisément écrire en fonction de λ , r et γ ,

$$R(y|x) = \begin{cases} \frac{1}{x} \gamma \left(\frac{y}{x} \right) & \text{si } x \leq 1, \\ \frac{\lambda}{r} (x-1)^{\lambda/r} \int_0^{y/x \wedge 1} \gamma(u) u^{\lambda/r} (y-u)^{-\lambda/r-1} du & \text{sinon.} \end{cases}$$

Comme on ne sait pas résoudre analytiquement cette équation sans plus d'hypothèses, on considère une approche numérique. Avec l'opérateur intégral

$$K : h \mapsto \int_1^\infty h(y)R(y|x)dy,$$

et en notant $s(x) = \int_0^1 R(y|x)dy$, l'équation se réécrit $\rho - K\rho = s$ et on peut approcher ρ par

$$\rho_M = \sum_{k=0}^M K^k s, \tag{7}$$

avec la propriété voulue de convergence dès lors que la norme d'opérateur de K est strictement inférieure à 1, ce qui est garanti lorsque

$$\int_0^1 \frac{\gamma(u)}{u} du < 1 + \frac{r}{\lambda}.$$

On peut suivre la même approche pour traiter $\eta_n(x)$ et voir que

$$\eta_n(x) = \int_1^\infty \eta_{n-1}(y)R(y|x)dy = K\eta_{n-1}(x).$$

En remarquant de plus que $\eta_1 = s$, on a $\eta_n = K^{n-1}s$. Finalement, le terme d'indice k de l'approximation (7) est $\eta_{k+1} = K^k s$ et les deux quantités peuvent être calculées en même temps.

On sait maintenant exprimer les quantités d'intérêt en fonction du noyau de transition R et donc des paramètres locaux du modèle. Ceci nous fournit une stratégie d'estimation en les remplaçant par leur version estimée.

Pour ce modèle, les durées inter-sauts $(S_n)_{1 \leq n \leq N}$ sont indépendantes et de même loi exponentielle de paramètre λ . De plus, les fractions survivantes de population $(Z_n/Z_n^-)_{1 \leq n \leq N}$ sont indépendantes, indépendantes des durées inter-sauts, et de même loi notée γ . Par conséquent, l'estimation des paramètres ne présente aucune difficulté : γ peut être estimée par une méthode à noyau, alors que le taux λ de la loi exponentielle est aussi l'inverse de sa moyenne et peut donc être estimé par l'inverse de la moyenne empirique des durées inter-sauts. On obtient ainsi un estimateur de l'opérateur K qui nous permet d'estimer à la fois $\eta_n(x)$ et $\rho(x)$. Tout l'enjeu est alors de montrer que les propriétés de convergence des estimateurs de γ et λ se reportent convenablement sur l'estimateur de K de telle sorte que les estimateurs des quantités d'intérêt vérifient les propriétés souhaitées de convergence. C'est ce qui est montré dans [2].

Le même modèle est considéré dans [13] mais avec une approche paramétrique : on suppose que γ est la densité de la loi bêta $\mathcal{B}(\alpha, 1)$, $\alpha > 0$, i.e. $\gamma(x) \propto x^{\alpha-1}$ pour tout $0 < x < 1$. Dans ce cas, dès lors que $\lambda < r\alpha$, on obtient une forme explicite pour $\eta(x)$,

$$\eta(x) = \frac{\Gamma(\alpha)}{\Gamma(\frac{\lambda}{r})\Gamma(\alpha - \frac{\lambda}{r} + 1)} x^{\frac{\lambda}{r} - \alpha} \mathcal{G}\left(\alpha - \frac{\lambda}{r}, 1 - \frac{\lambda}{r}, \alpha - \frac{\lambda}{r} + 1, \frac{1}{x}\right), \quad (8)$$

où Γ est la fonction Gamma d'Euler et \mathcal{G} est la fonction hypergéométrique de Gauss. Le paramètre α est facile à estimer : l'espérance de γ est $\alpha/(\alpha + 1)$, donc $\hat{\alpha}_N = \bar{\gamma}_N/(1 - \bar{\gamma}_N)$ (où $\bar{\gamma}_N$ désigne la moyenne empirique des $(Z_n/Z_n^-)_{1 \leq n \leq N}$) estime α . En remplaçant dans (8) α et λ par leur estimateur, on obtient aisément un estimateur de $\eta(x)$.

6.2 Approche Monte-Carlo

On reprend le modèle de pharmacocinétique [4] déjà présenté plus haut dans un cadre un peu simplifié. On décrit des expositions à un contaminant alimentaire par un processus de renouvellement $(N_t)_{t \in \mathbf{R}_+}$ (dont on note $(T_n)_{n \in \mathbf{N}}$ la suite des instants de saut qui correspondent aux instants d'exposition) et une suite $(U_n)_{n \in \mathbf{N}}$ de variables aléatoires positives indépendantes et de même loi (où U_n est la quantité de contaminant lors de l'exposition à l'instant T_n). On suppose de plus que, entre les instants T_n et T_{n+1} , le contaminant est éliminé par l'organisme à vitesse $r(\cdot|\theta_n)$ où θ_n est un paramètre du métabolisme, constant sur la fenêtre temporelle considérée et issu d'une suite de variables aléatoires indépendantes et identiquement distribuées. Alors la quantité X_t de contaminant dans l'organisme à l'instant t s'écrit

$$X_t = X_0 + \sum_{n=1}^{N_t} U_n - \sum_{n=1}^{N_t+1} \int_{T_{n-1}}^{\min(T_n, t)} r(X_s|\theta_n) ds.$$

Il est intéressant de remarquer que ce processus est déterministe par morceaux mais ne vérifie pas la propriété de Markov. Toutefois, en notant $\Theta_t = \sum_{n \in \mathbf{N}} \theta_n \mathbf{1}\{T_n \leq t < T_{n+1}\}$ et $A_t = t - T_{N_t}$, le processus de dimension 3 $(X_t, \Theta_t, A_t)_{t \in \mathbf{R}_+}$ est lui bien markovien. L'aléa intervient de manière indépendante et identiquement distribuée : l'estimation des lois sous-jacentes (des durées inter-sauts de $(N_t)_{t \in \mathbf{R}_+}$, de $(U_n)_{n \in \mathbf{N}}$ et de $(\theta_n)_{n \in \mathbf{N}}$) ne présente donc aucune difficulté, par exemple par des méthodes à noyau.

Dans le cadre de ce modèle, on regarde des caractéristiques de l'exposition au contaminant intéressantes pour les toxicologues, notamment :

- l'exposition maximale, $\sup_{0 \leq t \leq T} X_t$;
- l'exposition moyenne, $\frac{1}{T} \int_0^T X_t dt$;
- le temps passé au-dessus d'un seuil critique u , $\frac{1}{T} \int_0^T \mathbf{1}\{X_t \geq u\} dt$;
- le temps de premier passage au-dessus de ce seuil, $\inf\{t \geq 0 : X_t \geq u\}$.

On voudrait estimer la valeur moyenne de ces quatre caractéristiques. Dans ce but, commençons par supposer qu'on connaît exactement la loi du processus. On peut alors mettre en place une méthode de Monte-Carlo en générant autant de trajectoires $(X_t^{(j)})_{t \in \mathbf{R}_+}$, $1 \leq j \leq J$, que nécessaire pour estimer par leur équivalent empirique les quantités recherchées. Par exemple,

$$\frac{1}{JT} \sum_{j=1}^J \int_0^T \mathbf{1}\{X_t^{(j)} \geq u\} dt \tag{9}$$

estime l'espérance du temps passé au-dessus du seuil u .

Dans le cas évidemment plus réaliste où les paramètres du modèle sont inconnus, on ne peut plus procéder ainsi. L'approche considérée dans [4] consiste à générer des trajectoires $(\widehat{X}_t^{(j)})_{t \in \mathbf{R}_+}$, $1 \leq j \leq J$, à partir des lois estimées et à construire les estimateurs empiriques (comme (9) ci-dessus) à partir de celles-ci. Les auteurs montrent de plus que, sous des conditions de régularité raisonnables, on obtient des estimateurs consistants des quantités cibles.

Références

- [1] AZAÏS, R. A recursive nonparametric estimator for the transition kernel of a piecewise-deterministic markov process. *ESAIM : PS 18* (2014), 726–749.
- [2] AZAÏS, R., AND GENADOT, A. Semi-parametric inference for the absorption features of a growth-fragmentation model. *TEST 24*, 2 (2015), 341–360.
- [3] AZAÏS, R., AND MULLER-GUEUDIN, A. Optimal choice among a class of nonparametric estimators of the jump rate for piecewise-deterministic Markov processes. *Electron. J. Statist.* 10, 2 (2016), 3648–3692.
- [4] BERTAIL, P., CLÉMENÇON, S., AND TRESSOU, J. A storage model with random release rate for modeling exposure to food contaminants. *Mathematical Biosciences and Engineering* 5, 1 (2008), 35–60.
- [5] CLOEZ, B., DESSALLES, R., GENADOT, A., MALRIEU, F., MARGUET, A., AND YVINEC, R. Probabilistic and piecewise deterministic models in biology. *ESAIM : Procs 60* (2017), 225–245.
- [6] DAVIS, M. Piecewise-deterministic markov processes : A general class of non-diffusion stochastic models. *Journal of the Royal Statistical Society. Series B (Methodological)* 46, 3 (1984), 353–388.
- [7] DAVIS, M. *Markov Models and Optimization*. Monographs on statistics and applied probability. Chapman & Hall, 1993.
- [8] DELMAS, J., AND JOURDAIN, B. *Modèles aléatoires : Applications aux sciences de l'ingénieur et du vivant*. Mathématiques et Applications. Springer Berlin Heidelberg, 2006.
- [9] DOUMIC, M., HOFFMANN, M., KRELL, N., AND ROBERT, L. Statistical estimation of a growth-fragmentation model observed on a genealogical tree. *Bernoulli* 21, 3 (2015), 1760 – 1799.
- [10] DUFLO, M., AND WILSON, S. S. *Random Iterative Models*, 1st ed. Springer-Verlag, Berlin, Heidelberg, 1997.
- [11] FOATA, D., AND FUCHS, A. *Calcul des probabilités : cours, exercices et problèmes corrigés*. Sciences sup. Dunod, 2003.

- [12] FUJII, T. Nonparametric estimation for a class of piecewise-deterministic Markov processes. *Journal of Applied Probability* 50, 4 (2013), 931–942.
- [13] HENSHAW, K., RAMIREZ, J. M., FLORES-CONTRÓ, J. M., THOMANN, E. A., LOKE, S.-H., AND CONSTANTINESCU, C. On the impact of insurance on households susceptible to random proportional losses : An analysis of poverty trapping, 2023.
- [14] JACOBSEN, M. *Point Process Theory and Applications : Marked Point and Piecewise Deterministic Processes*. Probability and Its Applications. Birkhäuser Boston, 2005.
- [15] KRELL, N. Statistical estimation of jump rates for a piecewise deterministic markov processes with deterministic increasing motion and jump mechanism. *ESAIM : PS* 20 (2016), 196–216.
- [16] KRELL, N., AND SCHMISSER, É. Nonparametric estimation of jump rates for a specific class of piecewise deterministic Markov processes. *Bernoulli* 27, 4 (2021), 2362 – 2388.
- [17] LASOTA, A., MACKEY, M. C., AND TYRCHA, J. The statistical dynamics of recurrent biological events. *Journal of Mathematical Biology* 30 (1992), 775 – 800.
- [18] MEYN, S., TWEEDIE, R. L., AND GLYNN, P. W. *Markov Chains and Stochastic Stability*, 2 ed. Cambridge Mathematical Library. Cambridge University Press, 2009.
- [19] PARZEN, E. On estimation of a probability density function and mode. *The Annals of Mathematical Statistics* 33, 3 (1962), 1065–1076.
- [20] ROSENBLATT, M. Remarks on some nonparametric estimates of a density function. *The Annals of Mathematical Statistics* 27, 3 (1956), 832–837.
- [21] WOLVERTON, C., AND WAGNER, T. Asymptotically optimal discriminant functions for pattern classification. *IEEE Transactions on Information Theory* 15, 2 (1969), 258–265.
- [22] YVINEC, R. Piecewise deterministic Markov processes, applications in biology. Lecture, 2015.